TECHNICAL UNIVERSITY OF CRETE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING



Multi-Subject fMRI processing via clustering techniques

by

Margarita Antonia Psychountaki

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DIPLOMA DEGREE OF

ELECTRICAL AND COMPUTER ENGINEERING

July 2021

THESIS COMMITTEE

Professor Athanasios P. Liavas, *Thesis Supervisor*
Professor George N. Karystinos
Professor Michalis Zervakis

Abstract

Functional magnetic resonance imaging (fMRI) is a noninvasive method which provides significant insight into brain functionality. It measures brain activity by detecting changes associated with blood flow. The most common form of fMRI measures the blood-oxygen-level dependent (BOLD) signal, which is created by changes in blood flow in the brain. Since brain activity is intrinsic, any brain region will have spontaneous fluctuations in BOLD signal. Thus, even at rest, the BOLD signal reflects systematic fluctuations in the regional brain activity. It is widely believed that resting-state networks are the cause of these systematic fluctuations. Many studies acclaim that these networks are common, spatially, across subjects and do not overlap each other. That is, there is a common brain parcellation across subjects, based on functionality.

In this thesis, we study a method where the authors compute a common (over many subjects) brain parcellation, with no prior knowledge on the properties of the parcels. At first, the proposed method uses the subjects' resting state fMRI data and computes an orthonormal basis for the subjects' common spatial subspace via Generalized Canonical Correlation Analysis (gCCA). Then, an estimate of the common whole-brain parcellation is obtained as the solution of a semi-orthogonal nonnegative matrix factorization problem. We test this approach on synthetic and real data. Further, we study the preprocessing steps which are applied to raw fMRI data.

Acknowledgements

First of all, I would like to thank my thesis supervisor, Professor Athanasios Liavas, for his continuous guidance throughout this work. Also, my friends and colleagues P. Karakasis, J.Papagiannakos, N. Siaminou and C. Kolomvakis for all the help they provided during my thesis. Finally, I would like to thank my family and my close friends for their support and encouragement throughout my study years.

Table of Contents

Acknowledgements . . . 5

Table of Contents . . . 7

List of Abbreviations . . . 9

1 Introduction . . . 11
- 1.1 Related work . . . 11
- 1.2 Contribution . . . 12
- 1.3 Notation . . . 12
- 1.4 Outline . . . 12

2 Background . . . 15
- 2.1 What is the fMRI method ? . . . 15
- 2.2 MRI Physics and BOLD Imaging . . . 15
- 2.3 The Scanning Session . . . 17
- 2.4 Modeling the BOLD Response . . . 19
- 2.5 Experimental Design . . . 19

3 Implementation . . . 21
- 3.1 Data Preprocessing . . . 21
  - 3.1.1 Realignment . . . 21
  - 3.1.2 Slice-time correction . . . 28
  - 3.1.3 Co-registration . . . 30
  - 3.1.4 Segmentation . . . 32
  - 3.1.5 Spatial Normalization . . . 35
  - 3.1.6 Spatial Smoothing . . . 36
- 3.2 Brain parcellation . . . 36
  - 3.2.1 Estimating Common Spatial Subspace via gCCA . . . 37
  - 3.2.2 Estimating the numbers of clusters . . . 40
  - 3.2.3 Estimation of the common spatial factor $\mathbf{A}$ and temporal components $\mathbf{S}_k$ . . . 41

4 Evaluation . . . 49
- 4.1 Synthetic Data . . . 49
  - 4.1.1 Experimental evaluation of gap criterion . . . 50

4.1.2 Experimental evaluation of the estimation of the common spatial factor A . . . 57
4.2 Real Data . . . 75
4.2.1 Estimation of common subspace dimension . . . 75
4.2.2 Estimation of the common spatial factor A . . . 76

5 Conclusion and Future Work . . . 85

Appendices . . . 87

A Synthetic Data . . . 89
A.1 Correlation coefficients of temporal components $\mathbf{S}_k$ . . . 89
A.2 Figures for experimental evaluation of gap criterion . . . 93
A.2.1 Case 1: Common orthonormal matrix $\mathbf{A}$ with each subject having independent temporal components $\mathbf{S}_k$. . . . 93
A.2.2 Case 2: Common orthonormal matrix $\mathbf{A}$ with each subject having correlated temporal components $\mathbf{S}_k$. . . . 105
A.2.3 Case 3: Corrupted common orthonormal matrix $\mathbf{A}$ with each subject having correlated temporal components $\mathbf{S}_k$. . . . 111
A.3 Figures for experimental evaluation of the estimation of the common spatial factor A . . . 117
A.3.1 Case 1: Common orthonormal matrix $\mathbf{A}$ with each subject having independent temporal components $\mathbf{S}_k$. . . . 117
A.3.2 Case 2: Common orthonormal matrix $\mathbf{A}$ with each subject having correlated temporal components $\mathbf{S}_k$. . . . 235
A.3.3 Case 3: Corrupted common orthonormal matrix $\mathbf{A}$ with each subject having correlated temporal components $\mathbf{S}_k$. . . . 289

B Real Data . . . 343
B.1 Estimation of common subspace dimension . . . 344
B.2 Estimation of the common spatial factor A . . . 350
B.3 Clusters for common subspace dimension equal to 27 . . . 374
B.4 Clusters for common subspace dimension equal to 42 . . . 482
B.5 Correlation coefficients of temporal components S_k, $R = 27$ . . . 651
B.6 Correlation coefficients of temporal components S_k, $R = 42$ . . . 656
B.7 Comparison of estimated A matrices for common subspace dimension 27 and 42 . . . 660

Bibliography . . . 675

List of Abbreviations

ALS	Alternating Least Squares
r-fMRI	Resting state Functional Magnetic Resonance Imaging
fMRI	Functional Magnetic Resonance Imaging
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
CT	Computed Tomography
DCT	(Discrete Cosine Transformation)
CSF	(Cerebrospinal fluid)

Chapter 1

Introduction

Functional magnetic resonance imaging (fMRI) is a noninvasive method which provides significant insight into brain functionality. It measures brain activity by detecting changes associated with blood flow. Data associated with brain activity are collected during well designed sessions, where the subject can either execute tasks or lie without doing anything. The experimental design of each session is based on the purposes for which the study is conducted. Many studies attempt to provide a statistical inference on a whole-brain basis to describe brain responses in terms of spatial and temporal patterns by applying a wide range of unsupervised multivariate statistical methods on fMRI data.

The fMRI data that are collected when the subject is resting are commonly used by neuroscientists in order to analyze and interpret human brain activity. A question which is worth researching is if there are any similarities in subjects' brain activity when they are resting. In this thesis, we examine this problem by studying the methodology, which is proposed at [1]. At first, authors find a common brain parcellation over different subjects. These parcels do not overlap each other and no prior knowledge on their properties is required. More specifically, the proposed method uses subjects' resting state fMRI data to compute an orthonormal basis for the subjects' common spatial subspace, $\mathbf{G}$, via Generalized Canonical Correlation Analysis (gCCA). Then, they solve a semi-orthogonal nonnegative matrix factorization problem to obtain the common spatial factor $\mathbf{A}$ from $\mathbf{G}$, such that $\mathbf{A}^T\mathbf{A} = \mathbf{I}$ and $\mathbf{A} \geq 0$. Because of its properties, $\mathbf{A}$ practically defines the regions which are included in each parcel.

1.1 Related work

Generally, gCCA is commonly used in fMRI studies in order for a common spatial component over subjects to be found. Also, there is an extensive overview of fMRI clustering methods and the problem of data driven brain parcellation in [2], which refers to task related fMRI data (the subjects execute tasks in the duration of session). Another approach of the problem is provided at [3], where the authors solve the problem only for cortex areas. They sample some regions of interest on the brain and calculate the correlation between these regions and the fMRI time series at each spatial location, constructing a matrix for each subject. They binarize these matrices by keeping the top 10% of the correlations and average the binarized matrices. Then, they apply a clustering algorithm. In order to estimate the number of clusters, they apply a stability analysis, which suggests 7 and 17 parcels, for which they examine the results of clustering.

1.2 Contribution

In this thesis, we examined the problem of brain parcellation by studying the method which is proposed in [1]. Firstly, we studied the preprocessing steps, which are applied to raw fMRI data so that the fMRI data can be used for analysis. Namely, we comprehended which problems are solved by pre-processing steps, how these problems are modeled and how they are resolved in order to be reassured that the initial assumptions of [1] are not violated because of the data preprocessing. We investigated the behavior of the algorithms, which authors used, on synthetic and real data. We used three different synthetic data formats trying to model them as close as we could to the real data. For the synthetic data, we examined the behavior of rank criterion, based on which the authors decided the number of parcels, thoroughly. We also applied several algorithms to find $\mathbf{A}$, two of which will be presented here. The second algorithm is the method suggested by the authors. As for the real data, we excluded a small part of it, which was included to the initial work. We also studied the rank criterion for the case of real data. In this case, it was not clear which number of cluster should be used. Therefore, we used a threshold based on the behavior of the criterion for synthetic data and we also applied trimming to exclude outlier values in order to have a clearer insight. We decided to examine the brain parcellation for two numbers of parcels 27 and 42. Finally, we compared the two solutions.

1.3 Notation

In this work, scalars are denoted by a small letter, for example x, vectors are denoted by a small bold letter, for example $\mathbf{x}$ and matrices are denoted by a capital bold letter, for example $\mathbf{X}$. Also, $\mathbf{X}^T$ denotes the transpose matrix of $\mathbf{X}$ and $\mathbf{X} \geq \mathbf{0}$ denotes a matrix with non-negative elements. $\mathbb{R}$ denotes the set of real numbers, while $\mathbb{R}^{N\times M}$ denotes the set of $N \times M$ real matrices and $\mathbb{R}_{+}^{N\times M}$ denotes the set of $N \times M$ matrices, which have non-negative elements. Moreover, $\mathbf{I_N}$ denotes the $N \times N$ identity matrix. $||\mathbf{X}||_2$ and $||\mathbf{X}||_F$ denotes the euclidean and frobenius norm of matrix $\mathbf{X}$, respectively. Finally, $\mathbf{P}_S$ denotes the orthogonal projection onto a linear subspace S.

1.4 Outline

Chapter 2 provides useful information about the fMRI method. Namely, in this chapter, information about what fmri method is, how it works and details about the design of scanning sessions are provided.

Chapter 3 includes the main preprocessing steps which are applied in raw fMRI data, since the raw data which are collected from scanning session are not suitable for analysis. The six basic steps of preprocessing are realignment, slice time correction, coregistration, segmentation, normalization and smoothing. It also describes the methods which are applied in order to estimate the common brain parcellation over all subjects.

Chapter 4 presents the experimental results of the methods which are used for the brain parcellation, both on synthetic and real data.

Chapter 5 gives some closing remarks along with interesting directions for the future.

Chapter 2

Background

2.1 What is the fMRI method ?

Functional magnetic resonance imaging (fMRI) is a noninvasive method, which studies brain function and dominates brain mapping research. More specifically, fMRI is a technique which uses magnetic resonance imaging to measure brain activity based on changes in blood oxygenation and flow. The most common form of fMRI measures the blood-oxygen-level dependent (BOLD) signal. It takes advantage of the hypothesis that, when a set of neurons becomes active, a larger amount of blood rich to oxygen flows in this brain region. The signal which fMRI measures depends on these changes of oxygenation. The use of sophisticated techniques from signal and image processing and statistics are required for fMRI analysis. The output of these procedures is a map showing which brain regions responded to some particular manipulation of mental or perceptual functions. According to Ashby [4], the information which is provided by fMRI is useful for studying neural and congnitive architecture, but not for answering questions about the process. That lies on the fact that fMRI has poor temporal and limited spatial resolution. Because of its poor temporal resolution, fMRI can not identify small timing differences between different processes or cognitive stages reliably. On the other hand, spatial resolution of fMRI is adequate to localize brain activity at the level of main brain structure, but not at a smaller level as cortical column.

2.2 MRI Physics and BOLD Imaging

Magnetic resonance imaging (MRI) is a medical imaging technique which forms pictures of the anatomy and the physiological processes of the body.

The magnetic resonance (MR) scanner uses superconducting electromagnets to produce a static, uniform magnetic field of high strength. Some years ago, the standard field strength used in fMRI research was 1.5 Tesla (T). Some research centers have scanners considerably stronger than this (e.g., above 10 T).

The first step to produce and measure MR signal is to place the patient in a strong, static, uniform magnetic field, B_0. The nucleus of all the atoms in a patient's body are randomly oriented, but when the patient enters the strong static, uniform field, most of them align along or opposite the magnetic field and they start to precess around the direction of B_0. Despite the parallel alignment, the nuclei are spinning out of phase. The velocity of rotation around the field direction is the Larmor frequency. The average of magnetic dipole moment density is called net magnetization vector.

The second step is the application of a second radiofrequency (RF) magnetic field B_1, which is applied vertically to B_0. If the RF pulse's frequency is equal to Larmor frequency, then resonance occurs. This causes excitation of nucleus. Namely, turning the pulse on, the nuclei spin in-phase and align their magnetization along with the transversal plane. The absorption of RF pulses' energy by the nuclei causes a transition from higher to lower energy levels.

The third step is relaxation. Turning off the RF pulses, the nucleus relaxes to its original equilibrium alignment. Two separate processes take place during relaxation: T1 relaxation and T2 relaxation. The T1 relaxation is the process by which the net magnetization returns to its initial maximum value parallel to B_0. The T1 relaxation time is defined as the time needed to achieve 63% of the original longitudinal magnetization. The T2 and T2* relaxation describe what happens in transerval plane. When pulse is turned off the nucleus dephases, because of inhomogeneities of the main magnetic field (T2* decays) and spin-spin interactions (T2 decays). These inhomogeneities may be the result of intrinsic defects in the magnet itself or from susceptibility-induced field distortions produced by the tissue or other materials placed within the field. T2 relaxation time is defined as the time needed to dephase up to 37% of the original value. When the nucleus relaxes to its original equilibrium alignment, energy is released (transition from lower to higher energy levels), which is detected by the coils as the raw MR signal. The energy absorbed (and subsequently emitted) by the nuclei, induces a voltage that can be detected by a suitably tuned coil of wire, amplified, and displayed as the free-induction decay (FID).

Spatial resolution is provided by gradients, which are additional spatially linear variations in the static field strength. Gradients can be applied in any orthogonal direction using the three sets of gradient coils, Gx, Gy, and Gz, within the MR system. These slight disorders of static magnetic field in a predictable pattern cause the resonance frequency of nucleus to vary as a function of position. Moreover, faster or slower precession is detected as higher or lower MR signal. So, frequency measurements allow us to distinguish MR signals at different positions in space and reconstruct image in three dimensions. The software that controls all these magnetic fields is typically called "the pulse sequence".

The pulse sequence is run on the main computer by which the scanner is controlled. In most fMRI experiments, the stimuli which are presented to the subject are generated by a second computer which also records the subjects behavioral responses. Both computers have to be synchronized, in a way that the onset of each stimulus presentation occurs at a precisely controlled moment during image acquisition.

Two types of pulse sequences are commonly used, depending on whether the goal is structural or functional imaging. Structural MR is normally used to measure the density of water molecules, which differs, for example, in bone, gray matter, cerebrospinal fluid, and tumors. The vast majority of functional MRI experiments measure the blood oxygen-level dependent (BOLD) signal.

BOLD contrast results from changing regional blood concentrations of oxy- and deoxyhemoglobin. Hemoglobin in blood carries oxygen from the lungs or gills to the rest of the

body (i.e. the tissues). Oxyhemoglobin is formed during physiological respiration when oxygen binds to the heme component of the protein hemoglobin in red blood cells. Oxyhemoglobin has no unpaired electrons and is weakly diamagnetic. When oxygen is released, hemoglobin is without the bound oxygen and this form of hemoglobin is called deoxyhemoglobin and it is strongly paramagnetic. The presence of paramagnetic deoxyhemoglobin creates local magnetic field distortions in and around blood vessels. So, the regional T2 and T2* relaxation times of brain decrease as the fraction of deoxyhemoglobin increases. Shorter T2 and T2* relaxation times lead to smaller/weaker signals in regions which contain deoxyhemoglobin, in comparison with brain areas with more oxyhemoglobin, which have higher signal (and appear brighter).

The BOLD fMRI technique basically measures changes in the inhomogeneity of the magnetic field which are a result of changes in the level of oxygen present in the blood. It takes advantage of the assumption that, when a neuron becomes active, a larger amount of blood rich in oxygen flows in these brain regions. So, activation can be detected. The following figure illustrates the fact that, 6 seconds after the termination of neural activation, BOLD signal rises well above the baseline (increase of deoxyhemoglobin) and decays back over 20-25 seconds.

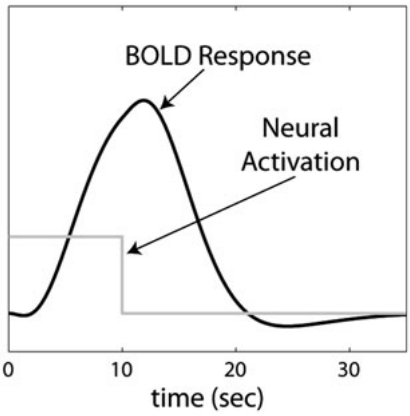


Figure 2.1: Illustration of the hypothetical BOLD response (black curve) to a constant 10 s neural activation (gray curve) [4].

2.3 The Scanning Session

An experimental session that collects fMRI data includes at least four different types of scans [4].

The first one to be completed in each session is the localizer. It is a quick structural scan ($1-2$ min) of low spatial resolution and is used only to locate the subject's brain in 3-dimensional space, in order to provide the knowledge required to optimize the location of the slices that will be taken through the brain in the high-resolution structural scan and in the functional scans that follow.

The second type of scan usually is the high-resolution structural scan. The spatial resolution and the pulse sequence which is used, define the time duration of session, which can last 8-10 minutes. The structural scan is crucial for fMRI data. The fMRI scanning

needs high temporal resolution, which leads to poor spatial resolution. The high resolution of structural images helps to counteract this fact, by aligning structural data with functional at preprocessing stage. This mapping allows the examination of observed spatial coordinates of activation during fMRI via aligned coordinates in the structural image.

The third type of scan usually is fMRI scan. It can range from 20−30 min to complete (in long run), or it can be broken down into 2 or 3 shorter runs, with brief rests in between. There are many parameter choices to make here, two of which are highly important for the subsequent fMRI data analysis.

One choice is repetition time, TR, which is the time between successive whole brain scans. Typical TRs range from 2 to 3 s, if the whole brain is scanned. However, TRs as low as 1 s are possible on many machines, especially if some parts of the brain are excluded from the scanning.

The second important choice is voxel size, which determines the spatial resolution of the functional data. The patient's brain occupies a certain volume. If we assign a coordinate system to the bore of the magnet, then we could identify any point in the subject's brain by a set of three coordinate values (x, y, z). By convention, the z direction runs down the length of the bore (from the feet to the head), and the x and y directions reference the plane that is created by taking a perpendicular cut to the z axis. The brain, of course, is a continuous medium, in the sense that neurons exist at (almost) every set of coordinate values inside the brain. FMRI data, however, are discrete. The analog-to-digital conversion is performed by dividing the brain into a set of cubes (or more accurately, rectangular right prisms). These cubes are called voxels. A typical voxel size might be 3 mm $\times$ 3 mm $\times$ 3.5 mm. In this case, in a typical human brain, 33 separate slices might be acquired each containing a 64 $\times$ 64 array of voxels for a whole brain total of 135.168 voxels. In each fMRI run, a BOLD response is recorded every TR seconds in each voxel. Thus, for example, in a 30 min run with a TR of 2 s, 135.168 BOLD responses could be recorded 900 separate time instants (i.e., 30 time instants per minute $\times$ 30 min), for a total of 121.651.200 BOLD values. This is an immense amount of data, and its sheer volume greatly contributes to the difficulties in data analysis.

Many studies stop after completion of the functional data acquisition, but some other types of scans are also common. A fourth common type of scan is the field map. We make the assumption that a scanner has a completely uniform magnetic field across its entire bore. Even if this were true, placing a human subject inside the bore distorts this field to some extent. These inhomogeneities in the magnetic field are corrected via a process known as shimming. If shimming is successful, the magnetic field will be uniform at the beginning of scanning. However, especially in less reliable machines, sometimes the distortions in the magnetic field will appear in the middle of the session. The field map, which takes only a minute or two to be collected, measures the homogeneity of the magnetic field at the moment when the map is created. Thus, the field map can be used during later data analysis to correct any possible nonlinear distortions in the strength of the magnetic field that develop during the course of the scanning session.

2.4 Modeling the BOLD Response

It is necessary to determine a model for how the neural activity and BOLD response are related, in order for us to be able to make inferences about neural activity by observing the BOLD response. Most fMRI applications make the assumption that this transformation from neural activation to BOLD response can be modeled as a linear and time-invariant system. As we know, a system is linear and time-invariant system if and only if the superposition principle is satisfied. By making the assumption that the superposition principle holds, we can determine the BOLD response to any neural activation from the results of one simple experiment. All we need to do, is to determine the impulse response function, $h(t)$, which in case of fMRI is known as hemodynamic response function (HRF). Now if we know $h(t)$ for any neural activation $N(t)$, we can calculate the BOLD response $B(t)$ to this activation by:

$$B(t) = \int_0^t N(\tau)h(t-\tau)d\tau. \tag{2.1}$$

The HRF plays an essential role for the analysis of fMRI data. The most common method to determine the HRF is the use of a specific mathematical function, based on how we expect the HRF to look like. The gamma function or the difference of two gamma functions are the most commonly used functions for determining HRF. In Figure(2.2) we illustrate a HRF, which was produced by SPM12 as the difference of two gamma functions.

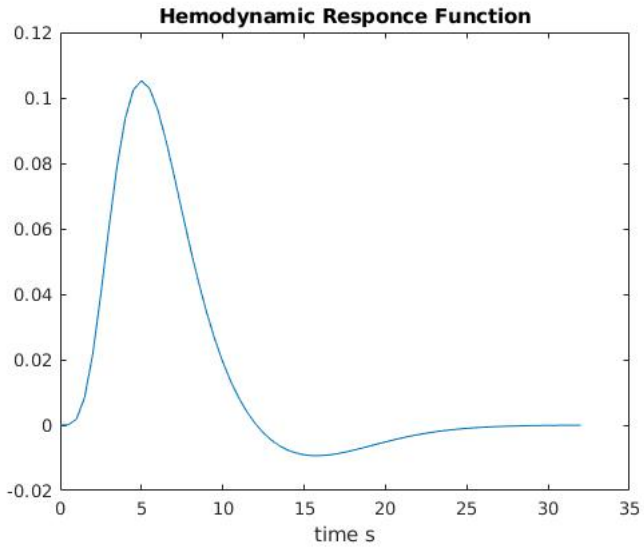


Figure 2.2: Illustration of the hemodynamic response function. We used the SPM12 to create HRF and illustrate it.

2.5 Experimental Design

There are three main experimental designs, which are used for fMRI experiments, block design, event-related design and free-running design. This classification is based on the way stimuli are presented to the subject.

In a block design, the functional run consists of a series of on-off blocks. Each block

may last from 20 s to a couple of minutes. Within an "on" block, subjects perform the same cognitive, perceptual, or motor task from the beginning of the block until the end. In a block design, two or more conditions are alternated in blocks, but in each block one stimulus is present. Within an "off" block, subjects rest. The main advantage of this technique is that it offers considerable statistical power. This happens because the amplitude of the HRF increases when multiple stimuli are presented in rapid succession. Another condition which is needed, in order for considerable statistical power to be succeeded, is to have the maximum amount of variability in the signal. This can happen if "on-off" blocks are alternated so that the HRF has enough time to return to baseline. However, in block design, the randomization of stimuli type is not possible within a block.

In an event-related design, the experiment consists of discrete trials. Each trial is usually one type of several stimuli. Each type of stimulus is repeated at least 20 times. The order of trials in duration of experiment is random. Event-related experiments allow greater flexibility and randomization than the block design. So, experiments are cleverer and less predictable. However, this method deals with lower signal-to-noise ratio (SNR), because the signal change in the BOLD fMRI signal which corresponds to a single stimulus presentation is small.

In a free-running design, events are presented to the subject continuously. For instance, subjects watch a movie or simply lie passively. These events are more natural than the above designs, but the result is more challenging to be analyzed.

Chapter 3

Implementation

3.1 Data Preprocessing

The aim of this chapter is to present the basic pre-processing steps. The pre-processing steps are applied to fMRI data independently of their experimental design in order to be suitable for analysis. Therefore, in this chapter, we present the purpose of each pre-processing step, the way the problems dealt by pre-proseccing steps are modeled and finally solved [5], [6], [7], [8],[9], [10]. We mainly focus on the pre-processing steps as they are implemented by SPM toolbox, which was used for the preprocessing of the real data.

3.1.1 Realignment

As mentioned above, the product of an fMRI session is a sequence of N 3D images, where N is the number of time points. In other words, the product is a 3D movie. However, in the main part of analysis of fMRI the data are represented as a 2D matrix. One of the dimensions is time points and the other is voxels. As it is obvious, a main assumption for the data is that the signal of each time course corresponds to only one voxel. The motion of subject's head during an fMRI session primarily offends against this assumption. This motion leads to misregistration among the images. Namely, the voxels of images do not correspond to the same part of real brain. Figure 3.1 illustrates it clearly. Moreover, the signal changes occurred by the head's motion can be comparable with any haemodynamic response. So, if the motion is synchronized with the task event, it might lead to false detection of activation. The complete elimination of head's movement during scanning is not possible. However, there are techniques which alleviate this artifact.

The answer to the problem of the misregistration of images is realignment. One image is chosen as the reference image. The realignment tries to estimate the best rotation and translation parameters in order to align each other image with the reference image. After that, the re-slicing, in which the estimated parameters are applied to the images, follows and then the data are re-sampled in the grid of the reference image.

Image registration looks for an optimal geometric transformation, which brings the images to a common coordinate system. In this case, the geometric transformation is a rigid body transformation. Rigid body transformations include rotation and translation, but they do not warp the image.

Translation: The shifting of a point, which belongs to $\mathbb{R}^3$, $\mathbf{x} = [x_1 x_2 x_3]^T$, at another point $\mathbf{x}' = [x_1' x_2' x_3']^T$, is equivalent to the addition of the desired amount to each axis. In

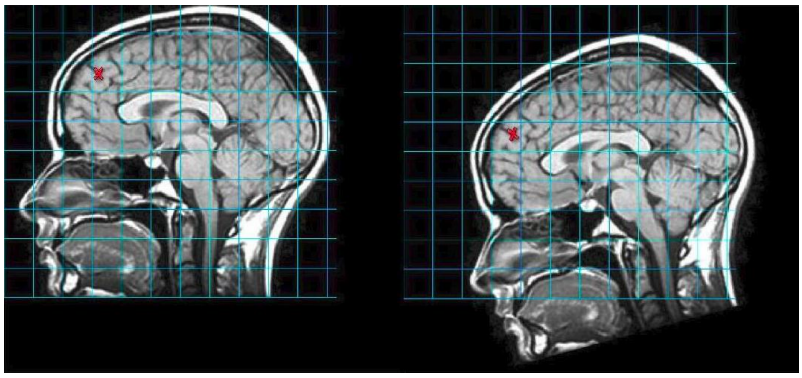


Figure 3.1: The misregistration between the two images is obvious. As the subject moves its head, although the information is found in the same part of the brain, it is registered in different locations in the pictures.

matrix form, we have:

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & q_1 \\ 0 & 1 & 0 & q_2 \\ 0 & 0 & 1 & q_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}.$$

Rotation: In the 3D case, the rotation can break down in three matrices.
A rotation of θ_1 radians around the first x axis is called a pitch, and its rotation matrix is

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_1) & \sin(\theta_1) & 0 \\ 0 & -\sin(\theta_1) & \cos(\theta_1) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

A rotation of θ_2 radians around the second y axis is called a roll, and its rotation matrix is:

$$\mathbf{R}_y = \begin{bmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

A rotation of θ_3 radians around the third z axis is called a yaw, and its rotation matrix is:

$$\mathbf{R}_z = \begin{bmatrix} \cos(\theta_3) & \sin(\theta_3) & 0 & 0 \\ -\sin(\theta_3) & \cos(\theta_3) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The rotation is performed as:

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ 1 \end{bmatrix} = \mathbf{R}_i \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix},$$

where $i = x,\ y,\ z$.

The combination of rotations can be applied by multiplication of $\mathbf{R}$ matrices in the correct order. If the order of multiplications changes, then the result will be different.

The estimation of these six parameters is necessary for the registration of the images. One way to parameterize the rigid body transformation matrix in terms of six parameters ($q_1,\ q_2,\ q_3,\ \theta_1,\ \theta_2,\ \theta_3$) is to express the transformation matrix $\mathbf{M}$ as follow:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & q_1 \\ 0 & 1 & 0 & q_2 \\ 0 & 0 & 1 & q_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \mathbf{R}_x \times \mathbf{R}_y \times \mathbf{R}_z =$$

$$\begin{bmatrix} \cos(\theta_3)\cos(\theta_2) & \cos(\theta_2)\sin(\theta_3) & \sin(\theta_2) & q_1 \\ -\sin(\theta_1)\sin(\theta_2)\cos(\theta_3) - \cos(\theta_1)\sin(\theta_3) & -\sin(\theta_1)\sin(\theta_2)\sin(\theta_3) + \cos(\theta_1)\cos(\theta_3) & \sin(\theta_1)\cos(\theta_2) & q_2 \\ -\cos(\theta_1)\sin(\theta_2)\cos(\theta_3) + \sin(\theta_1)\sin(\theta_3 & -\cos(\theta_1)\sin(\theta_2)\sin(\theta_3) - \sin(\theta_1)\cos(\theta_3) & \cos(\theta_1)\cos(\theta_2) & q_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We can extract the value of parameters as follows:

- $q_1 = m_{1,4}$
- $q_2 = m_{2,4}$
- $q_3 = m_{3,4}$
- $m_{1,3} = \sin(\theta_2) \Rightarrow \theta_2 = \sin^{-1}(m_{1,3})$
- $m_{2,3} = \sin(\theta_1)\cos(\theta_2) \Rightarrow \sin(\theta_1) = \frac{m_{2,3}}{\cos(\theta_2)}$
- $m_{3,3} = \cos(\theta_1)\cos(\theta_2) \Rightarrow \cos(\theta_1) = \frac{m_{3,3}}{\cos(\theta_2)}$
- $\theta_1 = \text{atan2}(\frac{m_{2,3}}{\cos(\theta_2)}, \frac{m_{3,3}}{\cos(\theta_2)})$
- $m_{1,1} = \cos(\theta_3)\cos(\theta_2) \Rightarrow \cos(\theta_3) = \frac{m_{1,1}}{\cos(\theta_2)}$
- $m_{1,2} = \sin(\theta_3)\cos(\theta_2) \Rightarrow \sin(\theta_3) = \frac{m_{1,2}}{\cos(\theta_2)}$
- $\theta_3 = \text{atan2}\left(\frac{m_{1,2}}{\cos(\theta_2)}, \frac{m_{1,1}}{\cos(\theta_2)}\right)$,

where atan2 is the four-quadrant inverse tangent.

The header of each fMRI image contains a 4×4 affine transformation matrix, which provides information about the mapping from the voxel coordinates (x_1, x_2, x_3), where the first voxel is at coordinate $(1, 1, 1)$, to coordinates in millimeters (y_1, y_2, y_3). Namely,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{bmatrix} = \begin{bmatrix} m_{1,1} & m_{1,2} & m_{1,3} & m_{1,4} \\ m_{2,1} & m_{2,2} & m_{2,3} & m_{2,4} \\ m_{3,1} & m_{3,2} & m_{3,3} & m_{3,4} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}.$$

The affine transformation, except for translation and rotation, also includes shearing and scaling. So, rigid body transformations are a subset of affine transformations.

Scaling: Scaling is used for changing the size of an image. Also, it is needed because the size of the voxels might not be equal in all dimensions or because the size of voxels can be different among images. The transformation matrix for scaling is:

$$\mathbf{Sc} = \begin{bmatrix} q_4 & 0 & 0 & 0 \\ 0 & q_5 & 0 & 0 \\ 0 & 0 & q_6 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Shearing: Shearing slants the shape of an object and its transformation matrix is:

$$\mathbf{S} = \begin{bmatrix} 1 & q_7 & q_8 & 0 \\ 0 & 1 & q_9 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The objective of rigid body transformation is to find the optimal transformation matrix, which maps the coordinate system of one image to another, in order to bring the homologous points of the two images as close as possible. Suppose two images $\mathbf{f}$, $\mathbf{g}$ with affine transformation matrices $\mathbf{M}_f$, $\mathbf{M}_g$, respectively. Then, if the optimal transformation matrix for rigid body transformation, denoted by $\mathbf{M}$, is multiplied with the transformation matrix of the image $\mathbf{f}$, the matrix $\mathbf{M}_f^{-1}\mathbf{M}^{-1}\mathbf{M}_g$ maps the voxels of image $\mathbf{g}$ to those of image $\mathbf{f}$. Namely, the $\mathbf{M}\mathbf{M}_f$ rotates and translates the $\mathbf{f}$ image. So, the $\mathbf{M}_f^{-1}\mathbf{M}^{-1}\mathbf{M}_g$ maps the coordinate system of image $\mathbf{g}$ to millimeter space and, after that, from millimeters it maps into space of rotated and translated image $\mathbf{f}$.

The most common way to estimate the optimal transformation matrix of rigid body transformation is by minimising the sum of squared differences between each of the images and a reference image. A slight improvement in the results can be achieved with the repetition of this procedure, but the reference image will be the mean of the realigned images.

So, the problem is

$$\min_{\mathbf{M}_1} \sum_{i=1}^{N} \mathbf{f}(\mathbf{M}_1 x_i) - \beta \mathbf{g}(x_i)^2,$$

where N is the number of voxels of image, $\mathbf{M}_1 = \mathbf{M}_f^{-1}\mathbf{M}^{-1}\mathbf{M}_g$, the $\mathbf{M}$ is constructed from rigid body transformation parameters, denoted by $\mathbf{q}$, $\mathbf{g}$ is the reference image, $\mathbf{f}$ the image which we want to align with $\mathbf{g}$, and β is a scaling constant.

A way to solve this problem is by using the Gauss-Newton algorithm. If we assume that the images are smooth (or that they can be rendered smooth), then we can use the first order Taylor approximation and the problem becomes:

$$\min_{\mathbf{\Delta M}_1} ||\mathbf{f}(\mathbf{M_1 x}) - \beta \mathbf{g}(\mathbf{x}) + \mathbf{J}_{\mathbf{M}_1} \mathbf{\Delta M}_1||_2^2,$$

where $\mathbf{J}_{\mathbf{M}_1}$ is the Jacobian matrix and is defined as:

$$\mathbf{J}_{\mathbf{M}_1} = \begin{bmatrix} \frac{\partial \mathbf{M}_1 x_1}{\partial q_1} & \frac{\partial \mathbf{M}_1 x_1}{\partial q_2} & \cdots & \frac{\partial \mathbf{M}_1 x_1}{\partial q_6} \\ \frac{\partial \mathbf{M}_1 x_2}{\partial q_1} & \frac{\partial \mathbf{M}_1 x_2}{\partial q_2} & \cdots & \frac{\partial \mathbf{M}_1 x_2}{\partial q_6} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \mathbf{M}_1 x_N}{\partial q_1} & \frac{\partial \mathbf{M}_1 x_N}{\partial q_2} & \cdots & \frac{\partial \mathbf{M}_1 x_N}{\partial q_6} \end{bmatrix}.$$

The Gauss-Newton is an iterative algorithm. For step n of the algorithm, the update of the parameters $\mathbf{q}$ are:

$$\mathbf{q}^n = \mathbf{q}^{n-1} - (\mathbf{J}_{\mathbf{M}_1}^T \mathbf{J}_{\mathbf{M}_1})^{-1} \mathbf{J}_{\mathbf{M}_1}^T (\mathbf{f}(\mathbf{M_1 x}) - \beta \mathbf{g}(\mathbf{x})).$$

After the estimation of the rigid body transformation matrix, we proceed to reslicing. The transformation matrix rotates and translates all images. Applying matrix $\mathbf{M}_f^{-1}\mathbf{M}_g$ (let $\mathbf{f}$ be the reference image), the coordinate system of the rotated and translated image $\mathbf{g}$ is mapped into the coordinate system of the reference image. We interpolate the value of image $\mathbf{g}$ at the new points.

Interpolation: Interpolation is a mathematical technique which estimates the values of a function at a set of points, by using the values of this function at a discrete set of points for which its values are known. There are several forms of interpolation.

The nearest neighbor: The simplest interpolation method is the nearest neighbor, in which the new data point takes the value of the nearest to it known data point. The advantage of this method is that the initial intensities of the voxels are preserved, but the image has a 'blocky' appearance.

Trilinear interpolation: Trilinear interpolation is an extension of bilinear interpolation, which in turn is an extension of linear interpolation. Linear interpolation on a set of data points (x_0, y_0), (x_1, y_1), ..., (x_n, y_n) is defined as the concatenation of linear interpolants, which are the straight line between two successive data points, for each successive pair of data points. So, from a discrete function we produce a continuous function, which is

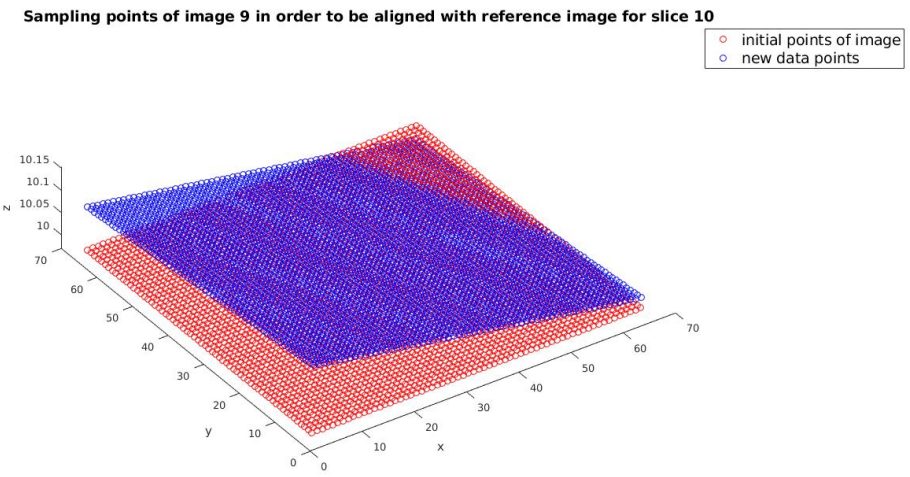


Figure 3.2: Illustration of new data points (blue spots) in which we have to interpolate the values of image in order to realign image 9 with the reference one. Red spots indicate the points for which the value of image is known.

not differentiable in general. Bilinear interpolation is the extension of linear interpolation in a function of two variables. Assume the following grid of pixels (Figure 3.3). We denote the value of a point a with coordinates (x_a, y_a) as v_a. We first apply linear interpolation in the x-direction. Namely, we estimate the value of two other points **k**, **l**. The coordinate of **k** in the x-axis is the same with point **u** and the coordinate in the y-axis is the same with points 5, 6, 7, 8. The coordinate of **l** in x-axis is the same with point **u** and the coordinate in y-axis is the same with points 9, 10, 11, 12. The values of these points are:

$$v_k = \frac{(x_6 - x_k)v_5 + (x_k - x_5)v_6}{x_6 - x_5}$$

$$v_l = \frac{(x_{10} - x_l)v_9 + (x_l - x_9)v_{10}}{x_{10} - x_9}.$$

We continue by interpolating in the y-direction to estimate the value of point **u**

$$v_u = \frac{(y_u - y_l)v_k + (y_k - y_u)v_l}{y_k - y_l}.$$

The extension of the approach to three dimensions is trivial. The result is less 'blocky', but high frequencies might be lost.

Polynomial Interpolation : We can use more voxels than the eight nearest neighbors (in 3D) to estimate the value at a point. A more precise approach uses a polynomial function instead of the straight line to connect the points in 1D. This can be extended to 3D. Given $n + 1$ distinct points $x_0, x_1, \ldots, x_n$ and corresponding values $v_0, v_1, \ldots, v_n$, there is a unique polynomial, the degree of which is a at the most n, that interpolates the data $(x_0, v_0), \ldots, (x_n, v_n)$. The nearest neighbor and trilinear interpolations are simply low-order polynomial.

Windowed Sinc Interpolation : The optimal way to reduce the artifacts of inter-

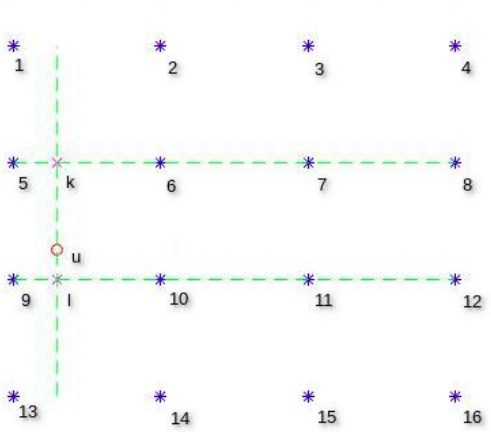


Figure 3.3: This is the grid of the two dimensional example which we examine for the bilinear interpolation. The values of data at points 1 to 16 are known. We want to estimate the value of data at point $\mathbf{u}$, v_u. The points k, l is supplementary for the estimation of value v_u.

polation when we apply a rigid body transformation is to apply it in Fourier domain. The interpolation method that gives results closest to a Fourier interpolation is the *sinc* interpolation. Actually, we convolve the image with a *sinc* function centered at the point to be resampled. Because of speed, we cannot use all voxels to sample one voxel. So, the sinc function is often truncated by modulating it with a Hanning window. Because the function is separable, the interpolation is performed sequentially in the three dimensions of the volume.

B − splines : According to [11], traditional interpolation can be expressed as:

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^q} f_{\mathbf{k}}\phi_{int}(\mathbf{x}-\mathbf{k}),$$

where $f(\mathbf{x})$ is the interpolated value at some coordinate $\mathbf{x}$ in a space of dimension q, $f_{\mathbf{k}}$ are the real values at known points at coordinates $\mathbf{k} = (k_1, k_2, \ldots, k_q) \in \mathbb{Z}^q$. Namely, the value of f at point $\mathbf{x}$ is a linear combination of sample $f_{\mathbf{k}}$. The weight of each sample is given by the $\phi_{int}(\mathbf{x}-\mathbf{k})$. For exact interpolation the $\phi_{int}(\mathbf{x}-\mathbf{k})$ at any of integer coordinate $\mathbf{k}$ should eliminate all other terms except for the origin, where its value should be 1. Sinc is a function of that kind of basis function.

Generalized interpolation is expressed as:

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^q} c_{\mathbf{k}}\phi_{int}(\mathbf{x}-\mathbf{k}).$$

In this case, the interpolation can be performed in two steps. First, we determine the values of coefficient $c_{\mathbf{k}}$ and, second, we calculate the value of f at each point $\mathbf{x}$ using the coefficients. This separation allows the basis function to have better properties than the previous case.

A family of basis function is made of B-splines β^n.

$$\beta^n(x) = \sum_{k=0}^{n+1} \frac{(-1)^k(n+1)}{(n+!-k)!k!} \left(\frac{n+1}{2} + x - k\right)_+^n,$$

for $x \in \mathbb{R}$, $n \in \mathbb{N}$ and x_+^n is given by:

$$x_+^n = \begin{cases} 0, & n = 0 \wedge x < 0, \\ \frac{1}{2}, & n = 0 \wedge x = 0, \\ 1, & n = 0 \wedge x > 0, \\ x_+^0 x^n, & n > 0. \end{cases}$$

An n^{th} degree B-spline uses a linear combination of $n + 1$ basis functions to compute an interpolated value. Moreover, a B-spline of order n is $(n - 1)$ times continuously differentiable. The nearest neighbor is an interpolation using B-spline of 0 degree and linear interpolation is an interpolation using B-spline of first order.

3.1.2 Slice-time correction

Another assumption which is made during the analysis of fMRI is that the entire 3-D volume is measured at each time point at the same time. However, the data of whole brain are not collected at once. In practice, the 3-D brain breaks into 2-D slices. As a result, the slices of the brain are collected at slightly different times. For example, if TR is 3.5 seconds (TR: the time which is needed for complete brain scanning) and the volume consists of 36 slices, the first and the last slice are collected with the difference of 3.4 seconds.

The most common way to cope with the above artifact is to shift the slices in time with respect to the reference slice. Namely, each of the slices is shifted proportionally to the temporal difference in the scan time from the reference image. In this way, all slices are measured at the same time point. As reference image is commonly used the first or middle slice. The temporal shifts depend on the order in which the slices of a volume are scanned. Usually, the odd slice numbers are recorded first followed by the even slice numbers. Also, the data-shifting requires interpolation (resampling) of the slice time courses. Methods for interpolation were described in the previous section. Two known tools for fMRI analysis are SPM (Statistical Parametric Mapping) and BrainVoyager. The first uses sinc interpolation for slice time correction and the second cubic splines interpolation.

An example of synthetic data will make more clear the above problem. In Figure 3.4, the expected response of a block design is illustrated. The duration of the block is 34,56 seconds, the sampling time TR is 3.5 seconds, the 3-D brain breaks into 35 2-D slices and the order of sampling is 1, 2, ..., 34, 35. The expected response is the result of convolution of the HRF with the stimulus function.

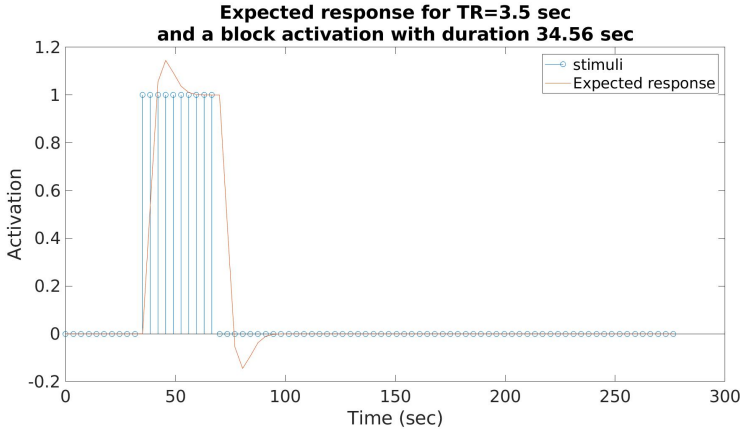


Figure 3.4: Block design and corresponding expected response.

Ideally, if there was not any kind of noise, the time series of each activated voxel would be identical to the expected response. However, the time series of two activated voxels, which belong to different slices, will differ because of their data being collected in slightly different instants. Figure 3.5 illustrates the above problem.

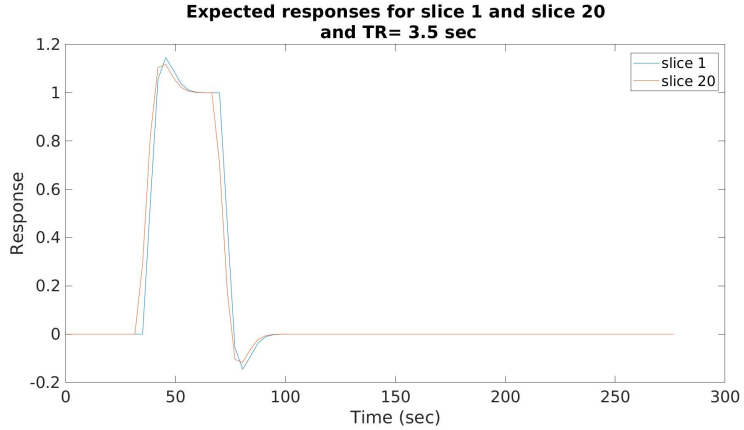


Figure 3.5: Time series of activated voxels which belong to slices 1 and 20.

The results of the application of slice time correction using SPM, are plotted in Figure 3.6.

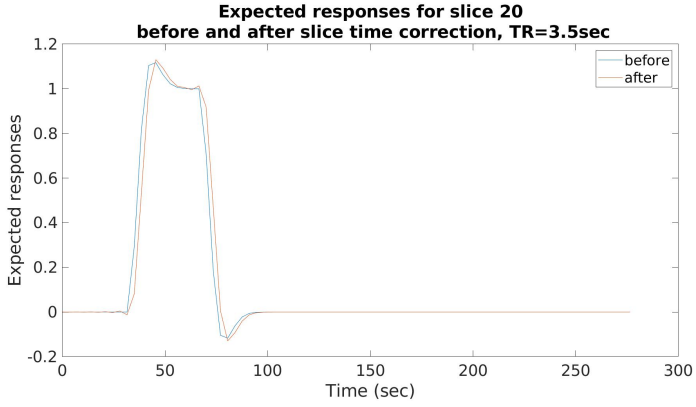


Figure 3.6: Time series of an activated voxel which belongs to slice 20 before and after the application of slice time correction.

Finally, the time course of a voxel is illustrated before and after the application of slice time correction in Figure 3.7. In this case, real data from an fMRI session are used.

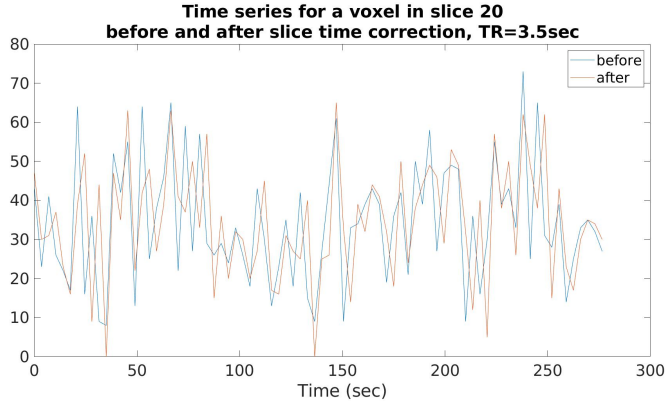


Figure 3.7: Time series of a voxel which belongs to a slice with sampling order twenty, before and after the application of slice time correction.

3.1.3 Co-registration

Co-registration is applied for alignment and overlay of fMRI data of a single subject with its own structural data. The structural data can be obtained either from an MRI session, or from totally different imaging modalities (PET, CT, etc). The above registration can enhance the information provided by the fMRI session. For example, in a single subject study, co-registration can localize the activation areas better by superimposing them on the structural image, which has a higher resolution than the fMRI one. Moreover, the spatial normalization can be more accurate if it is computed from a structural image. If co-registration is applied, any warp computed using the structural image can be applied to the corresponding functional one too.

In the case of the realignment, images we try to align, belong to the same modality, so they are similar. It is sufficient to align them by finding the rigid body transformation which minimizes the sum of square differences between the reference image and each and every of the rest images. In case of co-registration, the two images we try to align, belong to different imaging modalities, thereby there is not linear relationship between the images intensities and the minimization function is not really easy to be found.

The techniques for registration between different modalities are landmark, surface and voxel based. The most robust technique for the case in which functional scans are involved, is voxel based registration. A voxel based registration method optimizes a function measuring the similarity of all possible pairs of voxel gray-values. The most recent methods measure voxel similarity using information theory.

Co-registration pre-processing step attempts to find out the rigid body transformation $\mathbf{M_q}$ ($\mathbf{q}$ contains three translations and three rotations), which minimizes (maximizes) a voxel matching criterion. The basic assumption is that, regions which belong in similar tissues have similar gray-values in each image. So, a region with similar gray-values in one image will correspond to a region with similar gray-values in the other image. The gray-values between the two associated regions of the two images will probably be different.

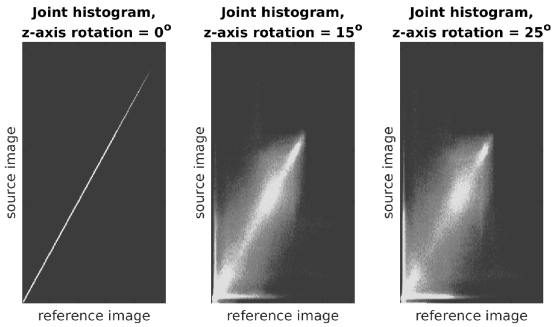


Figure 3.8: Joint histogram of an MR image with (a) itself (b) the rotated MR image by 15° over z-axis, (c) the rotated MR image by 25° over z-axis. We observe that the more misregistered the images are the more dispersion their joint histogram exhibits.

A joint histogram of two images, which is two-dimensional, is a tool which helps to the visualization of the relationship between the intensities of the corresponding voxels. One dimension is the gray values of the first image and the other dimension is the gray values of the other image. It includes the voxels of two images which overlap. The value of position (x, y) is the number of times the voxels of the first image have gray value x and the corresponding voxels of the other image have gray value y.

The joint histogram of two images changes when their alignment changes. When the two images are not registered, the histogram will exhibit dispersion and will be more blurry with low-intensity values. Figure 3.8 will make the above observation more clear. We use an MR image and rotate the image over z-axis by 15° and 25°. In Figure 3.8, we display the joint histogram of the MR image and (a) itself, (b) the rotated image by 15° over z-axis, (c) the rotated image by 25° over z-axis.

Using the above characteristic of a joint histogram, the first approach used entropy of the joint probability distribution of the combined gray-values of all common voxel pairs in the two images, which we want to match.

$$H(\mathbf{f}, \mathbf{g}) = -\sum_{\mathbf{s}} p_{\mathbf{f},\mathbf{g}}(\mathbf{f}(\mathbf{M_q s}), \mathbf{g}(\mathbf{s})) \log p_{\mathbf{f},\mathbf{g}}(\mathbf{f}(\mathbf{M_q s}), \mathbf{g}(\mathbf{s})),$$

where the $\mathbf{f}$, $\mathbf{g}$ are the images we want to register, $\mathbf{s}$ is a sample of coordinates of the image $\mathbf{g}$, $\mathbf{f}(\mathbf{M_q s})$ is the grey values of the rotated and translated image $\mathbf{f}$ at point $\mathbf{M_q s}$, $p_{\mathbf{f},\mathbf{g}}$ is the joint probability distribution (calculated by the joint histogram, which is normalized). When the joint histogram has few sharply defined and dominant peaks, the entropy is low in which case the images have been aligned correctly. Thereby, the $H(\mathbf{f}, \mathbf{g})$ should be minimized in order for the two images to be registered. However, above approach has poor results.

A more robust method used as voxel matching criterion is mutual information. Mutual information measures the dependence between the two images and is defined as follows:

$$I(\mathbf{f}, \mathbf{g}) = D_{KL}(P_{\mathbf{f},\mathbf{g}}(f, g) || P_{\mathbf{f}}(f) P_{\mathbf{g}}(g)) = \sum_{\mathbf{s}} p_{\mathbf{f},\mathbf{g}}(\mathbf{f}(\mathbf{M_q s}), \mathbf{g}(\mathbf{s})) \log \frac{p_{\mathbf{f},\mathbf{g}}(\mathbf{f}(\mathbf{M_q s}), \mathbf{g}(\mathbf{s}))}{p_{\mathbf{f}}(\mathbf{f}(\mathbf{M_q s}))(f) p_{\mathbf{g}}(\mathbf{g}(\mathbf{s}))},$$

where D_{KL} is the Kullback-Leibler divergence, the $\mathbf{f}$, $\mathbf{g}$ are the images we want to register, $\mathbf{s}$ is a sample of coordinates of the image $\mathbf{g}$, $\mathbf{f}(\mathbf{M_q s})$ is the grey-value of the rotated and translated image $\mathbf{f}$ at point $\mathbf{M_q s}$, $p_{\mathbf{f},\mathbf{g}}$ is the joint probability distribution (calculated by the joint histogram, which is normalized), $p_{\mathbf{f}}$, $p_{\mathbf{g}}$ are the marginal probabilities. The Kullback-Leibler divergence measures the dependence between two images. When the dependence between the gray values of the images is maximal then they are correctly aligned.

Summarizing the solution of the co-registration problem is:

$$\mathbf{q}* = \arg\max_{\mathbf{q}} \left(\sum_{\mathbf{s}} p_{\mathbf{f},\mathbf{g}}(\mathbf{f}(\mathbf{M_q s}), \mathbf{g}(\mathbf{s})) \log \frac{p_{\mathbf{f},\mathbf{g}}(\mathbf{f}(\mathbf{M_q s}), \mathbf{g}(\mathbf{s}))}{p_{\mathbf{f}}(\mathbf{f}(\mathbf{M_q s}))(f) p_{\mathbf{g}}(\mathbf{g}(\mathbf{s}))} \right).$$

Given the fact that the parameters $\mathbf{q}$ take continuous values, it's a necessity to interpolate the value of image $\mathbf{f}$ at points $\mathbf{M_q s}$ for the different $\mathbf{s}$. One way to deal with that is to use classical interpolation methods like nearest neighbors or trilinear interpolation. In this case, new values can lead to new intensity values in the histograms, which introduce errors. Another approach is the partial volume interpolation. It is akin to trilinear interpolation, but it does not calculate a new point as a weighted average of its neighbors and does not increment by one the entries of the histogram. Instead of that, it updates the histogram entries of neighbors by increasing them by the corresponding weight. This method also introduces errors. Many authors find trilinear interpolation sufficient.

A different approach is the use of Parzen window density estimation, which estimates a continuous PDF. The general form of a Parzen window is:

$$P(x, A) = \frac{1}{N_A} \sum_{x_a \in A} R(x - x_a),$$

where x_a is a point with known value, A contains all points with known value used to estimate the value of distribution at x, N_A is the number of known points and R a valid density function. Parzen windowing provides more reliability of probability estimation and in comparison with a histogram method it is possible to calculate derivatives for entropies. However, it is more computationally expensive.

As already mentioned, derivatives of cost function can be calculated, thereby a calculus-based optimization algorithm can solve the problem. However, most implementations of co-registration do not use derivatives. They use standard numerical optimization techniques such as Powell's direction set method. Powell's direction set method is an optimization method which minimizes the cost function in the N-dimensional space of parameters by a succession of one dimension minimizations. In the case of 3-D images, a rigid body transformation has six degrees of freedom, so the optimum is found in a six-dimensional parameter space. Also, an affine transformation has twelve degrees of freedom which corresponds to a twelve-dimensional parameter space.

3.1.4 Segmentation

The goal of segmentation preprocessing step is to divide the anatomical image into different tissue classes. The main tissue classes are grey matter, white matter and Cerebrospinal

fluid (CSF). The method used for segmenting the MR image is a modified Gaussian mixture model. The modified model combines three steps, segmentation, bias correction, and spatial normalization, in one model. Namely, this model is based on a Gaussian mixture model, which is extended in order to include a smooth intensity variation and non-linear registration with tissue probability map.

A Gaussian mixture model is a probabilistic model, that assumes that the distribution of data can be modeled by a mixture of a finite number of Gaussian distributions with unknown parameters. Suppose the finite number of Gaussian distributions is K and $k \in \{1, \ldots, K\}$, the k^{th} component has a mean μ_k and variance σ_k^2, for the univariate case, and a mean $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$, for the multivariate case. Also, the mixing proportion is defined as γ_k with the constraints $\sum_{i=1}^{K} \gamma_k = 1$ and $\gamma_k \geq 0$.

The prior probability of any data point y_i, irrespective of its value, belonging to the k_{th} Gaussian is : $p(c_i = k) = \gamma_k$, where c_i is a latent variable, the value of which indicates the cluster in which y_i belongs. The probability of a data point which belongs to the k_{th} Gaussian with parameters μ_k andσ_k^2, has value y_i is:

$$p(y_i|c_i = k, \mu_k, \sigma_k^2) = \frac{1}{(2\pi\sigma_k^2)^{\frac{1}{2}}} \exp(-\frac{(y_i - \mu_k)^2}{2\sigma_k^2}). \tag{3.1}$$

The probability of a data point, given the parameters of Gaussian distributions, has value y_i is:

$$p(y_i|\boldsymbol{\mu}, \boldsymbol{\sigma}^2) = \sum_{k=1}^{K} p(y_i|c_i = k, \mu_k, \sigma_k^2) p(c_i = k) = \sum_{k=1}^{K} \frac{\gamma_k}{(2\pi\sigma_k^2)^{\frac{1}{2}}} \exp(-\frac{(y_i - \mu_k)^2}{2\sigma_k^2}). \tag{3.2}$$

The probability for the entire dataset is:

$$p(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\sigma}^2) = \prod_{i=1}^{I} p(y_i|\boldsymbol{\mu}, \boldsymbol{\sigma}^2). \tag{3.3}$$

Fitting a mixture of Gaussians model involves maximizing the equation 3.3 which is equivalent to the minimization of cost function $\mathcal{E}$.

$$\mathcal{E} = -\log p(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\sigma}^2) = -\sum_{i=1}^{I} \log \sum_{k=1}^{K} \frac{\gamma_k}{(2\pi\sigma_k^2)^{\frac{1}{2}}} \exp(-\frac{(y_i - \mu_k)^2}{2\sigma_k^2}). \tag{3.4}$$

In our case, the data are univariate and the y_i is the intensity of voxel.

MR images are usually corrupted by a smooth, spatially varying artifact that modifies the intensity of the image (bias). This artifact may not be a problem for a visual check, however, it can hamper the automated processing of images. There are parametric and no-parametric techniques to deal with this problem. A parametric bias correction model is used in this approach. The observed signal y_i is an uncorrupted signal μ_i with added Gaussian noise n_i and the sum is scaled by some bias ρ_i: $y_i = \frac{(\mu_i + n_i)}{\rho_i}$.

Extra parameters for bias correction are included in the previous Gaussian mixture model. We define as $\rho_i(\boldsymbol{\beta})$ the field which models the variation of element i,where $\boldsymbol{\beta}$ is

the vector of the unknown parameters. For each k^{th} cluster its intensities are assumed to be normally distributed with mean $\frac{\mu_k}{\rho_i(\boldsymbol{\beta})}$ and variance $(\frac{\sigma_k}{\rho_i(\boldsymbol{\beta})})^2$. So, the probability of the i element which belongs to k^{th} cluster is:

$$p(y_i|c_i = k, \mu_k, \sigma_k^2, \boldsymbol{\beta}) = \frac{\rho_i(\boldsymbol{\beta})}{(2\pi\sigma_k^2)^{\frac{1}{2}}} \exp(-\frac{(\rho_i(\boldsymbol{\beta})y_i - \mu_k)^2}{2\sigma_k^2}). \tag{3.5}$$

The cost function ($\mathcal{E}$) becomes :

$$\mathcal{E} = -\sum_{i=1}^{I} \log \rho_i(\boldsymbol{\beta}) \sum_{k=1}^{K} \frac{\gamma_k}{(2\pi\sigma_k^2)^{\frac{1}{2}}} \exp(-\frac{(\rho_i(\boldsymbol{\beta})y_i - \mu_k)^2}{2\sigma_k^2}). \tag{3.6}$$

Apart from stationary prior probabilities based on mixing proportions γ_k, information which derives from tissue probability maps from other subjects' brain images can be included. Tissue probability maps are probability maps for different tissue classes constructed from a large number of brains which are registered into a common space. The model in equation 3.6 is modified in order to include the additional information. The prior probability of voxel i to belong in the k^{th} cluster is:

$$p(c_i = k) = \frac{\gamma_k b_{ik}}{\sum_{j=1}^{K} \gamma_k b_{ij}}, \tag{3.7}$$

where b_{ik} is the tissue probability of voxel i to belong in k^{th} class.

A single Gaussian distribution for each class is not a logical assumption, because a voxel can consist of more than one tissue types. So, it contains signal from a number of different tissues. Thereby, the number of Gaussians which is used to represent the intensity distribution for each tissue class can be greater than one.

The equation 3.7 can be improved by using parameters $\mathbf{a}$ which allow the deformation of tissue probability maps. Namely, the registration to a standard space is included to the generative model.

$$p(c_i = k|\mathbf{a}) = \frac{\gamma_k b_{ik}(\mathbf{a})}{\sum_{j=1}^{K} \gamma_k b_{ij}(\mathbf{a})}, \tag{3.8}$$

A way to parameterize the deformations of probability maps is by a linear combination of about a thousand cosine transform bases. This deformation means that inversion of the generative model can normalize images.

The objective function now is:

$$\mathcal{E} = -\sum_{i=1}^{I} \log \frac{\rho_i(\boldsymbol{\beta})}{\sum_{j=1}^{K} \gamma_k b_{ij}(\mathbf{a})} \rho_i(\boldsymbol{\beta}) \sum_{k=1}^{K} \frac{\gamma_k b_{ik}(\mathbf{a})}{(2\pi\sigma_k^2)^{\frac{1}{2}}} \exp(-\frac{(\rho_i(\boldsymbol{\beta})y_i - \mu_k)^2}{2\sigma_k^2}). \tag{3.9}$$

Regularization is needed in order for us to have more accurate results. Using regularization the following equation has to be minimized:

$$\mathcal{F} = -\log p(\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{a}|\boldsymbol{\mu}\boldsymbol{\sigma}^2) = \mathcal{E} - \log p(\boldsymbol{\beta}) - \log p(\mathbf{a}). \tag{3.10}$$

The problem 3.10 has not a closed-form solution. So, we use an iterated conditional modes (ICM) approach. It assigns initial points and then iterates until it finds a local minimum. Each iteration estimates all parameters by steps. In each step, the values of a group of parameters are fixed and the values of others are estimated by solving an optimization problem. The parameters of the Gaussian mixture model are estimated by using Expectation Maximization algorithm with the bias correction parameters and the deformation parameters being fixed. The bias correction parameters are estimated by using Levenberg-Marquardt (LM) algorithm with the values of others parameters being fixed. Finally, the deformation parameters are also estimated by using LM algorithm with the rest of parameters being fixed. The LM algorithm is ideal because the number of parameters is small.

3.1.5 Spatial Normalization

Spatial normalization is the preprocessing step which registers images of different individuals in a common coordinate system defined by a template image. In fMRI studies, this preprocessing step is essential for studying the results of an experiment over individuals. Methods used for the implementation of spatial normalization can be categorized into label based and intensity based. This section will illustrate the basic idea of an intensity based method, in order for the spatial normalization to be applied.

Firstly, the image and template have to be approximately registered, using a nine or twelve parameter affine transformation, by minimizing or maximizing some mutual information criterion. The next step is a nonlinear registration for correction of gross differences in head shape, which cannot be accounted by the previous affine registration alone. So parameters, which describe the global shape differences between the image and template should be determined. The nonlinear warps are modeled by a linear combination of low-frequency periodic basis functions. The spatial transformation from co-ordinates $\mathbf{x}_i$, to co-ordinates $\mathbf{y}_i$ is:

$$y_{1,i} = x_{1,i} + \sum_j q_{j,1} d_j(\mathbf{x}_i)$$

$$y_{2,i} = x_{2,i} + \sum_j q_{j,2} d_j(\mathbf{x}_i)$$

$$y_{3,i} = x_{3,i} + \sum_j q_{j,3} d_j(\mathbf{x}_i),$$

where $q_{j,k}$ is the j^{th} coefficient for dimension k and $d_j(x_i)$ is the j^{th} basis function at position $\mathbf{x}_i$. In our case, the basis functions used, are the lowest frequency components of the three dimensional DCT (Discrete Cosine Transformation).

In order to estimate these parameters, we solve the problem:

$$\min \sum (\mathbf{f}(\mathbf{y}_i) - w\mathbf{g}(\mathbf{x}_i))^2,$$

where $\mathbf{f}$ is the image and $\mathbf{g}$ is the template image. The images may be scaled differently, so an additional scalar parameter w is used.

The problem can be solved via Gauss-Newton algorithm. However, without regularization the algorithm can become unstable. Regularization can be achieved using Bayesian statistic. This framework attempts to estimate the parameters $\mathbf{q}$ of warping which maximize the posterior probability $p(\mathbf{q}|\mathbf{y})$ given the data $\mathbf{y}$. Namely, we are looking for the deformation which is more likely, given the data. This is equivalent to maximization of the quantity $p(\mathbf{y}|\mathbf{q})p(\mathbf{q})$. The term $p(\mathbf{y}|\mathbf{p})$ is related to the sum of squared differences between the template and deformed image. The MAP approach requires, some form of prior distribution for the parameters to be defined.There are many possible forms for modeling these priors.

So far, the model includes only one intensity scalar parameter. However, more parameters have to be considered if the intensity transformation is to be described. So, two sets of parameters can be used: those which describe spatial transformations and those for describing intensity transformations. So the difference function can be expressed in a generic form as:

$$\mathbf{f}(\mathbf{t}(\mathbf{x}_i, \mathbf{p_t})) - w(\mathbf{x}_i, \mathbf{p}_w),$$

where $\mathbf{f}$ is the image, $\mathbf{t}()$ is a vector function which describes the spatial transformations based on parameters $\mathbf{p_t}$, $w()$ is a scalar function which describes the intensity transformations based on parameters $\mathbf{p}_w$ and $\mathbf{x}_i$ represents the coordinates of the i^{th} sampled point. An example of function w could be: $p_{w1}\mathbf{g}(\mathbf{x_i}) + p_{w2}x_{1i}\mathbf{g}(\mathbf{x_i}) + p_{w3}x_{2i}\mathbf{g}(\mathbf{x_i}) + p_{w4}x_{3i}\mathbf{g}(\mathbf{x_i})$, where $\mathbf{g}$ is the template image.

3.1.6 Spatial Smoothing

Spatial smoothing is the final step in fMRI preprocessing. Spatial smoothing averages every voxel with a weighted sum of its neighbors. The weights are defined by a Gaussian kernel. The width of the kernel is expressed by Full Width at Half Maximum (FWHM). The main purposes of this preprocessing step are to improve the signal to noise ratio (SNR) and to deal with the differences between individual brains. The reduction of random noise will improve the capacity of the statistical tests to detect regions of activation. Also, after the spatial normalization the individual brains are registered to a common template. However, the variation of brains is substantial, so the activation areas seldom include the exact same voxels. The spatial smoothing increases the overlap area of the activated areas. However, there are trade-offs. The spatial smoothing decreases the spatial resolution. Moreover, if the kernel size is too large some activated areas may not be detected by the statistical tests.

3.2 Brain parcellation

This chapter focus on the presentation of a whole-brain parcellation methodology, which is introduced at [1]. The methodology performed on resting state fMRI data and aims to extract a common brain parcellation, to all subjects, into non-overlapping clusters, without any prior knowledge about clusters.

Given the hypothesis that, there is a common brain parcellation for all subjects, resting

state data for each subject k can be modeled as follows:

$$\mathbf{X}_k = \mathbf{A}\mathbf{S}_k + \mathbf{E}_k, \text{ for } k = 1, \cdots, K \tag{3.11}$$

where

- K is the number of subjects.
- $\mathbf{X}_k$ is the fMRI pre-processed data of k^{th} subject. More specifically, $\mathbf{X}_k \in \mathbb{R}^{N \times M}$, where N is the number of voxels, and M is the number of time points with $N \gg M$.
- $\mathbf{A} \in \mathbb{R}_+^{N \times R}$ with $R < M$ and $\mathbf{A}^T\mathbf{A} = \mathbf{I_R}$. It denotes the common spatial component related with the spontaneous fMRI activity. Namely, each column of $\mathbf{A}$ represents one of the R clusters. More specifically, the non zero elements of each column indicate the voxels which belong to this cluster.
- $\mathbf{S}_k \in \mathbb{R}^{M \times R}$, each column of $\mathbf{S}_k$ indicates the temporal component of each cluster, which is associated with the spontaneous fMRI activity. The term $\mathbf{S}_k$ may be different across subjects.
- $\mathbf{E}_k \in \mathbb{R}^{N \times M}$ represents the unmodeled fMRI signal of k^{th} subject and it can be considered as (strong) additive noise. The terms $\mathbf{E}_k$ are assumed as statistically independent from each other.

3.2.1 Estimating Common Spatial Subspace via gCCA

Canonical Correlation Analysis (CCA) is a method which estimates a linear common subspace of two sets of random variables. Generalization of CCA (gCCA) estimates a linear common subspace of more than two random vectors. It is a well-studied subject [12], [13], [14]. Our model 3.11 assumes that subjects share common information, and at the same time allows to each subject to retain its individual data features. We can estimate the common spatial subspace over all subjects via the MAX-VAR formulation of GCCA. The MAX-VAR formulation of GCCA is modeled as follows:

$$\begin{aligned} \underset{\{\mathbf{Q}_i\}_{i=1}^I, \mathbf{G}}{\text{minimize}} \quad & \sum_{i=1}^{I} ||\mathbf{X}_i\mathbf{Q}_i - \mathbf{G}||_F^2, \\ \text{subject to} \quad & \mathbf{G}^T\mathbf{G} = \mathbf{I} \end{aligned} \tag{3.12}$$

where I is the number of subjects, $\mathbf{X}_i \in \mathbb{R}^{N \times M}$ is the data of i^{th} subject, $\mathbf{Q}_i \in \mathbb{R}^{M \times R}$ correspond to the R canonical components of $\mathbf{X}_i$, $\mathbf{G} \in \mathbb{R}^{N \times R}$ corresponds to a common latent representation of our data.

The problem has a closed-form solution. More specifically, for a fixed $\mathbf{G}$ the problem becomes unconstrained.

$$\underset{\{\mathbf{Q}_i\}_{i=1}^I}{\text{minimize}} \quad \sum_{i=1}^{I} ||\mathbf{X}_i\mathbf{Q}_i - \mathbf{G}||_F^2. \tag{3.13}$$

Let $F : \mathbb{R}^{M\times R} \to \mathbb{R}_+$:

$$\begin{aligned} F(\mathbf{Q}_i) &= ||\mathbf{X}_i\mathbf{Q}_i - \mathbf{G}||_F^2 \\ &= \text{tr}((\mathbf{X}_i\mathbf{Q}_i - \mathbf{G})^T(\mathbf{X}_i\mathbf{Q}_i - \mathbf{G})) \\ &= \text{tr}(\mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i) - 2\text{tr}(\mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{G}) + \text{tr}(\mathbf{G}^T\mathbf{G}). \end{aligned} \tag{3.14}$$

The first order Taylor approximation of F is given by :

$$F(\mathbf{Q}_i + \boldsymbol{\Delta}) = F(\mathbf{Q}_i) + \text{tr}(\mathbf{W}^T\boldsymbol{\Delta}) + O(||\boldsymbol{\Delta}||_F^2). \tag{3.15}$$

Then

$$\nabla F(\mathbf{Q}_i) = \mathbf{W}. \tag{3.16}$$

So,

$$\begin{aligned} F(\mathbf{Q}_i + \boldsymbol{\Delta}) &= \text{tr}((\mathbf{Q}_i + \boldsymbol{\Delta})^T\mathbf{X}_i^T\mathbf{X}_i(\mathbf{Q}_i + \boldsymbol{\Delta})) \\ &\quad - 2\text{tr}((\mathbf{Q}_i + \boldsymbol{\Delta})^T\mathbf{X}_i^T\mathbf{G}) + \text{tr}(\mathbf{G}^T\mathbf{G}) \\ &= \text{tr}(\mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i + \mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{X}_i\boldsymbol{\Delta} + \boldsymbol{\Delta}^T\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i + \boldsymbol{\Delta}^T\mathbf{X}_i^T\mathbf{X}_i\boldsymbol{\Delta}) \\ &\quad - 2\text{tr}(\mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{G} + \boldsymbol{\Delta}^T\mathbf{X}_i^T\mathbf{G}) + \text{tr}(\mathbf{G}^T\mathbf{G}) \\ &= \text{tr}(\mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i) + \text{tr}(\mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{X}_i\boldsymbol{\Delta}) + \text{tr}(\boldsymbol{\Delta}^T\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i) \\ &\quad + \text{tr}(\boldsymbol{\Delta}^T\mathbf{X}_i^T\mathbf{X}_i\boldsymbol{\Delta}) - 2\text{tr}(\mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{G}) - 2\text{tr}(\boldsymbol{\Delta}^T\mathbf{X}_i^T\mathbf{G})) \\ &\quad + \text{tr}(\mathbf{G}^T\mathbf{G}) \\ &= F(A) + 2\text{tr}(\mathbf{Q}_i^T\mathbf{X}_i^T\mathbf{X}_i\boldsymbol{\Delta}) - 2\text{tr}(\mathbf{G}^T\mathbf{X}_i\boldsymbol{\Delta})) + O(||\boldsymbol{\Delta}||_F^2) \\ &= F(A) + \text{tr}((2\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i^T - 2\mathbf{X}_i^T\mathbf{G})^T\boldsymbol{\Delta}) + O(||\boldsymbol{\Delta}||_F^2). \end{aligned} \tag{3.17}$$

So, the first derivative of function F is:

$$\nabla F(\mathbf{Q}_i) = 2\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i - 2\mathbf{X}_i^T\mathbf{G}. \tag{3.18}$$

The second order Taylor approximation of F is given by:

$$F(\mathbf{Q}_i + \boldsymbol{\Delta}) = F(\mathbf{Q}_i) + \text{tr}(\mathbf{W}^T\boldsymbol{\Delta}) + \frac{1}{2}\text{vec}(\boldsymbol{\Delta})^T\nabla^2 F(\mathbf{Q}_i)\text{vec}(\boldsymbol{\Delta}) + O(||\boldsymbol{\Delta}||_F^3), \tag{3.19}$$

where $\nabla^2 F(\mathbf{Q}_i)$ is the Hessian of F and is defined as:

$$\nabla^2 F(\mathbf{Q}_i) = \frac{\partial^2 F(\mathbf{Q}_i)}{\partial\text{vec}(\mathbf{Q}_i)\partial\text{vec}(\mathbf{Q}_i)^T} \in \mathbb{R}^{MR\times MR}. \tag{3.20}$$

So,

$$\begin{aligned} F(\mathbf{Q}_i + \mathbf{\Delta}) &= F(A) + \text{tr}((2\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i^T - 2\mathbf{X}_i^T\mathbf{G})^T\mathbf{\Delta}) + \text{tr}(\mathbf{\Delta}^T\mathbf{X}_i^T\mathbf{X}_i\mathbf{\Delta}) \\ &= F(A) + \text{tr}((2\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i^T - 2\mathbf{X}_i^T\mathbf{G})^T\mathbf{\Delta}) + \text{tr}(\mathbf{\Delta}^T\mathbf{X}_i^T\mathbf{X}_i\mathbf{\Delta}\mathbf{I}_R) \\ &= F(A) + \text{tr}((2\mathbf{X}_i^T\mathbf{X}_i\mathbf{Q}_i^T - 2\mathbf{X}_i^T\mathbf{G})^T\mathbf{\Delta}) + \text{vec}(\mathbf{\Delta})^T(\mathbf{I}_R \otimes \mathbf{X}_i^T\mathbf{X})\text{vec}(\mathbf{\Delta}). \end{aligned} \tag{3.21}$$

So, the Hessian of function F is :

$$\nabla^2 F(\mathbf{Q}_i) = 2(\mathbf{I}_R \otimes \mathbf{X}_i^T\mathbf{X}). \tag{3.22}$$

We know that the identity matrix $\mathbf{I}_R$ has R eigenvalues, which are equal to 1 and the matrix $\mathbf{X}_i^T\mathbf{X}_i$ has M non-negative eigenvalues ($\mathbf{X}_i$ is a full rank matrix). We also know that the set of eigenvalues of $\mathbf{I}_R \otimes \mathbf{X}_i^T\mathbf{X}$ is $\{\lambda_i\mu_j, \quad i = 1, \ldots, R \text{ and } j = 1, \ldots, M\}$, where λ_i and μ_i denote the eigenvalues of the matrix $\mathbf{I}_R$ and $\mathbf{X}_i^T\mathbf{X}_i$ correspondingly. Thereby, the $\nabla^2 F(\mathbf{Q}_i)$ has MR positive eigenvalues and so it is a positive definite matrix. So, the terms $||\mathbf{X}_i\mathbf{Q}_i - \mathbf{G}||_F^2$ are convex and have a unique minimum, which is $\mathbf{Q}_i = (\mathbf{X}_i^T\mathbf{X}_i)^{-1}\mathbf{X}_i^T\mathbf{G}$. Also, given the fact that, the terms $||\mathbf{X}_i\mathbf{Q}_i - \mathbf{G}||_F^2$ are convex the problem 3.13 is convex too. The minimization of which is achieved when $\mathbf{Q}_i = (\mathbf{X}_i^T\mathbf{X}_i)^{-1}\mathbf{X}_i^T\mathbf{G}$.

Now, given the $\mathbf{Q}_i$ and using the relationship $||\mathbf{A}||_F^2 = \text{tr}(\mathbf{A}^T\mathbf{A})$, the problem 3.12 is transformed to:

$$\begin{aligned} \max_{\mathbf{G}} \quad & \text{tr}(\mathbf{G^T}(\sum_{i=1}^{I} \mathbf{X}_i(\mathbf{X}_i^T\mathbf{X}_i)^{-1}\mathbf{X}_i^T)\mathbf{G}), \\ \text{subject to} \quad & \mathbf{G}^T\mathbf{G} = \mathbf{I} \end{aligned} \tag{3.23}$$

Let matrices $\mathbf{M} = \sum_{i=1}^{I} \mathbf{X}_i(\mathbf{X}_i^T\mathbf{X}_i)^{-1}\mathbf{X}_i^T$ and $\mathbf{B} = \mathbf{G^T}\mathbf{M}\mathbf{G}$.

From Poincare separation theorem we know that:

$$\lambda_j^{\mathbf{M}} \geq \lambda_j^{\mathbf{B}}, \quad j = 1, \ldots, R, \tag{3.24}$$

where $\lambda_j^{\mathbf{M}} \quad j = 1, \ldots, N$ and $\lambda_j^{\mathbf{B}}, \quad j = 1, \ldots, R$ are eigenvalues of matrices $\mathbf{M}$ and $\mathbf{B}$, respectively (in descending order).

From equation 3.24:

$$\sum_{j=1}^{R} \lambda_j^{\mathbf{M}} \geq \sum_{j=1}^{R} \lambda_j^{\mathbf{B}} = \text{tr}(\mathbf{B}) = \text{tr}(\mathbf{G^T}\mathbf{M}\mathbf{G}). \tag{3.25}$$

The upper bound is met with equality, when $\mathbf{G}$ is chosen to have the first R eigenvectors of $\mathbf{M}$. So, the solution of problem 3.23 is the first R eigenvectors of the matrix $\mathbf{M}$, $\mathbf{G}_{opt} = \mathbf{U}_{\mathbf{M}}(:, 1:R)$.

In order to find $\mathbf{G}_{opt}$, the matrix $\mathbf{M}$ have to be calculated and be decomposed. The dimensions of the matrix $\mathbf{M}$ are $N \times N$, if N is large, then the calculation of $\mathbf{G}_{opt}$ is impossible. Howbeit, we can compress the data and compute the $\mathbf{G}_{opt}$.
Let matrix $\mathbf{Y} \in \mathbb{R}^{N \times IM}$ to be the concatenation of I $\mathbf{X}_i$. Namely,

$$\mathbf{Y} = [\mathbf{X}_1\mathbf{X}_2 \dots \mathbf{X}_I]. \tag{3.26}$$

$\mathbf{Y}$ can be factorized as follows :

$$\mathbf{Y} = \mathbf{U_Y}\mathbf{V_Y}, \tag{3.27}$$

where $\mathbf{U_Y} \in \mathbb{R}^{N\times IM}$ with $\mathbf{U_Y^T}\mathbf{U_Y} = \mathbf{I_{IM}}$ and $\mathbf{V_Y} \in \mathbb{R}^{IM\times IM}$.

From equations 3.26 and 3.27, we can conclude the follows:

$$\text{col}(\mathbf{X}_i) \subseteq \text{col}(\mathbf{Y}) = \bigcup_{i=1}^{I} \mathbf{X}_i \subseteq \text{col}(\mathbf{U_Y}). \tag{3.28}$$

So, $\mathbf{X}_i$ can be factorized as:

$$\mathbf{X}_i = \mathbf{U_Y}\mathbf{H}_i, \tag{3.29}$$

where $\mathbf{H}_i \in \mathbb{R}^{IM\times M}$. Also,

$$\mathbf{H}_i = \mathbf{U_Y^T}\mathbf{X}_i. \tag{3.30}$$

Using equation 3.29, the M will be:

$$\mathbf{M} = \mathbf{U_Y}\left(\sum_{i=1}^{I} \mathbf{H}_i(\mathbf{H}_i^T\mathbf{H}_i)^{-1}\mathbf{H}_i^T\right)\mathbf{U_Y^T}. \tag{3.31}$$

We assign $\mathbf{M}_1 = \sum_{i=1}^{I} \mathbf{H}_i(\mathbf{H}_i^T\mathbf{H}_i)^{-1}\mathbf{H}_i^T$. We can also see that $\mathbf{M}$ and $\mathbf{M}_1$ are symmetric matrices, so we can rewrite them as:

$$\mathbf{M} = \mathbf{U_M}\mathbf{\Lambda_M}\mathbf{U_M^T} \quad \text{and} \quad \mathbf{M}_1 = \mathbf{U_{M_1}}\mathbf{\Lambda_{M_1}}\mathbf{U_{M_1}^T}, \tag{3.32}$$

From equations 3.31 and 3.32:

$$\mathbf{M} = \mathbf{U_Y}\mathbf{U_{M_1}}\mathbf{\Lambda_{M_1}}\mathbf{U_{M_1}^T}\mathbf{U_Y^T}. \tag{3.33}$$

So,

$$\mathbf{U_M} = \mathbf{U_Y}\mathbf{U_{M_1}}. \tag{3.34}$$

Thereby,

$$\mathbf{G}_{opt} = \mathbf{U_Y}\mathbf{U_{M_1}}(:, 1:R). \tag{3.35}$$

If there were no noise, then

$$\text{col}(\mathbf{G}_{opt}) = \text{col}(\mathbf{A}). \tag{3.36}$$

Namely, by solving the MAX-VAR we would find the column space of the common subspace. However, the real data have nonzero noise and the equation 3.36 is approximate.

3.2.2 Estimating the numbers of clusters

In the previous sections, we assume that the real number of clusters (which is equivalent to the common subspace dimension) R is known, but that seldom happens. However, by

using the data, we can estimate the number of clusters. A way to do this is by dividing our data set into two sets. By using the gCCA method we can estimate the column space of these sets and after that, we can apply a criterion function, that uses the distance between the column spaces of the two sets, to come to a conclusion about the number of clusters.

Let us consider that data are noiseless and we divide our data set of K elements to two sets $\{\mathbf{S}_1 \quad \mathbf{S_2}\}$ randomly. If we assume that R is the real number of clusters and estimate the column space of two sets, then $\text{col}(\mathbf{G}_{opt,1}) = \text{col}(\mathbf{G}_{opt,2})$. However, the real data set is not noiseless. In this case, the two column spaces will not be equal, but they will be close to each other. One way to calculate the distance between two linear subspace $\mathbf{Y}_1$, $\mathbf{Y}_2$ is to compute their gap, which is defined at [15] as:

$$\rho_{g,2}(\mathbf{Y}_1, \mathbf{Y}_2) = ||\mathbf{P}_{\mathbf{Y}_1} - \mathbf{P}_{\mathbf{Y}_2}||_2 = \sin(\theta_1), \tag{3.37}$$

where $\mathbf{P}_{\mathbf{Y}_1}$ is the projection matrix of $\mathbf{Y}_1$, $\mathbf{P}_{\mathbf{Y}_2}$ is the projection matrix of $\mathbf{Y}_2$ and θ_1 is the largest canonical angle between $\mathbf{Y}_1$ and $\mathbf{Y}_2$.

In the sequel, we assume that $||\mathbf{E}_k|| = O(\epsilon)$, for $k = 1, \ldots, K$ with ϵ a small positive number. We have to examine three cases:

1. $R_{est} = R$
   In this case, the gCCA method returns two orthonormal bases whose vectors span approximately the same subspace and so the gap will be $O(\epsilon)$.

2. $R_{est} > R$
   In this case, the gCCA method except for the $R-$dimensional common spatial subspace attempts to model "common noise". However, since we assume that the $\mathbf{E}_k$ are statistically independent among subjects and $N \gg M$, we have no reason to believe that there is common noise subspace, thus we expect that the gap will be close to 1.

3. $R_{est} < R$
   In this case, we underestimate the dimension of common spatial subspace, and we expect that the gap value will be range from $O(\epsilon)$ to 1, because each gCCA will randomly estimate a "subspace" of the common spatial subspace.

As we can see, the gap criterion can be useful for the estimation of the number of clusters. We examine the accuracy of the criterion more extensively via experiments in subsection 4.1.1.

3.2.3 Estimation of the common spatial factor A and temporal components $\mathbf{S}_k$

We have already estimated the common spatial subspace $\mathbf{G}_{opt}$ and so, we have an estimation of the column subspace of common spatial factor $\mathbf{A}$. Now, we want to obtain from the orthonormal basis $\mathbf{G}_{opt}$ the common spatial factor $\mathbf{A}$. From equation 3.36, we

know that in the noiseless case the $\mathbf{G}_{opt}$ can be expressed as:

$$\mathbf{G}_{opt} = \mathbf{A}\mathbf{B}^T, \tag{3.38}$$

where $\mathbf{B} \in \mathbb{R}^{R\times R}$. However, there is noise, so we will try to solve the following semi-Orthogonal Nonnegative Matrix Factorization (s-ONMF) problem in order to estimate factor $\mathbf{A}$.

$$\begin{aligned} &\underset{\mathbf{A},\mathbf{B}}{\text{minimize}} \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2. \\ &\text{subject to} \quad \mathbf{A}^T\mathbf{A} = \mathbf{I}, \quad \mathbf{A} \geq \mathbf{0} \end{aligned} \tag{3.39}$$

The above problem is not convex and has no closed-form solution, so we cannot guarantee that the algorithms, which we use to deal with it, will not converge in local minima. There are several ways to solve that problem, two of which we will present in this subsection. Both of the algorithms are implemented via a penalty method.

Penalty methods are a certain class of algorithms for solving constrained optimization problems. These methods try to approximate a constrained optimization problem with an unconstrained one and then apply standard search techniques to obtain solutions. The approximation is achieved by adding a term to the objective function that prescribes a high cost for violation of the constraints.

Penalty method for semi-Orthogonal Nonnegative Matrix Factorization (s-ONMF)

In the case of the problem 3.39, the problem can be expressed by:

$$\begin{aligned} &\underset{\mathbf{A},\mathbf{B}}{\text{minimize}} \ \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2 + \frac{c_k}{2}||\mathbf{A}^T\mathbf{A} - \mathbf{I}||_F^2, \\ &\text{subject to} \quad \mathbf{A} \geq 0 \end{aligned} \tag{3.40}$$

Firstly, we examine the problem 3.40 considering that $\mathbf{B}$ is fixed. So, the problem becomes:

$$\begin{aligned} &\underset{\mathbf{A}}{\text{minimize}} \ \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2 + \frac{c_k}{2}||\mathbf{A}^T\mathbf{A} - \mathbf{I}||_F^2. \\ &\text{subject to} \quad \mathbf{A} \geq 0 \end{aligned} \tag{3.41}$$

Let $F, f_1, f_2 : \mathbb{R}^{N\times R} \to \mathbb{R}_+$:

$$\begin{aligned} F(\mathbf{A}) &= \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2 + \frac{c_k}{2}||\mathbf{A}^T\mathbf{A} - \mathbf{I}||_F^2 \\ &= f_1(\mathbf{A}) + f_2(\mathbf{A}). \end{aligned} \tag{3.42}$$

f_1 and f_2 can be expressed as:

$$\begin{aligned} f_1(\mathbf{A}) &= \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2 \\ &= \frac{1}{2}\text{tr}(\mathbf{B}\mathbf{A}^T\mathbf{A}\mathbf{B}^T) - \text{tr}(\mathbf{G}_{opt}^T\mathbf{A}\mathbf{B}^T) + \frac{1}{2}\text{tr}(\mathbf{G}_{opt}^T\mathbf{G}_{opt}) \end{aligned} \tag{3.43}$$

$$\begin{aligned} f_2(\mathbf{A}) &= \frac{c_k}{2}||\mathbf{A}^T\mathbf{A} - \mathbf{I}||_F^2 \\ &= \frac{c_k}{2}\text{tr}(\mathbf{A}^T\mathbf{A}\mathbf{A}^T\mathbf{A}) - c_k\text{tr}(\mathbf{A}^T\mathbf{A}) + \frac{c_k}{2}\text{tr}(\mathbf{I}). \end{aligned} \tag{3.44}$$

The first order Taylor approximation of f_1, f_2 is given by :

$$f_i(\mathbf{A} + \mathbf{\Delta}) = f_i(\mathbf{A}) + \text{tr}(\mathbf{W}^T\mathbf{\Delta}) + O(||\mathbf{\Delta}||_F^2), \quad i = 1, 2 \tag{3.45}$$

We have:

$$\begin{aligned} f_1(\mathbf{A} + \mathbf{\Delta}) &= \frac{1}{2}\text{tr}(\mathbf{B}(\mathbf{A} + \mathbf{\Delta})^T(\mathbf{A} + \mathbf{\Delta})\mathbf{B}^T) - \text{tr}(\mathbf{G}_{opt}^T(\mathbf{A} + \mathbf{\Delta})\mathbf{B}^T) + \frac{1}{2}\text{tr}(\mathbf{G}_{opt}^T\mathbf{G}_{opt}) \\ &= f_1(\mathbf{A}) + \text{tr}((\mathbf{A}\mathbf{B}^T\mathbf{B} - \mathbf{G}_{opt}\mathbf{B})^T\mathbf{\Delta}) + O(||\Delta||_F^2) \end{aligned} \tag{3.46}$$

$$\begin{aligned} f_2(\mathbf{A} + \mathbf{\Delta}) &= \frac{c_k}{2}\text{tr}((\mathbf{A} + \mathbf{\Delta})^T(\mathbf{A} + \mathbf{\Delta})(\mathbf{A} + \mathbf{\Delta})^T(\mathbf{A} + \mathbf{\Delta})) \\ &\quad - c_k\text{tr}((\mathbf{A} + \mathbf{\Delta})^T(\mathbf{A} + \mathbf{\Delta})) + \frac{c_k}{2}\text{tr}(\mathbf{I}) \\ &= f_2(\mathbf{A}) + \text{tr}(2c_k(\mathbf{A}\mathbf{A}^T\mathbf{A} - \mathbf{A})^T\mathbf{\Delta}) + O(||\Delta||_F^2). \end{aligned} \tag{3.47}$$

So,

$$\nabla f_1(\mathbf{A}) = \mathbf{A}\mathbf{B}^T\mathbf{B} - \mathbf{G}_{opt}\mathbf{B} \quad \textit{and} \quad \nabla f_2(\mathbf{A}) = 2c_k(\mathbf{A}\mathbf{A}^T\mathbf{A} - \mathbf{A}). \tag{3.48}$$

The second order Taylor approximation of f_1, f_2 is given by:

$$f_i(\mathbf{A}+\mathbf{\Delta}) = f_i(\mathbf{A})+\text{tr}(\mathbf{W}^T\mathbf{\Delta})+\frac{1}{2}\text{vec}(\mathbf{\Delta})^T\nabla^2 f_i(\mathbf{A})\text{vec}(\mathbf{\Delta})+O(||\mathbf{\Delta}||_F^3), \quad i = 1, 2. \tag{3.49}$$

We have:

$$\begin{aligned} f_1(\mathbf{A} + \mathbf{\Delta}) &= f_1(\mathbf{A}) + \text{tr}((\mathbf{A}\mathbf{B}^T\mathbf{B} - \mathbf{G}_{opt}\mathbf{B})^T\mathbf{\Delta}) + \frac{1}{2}\text{tr}(\mathbf{B}\mathbf{\Delta}^T\mathbf{\Delta}\mathbf{B}^T) \\ &= f_1(\mathbf{A}) + \text{tr}((\mathbf{A}\mathbf{B}^T\mathbf{B} - \mathbf{G}_{opt}\mathbf{B})^T\mathbf{\Delta}) + \frac{1}{2}\text{tr}(\mathbf{\Delta}^T\mathbf{\Delta}\mathbf{B}^T\mathbf{B}) \\ &= f_1(\mathbf{A}) + \text{tr}((\mathbf{A}\mathbf{B}^T\mathbf{B} - \mathbf{G}_{opt}\mathbf{B})^T\mathbf{\Delta}) + \frac{1}{2}\text{tr}(\mathbf{\Delta}^T\mathbf{I}_N\mathbf{\Delta}\mathbf{B}^T\mathbf{B}) \\ &= f_1(\mathbf{A}) + \text{tr}((\mathbf{A}\mathbf{B}^T\mathbf{B} - \mathbf{G}_{opt}\mathbf{B})^T\mathbf{\Delta}) + \frac{1}{2}\text{vec}(\mathbf{\Delta})^T(\mathbf{B}^T\mathbf{B} \otimes \mathbf{I}_N)\text{vec}(\mathbf{\Delta}) \end{aligned} \tag{3.50}$$

$$
\begin{aligned}
f_2(\mathbf{A}+\mathbf{\Delta}) &= \frac{c_k}{2}\mathrm{tr}((\mathbf{A}+\mathbf{\Delta})^T(\mathbf{A}+\mathbf{\Delta})(\mathbf{A}+\mathbf{\Delta})^T(\mathbf{A}+\mathbf{\Delta})) \\
&\quad - c_k\mathrm{tr}((\mathbf{A}+\mathbf{\Delta})^T(\mathbf{A}+\mathbf{\Delta})) + \frac{c_k}{2}\mathrm{tr}(\mathbf{I}) \\
&= f_2(\mathbf{A}) + \mathrm{tr}(2c_k(\mathbf{A}\mathbf{A}^T\mathbf{A}-\mathbf{A})^T\mathbf{\Delta}) + c_k(\mathrm{tr}(\mathbf{A}^T\mathbf{A}\mathbf{\Delta}^T\mathbf{\Delta}) \\
&\quad + \mathrm{tr}(\mathbf{A}^T\mathbf{\Delta}\mathbf{A}^T\mathbf{\Delta}) + \mathrm{tr}(\mathbf{\Delta}^T\mathbf{A}\mathbf{A}^T\mathbf{\Delta}\mathbf{I}_R)) - c_k\mathrm{tr}(\mathbf{\Delta}^T\mathbf{\Delta}) \\
&= f_2(\mathbf{A}) + \mathrm{tr}(2c_k(\mathbf{A}\mathbf{A}^T\mathbf{A}-\mathbf{A})^T\mathbf{\Delta}) + \frac{2c_k}{2}(\mathrm{vec}(\mathbf{\Delta})^T(\mathbf{A}^T\mathbf{A}\otimes\mathbf{I}_N)\mathrm{vec}(\mathbf{\Delta}) \\
&\quad + \mathrm{vec}(\mathbf{\Delta}^T)^T(\mathbf{A}\otimes\mathbf{A}^T)\mathrm{vec}(\mathbf{\Delta}) + \mathrm{vec}(\mathbf{\Delta})^T(\mathbf{I}_R\otimes\mathbf{A}\mathbf{A}^T)\mathrm{vec}(\mathbf{\Delta}) \\
&\quad - \frac{1}{2}\mathrm{vec}(\mathbf{\Delta})^T\mathbf{I}_{NR}\mathrm{vec}(\mathbf{\Delta})) \\
&= f_2(\mathbf{A}) + \mathrm{tr}(2c_k(\mathbf{A}\mathbf{A}^T\mathbf{A}-\mathbf{A})^T\mathbf{\Delta}) \\
&\quad + \frac{1}{2}\mathrm{vec}(\mathbf{\Delta})^T 2c_k(\mathbf{A}^T\mathbf{A}\otimes\mathbf{I}_N + \mathbf{K}^T(\mathbf{A}\otimes\mathbf{A}^T) + \mathbf{I}_R\otimes\mathbf{A}\mathbf{A}^T - \frac{1}{2}\mathbf{I}_{NR})\mathrm{vec}(\mathbf{\Delta}),
\end{aligned}
$$

where K is a $NR\times RN$ matrix, which is called commutation matrix and transforms $\mathrm{vec}(\mathbf{\Delta})$ into $\mathrm{vec}(\mathbf{\Delta}^T)$: $K\mathrm{vec}(\mathbf{\Delta}) = \mathrm{vec}(\mathbf{\Delta}^T)$
So,

$$\nabla^2 f_1(\mathbf{A}) = (\mathbf{B}^T\mathbf{B}\otimes\mathbf{I}_N) \tag{3.51}$$

and

$$\nabla^2 f_2(\mathbf{A}) = 2c_k(\mathbf{A}^T\mathbf{A}\otimes\mathbf{I}_N + \mathbf{K}^T(\mathbf{A}\otimes\mathbf{A}^T) + \mathbf{I}_R\otimes\mathbf{A}\mathbf{A}^T - \frac{1}{2}\mathbf{I}_{NR}). \tag{3.52}$$

From 3.51:

$$||\nabla^2 f_1(\mathbf{A})||_2 \leq \lambda_{max}(\mathbf{B}^T\mathbf{B}\otimes\mathbf{I}_N) = \lambda_{max}(\mathbf{B}^T\mathbf{B}) = L_1. \tag{3.53}$$

From 3.52:

$$
\begin{aligned}
||\nabla^2 f_2(\mathbf{A})||_2 &\leq 2c_k(||\mathbf{A}^T\mathbf{A}\otimes\mathbf{I}_N + \mathbf{I}_R\otimes\mathbf{A}\mathbf{A}^T - \frac{1}{2}\mathbf{I}_{NR}||_2 + ||\mathbf{K}^T||_2||(\mathbf{A}\otimes\mathbf{A}^T)||_2) \\
&\leq 2c_k(\lambda_{max}(\mathbf{A}^T\mathbf{A}) + \lambda_{max}(\mathbf{A}^T\mathbf{A}) - \frac{1}{2} + \lambda_{max}(\mathbf{A}^T\mathbf{A})) \\
&= 2c_k(3\lambda_{max}(\mathbf{A}^T\mathbf{A}) - \frac{1}{2}) = L_2.
\end{aligned} \tag{3.54}
$$

From 3.52,3.51:

$$
\begin{aligned}
||\nabla^2 F(\mathbf{A})||_2 &\leq ||\nabla^2 f_1(\mathbf{A})||_2 + ||\nabla^2 f_2(\mathbf{A})||_2 \\
&= \lambda_{max}(\mathbf{B}^T\mathbf{B}) + 2c_k(3\lambda_{max}(\mathbf{A}^T\mathbf{A}) - \frac{1}{2}) = L.
\end{aligned} \tag{3.55}
$$

We can see that, $F(\mathbf{A})$ is twice differentiable and its Hessian matrix has a bounded spectral norm by a number L, so $F(\mathbf{A})$ has a Lipschitz continuous first derivative with Lipschitz constant L. However, we cannot conclude that $F(\mathbf{A})$ is a convex function because of the $f_2(\mathbf{A})$ term.

Now we consider that $\mathbf{A}$ is fixed. The problem is unconstrained.

$$\underset{\mathbf{B}}{\text{minimize}}\ \frac{1}{2}||\mathbf{A}\mathbf{B}^T - \mathbf{G}_{opt}||_F^2, \tag{3.56}$$

Algorithm 1 Penalty method for s-0NMF algorithm

```
procedure PENALTY METHOD(G_opt, c_k)
    initialize A_0, B_0, k = 0
    while terminating condition is FALSE do
        A_k = Projected Gradient (A_{k-1}, B_{k-1}, G_opt, c_k)
        B_k = (A_k^T A_k)^{-1} A_k^T G_opt
        k = k + 1
    end while
    return A_k
end procedure
```

Let $G : \mathbb{R}^{R\times R} \to \mathbb{R}_+$:

$$\begin{aligned} G(\mathbf{B}) &= \frac{1}{2}||\mathbf{A}\mathbf{B}^T - \mathbf{G}_{opt}|||_F^2 + \frac{c_k}{2}||\mathbf{A}^T\mathbf{A} - \mathbf{I}||_F^2 \\ &= \frac{1}{2}\mathrm{tr}(\mathbf{B}\mathbf{A}^T\mathbf{A}\mathbf{B}^T) - \mathrm{tr}(\mathbf{G}_{opt}^T\mathbf{A}\mathbf{B}^T) + \frac{1}{2}\mathrm{tr}(\mathbf{G}_{opt}^T\mathbf{G}_{opt}). \end{aligned} \tag{3.57}$$

The first derivative of $\mathbf{G}$ is:

$$\nabla G(\mathbf{B}) = \mathbf{B}^T\mathbf{A}^T\mathbf{A} - \mathbf{G}_{opt}^T\mathbf{A}. \tag{3.58}$$

The Hessian of $\mathbf{G}$ is:

$$\nabla^2 G(\mathbf{B}) = \mathbf{I}_R \otimes \mathbf{A}^T\mathbf{A}. \tag{3.59}$$

The Hessian has non-negative eigenvalues, so the G is a convex function. A solution of the problem 3.56 problem is $\mathbf{B} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{G}_{opt}$.

The problem 3.40 is solved by alternating least squares method. It is a two-step iterative optimization process. More specifically, we consider random initial matrices $\mathbf{A}_0, \mathbf{B}_0$. In each iteration, we estimate the matrix $\mathbf{A}$ by solving the problem 3.41 using a projected gradient method and considering that the matrix $\mathbf{B}$ is fixed and after that, we estimate the problem 3.56 considering that the matrix $\mathbf{A}$ is fixed. The problem 3.56 has a closed-form solution as we have already proved. The corresponding algorithms are described in Algorithm 1 and Algorithm 2.

Accelerated penalty method for semi-Orthogonal Nonnegative Matrix Factorization (s-ONMF)

In this section, we modify the cost function as proposed in [16]. The constraint $\mathbf{A}^T\mathbf{A} = \mathbf{I}$ can be divided into two parts. First part includes the constrains $\mathbf{a}_i^T\mathbf{a}_i = 1$, which is not important, because we can scale the columns of $\mathbf{A}$ and $\mathbf{B}$ without increasing the error. Second part is the constrains with form $\mathbf{a}_i^T\mathbf{a}_j = 0$ for any $i \neq j$. So the problem can be

Algorithm 2 Projected Gradient for **A** estimation

procedure PROJECTED GRADIENT($\mathbf{A}_0$,$\mathbf{B}$, $\mathbf{G}_{opt}$,c_k)
  initialize $k = 0$
  while terminating condition is FALSE **do**
    Compute $\nabla F(\mathbf{A}_{k-1})$
    $a_k = \frac{1}{\lambda_{max}(\mathbf{B}^T\mathbf{B})+2c_k(3\lambda_{max}(\mathbf{A}_{k-1}^T\mathbf{A}_{k-1})-\frac{1}{2}\mathbf{I}_R)}$
    $\mathbf{A}_k = \max(0,\ \mathbf{A}_{k-1} - a_k\nabla F(\mathbf{A}_{k-1}))$
    $k = k + 1$
  end while
  return $\mathbf{A}_k$
end procedure

expressed as:

$$\begin{aligned} &\underset{\mathbf{A},\mathbf{B}}{\text{minimize}}\ \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2 + \frac{c_k}{2}\sum_{i=1}^{R}\sum_{j\neq i=1}\mathbf{a}_i^T\mathbf{a}_j. \\ &\text{subject to} \quad \mathbf{A} \geq 0 \end{aligned} \tag{3.60}$$

The problem 3.60 can also be written as:

$$\begin{aligned} &\underset{\mathbf{A},\mathbf{B}}{\text{minimize}}\ \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2 + \frac{c_k}{2}\text{tr}(\mathbf{Q}\mathbf{A}^T\mathbf{A}), \\ &\text{subject to} \quad \mathbf{A} \geq 0 \end{aligned} \tag{3.61}$$

where $\mathbf{Q} = \mathbf{1}_R\mathbf{1}_R^T - \mathbf{I}_R$.

Firstly, we examine the problem 3.61 considering that $\mathbf{B}$ is fixed. So, the problem becomes:

$$\begin{aligned} &\underset{\mathbf{A}}{\text{minimize}}\ \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2 + \frac{c_k}{2}\text{tr}(\mathbf{Q}\mathbf{A}^T\mathbf{A}), \\ &\text{subject to} \quad \mathbf{A} \geq 0 \end{aligned} \tag{3.62}$$

Let $F : \mathbb{R}^{N\times R} \to \mathbb{R}_+$:

$$\begin{aligned} F(\mathbf{A}) &= \frac{1}{2}||\mathbf{G}_{opt} - \mathbf{A}\mathbf{B}^T||_F^2 + \frac{c_k}{2}\text{tr}(\mathbf{Q}\mathbf{A}^T\mathbf{A}) \\ &= \frac{1}{2}\text{tr}(\mathbf{G}_{opt}^T\mathbf{G}_{opt}) - \text{tr}(\mathbf{B}\mathbf{A}^T\mathbf{G}_{opt}) + \frac{1}{2}\text{tr}(\mathbf{B}\mathbf{A}^T\mathbf{A}\mathbf{B}^T) + \frac{c_k}{2}\text{tr}(\mathbf{A}\mathbf{Q}\mathbf{A}^T). \end{aligned} \tag{3.63}$$

The first derivative of $\mathbf{F}$ is:

$$\nabla F(\mathbf{A}) = \mathbf{A}\mathbf{B}^T\mathbf{B} - \mathbf{G}_{opt}\mathbf{B} + c_k\mathbf{A}\mathbf{Q}. \tag{3.64}$$

The Hessian of $\mathbf{G}$ is:

$$\nabla^2 F(\mathbf{A}) = \mathbf{B}\mathbf{B}^T \otimes \mathbf{I}_N + c_k\mathbf{Q} \otimes \mathbf{I}_N. \tag{3.65}$$

For each $\mathbf{A}_1 \in \mathbb{R}_+^{N\times R}$:

$$\begin{aligned}\text{vec}(\mathbf{A}_1)^T(\mathbf{B}\mathbf{B}^T \otimes \mathbf{I}_N + c_k\mathbf{Q} \otimes \mathbf{I}_N)\text{vec}(\mathbf{A}_1) &= \text{tr}(\mathbf{A}_1^T\mathbf{A}_1\mathbf{B}^T\mathbf{B}) + c_k\text{tr}(\mathbf{A}_1^T\mathbf{A}_1\mathbf{Q}) \\ &= \|\mathbf{B}\mathbf{A}_1^T\|_F^2 + c_k\text{tr}(\mathbf{A}_1^T\mathbf{A}_1\mathbf{Q}) \\ &\geq 0.\end{aligned} \tag{3.66}$$

The trace of matrix $\mathbf{A}_1^T\mathbf{A}_1\mathbf{Q}$ is non-negative, because the matrices $\mathbf{A}_1$, $\mathbf{Q}$ have non-negative elements. So, the function F is convex for the set of $\mathbf{A} \in \mathbb{R}+^{N\times R}$ and considering that $\mathbf{B}$ is fixed. Also, we can observe, that: $\|\nabla F(\mathbf{A})\|_2 \leq \lambda_{max}(\mathbf{B}\mathbf{B}^T + c_k\mathbf{Q}) = L_A$. Thereby, the function F has a Lipschitz continuous first derivative with Lipschitz constant L_A.

Now we consider that $\mathbf{A}$ is fixed. The problem is unconstrained.

$$\underset{\mathbf{B}}{\text{minimize}}\ \frac{1}{2}||\mathbf{A}\mathbf{B}^T - \mathbf{G}_{opt}||_F^2, \tag{3.67}$$

Let $G : \mathbb{R}^{R\times R} \to \mathbb{R}_+$:

$$\begin{aligned}G(\mathbf{B}) &= \frac{1}{2}||\mathbf{A}\mathbf{B} - \mathbf{G}_{opt}||_F^2 \\ &= \frac{1}{2}\text{tr}(\mathbf{B}\mathbf{A}^T\mathbf{A}\mathbf{B}^T) - \text{tr}(\mathbf{G}_{opt}^T\mathbf{A}\mathbf{B}^T) + \frac{1}{2}\text{tr}(\mathbf{G}_{opt}^T\mathbf{G}_{opt}).\end{aligned} \tag{3.68}$$

The first derivative of $\mathbf{G}$ is:

$$\nabla G(\mathbf{B}) = \mathbf{B}\mathbf{A}^T\mathbf{A} - \mathbf{G}_{opt}^T\mathbf{A}. \tag{3.69}$$

The Hessian of $\mathbf{G}$ is:

$$\nabla^2 G(\mathbf{B}) = \mathbf{I}_R \otimes \mathbf{A}^T\mathbf{A}. \tag{3.70}$$

The Hessian has non-negative eigenvalues, so G is a convex function. A solution of the problem 3.67 problem is $\mathbf{B} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{G}_{opt}^T\mathbf{A}$.

We solve a sequence of subproblems, which have the form of the problem 3.60. In each iteration, we increase the value of penalty constant c_k and solve the corresponding subproblem. In the first iteration we consider that initial points for the problem 3.60 are random matrices, for the rest iterations the initial points will be the solutions of previous subproblem. The subproblem 3.60 is solved by using alternating least squares. It is a two-step iterative optimization process. More specifically, in each iteration we estimate the matrix $\mathbf{A}$ by solving the problem 3.61 using a projected gradient method and considering that the matrix $\mathbf{B}$ is fixed and after that we estimate the problem 3.67 considering that the matrix $\mathbf{A}$ is fixed. The problem 3.67 has a closed-form solution as we have already proved. The corresponding algorithms are described in Algorithm 3 and Algorithm 4.

Algorithm 3 Accelerated Penalty Method for s-0NMF algorithm

procedure ACCELERATED PENALTY METHOD($\mathbf{G}_{opt}$,$\mathbf{A}_0$, $\mathbf{B}_0$ c_k, a)
  while terminating condition is FALSE **do**
    while terminating condition is FALSE **do**
      $\mathbf{A}_k$ = Nesterov Projected Gradient ($\mathbf{A}_{k-1}$,$\mathbf{B}_{k-1}$, $\mathbf{G}_{opt}$,c_k, α)
      normalization of matrix $\mathbf{A}_k$
      $\mathbf{B}_k = (\mathbf{A}_k^T\mathbf{A}_k)^{-1}\mathbf{G}_{opt}^T\mathbf{A}_k$
      $k = k + 1$
    end while
    $c_k = a * c_k$
    $k = 0$
    $\mathbf{A}_0 = \mathbf{A}_k$
    $\mathbf{B}_0 = \mathbf{B}_k$
  end while
  return $\mathbf{A}_k$
end procedure

Algorithm 4 Nesterov Accelerated Projected Gradient for $\mathbf{A}$ estimation

procedure NESTEROV PROJECTED GRADIENT($\mathbf{A}_0$,$\mathbf{B}$, $\mathbf{G}_{opt}$,c_k,α_0)
  initialize $k = 1$,$\mathbf{X}_0 = A_0$
  $q = \frac{\lambda_{max}(\mathbf{B}\mathbf{B}^T\otimes\mathbf{I}_N+c_k\mathbf{Q}\otimes\mathbf{I}_N)}{\lambda_{min}(\mathbf{B}\mathbf{B}^T\otimes\mathbf{I}_N+c_k\mathbf{Q}\otimes\mathbf{I}_N)}$
  while terminating condition is FALSE **do**
    Compute $\nabla F(\mathbf{A}_{k-1})$
    $a_k = \frac{1}{\lambda_{max}(\mathbf{B}\mathbf{B}^T\otimes\mathbf{I}_N+c_k\mathbf{Q})}$
    $\mathbf{X}_k = \max(0,\ \mathbf{A}_{k-1} - a_k\nabla F(\mathbf{A}_{k-1}))$
    Compute $\alpha_k \in (0,1)$ as the solution of equation: $\alpha_k^2 - (1-\alpha_k)\alpha_{k-1}^2 - q\alpha_k = 0$
    $\beta_{k-1} = \frac{\alpha_{k-1}(1-\alpha_{k-1})}{\alpha_{k-1}^2+\alpha_k}$
    $\mathbf{A}_k = \mathbf{X}_k + \beta_{k-1}(\mathbf{X}_k - \mathbf{X}_{k-1})$
    $k = k + 1$
  end while
  return $\mathbf{A}_k$
end procedure

Estimation of the temporal components $\mathbf{S}_k$

Finally, after the estimation of the common subspace factor $\mathbf{A}$, we can estimate the temporal components $\mathbf{S_k}$, for $k = 1, \ldots, K$. Each of the term $\mathbf{S}_k$ is the solution of the least square problem:

$$\begin{aligned} &\underset{\mathbf{A},\mathbf{S}_k}{\text{minimize}}\ \frac{1}{2}||\mathbf{X}_k - \mathbf{A}\mathbf{S}_k^T||_F^2, \\ &\text{subject to}\quad \mathbf{A}^T\mathbf{A} = \mathbf{I},\quad \mathbf{A} \geq \mathbf{0} \end{aligned} \tag{3.71}$$

where $\mathbf{A}$ is fixed and equal to the $\mathbf{A}_{est}$, which is the solution of the factorization problem. So, $\mathbf{S}_k = (\mathbf{A}_k^T\mathbf{A}_k)^{-1}\mathbf{A}_k^T\mathbf{X}_k$.

Chapter 4

Evaluation

4.1 Synthetic Data

In this section, we examine the effectiveness of the proposed model on synthetic data. We generate our data based on the 3.11 model for three different cases.

1. **Case 1: Common orthonormal matrix A with each subject having independent temporal components $\mathbf{S}_k$.**

$$\mathbf{X}_k = \mathbf{A}\mathbf{S}_k^T + \beta_k \mathbf{E}_k, \tag{4.1}$$

where

- $\mathbf{A}$ meets the orthonormality constraints $\mathbf{A}^T\mathbf{A} = \mathbf{I}_R$. More specifically, each row of $\mathbf{A}$ has only one nonzero element the position of which is chosen randomly. The values of nonzero elements are drawn from distribution $\mathcal{U}[0,1]$. Next, we normalize each column of $\mathbf{A}$.
- $\mathbf{E}_k$, for $k = 1, \ldots, K$, has i.i.d $\mathcal{N}(0,1)$ elements.
- $\mathbf{S}_k(:,r)$, for $k = 1, \ldots, K$ and $r = 1, \ldots, R$, has i.i.d $\mathcal{N}(0,\sigma_r^2)$ elements. The value of σ_r^2 can vary.
- β_k is a scalar, which determines the SNR.

We define the SNR as:

$$\mathrm{SNR} = \frac{\sum_{\mathrm{k}=1}^{\mathrm{K}} ||\mathbf{A}\mathbf{S}_{\mathrm{k}}^{\mathrm{T}}||_{\mathrm{F}}^2}{\beta_{\mathrm{k}}^2 \sum_{\mathrm{k}=1}^{\mathrm{K}} ||\mathbf{E}_{\mathrm{k}}||_{\mathrm{F}}^2}. \tag{4.2}$$

2. **Case 2: Common orthonormal matrix A with each subject having correlated temporal components $\mathbf{S}_k$.**

$$\mathbf{X}_k = \mathbf{A}\mathbf{S}_k^T + \beta_k \mathbf{E}_k, \tag{4.3}$$

where

- $\mathbf{A}$ meets the orthonormality constraints $\mathbf{A}^T\mathbf{A} = \mathbf{I}_R$. More specifically, each row of $\mathbf{A}$ has only one nonzero element the position of which is chosen randomly. The values of nonzero elements are drawn from distribution $\mathcal{U}[0,1]$. Next, we normalize each column of $\mathbf{A}$.
- $\mathbf{E}_k$, for $k = 1, \ldots, K$, has i.i.d $\mathcal{N}(0,1)$ elements.

- $\mathbf{S}_k(:, r)$ is correlated to other $R-1$ columns of the temporal component of the k^{th} subject, for $k = 1, \ldots, K$ and $r = 1, \ldots, R$. The correlation among temporal components is provided in section A.1.
- β_k is a scalar, which determines the SNR.

3. **Case 3: Corrupted common orthonormal matrix A with each subject having correlated temporal components $\mathbf{S}_k$.**

$$\mathbf{X}_k = \mathbf{A}_k \mathbf{S}_k^T + \beta_k \mathbf{E}_k, \tag{4.4}$$

where

- $\mathbf{A}_k$, for $k = 1, \ldots, K$, is a corrupted version of a matrix $\mathbf{A}$, which meets the orthonormality constraints $\mathbf{A}^T\mathbf{A} = \mathbf{I}_R$. Each row of $\mathbf{A}$ has only one nonzero element the position of which is chosen randomly. The values of nonzero elements are drawn from distribution $\mathcal{U}[0, 1]$. Next, we normalize each column of $\mathbf{A}$. We generate $\mathbf{A}_k$ by choosing randomly $10 - 15\%$ of the non-negative elements of each column of matrix $\mathbf{A}$ and assign them to another column randomly. Finally, we normalize the columns of $\mathbf{A}_k$.
- $\mathbf{E}_k$, for $k = 1, \ldots, K$, has i.i.d $\mathcal{N}(0, 1)$ elements.
- $\mathbf{S}_k(:, r)$ is correlated to other $R-1$ columns of the temporal component of the k^{th} subject, for $k = 1, \ldots, K$ and $r = 1, \ldots, R$. The correlation among temporal components is provided in section A.1.
- β_k is a scalar, which is vary depending the SNR.

We define the SNR as:

$$\mathrm{SNR} = \frac{\sum_{\mathrm{k}=1}^{\mathrm{K}} ||\mathbf{A}_\mathrm{k}\mathbf{S}_\mathrm{k}^\mathrm{T}||_\mathrm{F}^2}{\beta_\mathrm{k}^2 \sum_{\mathrm{k}=1}^{\mathrm{K}} ||\mathbf{E}_\mathrm{k}||_\mathrm{F}^2}. \tag{4.5}$$

For our experiments, we set $N = 10^5$, $M = 80$, $K = 25$, and $R = 20$.

4.1.1 Experimental evaluation of gap criterion

We examine how accurately the gap criterion can estimate the number of clusters for different cases of noise and for the three cases of synthetic data. More specifically, we validate the criterion for noiseless case and for various values of the SNR.

As mentioned above, we have to partition our data into two sets in order to apply the gap criterion. We apply the criterion 3.37 for 25 different random partitions of our data in order to examine if the criterion is affected by the way into which the data are partitioned. We repeat the above procedure for various noise cases. We present the results for noiseless case and the case in which the SNR is equal to 0 dB. The rest of the cases are given in section A.2. Finally, it is worth being examined how close the column spaces of $\mathbf{A}$ and estimated $\mathbf{G}_{opt}$ are for the true common subspace dimension, $R = 20$, and for different cases of noise.

1. **Case 1: Common orthonormal matrix A with each subject having independent temporal components $\mathbf{S}_k$.**

   In this case, we also examine the behavior of the criterion when the rank-one terms of $\mathbf{AS}_k^T$ are of equal power and when the powers of the rank-one terms of $\mathbf{AS}_k^T$ are unequal.

   (a) Noiseless case

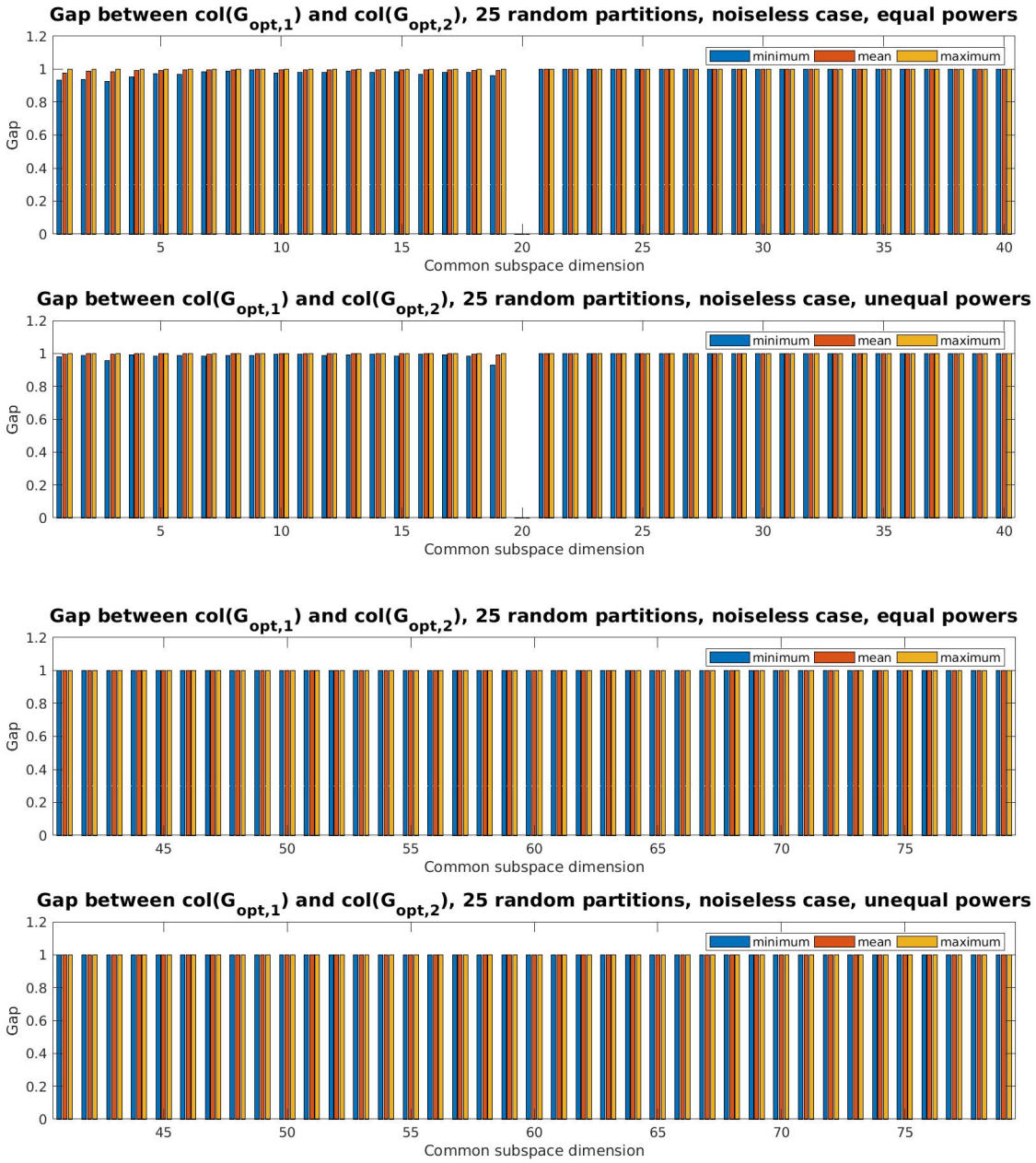


Figure 4.1: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for noiseless case. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension. **First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

(b) SNR=0 dB

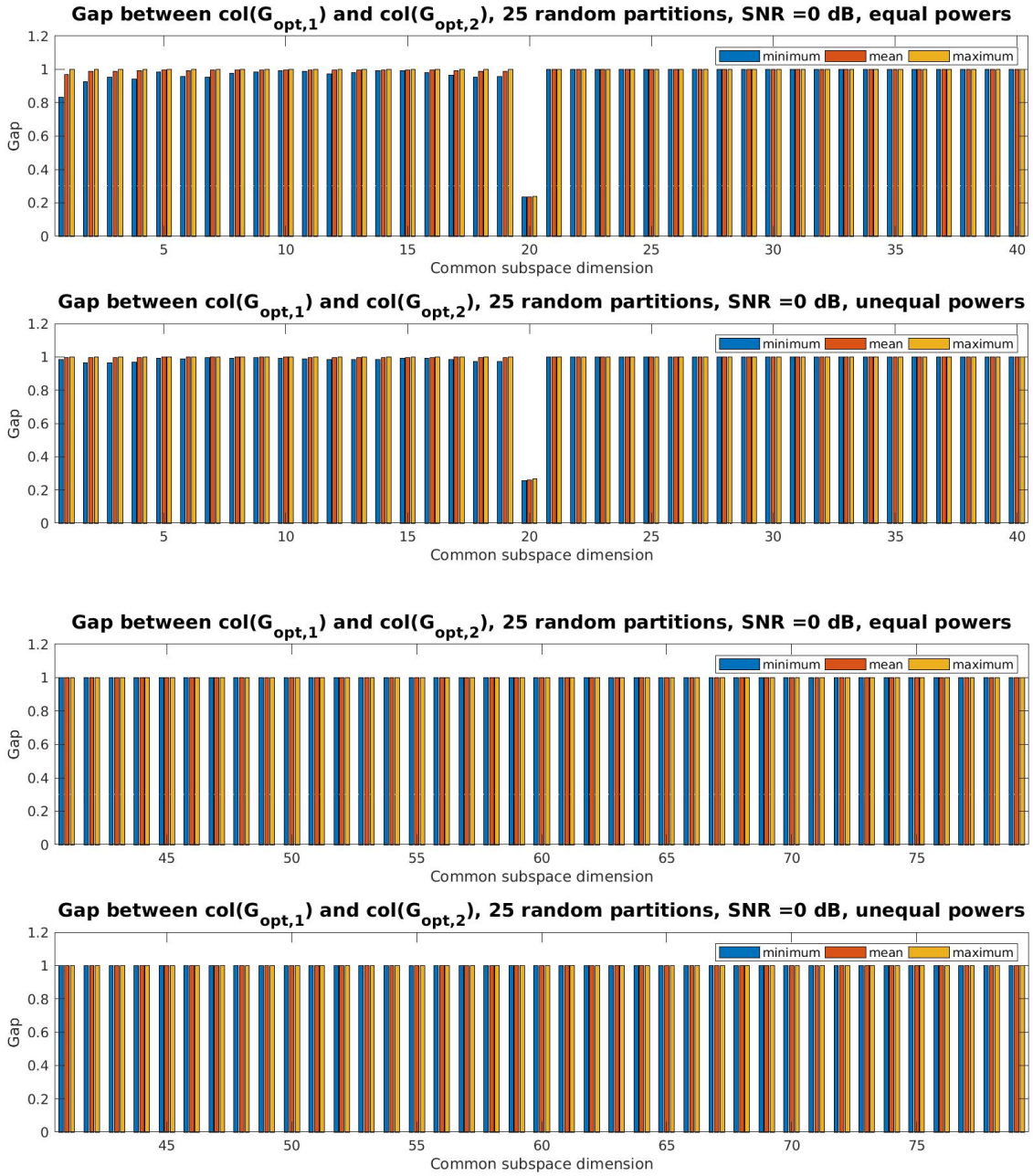


Figure 4.2: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension, for SNR=0 dB. **First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

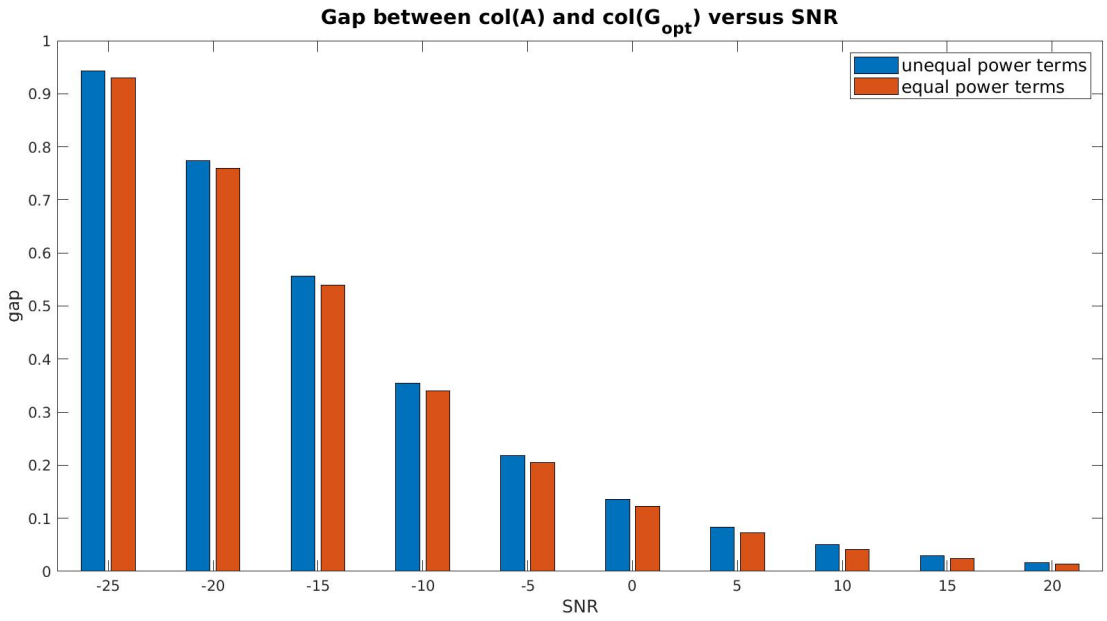


Figure 4.3: We compute the gap between column space of $\mathbf{A}$ and column space of $\mathbf{G}_{opt}$, for R=20, versus SNR. Red color: equal power terms. Blue color: unequal power terms.

The images for the rest of noise cases are given in the subsection A.2.1. We gather the results of all experiments for each SNR and each common subspace dimension in Tables A.1, A.2, A.3, A.4, A.6, A.5. We see that the experimental results and the theory are consistent. Namely, we observe that, the value of the gap between $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ for common subspace dimensions smaller than the true dimension ranges. Also, when the common subspace dimension is greater than the true dimension, the value of the gap is really close to 1 and when the common subspace dimension is equal to true dimension the value of the gap depends on the power of noise. More specifically, for noiseless case the value of the gap is 0 and as the power of noise is increasing the value of the gap is also increasing. We note that, when SNR is equal to -25 dB the criterion cannot indicate correctly the true common subspace dimension. Finally, for the case in which the rank-one terms are of equal power, the values of the gap for common subspace dimension equal to true dimension are slightly smaller than the case in which the rank-one terms are of unequal power. From Figure 4.3, we also observe that, even if we can estimate correctly the common subspace dimension for SNR greater than -25 dB, the maximum minimum distance of col($\mathbf{A}$) from col($\mathbf{G}_{opt}$) is quite large for SNR smaller than -5 dB. This means that, at least the pair of principal vectors, which corresponds to the greatest canonical angle, is not really close. So, the two subspaces differ and we cannot guarantee that the col($\mathbf{G}_{opt}$) is a good estimate of col($\mathbf{A}$). Moreover, the estimation of col($\mathbf{A}$) is slightly more accurate for the case in which the rank-one terms are of equal power.

2. **Case 2: Common orthonormal matrix A with each subject having correlated temporal components $\mathbf{S}_k$.**

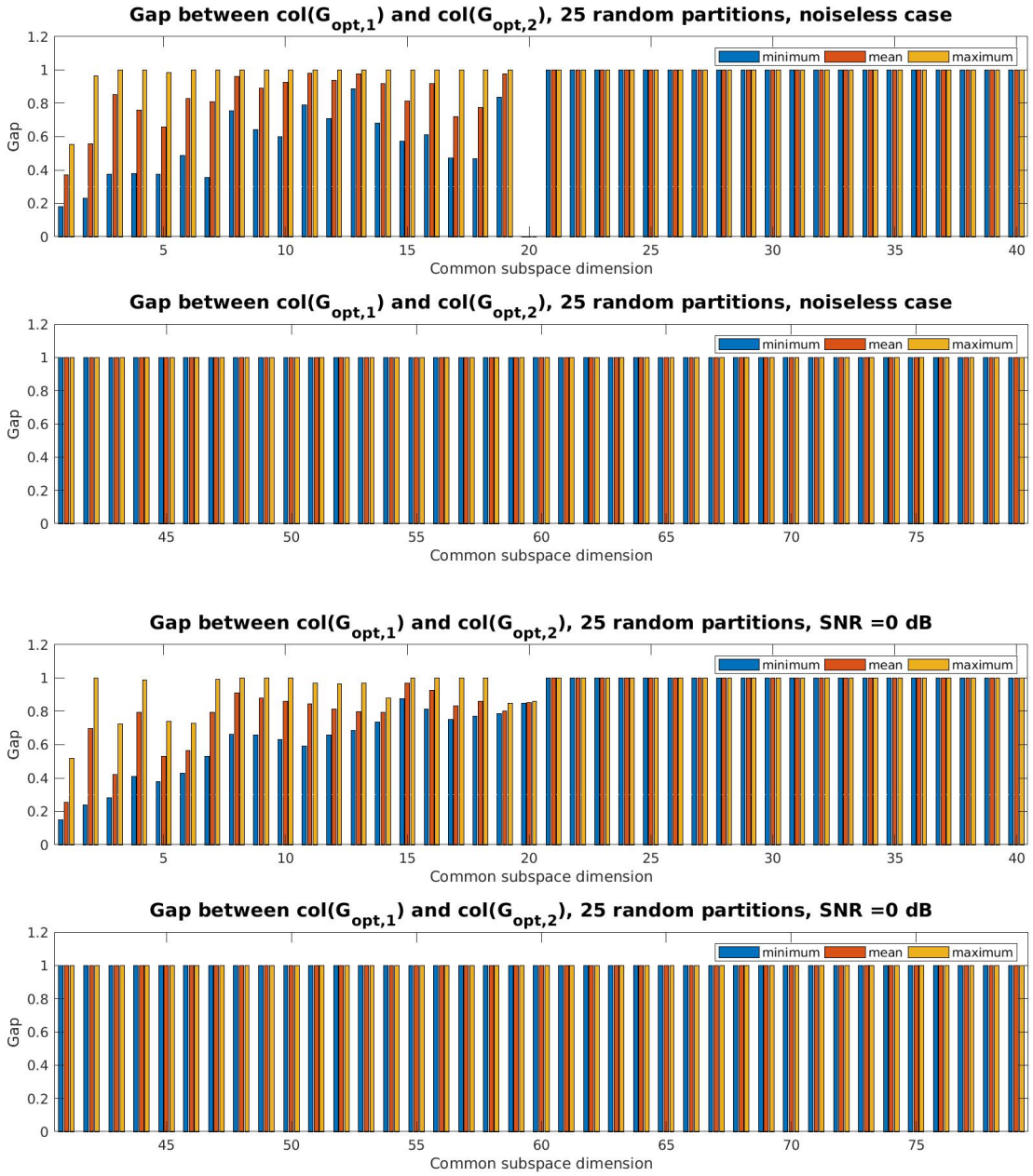


Figure 4.4: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First and Second image**: noiseless case. **Third and Fourth image**: SNR=0 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

The images for the rest of noise cases are provided in the subsection A.2.2. We gather the results of all experiments for each SNR and each common subspace dimension in Tables A.7, A.9, A.8. We see that the experimental results and the theory are consistent. Namely, we observe that, the value of the gap between $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ for common subspace dimensions smaller than the true dimension ranges. Also, when the common subspace dimensions are greater than the true dimension, the value of the gap is really close to 1 and when the common subspace

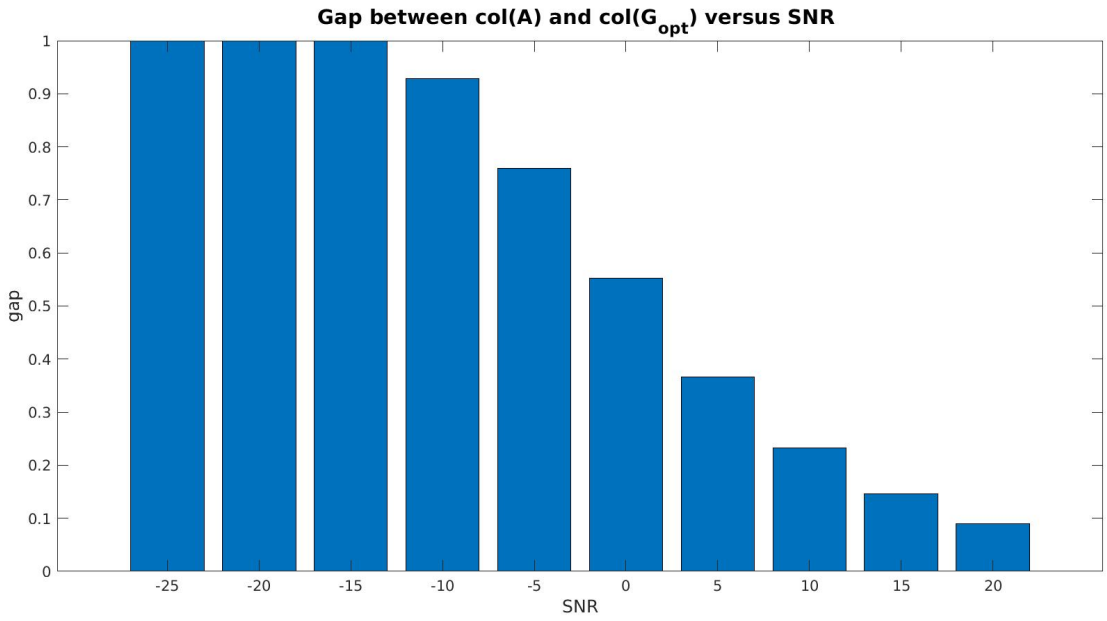


Figure 4.5: We compute the gap between column space of $\mathbf{A}$ and column space of $\mathbf{G}_{opt}$, for R=20, versus SNR.

dimension is equal to true dimension, the value of the gap depends on the power of noise. More specifically, for noiseless case the value of the gap is 0 and as the power of noise is increasing the value of the gap is also increasing. We note that, when SNR is equal or smaller to -10 dB the criterion cannot indicate correctly the true common subspace dimension. In the case of SNR$=-10$ dB, the mean gap for common subspace dimension equal to 20 is smaller than 1, but there is a partition for which gap for common subspace dimension equal to 20 (maximum value of gap for dimension equal to 20 is equal to 1) is equal to 1 and therefore, we cannot conclude that common subspace dimension is 20. From Figure 4.5, we also observe that even if we can estimate correctly the common subspace dimension for SNR greater than -10 dB, the maximum minimum distance of col($\mathbf{A}$) from col($\mathbf{G}_{opt}$) is quite large for SNR smaller than 0 dB. This means that, at least the pair of principal vectors, which corresponds to the greatest canonical angle, is not really close. So, the two subspaces differ and we cannot guarantee that the col($\mathbf{G}_{opt}$) is a good estimate of col($\mathbf{A}$).

3. **Case 3: Corrupted common orthonormal matrix A with each subject having correlated temporal components $\mathbf{S}_k$.**

The images for the rest of noise cases are given in the subsection A.2.3. We gather the results of all experiments for each SNR and each common subspace dimension in Tables A.10, A.12, A.11. In this case, the subjects do not share the exactly same matrix $\mathbf{A}$. Each $\mathbf{A_k}$ differs randomly $10-15\%$ from matrix $\mathbf{A}$. So, the experimental results and the theory are not complete consistent. Namely, we observe that, the value of the gap between $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ for noiseless case and for common subspace dimension 20 are 0.23, which is expected. However, for common subspace dimensions smaller than the true dimension the value of gap ra-

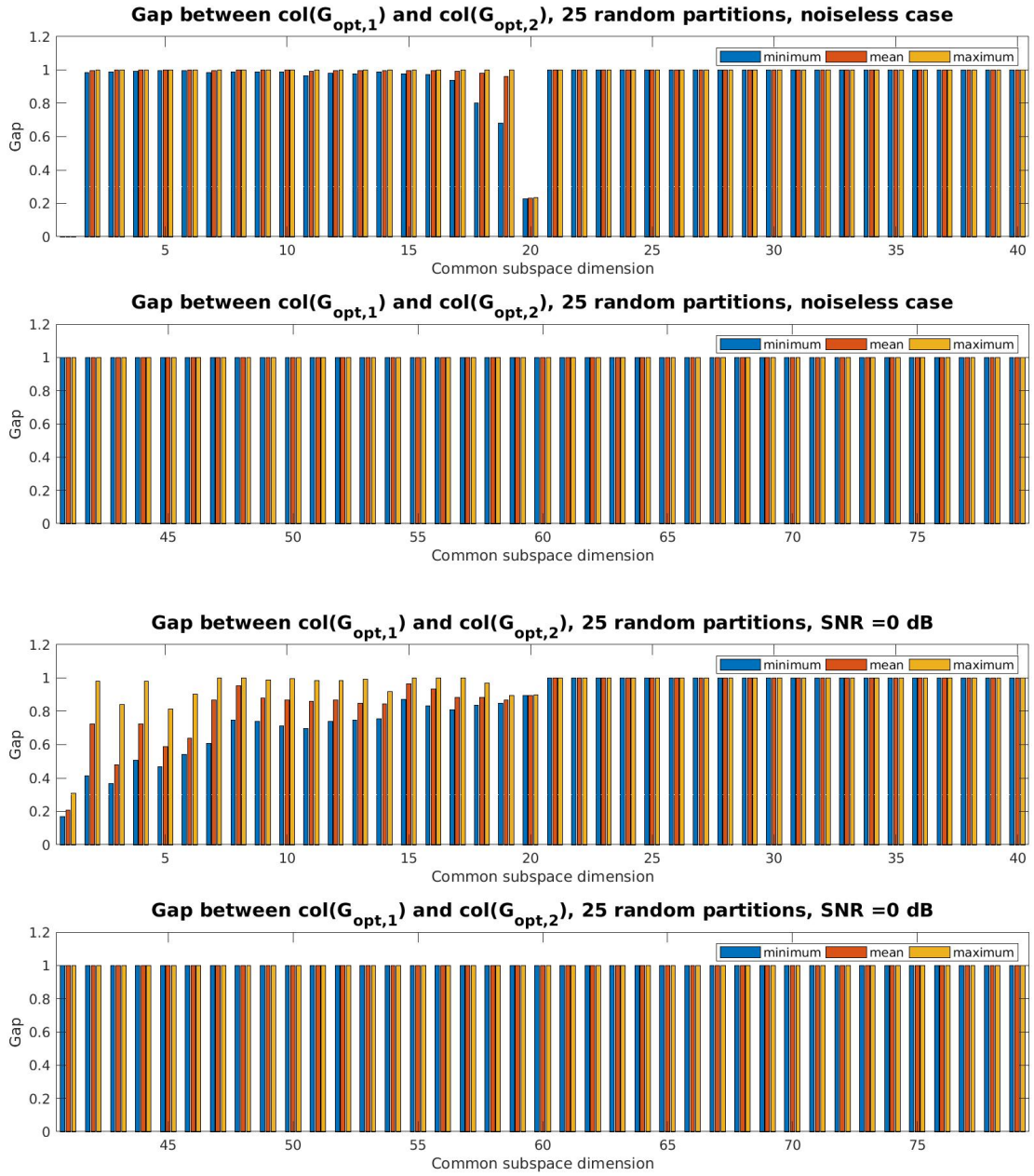


Figure 4.6: We consider 25 random partitions of data. For each and every of the partitions, we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension. **First and Second image**: noiseless case. **Third and Fourth image**: SNR=0 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

nges, for common subspace dimensions greater than the true dimension, the value of gap is really close to 1 and for common subspace dimension equal to true dimension the value of the gap is depends on the power of noise. More specifically, as the power of noise is increasing the value of the gap is also increasing. We note that, when SNR is equal or smaller to -10 dB the criterion cannot indicate correctly the true common subspace dimension. In the case of SNR$=-10$ dB, the mean gap for common subspace dimension equal to 20 is smaller than 1, but there is a partition

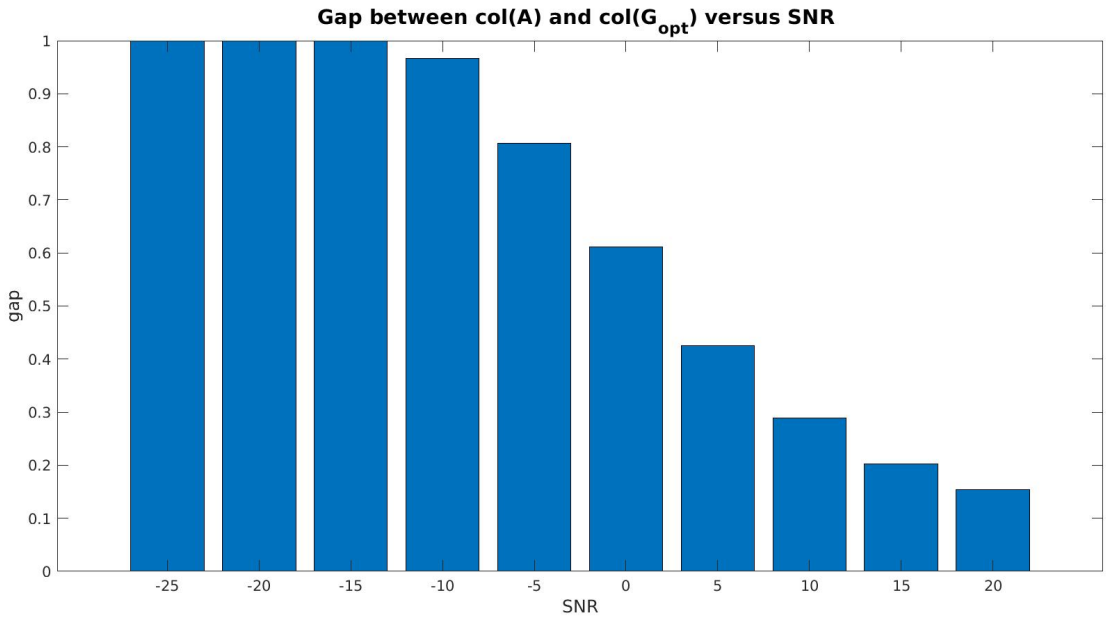


Figure 4.7: We compute the gap between column space of $\mathbf{A}$ and column space of $\mathbf{G}_{opt}$, for R=20, versus SNR.

for which gap for common subspace dimension equal to 20 (maximum value of gap for dimension equal to 20 is equal to 1) is equal to 1 and therefore we cannot conclude that common subspace dimension is 20. From Figure 4.7, we also observe that, even if we can estimate correctly the common subspace dimension for SNR greater than -10 dB, the maximum minimum distance of col($\mathbf{A}$) from col($\mathbf{G}_{opt}$) is quite large for SNR smaller than 10 dB. This means that, at least the pair of principal vectors, which corresponds to the greatest canonical angle, is not really close. So, the two subspaces differ and we cannot guarantee that the col($\mathbf{G}_{opt}$) is a good estimate of col($\mathbf{A}$)

4.1.2 Experimental evaluation of the estimation of the common spatial factor A

In this subsection, we examine the performance of the two algorithms 1, 3 for the case of the semi-Orthogonal Non-negative Matrix Factorization (s-ONMF) problem 3.38. More specifically, we investigate via experiments how efficiently these algorithms estimate the common subspace factor $\mathbf{A}$ of matrix $\mathbf{G}_{opt}$. To this end, we solve the problem five times via each algorithm. The initial starting points for each of the total ten experiments are different as they are generated randomly. We do that in order to examine how the initial points affect the solution. So, for each algorithm, we examine for each of the five solutions how close it is to the rest of them. Moreover, we calculate the relative error using the formula:

$$\frac{||\mathbf{G}_{opt} - \mathbf{A}_{est}\mathbf{B}_{est}^T||_F}{||\mathbf{G}_{opt}||_F}. \tag{4.6}$$

Also, we choose the best and worst solution based on relative error to examine how close they are to the real common spatial factor $\mathbf{A}$. Moreover, for the best solution we also estimate the temporal components $\mathbf{S}_k$ for each subject. Finally, we project the data $\mathbf{X}_k$

on the common subspace $\mathbf{G}_{opt}$.

We run these experiments for noiseless data and data with SNR= 0 dB, for each of the three synthetic data cases. For the case in which SNR= 0 dB, for each cluster r and subject k, for $r = 1, \ldots, R$ and $k = 1, \ldots, K$, we calculate the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We then calculate the correlation coefficients again, this time for the projected data on common subspace $\mathbf{G}_{opt}$. For each cluster r, for $r = 1, \ldots, R$, we gather the coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial and projected data. Given the fact that $\mathbf{A}$ indicates to which cluster each voxel belongs, we follow the aforementioned procedure for the best $\mathbf{A}_{est}$ for each of the two algorithms and compare the results.

1. **Case 1: Common orthonormal matrix A with each subject having independent temporal components $\mathbf{S}_k$.**

   (a) **Penalty method**

   Firstly, we examine for each solution how close it is to the rest of other four solutions. To this end, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 5$, we calculate the correlation coefficients between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 5$ and $j \neq i$, with which the column r has the greatest correlation coefficient. Also, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 5$, we calculate the intersection of non-zero elements between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 5$ and $j \neq i$, with which the column r has the greatest intersection. We examine if the correspondence between columns is the same for the two cases (correlation coefficients, intersection). We repeat the aforementioned steps for noiseless data and data with SNR= 0 dB. We illustrate the results between solution 1 and 2 in Figure 4.8 and the rest of them in the subsection A.3.1.

   We observe that the five solutions, even for the noiseless data, are different. Also, we see that, more than one columns of a solution j can associate with the same column of solution i. Moreover, a column of the solution i can associate with a column r_1 of the solution j based on correlation coefficients and to a different column r_2 of the solution j based on the intersection. Finally, if the column r of solution i is associated with the column r_1 of the solution j, it is not necessary for the column r_1 of solution j to associate with the column r of solution i.

   After that, we calculate relative error of solutions (Figure 4.9).

   For noiseless case, the solution 1 has the smallest relative error and thus, we consider it as the best solution. On the other hand, the solution 2 has the biggest relative error and thus, we consider it as the worst solution. For SNR= 0 dB, the best solution is the solution 3 and the worst is the solution

4. For each and every of them, we examine how close they are with the true common spatial factor **A**. More specifically, for every column r of the true **A**, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each and every column of the solutions 1, 2 for noiseless case and the solutions 3, 4 for SNR= 0 dB. We illustrate for each column r of the true **A**, the number of elements of its intersection with all columns of the solution and the number of the non-zero elements of each column of the solutions. The images are provided in the subsection A.3.1 and here we provide the corresponding Tables.

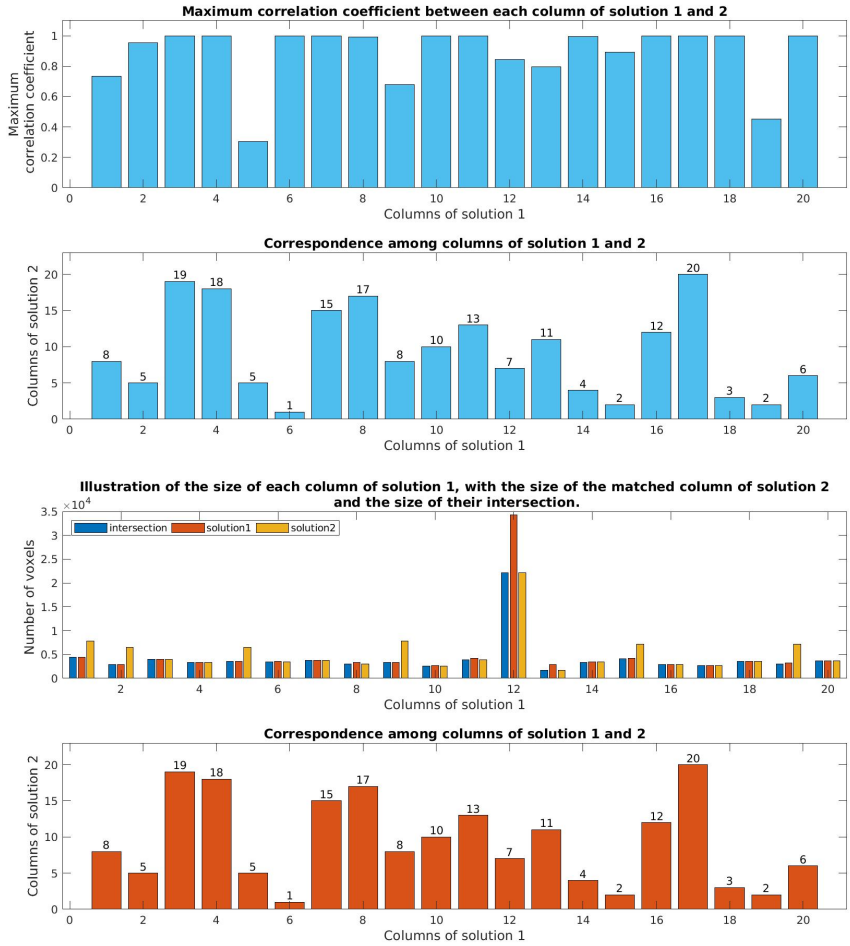


Figure 4.8: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates to which column of the solution 2 each column of the solution 1 corresponds based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates to which column of the solution 2 each column of the solution 1 corresponds based on intersection. Solutions 1, 2 correspond to penalty method and noiseless data.

		COLUMN OF $\mathbf{A}_{est}$																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
	1	0	0	0	3	0	0	0	0	0	2	4164	0	0	0	0	0	0	0	0	0	4169
	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4219	0	1	0	0	0	4220
	3	0	0	0	0	3569	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3569
	4	0	0	0	0	0	0	0	0	0	26	0	0	0	3391	0	0	0	0	152	0	3569
C	5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3082	0	3083
O	6	0	0	0	0	0	3518	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3518
L	7	0	0	0	0	0	0	0	0	3377	0	0	0	0	0	0	0	0	0	0	0	3377
U	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3504	0	0	3504
M	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3653	3653
N	10	0	0	0	0	0	0	3816	0	0	0	0	0	0	0	0	0	0	0	0	0	3816
	11	4437	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	4441
O	12	0	2915	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	2918
F	13	0	0	0	0	0	3	0	0	0	2693	0	0	0	0	0	0	0	0	0	0	2696
	14	0	0	0	3341	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3341
A	15	0	0	0	0	0	0	0	3367	0	0	0	0	0	0	0	0	0	0	0	0	3367
	16	0	0	4009	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	4010
	17	0	0	0	0	0	0	0	0	0	0	0	0	2862	0	0	0	0	0	0	1	2863
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2682	0	1	0	2683
	19	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2873	0	0	0	0	2874
	20	0	0	0	0	0	0	1	0	0	0	0	34327	1	0	0	0	0	0	0	0	34329
		4437	2915	4009	3344	3569	3521	3817	3367	3377	2723	4164	34327	2866	3391	4219	2873	2688	3504	3235	3654	

Table 4.1: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 1 for noiseless case. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 1 for noiseless case.

		COLUMN OF $\mathbf{A}_{est}$																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
	1	0	0	0	0	0	0	0	0	316	0	0	0	3850	0	0	0	3	0	0	0	4169
	2	0	4144	0	0	0	0	0	0	0	0	0	0	0	0	0	57	19	0	0	0	4220
	3	0	0	0	0	3561	0	0	0	0	0	3	0	0	1	0	4	0	0	0	0	3569
	4	0	0	0	3492	0	0	0	0	0	0	0	0	0	0	0	77	0	0	0	0	3569
C	5	0	3035	0	0	0	0	0	0	0	0	0	0	0	48	0	0	0	0	0	0	3083
O	6	3392	0	0	0	0	0	0	0	126	0	0	0	0	0	0	0	0	0	0	0	3518
L	7	0	0	0	0	0	0	0	3338	2	0	0	0	0	37	0	0	0	0	0	0	3377
U	8	0	0	3504	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3504
M	9	0	0	0	0	0	3653	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3653
N	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3815	0	1	0	0	0	3816
	11	0	0	0	0	0	0	0	4441	0	0	0	0	0	0	0	0	0	0	0	0	4441
O	12	0	0	0	0	2918	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2918
F	13	0	0	0	0	0	0	0	0	0	2614	0	0	0	0	0	82	0	0	0	0	2696
	14	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3340	0	0	3341
A	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	338	3029	0	0	0	3367
	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4010	0	4010
	17	0	0	0	0	0	0	0	0	2	0	1690	0	0	1171	0	0	0	0	0	0	2863
	18	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	2653	2683
	19	0	0	0	0	0	0	0	0	0	0	1	2873	0	0	0	0	0	0	0	0	2874
	20	0	0	0	0	0	0	22097	0	38	0	0	0	0	0	0	12194	0	0	0	0	34329
		3392	7179	3504	3492	6479	3653	22097	7779	514	2614	1695	2873	3850	1257	3815	12752	3052	3340	4010	2653	

Table 4.2: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 2 for noiseless case. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 2 for noiseless case.

	COLUMN OF $\mathbf{A}_{est}$																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	4	7	9	8	5	6	8	5	4	4	4	6	5	11	1	4057	4	3	10	8	4169
2	13	3	4110	9	7	2	4	7	10	6	6	1	5	9	5	6	1	4	8	4	4220
3	2	4	3	1	4	1	5	3	9	2	6	1	5	6	4	1	2	3501	5	4	3569
4	7	8	8	3469	2	6	6	5	5	4	5	7	5	8	6	1	1	4	8	4	3569
C 5	5	6	5	8	3	1	4	4	5	11	3	1	3008	1	3	3	2	4	1	5	3083
O 6	27	21	26	42	22	3	21	18	23	11	14	19	18	28	23	11	3127	18	18	28	3518
L 7	4	3296	4	3	4	4	3	7	7	9	5	5	1	3	4	4	2	5	4	3	3377
U 8	6	5	5	3	2	0	3	1	5	7	3	3	5	3440	1	4	1	3	4	3	3504
M 9	2	0	3	8	7	12	7	5	8	3555	2	5	3	5	4	4	4	9	3	7	3653
N 10	2	4	3	10	7	11	6	4	3	6	3721	3	1	5	6	6	3	6	6	3	3816
11	7	7	11	8	4	14	4291	8	9	13	8	2	11	5	8	7	3	10	7	8	4441
O 12	5	4	2	5	2	4	2	2	4	4	0	2864	3	2	2	4	1	4	2	2	2918
F 13	2641	1	3	4	0	5	4	1	4	5	3	1	1	4	6	2	4	0	5	2	2696
14	7	8	4	13	0	7	5	4	3249	9	6	6	3	5	3	1	2	3	0	6	3341
A 15	4	6	3	0	7	0	9	6	3	3	4	4	4	5	3296	1	0	2	1	9	3367
16	4	7	2	7	8	1070	10	6	5	7	6	2	6	5	3	6	4	6	2	2844	4010
17	1	1	8	5	3	1	4	2800	5	4	3	4	6	2	6	0	2	3	1	4	2863
18	6	3	0	2	2	6	6	2	2	4	3	2	5	3	3	3	4	3	2622	2	2683
19	3	6	2	6	2809	8	3	3	4	6	5	2	2	2	1	4	1	1	1	5	2874
20	149	177	137	190	153	239	176	137	184	121	143	114	160	169	161	117	31381	155	112	154	34329
	2899	3574	4348	3801	3051	1400	4577	3028	3548	3791	3950	3052	3257	3718	3546	4242	34549	3744	2820	3105	

Table 4.3: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 3 for SNR= 0 dB. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 3 for SNR= 0 dB.

	COLUMN OF $\mathbf{A}_{est}$																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	2	4	7	14	2	3838	0	9	12	6	4	8	10	6	218	5	6	4	3	11	4169
2	6	4	4098	3	4	6	5	6	10	6	5	7	9	4	17	3	9	3	10	5	4220
3	5	0	3	3	3	1	5	38	6	3	3470	5	3	3	3	1	7	4	6	0	3569
4	4	4	11	8	3	1	7	53	8	7	2	7	7	2	12	2	7	4	5	3415	3569
C 5	3	7	8	3	2	4	7	4	7	5	9	10	0	10	10	0	16	2956	8	14	3083
O 6	4	1	5	1	1	1	6	314	3	2645	5	2	8	3	488	1	12	5	4	9	3518
L 7	10	9	9	5	5	8	11	10	4	3	9	7	8	3193	44	3	5	8	11	15	3377
U 8	1	1	4	1	4	4	2	25	3420	3	3	1	4	5	11	1	5	2	5	2	3504
M 9	10	2	4	4	3536	3	3	5	4	4	9	8	3	2	34	1	3	3	8	7	3653
N 10	5	4	4	21	3	4	7	32	6	5	4	7	5	3668	16	1	4	3	4	13	3816
11	3	6	11	4	7	7	9	7	8	9	9	4300	7	6	12	4	10	7	6	9	4441
O 12	4	2820	5	4	1	3	3	5	4	5	3	4	8	3	25	0	4	3	4	10	2918
F 13	2334	12	15	28	19	12	20	32	13	14	7	29	22	15	46	2	17	20	21	18	2696
14	1	5	6	4	7	2	5	29	7	3	3	6	1	3	15	1	4	6	3222	11	3341
A 15	7	7	3	1	3	2	3299	4	7	3	3	8	1	2	6	3	2	1	2	3	3367
16	4	5	2	16	4	9	4	104	8	11	8	11	5	10	25	5	3	3757	7	12	4010
17	4	2746	6	3	2	2	7	8	4	6	7	7	3	6	26	2	5	6	7	6	2863
18	4	3	0	2	2	3	3	197	4	2	3	6	2427	0	7	1	7	6	2	4	2683
19	2771	3	1	1	2	5	1	8	4	4	2	6	2	5	39	2	2	3	5	8	2874
20	102	94	99	8936	93	63	132	341	136	126	117	145	72	91	1135	16756	5508	101	119	163	34329
	5284	5737	4301	9062	3703	3978	3536	1231	3675	2870	3682	4584	2605	7037	2189	16794	5636	6902	3459	3735	

Table 4.4: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 4 for SNR= 0 dB. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 4 for SNR= 0 dB.

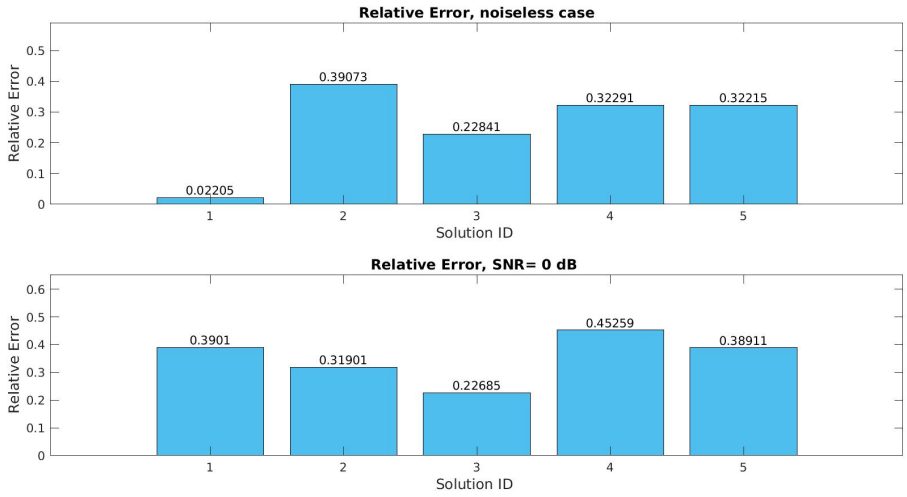


Figure 4.9: **First image**: Relative error of solutions for Penalty method and noiseless data. **Second image**: Relative error of solutions for Penalty method and SNR= 0 dB.

For the noiseless case, the solution 1, which has the smallest relative error, is really close to the true $\mathbf{A}$. The 99799 of 100000 of the non-zero elements of the solution 1 are placed in the correct cluster. However, for solution 2 (worst solution) only the 76208 of 100000 of its non-zero elements are placed in the correct cluster, even if there is no noise. We note that, the column 2 of the true $\mathbf{A}$ has been broken down into two columns of the $\mathbf{A}_{est}$ and the columns 7, 16. Moreover, we can see that column 2 of the $\mathbf{A}_{est}$ combines the columns 2, 5 of the true $\mathbf{A}$. For SNR= 0 dB, the solution 3 (best solution) is close to the true $\mathbf{A}$. The 90957 of 100000 of the non-zero elements of the solution 3 are placed in the correct cluster. However, the column 17 of the $\mathbf{A}_{est}$ combines the columns 2, 5 of the true $\mathbf{A}$. For solution 4 (worst solution) only the 67522 of 100000 of its non-zero elements are placed in the correct cluster. In this case, the column 20 of the true $\mathbf{A}$ has been broken down into the columns 4, 15, 16 and 17 of $\mathbf{A}_{est}$. Also, the column 18 of the $\mathbf{A}_{est}$ combines the columns 5, 16 of the true $\mathbf{A}$, the column 1 of the $\mathbf{A}_{est}$ combines the columns 13, 19 of the true $\mathbf{A}$, the column 2 of the $\mathbf{A}_{est}$ combines the columns 12, 17 of the true $\mathbf{A}$, the column 14 of the $\mathbf{A}_{est}$ combines the columns 7, 10 of the true $\mathbf{A}$. Finally, the column 8 of the $\mathbf{A}_{est}$ is comprised of small parts of several columns of the true $\mathbf{A}$. So, we can conclude that the initial points affect the penalty method and can lead to local minima which are not close enough to the true $\mathbf{A}$.

(b) **Accelerated Penalty method**

We examine for each solution how close it is to the rest of other four solutions. To this end, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 5$, we calculate the correlation coefficients between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for

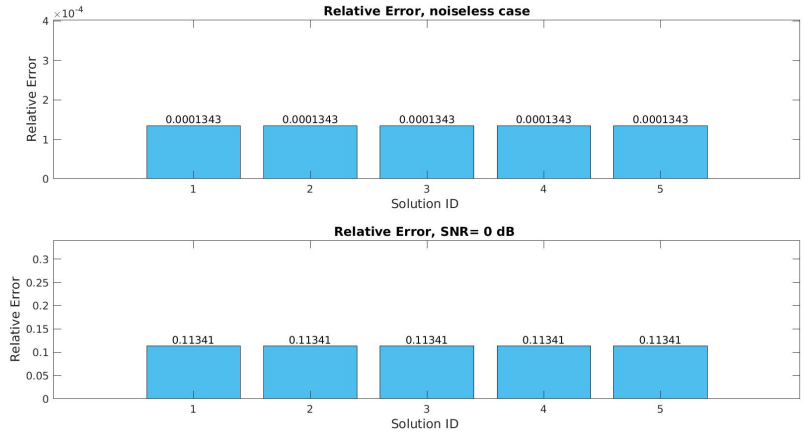


Figure 4.10: **First image**: Relative error of solutions for Accelerated Penalty method and noiseless data. **Second image**: Relative error of solutions for Accelerated Penalty method and SNR= 0 dB.

$j = 1, \ldots, 5$ and $j \neq i$, with which the column r has the greatest correlation coefficient. Also, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 5$, we calculate the intersection of non-zero elements between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 5$ and $j \neq i$, with which the column r has the greatest intersection. We also examine if the correspondence between columns is the same for the two cases (correlation coefficients, intersection). We repeat the aforementioned steps for noiseless data and data with SNR= 0 dB. We illustrate the results between solution 1 and 2 in Figure 4.11 and the rest of them in the subsection A.3.1.

We can observe that five solutions are the same, for both noiseless case and case in which SNR=0 dB. So, it seems that, the initial points do not affect the method.

After that, we calculate relative error of solutions (Figure 4.10). The relative error for five solutions is the same, as the solutions are identical. So, we choose the solution 1 for each case of noise and examine how close it is to the true common spatial factor $\mathbf{A}$. More specifically, for every column r of the true $\mathbf{A}$, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each and every column of the solutions 1 for noiseless case and the solutions 1 for SNR= 0 dB. We illustrate for each column r of the true $\mathbf{A}$ the number of elements of its intersection with all columns of the solution and the number of the non-zero elements of each column of the solution. The images are provided in the subsection A.3.1 and here we provide the corresponding Tables 4.5 and 4.6.

For the noiseless case, all solutions, which are identical, are exactly the same with the true common spatial factor $\mathbf{A}$. For SNR= 0 dB, the five solutions,

which are again identical, are really close to the the true common spatial factor **A**. The 95920 of 100000 of the non-zero elements of the solutions are placed in the correct cluster. So, it seems that the initial points do not affect the accelerated penalty method and lead to solutions which are really close to the true factor **A**.

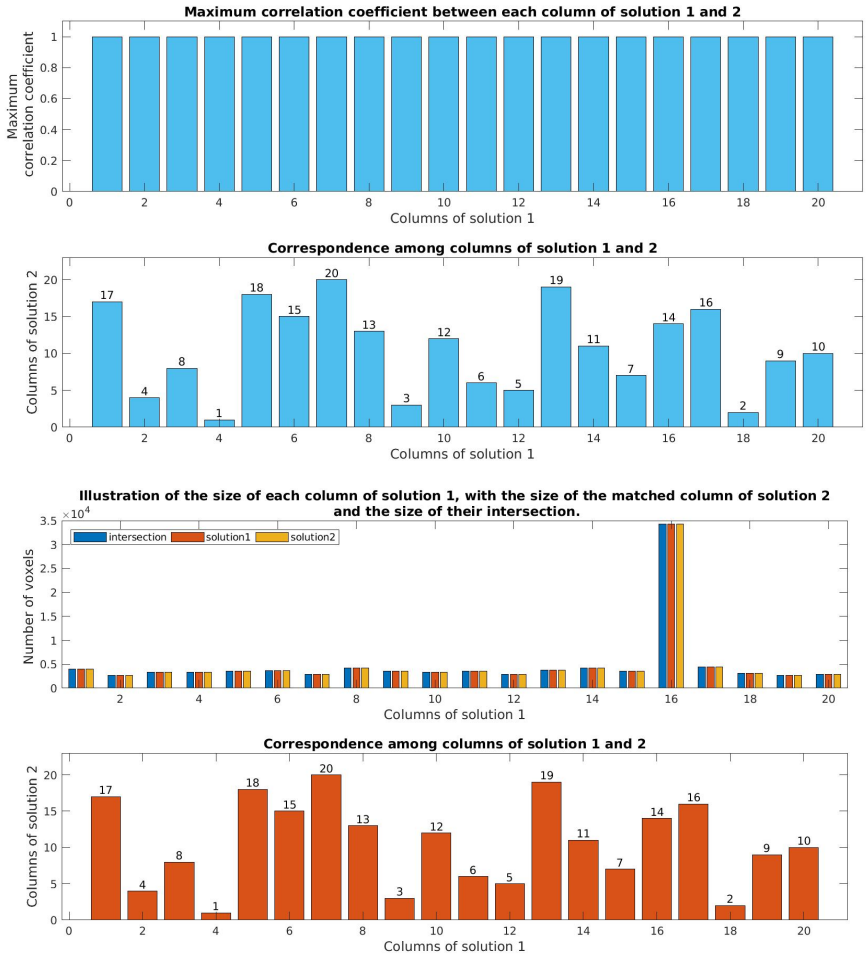


Figure 4.11: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates to which column of the solution 2 each column of the solution 1 corresponds based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates to which column of the solution 2 each column of the solution 1 corresponds based on intersection. Solutions 1, 2 correspond to accelerated penalty method and noiseless data.

COLUMN OF $\mathbf{A}_{est}$																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	4169	0	0	0	0	0	0	4169
2	0	0	0	0	0	0	0	4220	0	0	0	0	0	0	0	0	0	0	0	0	4220
3	0	0	0	0	0	0	0	0	3569	0	0	0	0	0	0	0	0	0	0	0	3569
4	0	0	0	0	3569	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3569
C 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3083	0	0	3083
O 6	0	0	0	0	0	0	0	0	0	0	3518	0	0	0	0	0	0	0	0	0	3518
L 7	0	0	0	0	0	0	0	0	0	3377	0	0	0	0	0	0	0	0	0	0	3377
U 8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3504	0	0	0	0	0	3504
M 9	0	0	0	0	0	3653	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3653
N 10	0	0	0	0	0	0	0	0	0	0	0	0	3816	0	0	0	0	0	0	0	3816
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4441	0	0	0	4441
O 12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2918	2918
F 13	0	2696	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2696
14	0	0	3341	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3341
A 15	0	0	0	3367	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3367
16	4010	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4010
17	0	0	0	0	0	0	2863	0	0	0	0	0	0	0	0	0	0	0	0	0	2863
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2683	0	2683
19	0	0	0	0	0	0	0	0	0	0	0	2874	0	0	0	0	0	0	0	0	2874
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	34329	0	0	0	0	34329
	4010	2696	3341	3367	3569	3653	2863	4220	3569	3377	3518	2874	3816	4169	3504	34329	4441	3083	2683	2918	

Table 4.5: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 1 of accelerated penalty method and for noiseless case. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 1 for noiseless case.

COLUMN OF $\mathbf{A}_{est}$																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	6	4059	5	5	8	7	3	4	5	4	4	4	12	9	9	3	1	5	10	6	4169
2	5	5	4	7	5	4	5	6	0	3	5	9	6	4109	7	14	6	7	7	6	4220
3	1	0	4	4	5	3	3	3	1	3502	3	9	6	3	5	2	4	5	1	5	3569
4	2	1	4	2	5	12	4	7	5	3	3	6	9	8	8	8	6	5	3468	3	3569
C 5	2	3	3	3	4	6	4	3	1	4	2	4	3	6	1	6	3	3015	7	3	3083
O 6	0	2	6	4	3	3	3	3438	4	7	1	7	6	5	5	3	6	2	7	6	3518
L 7	3	4	3	4	2	3303	7	4	5	5	3	5	2	4	3	5	3	1	4	7	3377
U 8	1	4	0	1	2	5	3	3	2	3	4	5	3447	4	3	5	1	4	3	4	3504
M 9	4	4	4	7	8	0	5	5	6	8	3561	9	6	2	3	2	4	4	7	4	3653
N 10	4	6	0	8	6	4	5	2	2	6	2	4	6	2	6	4	6	1	11	3731	3816
11	4	7	3	3	4304	8	9	8	3	8	7	9	8	11	7	7	8	10	9	8	4441
O 12	1	3	2	3	2	4	2	3	2864	4	2	4	2	2	2	6	2	4	6	0	2918
F 13	6	1	1	2	5	2	0	3	1	0	2	3	3	2	4	2646	6	2	4	3	2696
14	2	1	4	0	7	10	6	3	6	4	7	3252	5	4	0	4	2	4	14	6	3341
A 15	1	0	6	7	8	5	6	3	4	2	3	3	4	4	1	5	3297	3	1	4	3367
16	4	5	3894	8	10	6	5	11	4	7	4	6	7	2	3	6	3	7	10	8	4010
17	3	0	4	3	4	2	2803	4	4	3	1	4	2	5	1	1	5	5	5	4	2863
18	3	3	2	2	6	4	2	5	3	3	1	2	3	0	2625	7	3	4	2	3	2683
19	2	4	2	2817	4	6	3	3	2	1	0	4	3	1	2	4	1	3	6	6	2874
20	31785	102	81	147	168	175	124	148	94	136	92	168	154	115	91	146	137	142	185	139	34329
	31839	4214	4032	3037	4566	3569	3002	3666	3016	3713	3707	3517	3694	4298	2786	2884	3504	3233	3767	3956	

Table 4.6: Each cell (i,j) of the table shows the size of intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 1 of accelerated penalty method and for SNR= 0 dB. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 1 for SNR= 0 dB.

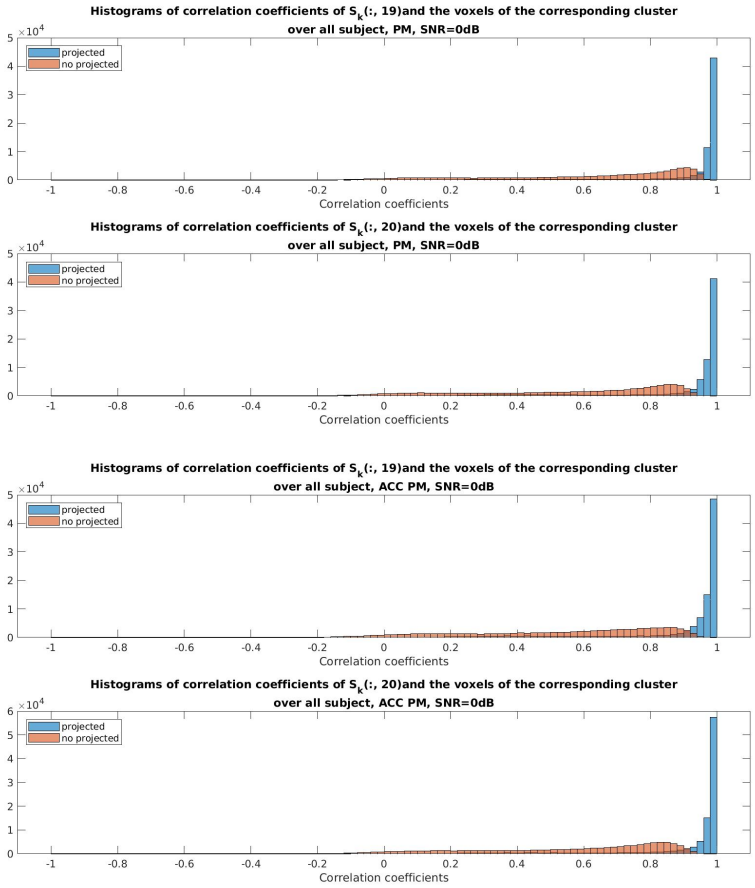


Figure 4.12: For the column r and subject k, for $r = 19,\ 20$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 19 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 20 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 19 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 20 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

Finally, for SNR=0 dB, we choose the best solution for each of the two methods. For each of the two solutions, we follow the following procedure. For each cluster r and subject k, for $r = 1, \ldots, R$ and $k = 1, \ldots, K$, we calculate the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. After that, we calculate the correlation coefficients again, this time for the projected data on common subspace $\mathbf{G}_{opt}$. For each cluster r, for $r = 1, \ldots, R$, we gather the coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial and projected data.

We illustrate the results for clusters $19, 20$ in Figure 4.12 and the rest of them are provided in the subsection A.3.1.

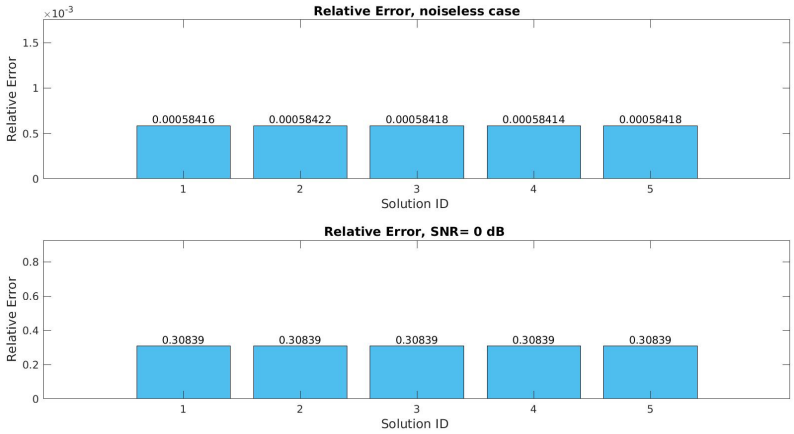


Figure 4.13: **First image**: Relative error of solutions for Accelerated Penalty method and noiseless data. **Second image**: Relative error of solutions for Accelerated Penalty method and SNR= 0 dB.

If there were no noise, then the histograms before and after the projection would be identical and all correlation coefficients would be equal to one. Here SNR=0 dB, so the noise and the common information are of equal power and we are expecting the histogram to be shifted to the left. We observe that histograms of initial data are shifted to the left and they have also a wider range of values. After the projection, which removes the noise depending on how close the col($\mathbf{G}_{opt}$) and col($\mathbf{A}$) are, we see that the histograms are more concentrated close to one. Also, we can observe that for cluster 17, which for penalty method combines the columns 2, 5 of the true $\mathbf{A}$, the histogram of projected data is more spread.

2. **Case 2: Common orthonormal matrix A with each subject having correlated temporal components $\mathbf{S}_k$.**

In this case, we examine the results of the factorization of $\mathbf{G}_{opt}$ only for accelerated penalty method, because in previous case we observed that penalty method is affected by initial points and its accuracy is not as good as the accuracy of accelerated penalty method.

Firstly, we examine for each solution how close it is to the rest of the other four solutions. To this end, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 5$, we calculate the correlation coefficients between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 5$ and $j \neq i$, with which the column r has the greatest correlation coefficient. Also, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 5$, we calculate the intersection of non-zero elements between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 5$ and $j \neq i$, with which the column r has the greatest intersection. We examine if the correspondence between columns is the same for the two cases (correlation coefficients,

intersection). We repeat the aforementioned steps for noiseless data and data with SNR= 0 dB. We illustrate here the results between solution 1 and 2 in Figure 4.14 and the rest of them in the subsection A.3.2.

We observe that five solutions are the same, for both noiseless case and case in which SNR=0 dB. So, it seems that, the initial points do not affect the method. After that, we calculate the relative error of the solutions (Figure 4.13).

The relative error for the five solutions is the same, as the solutions are identical. So, we choose the solution 1 for each case of noise and examine how close it is to the true common spatial factor $\mathbf{A}$. More specifically, for every column r of the true $\mathbf{A}$, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each and every column of the solutions 1 for noiseless case and the solutions 1 for SNR= 0 dB. We illustrate for each column r of the true $\mathbf{A}$ the number of elements of its intersection with all columns of the solution and the number of the non-zero elements of each column of the solution. The images are provided at A.3.2 and we provide, here, the corresponding Tables 4.7, 4.8.

For the noiseless case, all solutions, which are identical, are exactly the same with the true common spatial factor $\mathbf{A}$. For SNR= 0 dB, the five solutions, which are again identical, are close to the true common spatial factor $\mathbf{A}$. The 86399 of 100000 of the non-zero elements of the solutions are placed in the correct cluster. We observe that, column 20 of the true $\mathbf{A}$ is broken down into a main part and other smaller parts. More specifically, we can see that, most columns of $\mathbf{A}_{est}$ are comprised of a piece of column 20 of the true $\mathbf{A}$, the size of which is comparable to the size of the columns. So, it seems that the initial points do not affect the accelerated penalty method and lead to solutions which are close to the true factor $\mathbf{A}$. We note that, the noise here seems to affect the estimation of common spatial factor $\mathbf{A}$ more than in the case 1. This is because the temporal components $\mathbf{S}_k$ for some subjects are highly correlated, so the same amount of noise can lead to more erroneously placed voxels.

Finally, for SNR=0 dB and solution 1, we follow the following procedure. For each cluster r and subject k, for $r = 1, \ldots, R$ and $k = 1, \ldots, K$, we calculate the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. After that, we calculate the correlation coefficients again, this time for the projected data on common subspace $\mathbf{G}_{opt}$. For each cluster r, for $r = 1, \ldots, R$, we gather the coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial and projected data. The images are provided in the subsection A.3.2.

If there were no noise, then the histograms before and after the projection would be identical and all correlation coefficients would be equal to one. Here SNR=0 dB, so the noise and the common information are of equal power and we are expecting the histogram to be shifted to the left. We observe that, histograms of initial data are shifted to the left and they have also a wider range of values. After the projection, which removes the noise depending on how close the col($\mathbf{G}_{opt}$) and col($\mathbf{A}$)

are, we see that the histograms are more concentrated close to one.

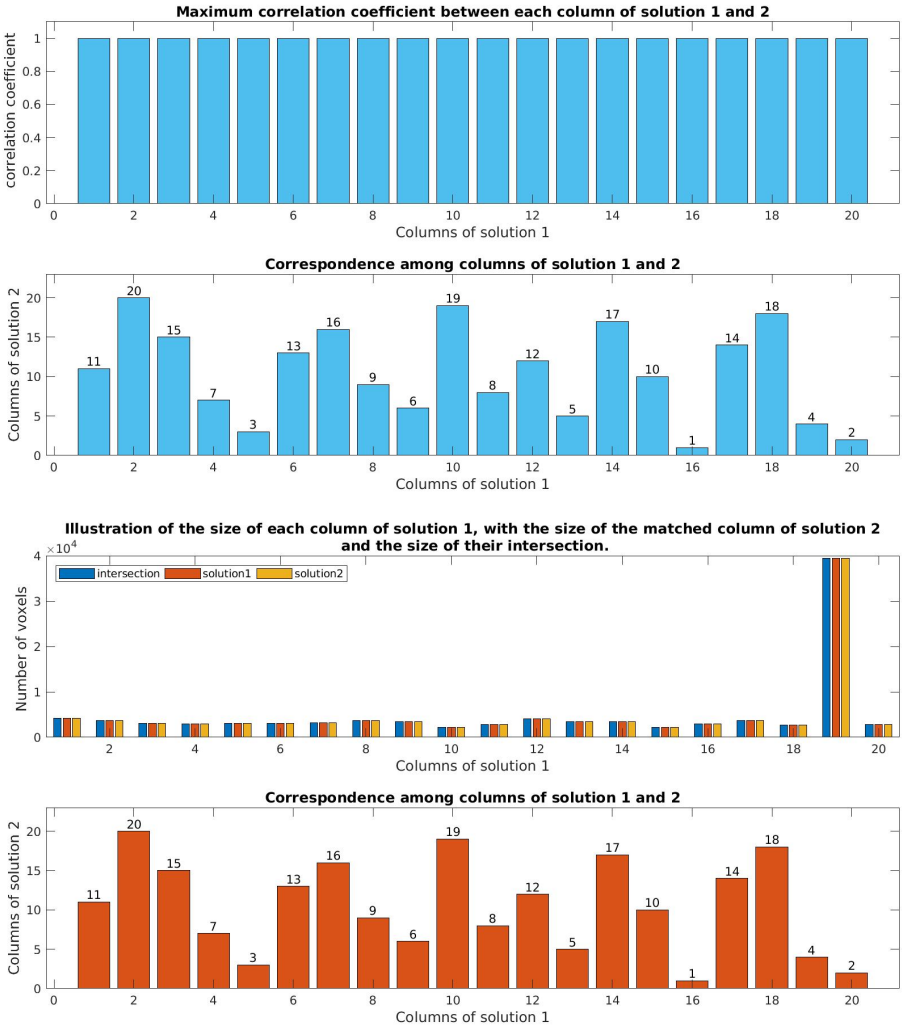


Figure 4.14: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates to which column of the solution 2 each column of the solution 1 corresponds based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates to which column of the solution 2 each column of the solution 1 corresponds based on intersection. Solutions 1, 2 correspond to accelerated penalty method and noiseless data.

COLUMN OF $\mathbf{A}_{est}$																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	3399	0	0	0	0	0	0	3399
2	0	3680	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3680
3	0	0	0	0	0	0	0	0	0	2156	0	0	0	0	0	0	0	0	0	0	2156
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2950	0	0	0	0	2950
C 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2788	2788
O 6	4185	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4185
L 7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2209	0	0	0	0	0	2209
U 8	0	0	0	0	0	0	3221	0	0	0	0	0	0	0	0	0	0	0	0	0	3221
M 9	0	0	0	0	0	0	0	0	0	0	2812	0	0	0	0	0	0	0	0	0	2812
N 10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2719	0	0	2719
11	0	0	3023	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3023
O 12	0	0	0	0	0	0	0	0	0	0	0	4075	0	0	0	0	0	0	0	0	4075
F 13	0	0	0	2901	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2901
14	0	0	0	0	0	0	0	3718	0	0	0	0	0	0	0	0	0	0	0	0	3718
A 15	0	0	0	0	0	0	0	0	0	0	0	0	3403	0	0	0	0	0	0	0	3403
16	0	0	0	0	0	3015	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3015
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3711	0	0	0	3711
18	0	0	0	0	3098	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3098
19	0	0	0	0	0	0	0	0	3483	0	0	0	0	0	0	0	0	0	0	0	3483
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	39454	0	39454
	4185	3680	3023	2901	3098	3015	3221	3718	3483	2156	2812	4075	3403	3399	2209	2950	3711	2719	39454	2788	

Table 4.7: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 1 of accelerated penalty method and for noiseless case. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 1 for noiseless case.

COLUMN OF $\mathbf{A}_{est}$																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	7	6	7	3	18	14	20	49	7	0	16	1	13	3186	12	11	10	14	0	5	3399
2	8	6	9	6	15	13	19	35	8	0	24	0	3412	29	16	27	14	30	2	7	3680
3	4	3	5	2	10	4	7	25	6	1	2035	0	9	8	7	12	5	11	0	2	2156
4	7	5	9	2	14	6	23	31	2760	2	6	0	4	18	13	15	8	20	0	7	2950
C 5	14	9	5	7	13	14	11	12	7	2608	7	0	10	21	11	13	7	15	0	4	2788
O 6	19	8	13	5	34	18	27	43	19	2	21	0	8	37	20	26	20	24	2	3839	4185
L 7	3	2087	5	3	7	7	11	14	7	2	5	0	8	19	6	3	7	14	0	1	2209
U 8	10	5	18	4	12	8	20	33	8	0	22	1	8	23	14	12	2996	22	0	5	3221
M 9	8	9	11	4	2626	9	9	28	8	2	10	0	9	16	15	13	7	23	0	5	2812
N 10	3	3	12	4	14	5	11	19	7	0	14	0	4	23	2549	9	15	26	0	1	2719
11	6	5	7	2838	18	7	8	23	8	0	19	0	17	17	11	10	7	15	2	5	3023
O 12	20	9	19	8	32	11	28	26	9	2	25	1	18	29	11	3774	16	28	1	8	4075
F 13	5	8	10	3	19	2728	20	26	9	0	12	0	5	22	3	10	4	15	1	1	2901
14	15	11	3418	7	15	14	39	26	11	2	28	0	12	40	19	15	9	30	1	6	3718
A 15	4	11	21	1	18	10	23	19	11	2	15	0	8	33	25	14	16	3166	1	5	3403
16	2813	5	4	1	21	11	12	24	5	4	14	0	8	28	13	15	8	22	0	7	3015
17	13	5	9	6	27	10	21	42	11	2	14	0	10	32	10	14	16	26	3438	5	3711
18	18	4	6	4	16	10	16	2854	10	1	17	0	15	47	12	20	12	26	2	8	3098
19	7	3	7	6	20	11	15	28	12	1	14	3246	18	25	16	17	12	21	0	4	3483
20	367	313	582	128	733	370	30026	1190	349	43	673	5	665	984	619	587	415	1158	29	218	39454
	3351	2515	4177	3042	3682	3280	30366	4547	3272	2674	2991	3254	4261	4637	3402	4617	3604	4706	3479	4143	

Table 4.8: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 1 of accelerated penalty method and for SNR=0 dB. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 1 for noiseless case.

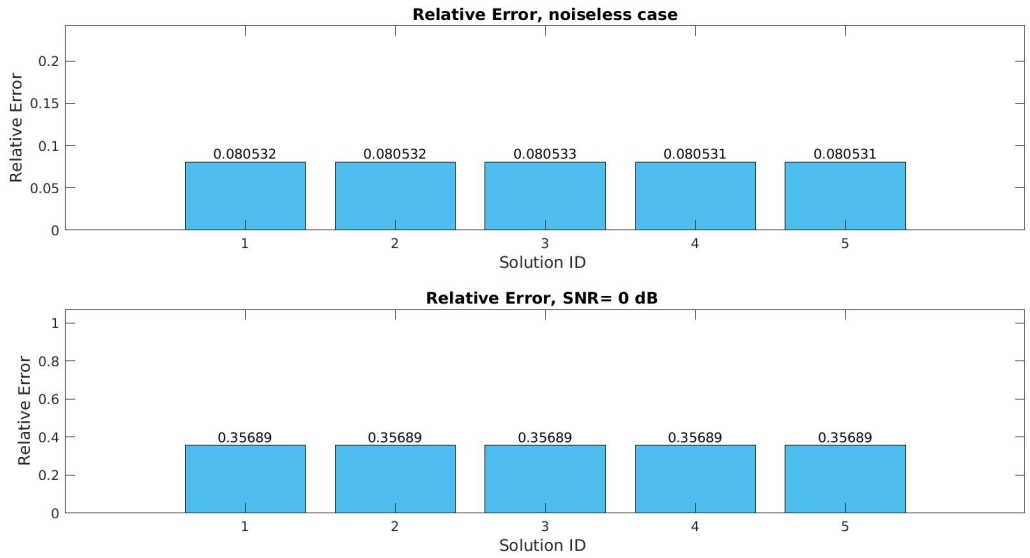


Figure 4.15: **First image**: Relative error of solutions for Accelerated Penalty method and noiseless data. **Second image**: Relative error of solutions for Accelerated Penalty method and SNR= 0 dB.

3. **Case 3: Corrupted common orthonormal matrix A with each subject having correlated temporal components $\mathbf{S}_k$.**

Also, in this case, we examine the results of the factorization of $\mathbf{G}_{opt}$ only for accelerated penalty method, because in case 1 we observed that, penalty method is affected by initial points and its accuracy is not as good as the accuracy of accelerated penalty method.

Firstly, we examine for each solution how close it is to the rest of other four solutions. To this end, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 5$, we calculate the correlation coefficients between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 5$ and $j \neq i$, with which the column r has the greatest correlation coefficient. Also, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 5$, we calculate the intersection of non-zero elements between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 5$ and $j \neq i$, with which the column r has the greatest intersection. We examine if the correspondence between columns is the same for the two cases (correlation coefficients, intersection). We repeat the aforementioned steps for noiseless data and data with SNR= 0 dB. We illustrate the results between solution 1 and 2 in Figure 4.16 and the rest of them in the subsection A.3.3.

We observe that, the five solutions are the same, for both noiseless case and case in which SNR=0 dB. So, it seems that, the initial points do not affect the method.

After that, we calculate the relative error of the solutions (Figure 4.15).

The relative error for the five solutions is the same, as the solutions are identical. So, we choose the solution 1 for each case of noise and examine how close it is to the true common spatial factor $\mathbf{A}$. More specifically, for each column r of the true $\mathbf{A}$, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each and every column of the solutions 1 for noiseless case and the solutions 1 for SNR= 0 dB. We illustrate for each column r of the true $\mathbf{A}$ the number of elements of its intersection with all columns of the solution and the number of the non-zero elements of each column of the solution. The images are provided at A.3.3 and we provide, here, the corresponding Tables 4.9, 4.10.

For the noiseless case, all solutions, which are identical, are really close to the true common spatial factor $\mathbf{A}$. The 99927 of 100000 of the non-zero elements of the solutions are placed in the correct cluster. For SNR= 0 dB, the five solutions, which are again identical, are close to the the true common spatial factor $\mathbf{A}$. The 85353 of 100000 of the non-zero elements of the solutions are placed in the correct cluster. We observe that, column 20 of the true $\mathbf{A}$ is broken down into a main part and other smaller parts. More specifically, we can see that most columns of $\mathbf{A}_{est}$ are comprised of a piece of column 20 of the true $\mathbf{A}$, the size of which is comparable to the size of the columns. So, it seems that, the initial points do not affect the accelerated penalty method and lead to solutions which are close to the true factor $\mathbf{A}$. We note that, the noise here seems to affect the estimation of common spatial factor $\mathbf{A}$ more than in the cases 1,2. This is because the temporal components $\mathbf{S}_k$ for some subjects are highly correlated and each subject has its own factor $\mathbf{A_k}$ which differs randomly $10 - 15\%$ from the true factor $\mathbf{A}$, so the same amount of noise can lead to more erroneously placed voxels.

Finally, for SNR=0 dB and solution 1, we follow the following procedure. For each cluster r and subject k, for $r = 1, \ldots, R$ and $k = 1, \ldots, K$, we calculate the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. After that, we calculate the correlation coefficients again, this time for the projected data on common subspace $\mathbf{G}_{opt}$. For each cluster r, for $r = 1, \ldots, R$, we gather the coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial and projected data. The images are provided in the subsection A.3.3.

If there were no noise, then the histograms before and after the projection would be identical and all correlation coefficients would be equal to one. Here SNR=0 dB, so the noise and the common information are of equal power and we are expecting the histogram to be shifted to the left. We observe that histograms of initial data are shifted to the left and they have also a wider range of values. After the projection, which removes the noise depending on how close the col($\mathbf{G}_{opt}$) and col($\mathbf{A}$) are, we see that the histograms are more concentrated close to one.

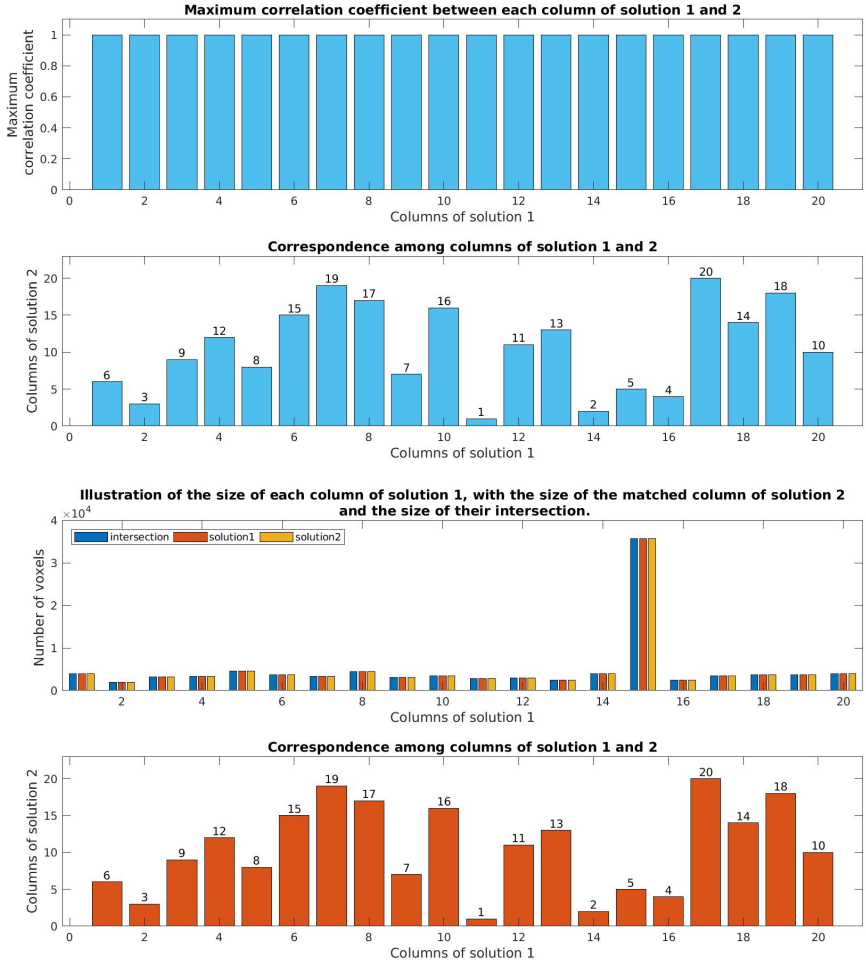


Figure 4.16: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates to which column of the solution 2 each column of the solution 1 corresponds based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates to which column of the solution 2 each column of the solution 1 corresponds based on intersection. Solutions 1, 2 correspond to accelerated penalty method and noiseless data.

		COLUMN OF $\mathbf{A}_{est}$																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
	1	0	0	0	0	0	0	3350	0	0	0	0	0	0	0	0	0	0	0	0	0	3350
	2	0	1895	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1895
	3	0	0	0	3322	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3324
	4	0	0	0	0	4521	0	0	0	0	0	0	1	2	0	1	0	0	0	1	0	4526
C	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3400	0	0	0	3401
O	6	0	0	0	0	0	0	0	0	0	3471	0	1	0	0	0	0	0	0	0	0	3472
L	7	0	0	0	0	0	0	0	0	0	0	2808	0	0	0	0	0	0	0	0	0	2808
U	8	0	0	0	0	0	0	0	0	0	0	0	0	0	3948	0	0	0	0	0	0	3948
M	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3660	0	3660
N	10	0	0	0	0	0	0	0	4459	0	0	0	0	0	0	0	0	0	0	0	0	4459
	11	1	0	3224	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3226
O	12	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3925	3926
F	13	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2385	0	0	0	0	2386
	14	0	0	0	0	0	0	0	0	1	0	0	2913	0	0	0	0	0	0	0	0	2914
A	15	0	0	0	0	0	0	0	0	0	0	0	0	2459	0	0	0	0	0	0	0	2459
	16	0	0	0	0	1	3734	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3735
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3739	0	0	3739
	18	3963	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3964
	19	0	0	0	0	0	0	0	0	3019	0	0	0	0	0	0	0	0	0	0	0	3019
	20	3	3	3	1	2	6	2	2	4	1	3	4	3	3	35732	4	4	1	2	6	35789
		3967	1898	3229	3323	4525	3741	3352	4461	3024	3472	2811	2920	2464	3951	35733	2389	3405	3741	3663	3931	

Table 4.9: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 1 of accelerated penalty method and for noiseless case. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 1 for noiseless case.

		COLUMN OF $\mathbf{A}_{est}$																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
	1	9	28	0	3	75	20	3	14	11	20	6	2	13	29	16	3045	24	10	3	19	3350
	2	3	13	0	0	6	9	2	8	5	16	1783	0	1	9	2	14	11	3	4	6	1895
	3	13	24	0	0	29	7	8	27	9	3069	13	2	12	21	10	24	28	5	4	19	3324
	4	20	29	0	3	50	20	11	28	4111	25	16	3	18	30	19	39	49	9	19	27	4526
C	5	11	24	1	6	25	13	10	26	14	13	8	3135	15	20	8	27	20	4	8	13	3401
O	6	19	13	1	2	29	14	11	7	17	20	12	3	7	34	8	37	27	5	3196	10	3472
L	7	13	15	0	1	16	9	2607	11	8	19	9	1	7	21	12	17	16	4	4	18	2808
U	8	3600	25	0	0	47	19	12	17	15	28	10	4	17	37	12	36	31	6	8	24	3948
M	9	16	25	0	2	39	17	12	21	13	14	11	3	17	32	12	33	3356	8	15	14	3660
N	10	23	33	0	0	37	26	11	27	18	32	16	2	10	49	22	43	29	3	6	4072	4459
	11	13	12	0	1	26	12	8	17	4	11	18	1	7	28	13	24	21	2985	9	16	3226
O	12	20	24	1	1	44	18	11	3589	22	30	20	2	11	26	17	39	24	3	6	18	3926
F	13	7	17	0	0	12	11	7	12	5	12	10	2	2215	18	5	20	24	2	3	4	2386
	14	7	17	1	2	20	2663	7	19	11	29	11	0	13	26	15	31	21	1	2	18	2914
A	15	10	14	0	0	21	20	5	4	2	14	9	3	8	2292	6	15	13	0	6	17	2459
	16	13	26	1	0	31	20	15	29	11	19	13	1	13	38	3426	29	16	7	9	18	3735
	17	9	20	0	3448	21	13	6	23	15	24	15	1	13	23	13	28	28	10	12	17	3739
	18	16	20	0	3	3608	12	16	26	8	32	16	2	15	34	21	76	34	4	5	16	3964
	19	10	20	2806	0	26	8	6	16	13	6	12	3	13	22	7	24	12	5	4	6	3019
	20	429	26347	6	29	1076	638	329	580	347	648	698	32	368	1183	364	973	795	126	203	618	35789
,		4261	26746	2817	3501	5238	3569	3097	4501	4659	4081	2706	3202	2793	3972	4008	4574	4579	3200	3526	4970	

Table 4.10: Each cell (i,j) of the table shows the intersection of the non-zero elements between the column i of the true common spatial factor $\mathbf{A}$ and the column j of the solution 1 of accelerated penalty method and for SNR=0 dB. The last column of table shows the number of the non-zero elements for each column of the true common spatial factor $\mathbf{A}$ and the last row of the table shows the number of th non-zero elements for each column of the solution 1 for noiseless case.

4.2 Real Data

The above methods are also applied in real Resting State fMRI data. The data was acquired during the period $2015-2016$ from 31 healthy subjects and was recorded in the MRI Unit, University Hospital of Heraklion.

Brain MRI examinations were performed on a clinical, upgraded 1.5 T Siemens Vision/Sonata scanner (Erlangen,Germany), Gradient strength: $45mT/m$, Gradient slew rate: $200mT/m/ms$ and a standard four-channel head array coil (minimum voxel dimensions: $70\mu m \times 70\mu m \times 300\mu m$). Resting-state fMRI was derived from a $T2$-weighted, fat-saturated 2D-FID-EPI sequence with repetition time (TR) $3500ms$, echo time (TE)$50ms$, field of view (FOV) $192\times192\times108$ (x, y, z), and acquisition voxel size $3\times3\times3mm$. Whole brain scans consist of 36 transverse slices with $3.0mm$ slice thickness and no interslice gap. Each BOLD timeseries consists of 150 dynamic volumes.

The fMRI data was pre-processed via SPM (Statistical Parametric Mapping). More specifically, realignment, co-registration, normalization and smoothing were applied to data. After these pre-processing steps, the first 5 volumes from each subject were discarded as it is usual in fMRI studies. Next, we applied a binary mask, which removes the voxels which correspond to areas of the $3D$ images which are likely to be background, bone or soft tissue. Then, after subtraction of the mean value from each and every voxel of each subject's brain, we removed the voxels, of which time-series had variance smaller than two for more than the half of the subjects. These voxels were located in the first and last slices of the $3D$ images. Finally, linear detrending was applied to data. Linear detrending approximates and removes a linear trend from data using a least squares best-fit line. Namely, we subtract the mean and the linear trend from each and every voxel of each subject's brain.

4.2.1 Estimation of common subspace dimension

In order to estimate the dimension of common subspace, we apply the gap criterion 3.37. We apply it to 25 different random partitions of our data. The results are provided in Figure 4.17.

Here, it is not easy to decide which is the common subspace dimension. We set a threshold value equal to 0.99 and we consider that the gap values which are greater than the threshold are equal to 1. From the Table 4.11, we observe that, the mean value is greater than 0.99 for common subspace dimension greater than 27. However, for common subspace dimension equal to 27 the maximum value is already 1, which may mean that either there is a partition for which the common subspace dimension is greater than 27 or there is a partition for which the common subspace dimension is smaller than 27. Also, the minimum value becomes greater than 0.99 for common subspace dimension greater than 42. So, there is a partition for which the common subspace dimension is equal to 42. Therefore, we estimate the common subspace dimension for each partition. The estimated common subspace dimension of each partition is given in section B.6. Also, the values of gap versus common subspace dimension for each partition are provided in the Tables B.3, B.4, B.5. Finally, we apply trimming. More specifically, we discard the 3 partitions which have the smallest common subspace dimension and the 3 partitions which have the largest

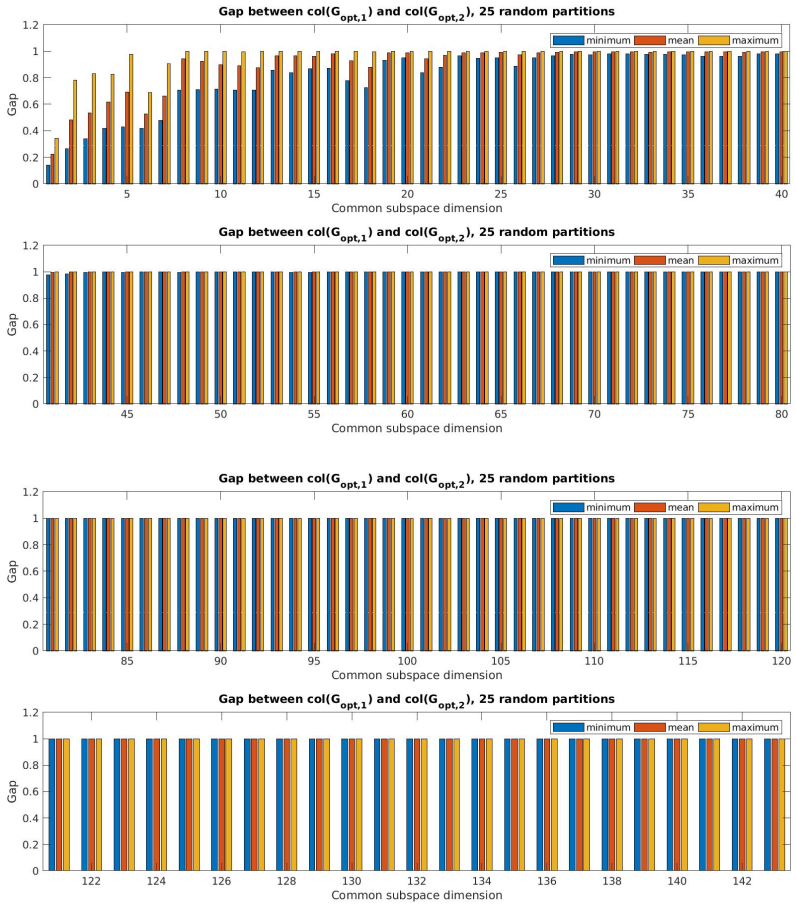


Figure 4.17: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension.

one and then we recalculate the minimum, mean and maximum value for each dimension (B.2). Moreover, for each dimension of common subspace, we remove the 3 smallest and the 3 largest values of gap from 25 values and then we recalculate the minimum, mean and maximum value for each dimension (Table: B.1) .

We observe that, the estimated common subspace dimension differs for the different partitions. Actually, the estimated common subspace dimension ranges from 21 to 42. Also, for both cases of trimming, the mean value of gap is greater to 0.99 for common subspace dimension greater to 27. So, we decide to solve the factorization problem for common subspace dimensions equal to 27 and 42. We examine the dimension 42 because it is the greatest of the estimated dimensions. So, there is possibility the rest of partitions to be divided into a way, which 'hides' a part of the common information.

4.2.2 Estimation of the common spatial factor A

We compute common spatial factor $\mathbf{A}$ for common subspace dimensions 27 and 42 via accelerated penalty method. We solve the factorization problem 4 times for each common subspace dimension. After that, we examine how close the solutions are for each common

		minimum gap	mean gap	maximum gap			minimum gap	mean gap	maximum gap
	1	0.13956	0.2238	0.34155		2	0.26507	0.48124	0.7836
	3	0.33918	0.53478	0.83025		4	0.41754	0.61604	0.82677
	5	0.43038	0.69118	0.97595		6	0.41844	0.52628	0.68759
	7	0.4789	0.66247	0.90636		8	0.70723	0.94385	0.99981
	9	0.70993	0.92525	0.9999		10	0.71515	0.89903	1
	11	0.70802	0.88951	0.99685		12	0.70587	0.87722	0.99752
	13	0.85745	0.96686	0.99998		14	0.8366	0.96456	0.99977
	15	0.86687	0.96037	0.99903		16	0.87117	0.98153	0.99999
	17	0.77777	0.92713	0.99976		18	0.72387	0.88082	0.99696
C	19	0.9336	0.98718	0.99999	C	20	0.95048	0.98975	0.99994
O	21	0.83791	0.94248	0.99866	O	22	0.87917	0.97059	0.99964
M	23	0.96594	0.98925	1	M	24	0.9482	0.98995	0.99997
M	25	0.9518	0.99031	1	M	26	0.88788	0.97207	0.99995
O	27	0.95237	0.98741	1	O	28	0.96582	0.99301	0.99983
N	29	0.97827	0.99376	1	N	30	0.97419	0.99422	0.99998
	31	0.97896	0.9959	0.99995		32	0.98199	0.99628	1
S	33	0.97727	0.99355	0.99987	S	34	0.97531	0.99405	1
U	35	0.97484	0.99603	0.99999	U	36	0.96307	0.99411	1
B	37	0.96353	0.99517	0.99995	B	38	0.96147	0.99177	1
S	39	0.97968	0.99393	0.99997	S	40	0.98067	0.99566	1
P	41	0.97743	0.9959	0.99999	P	42	0.98444	0.99776	0.99999
A	43	0.99327	0.99871	1	A	44	0.99747	0.99895	0.99998
C	45	0.99485	0.99936	1	C	46	0.99782	0.99948	1
E	47	0.99695	0.99955	1	E	48	0.99373	0.99923	1
	49	0.99788	0.99957	1		50	0.99832	0.99968	1
D	51	0.99779	0.99947	1	D	52	0.99852	0.99973	1
I	53	0.99863	0.99964	1	I	54	0.99523	0.99951	1
M	55	0.9951	0.9994	1	M	56	0.99799	0.99954	0.99999
E	57	0.99906	0.99973	1	E	58	0.99729	0.99941	1
N	59	0.9979	0.99957	1	N	60	0.99845	0.99961	1
S	61	0.99847	0.99962	1	S	62	0.99914	0.99979	1
I	63	0.99877	0.99974	1	I	64	0.99871	0.99975	1
0	65	0.99891	0.99979	1	0	66	0.99868	0.99983	1
N	67	0.9991	0.99985	1	N	68	0.99939	0.99987	1
	69	0.99924	0.99982	1		70	0.9994	0.99987	1
	71	0.99907	0.99988	1		72	0.99951	0.99989	1
	73	0.99942	0.99989	1		74	0.99933	0.99992	1
	75	0.99939	0.9999	1		76	0.99926	0.99988	1
	77	0.99976	0.99994	1		78	0.9996	0.9999	1
	79	0.99978	0.99993	1		80	0.99977	0.99994	1

		minimum gap	mean gap	maximum gap			minimum gap	mean gap	maximum gap
	81	0.99986	0.99997	1		82	0.9998	0.99995	1
C	83	0.99972	0.99994	1	C	84	0.9998	0.99995	1
O	85	0.99937	0.99991	1	O	86	0.99967	0.99996	1
M	87	0.99977	0.99996	1	M	88	0.99972	0.99995	1
M	89	0.99965	0.99995	1	M	90	0.99987	0.99998	1
O	91	0.99981	0.99997	1	O	92	0.99974	0.99995	1
N	93	0.99974	0.99997	1	N	94	0.99972	0.99994	1
	95	0.99959	0.99995	1		96	0.9998	0.99996	1
S	97	0.9998	0.99997	1	S	98	0.99988	0.99998	1
U	99	0.99988	0.99997	1	U	100	0.99989	0.99997	1
B	101	0.9999	0.99998	1	B	102	0.99993	0.99998	1
S	103	0.99989	0.99997	1	S	104	0.99981	0.99997	1
P	105	0.99984	0.99996	1	P	106	0.99968	0.99995	1
A	107	0.99974	0.99997	1	A	108	0.99987	0.99997	1
C	109	0.99989	0.99996	1	C	110	0.99989	0.99997	1
E	111	0.99989	0.99998	1	E	112	0.99988	0.99997	1
	113	0.99978	0.99997	1		114	0.99991	0.99998	1
D	115	0.99978	0.99997	1	D	116	0.9998	0.99998	1
I	117	0.99994	0.99999	1	I	118	0.99989	0.99998	1
M	119	0.99984	0.99998	1	M	120	0.99991	0.99997	1
E	121	0.9999	0.99998	1	E	122	0.99989	0.99998	1
N	123	0.99978	0.99997	1	N	124	0.9999	0.99997	1
S	125	0.99983	0.99997	1	S	126	0.99981	0.99997	1
I	127	0.99975	0.99997	1	I	128	0.99989	0.99997	1
0	129	0.99995	0.99998	1	0	130	0.99991	0.99998	1
N	131	0.99996	0.99999	1	N	132	0.99996	0.99999	1
	133	0.99992	0.99998	1		134	0.99993	0.99999	1
	135	0.99995	0.99999	1		136	0.9999	0.99999	1
	137	0.99989	0.99998	1		138	0.99993	0.99999	1
	139	0.99994	0.99999	1		140	0.99995	0.99999	1
	141	0.99994	0.99999	1		142	0.99993	0.99999	1
	143	0.99992	0.99998	1					

Table 4.11: Table shows the minimum, mean and maximum values of the gap over 25 random partitions of data for each dimension of common subspace.

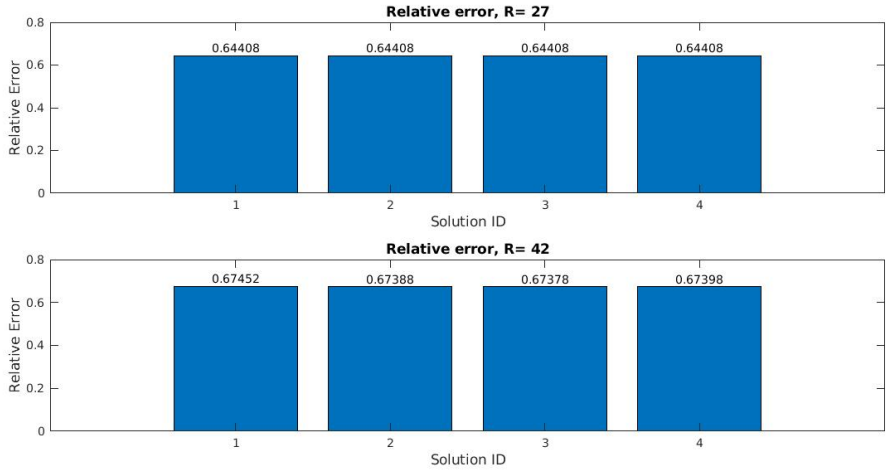


Figure 4.18: **First image**: Relative error of solutions for common subspace dimension equal to 27. **Second image**: Relative error of solutions for common subspace dimension equal to 42.

subspace dimension. For each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 4$, we calculate the correlation coefficients between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 4$ and $j \neq i$, with which the column r has the greatest correlation coefficient. Also, for each column r of each solution i, for $r = 1, \ldots, R$ and $i = 1, \ldots, 4$, we compute the intersection of non-zero elements between the column r of the solution i and each and every column of the rest of the solutions. We associate the column r of the solution i with the column of the solution j, for $j = 1, \ldots, 4$ and $j \neq i$, with which the column r has the greatest intersection. We examine if the correspondence between columns is the same for the two cases (correlation coefficients, intersection). The corresponding figures are provided in the section B.2.

We observe that, for common subspace dimension equal to 27 the four solutions are identical. However, in the case of the common subspace dimension 42, the solutions are different. More specifically, more than one columns of a solution j may associate with the same column of solution i. Moreover, a column of the solution i may associate with a column r_1 of the solution j based on correlation coefficients and to a different column r_2 of the solution j based on the intersection. Finally, if the column r of solution i is associated with the column r_1 of the solution j, it is not necessary the column r_1 of solution j to associate with the column r of solution i.

After that, we calculate the relative error of each solution for two common subspace dimensions using the formula:

$$\frac{||\mathbf{G}_{opt} - \mathbf{A}_{est}\mathbf{B}_{est}^T||_F}{||\mathbf{G}_{opt}||_F}.$$

We illustrate them in Figure 4.18.

We choose the solution 1 for common subspace dimension equal to 27 and the solution 3 for common subspace dimension 42, because it has the smallest relative error. We observe

	R=27	R=42		R=27	R=42
Subject1	-4.4059	-4.0721	Subject2	0.25752	0.74293
Subject3	-0.23917	0.21254	Subject4	-3.8589	-3.4057
Subject5	1.5214	1.6853	Subject6	0.48957	1.0338
Subject7	2.9306	3.2846	Subject8	-5.3062	-4.9389
Subject9	2.7585	3.0181	Subject10	-5.6537	-4.9842
Subject11	-6.8166	-6.1606	Subject12	0.49246	0.89077
Subject13	0.26362	0.79567	Subject14	-2.061	-1.7655
Subject15	2.1439	2.338	Subject16	0.87372	1.241
Subject17	-2.03	-1.6448	Subject18	-3.0221	-2.5166
Subject19	-4.3264	-3.8858	Subject20	-2.2719	-1.8074
Subject21	-3.4479	-3.2586	Subject22	-0.82121	-0.46208
Subject23	2.1953	2.322	Subject24	-3.6106	-3.1334
Subject25	-0.1377	0.1455	Subject26	-6.1322	-5.3978
Subject27	-1.8795	-1.636	Subject28	-3.3077	-2.9865
Subject29	0.28014	0.58457	Subject30	-3.362	-2.8788
Subject31	-0.31462	-0.094438	Total SNR	-0.65267	-0.32049

Table 4.12: The table shows the SNR for each and every subject and the total SNR for common subspace dimension equal to 27 and 42. The label SNR in x-axis corresponds to total SNR.

that, for common subspace dimension 42 the relative errors are real close, even if the solutions are not really similar. Also, the relative error for common subspace dimension equal to 42, is greater than the relative error common subspace dimension equal to 27.

For these solutions, we calculate the SNR of each subject k via

$$\text{SNR} = \frac{||\mathbf{A}^{\text{est}}(\mathbf{S}_{\text{k}}^{\text{est}})^{\text{T}}||_{\text{F}}^2}{\sum_{\text{k=1}}^{\text{K}} ||\mathbf{E}_{\text{k}}^{\text{est}}||_{\text{F}}^2}$$

and the total SNR using the formula

$$\text{SNR} = \frac{\sum_{\text{k=1}}^{\text{K}} ||\mathbf{A}^{\text{est}}\mathbf{S}_{\text{k}}^{\text{est}})^{\text{T}}||_{\text{F}}^2}{\sum_{\text{k=1}}^{\text{K}} ||\mathbf{E}_{\text{k}}^{\text{est}}||_{\text{F}}^2}.$$

We estimate the noise as

$$\mathbf{E}_k^{est} = \mathbf{X}_k - \mathbf{A}^{est}(\mathbf{S}_k^{est})^T.$$

We illustrate the results in Figures4.19, 4.20 and Table 4.12.

We observe that, subjects have different SNR values. For common subspace dimension equal to 27 the SNR values range from -6.8166 to 2.9306. For common subspace dimension equal to 42 the SNR values range from -6.1606 to 3.32846. The different SNR values indicate that the estimated $\mathbf{A}$ does not fit the same way to all subjects. The greater the SNR value, the better the $\mathbf{A}$ fits to the subject. Also, we observe that, for common

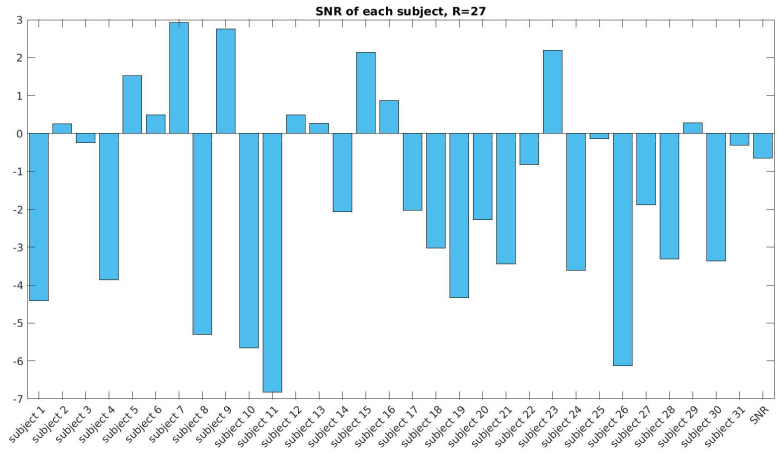


Figure 4.19: The image illustrates the SNR for each and every subject and the total SNR for common subspace dimension equal to 27. The label SNR in x-axis corresponds to total SNR.

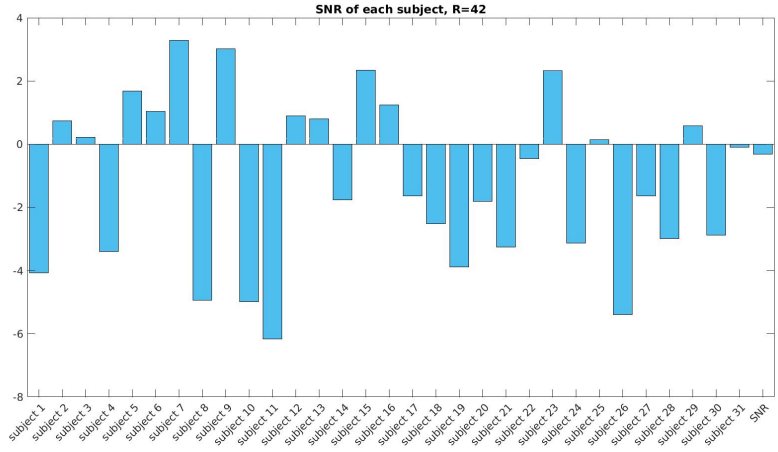


Figure 4.20: The image illustrates the SNR for each and every subject and the total SNR for common subspace dimension equal to 42. The label SNR in x-axis corresponds to total SNR.

subspace dimension 42 the SNR value of each subject is greater than it is for common subspace dimension 27.

We illustrate the sorted clusters for both common subspace dimensions in Figures 4.22 and 4.21. We observe that the size of clusters ranges.

For each common subspace dimension, we depict the clusters on the brain in Figures 4.23 and 4.24.For both cases of common subspace dimension, clusters seem to be logically formed, because as we can see in Figures 4.24, 4.23 they are spatially compact.

After that, for each cluster and for both common subspace dimensions, we illustrate the cluster on the brain and the temporal components, which corresponds to this cluster, for all subjects. Also, for each cluster r and subject k, for $r = 1, \ldots, R$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. After that, we compute the correlation coefficients again, this time for the projected data on common subspace $\mathbf{G}_{opt}$. For each cluster r, for $r = 1, \ldots, R$, we gather the coefficients over all subjects and illustrate the histograms of the correlation

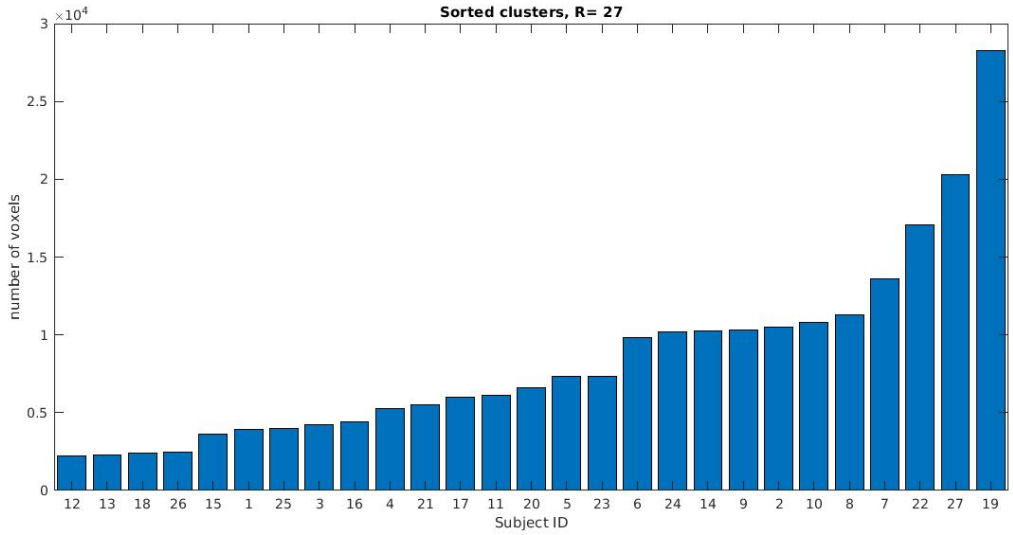


Figure 4.21: Sorted clusters for common subspace dimension equal to 27.

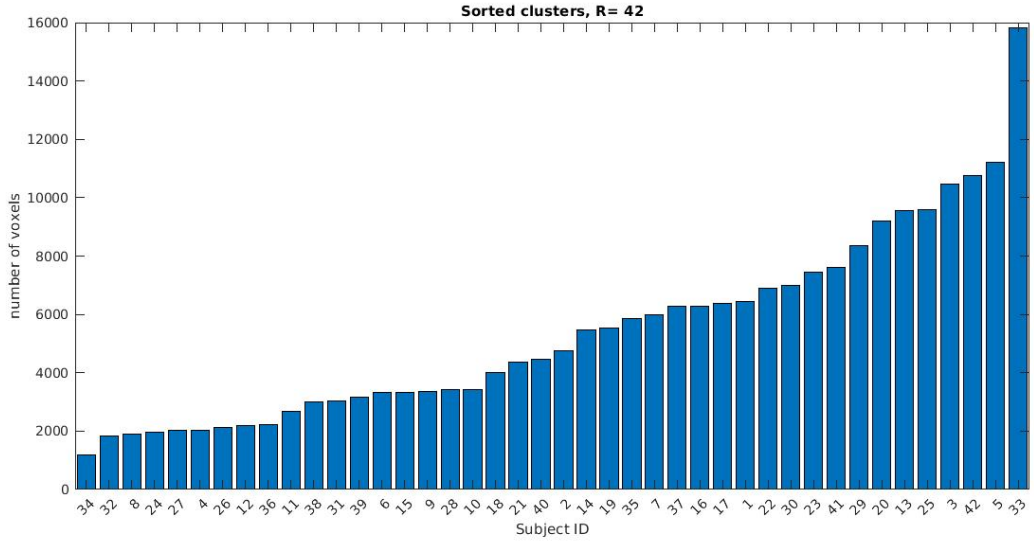


Figure 4.22: Sorted clusters for common subspace dimension equal to 42.

coefficients for initial and projected data. The images are provided in the sections B.3 and B.4. We observe that, the histograms of initial data are shifted to the left and they have also a wider range of values. After the projection, which removes the noise depending on how close the $col(\mathbf{G}_{opt})$ and $col(\mathbf{A})$ are, we see that the histograms are more concentrated close to one.

Also, we illustrate the correlation coefficients between temporal components $\mathbf{S}_k$ of each subject, for $k = 1, \ldots, K$. The figures are given in the sections B.5 and B.6. We can see that the temporal components of each subject are correlated.

Finally, we compare the results of clustering for the two dimensions (27, 42). For each column r of the $\mathbf{A}_{27}^{est}$, for $r = 1, \ldots, 27$, we calculate the intersection of non-zero elements between the column r and each and every column of the $\mathbf{A}_{42}^{est}$. We illustrate for each column r of the $\mathbf{A}_{27}^{est}$ the number of elements of its intersection with all columns of the

$\mathbf{A}_{42}^{est}$ and the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$. The images are provided in the section B.7. We observe that for the majority of clusters of $\mathbf{A}_{42}^{est}$, the main part of them belong to one cluster of $\mathbf{A}_{27}^{est}$. However, for some clusters of $\mathbf{A}_{42}^{est}$, we see that only small parts of them corresponds to clusters of $\mathbf{A}_{27}^{est}$.

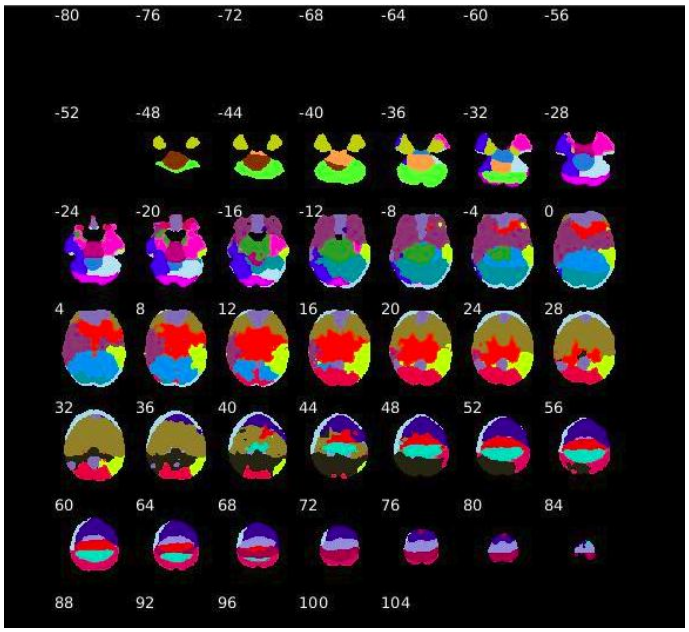


Figure 4.23: Whole brain parcellation for common subspace dimension equal to 27.

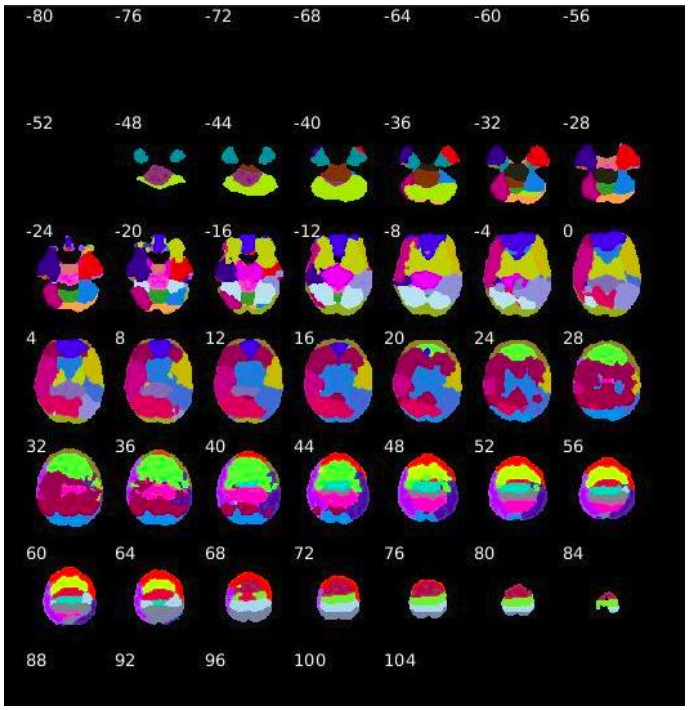


Figure 4.24: Whole brain parcellation for common subspace dimension equal to 42.

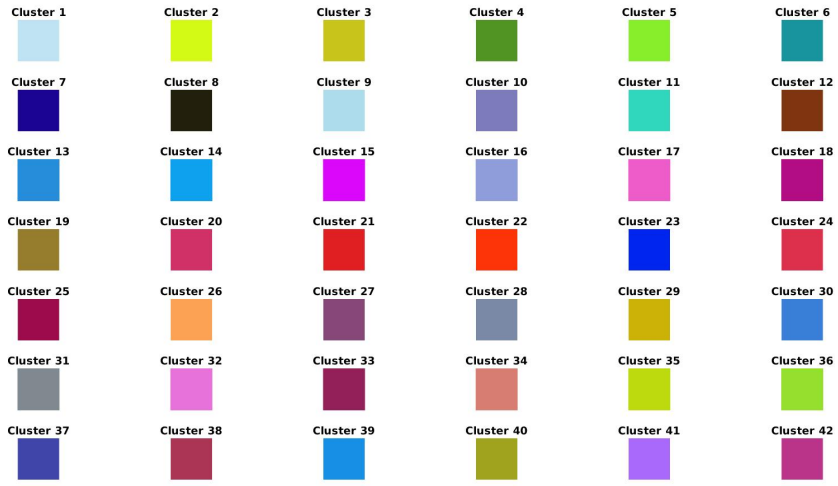


Figure 4.25: Corresponding colors for each cluster.

Chapter 5

Conclusion and Future Work

In this thesis, we examined the problem of the whole-brain parcellation using resting state fMRI data as it was modeled and implemented in [1]. Firstly, we studied and comprehended the basic pre-processing steps which are applied to raw fMRI data so that the fMRI data can be used for analysis. In this way, we ensured that the initial assumptions, which are made for the modeling of the whole-brain parcellation problem, are not violated because of the data pre-processing. After that, we examined experimentally the performances of the proposed methods using synthetic and real data. We investigated the methods with three cases of synthetic data trying to model synthetic data as close as we could to the real ones.

As for the estimation of common subspace dimension, we observed that for all cases of synthetic data using the gap criterion, we could estimate correctly the true common subspace dimension clearly even for small SNR values. When the structure of the sunthetic data was far from the assumed one (for example, different $\mathbf{A}_k$, for $k = 1, \ldots, K$ instead of same $\mathbf{A}$) the efficiency of the gap criterion was worsened, as it was expected. However, in real data it was not so clear from the criterion which was the common subspace dimension. Thus, we chose two common subspace dimensions equal to 27 and 42.

As for the factorization of the orthonormal basis of common subspace $\mathbf{G}_{opt}$, we applied two methods. Each of them tries to minimize a different cost function using a penalty method. We used the first method, which we call penalty method, to solve the factorization problem for the first case of synthetic data. More specifically, we solved the problem for noiseless data and data with SNR=0 dB. We observed that the first method even for the noiseless case cannot find the true factor $\mathbf{A}$ and it converges to local minima, which are highly depended on the initial points. The different solutions, which correspond to the same amount of noise, usually have different relative errors. On the other hand, using the second method, which we call accelerated penalty method, we solved the factorization problem for all cases of synthetic data without noise and for the case in which SNR=0 dB and observed that it converges to the same solution independently to the initials points. After that, we used it to estimate the common spatial factor $\mathbf{A}$ for real data and common subspace dimensions 27 and 42. For the case in which common subspace is equal to 27, the solutions are identical for the different initial points. However, for common subspace equal to 42, the method converges to different solutions for different initial points. However, the relative error of these solutions were really close.

For the real data and for both cases of common subspace dimension, clusters seem to be logically formed, since they are spatially compact. Also, from the histograms, we observed that the projection to common subspace removes the noise from subjects proportionally

to how close the $col(\mathbf{G}_{opt})$ and $col(\mathbf{A})$ are. Finally, the way clusters change from common subspace dimension 27 to 42 seems to have a coherence.

As a future work, we could try to find a more accurate criterion for the estimation of common subspace, which does not only use the information of the maximum canonical angle between two subspaces. Moreover, an interesting problem in fMRI field is the classification of healthy and mentally ill subjects based on their Resting state fMRI. Thus, we could examine how the brain parcellation of healthy subjects differs from the unhealthy ones and if it can be used as a criterion for classification between healthy and unhealthy subjects. Apart from that, it also can be used for the classification of different mental illnesses.

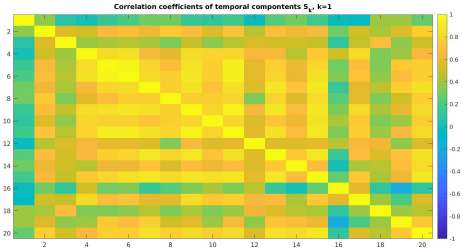
Appendices

Appendix A

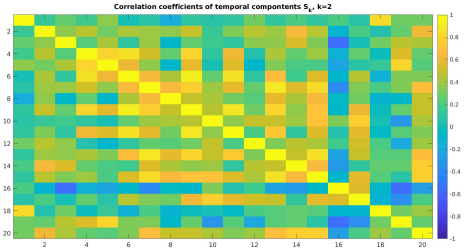
Synthetic Data

A.1 Correlation coefficients of temporal components $\mathbf{S}_k$

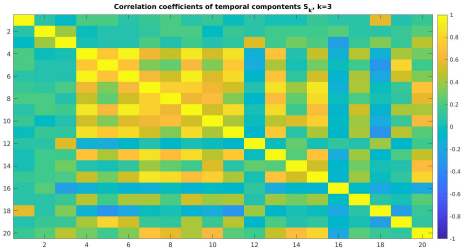
In this section, we illustrate the correlation coefficients of temporal components $\mathbf{S}_k$, for $k = 1, ..., 25$, for synthetic data.



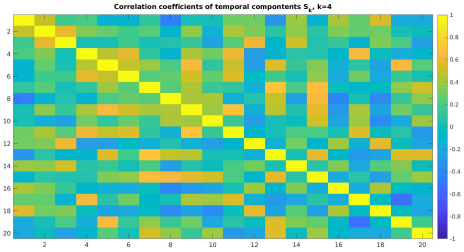
(a) Correlation coefficients of temporal components $\mathbf{S}_1$.



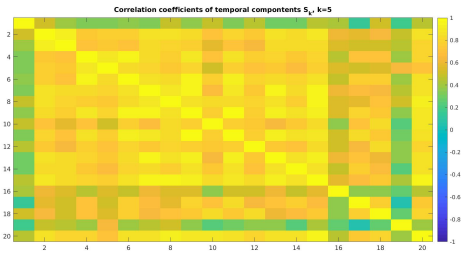
(b) Correlation coefficients of temporal components $\mathbf{S}_2$.



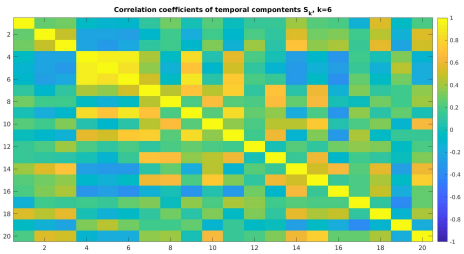
(a) Correlation coefficients of temporal components $\mathbf{S}_3$.



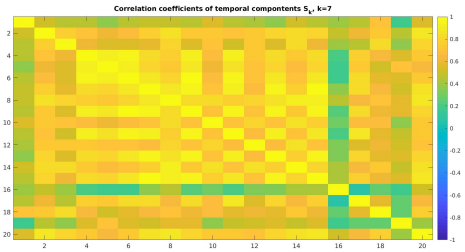
(b) Correlation coefficients of temporal components $\mathbf{S}_4$.



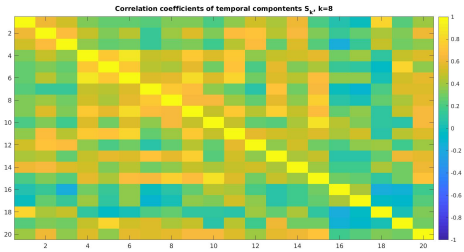
(a) Correlation coefficients of temporal components $\mathbf{S}_5$.



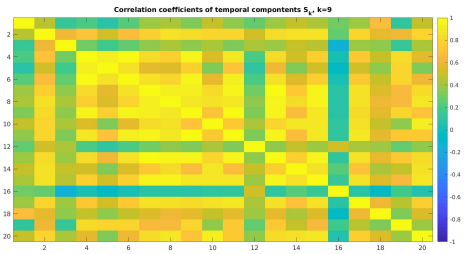
(b) Correlation coefficients of temporal components $\mathbf{S}_6$.



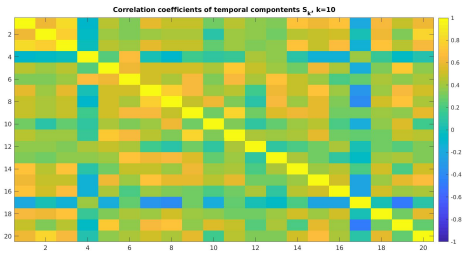
(a) Correlation coefficients of temporal components $\mathbf{S}_7$.



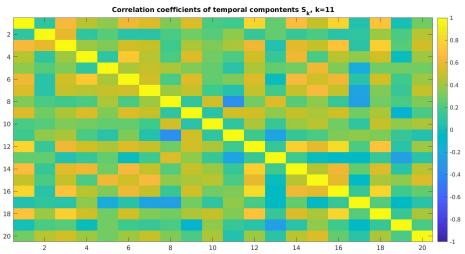
(b) Correlation coefficients of temporal components $\mathbf{S}_8$.



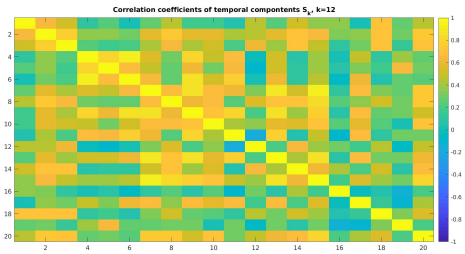
(a) Correlation coefficients of temporal components $\mathbf{S}_9$.



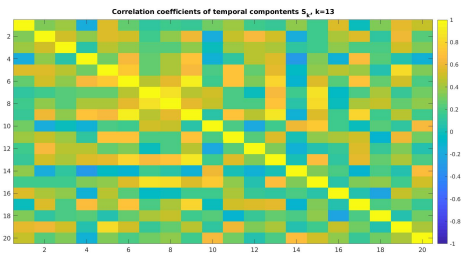
(b) Correlation coefficients of temporal components $\mathbf{S}_{10}$.



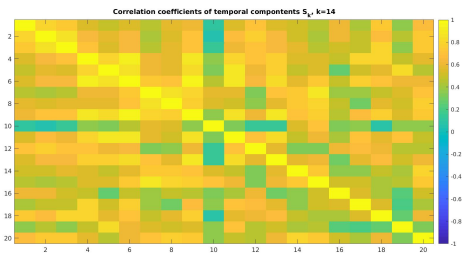
(a) Correlation coefficients of temporal components $\mathbf{S}_{11}$.



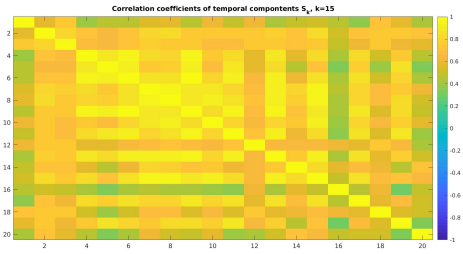
(b) Correlation coefficients of temporal components $\mathbf{S}_{12}$.



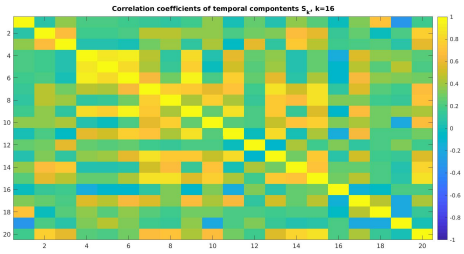
(a) Correlation coefficients of temporal components $\mathbf{S}_{13}$.



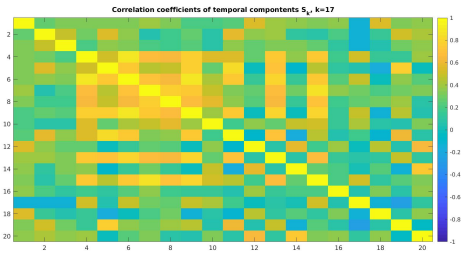
(b) Correlation coefficients of temporal components $\mathbf{S}_{14}$.



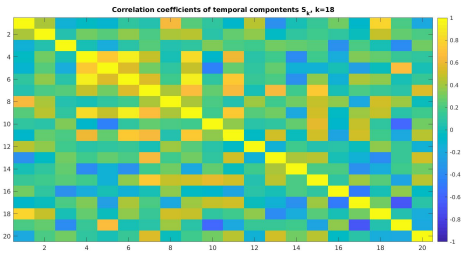
(a) Correlation coefficients of temporal components $\mathbf{S}_{15}$.



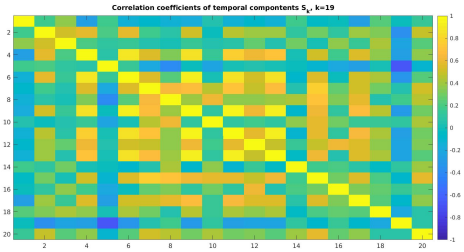
(b) Correlation coefficients of temporal components $\mathbf{S}_{16}$.



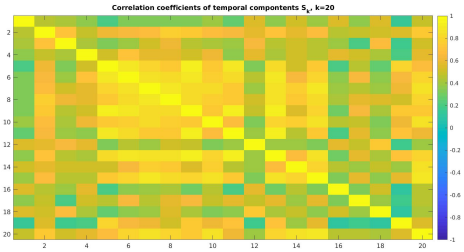
(a) Correlation coefficients of temporal components $\mathbf{S}_{17}$.



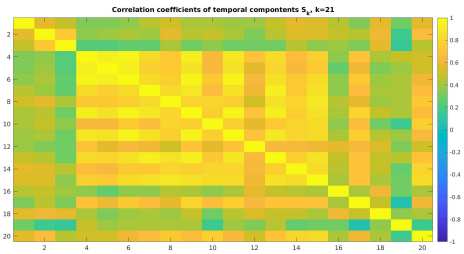
(b) Correlation coefficients of temporal components $\mathbf{S}_{18}$.



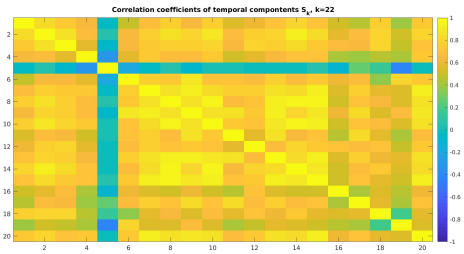
(a) Correlation coefficients of temporal components $\mathbf{S}_{19}$.



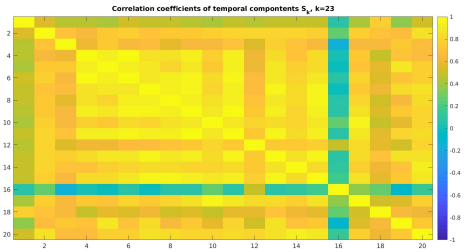
(b) Correlation coefficients of temporal components $\mathbf{S}_{20}$.



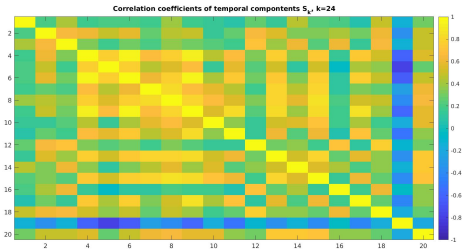
(a) Correlation coefficients of temporal components $\mathbf{S}_{21}$.



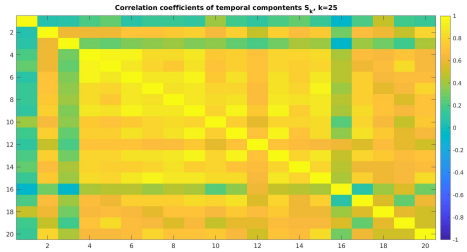
(b) Correlation coefficients of temporal components $\mathbf{S}_{22}$.



(a) Correlation coefficients of temporal components $\mathbf{S}_{23}$.



(b) Correlation coefficients of temporal components $\mathbf{S}_{24}$.



(a) Correlation coefficients of temporal components $\mathbf{S}_{25}$.

A.2 Figures for experimental evaluation of gap criterion

A.2.1 Case 1: Common orthonormal matrix A with each subject having independent temporal components S_k.

1. SNR=20 dB

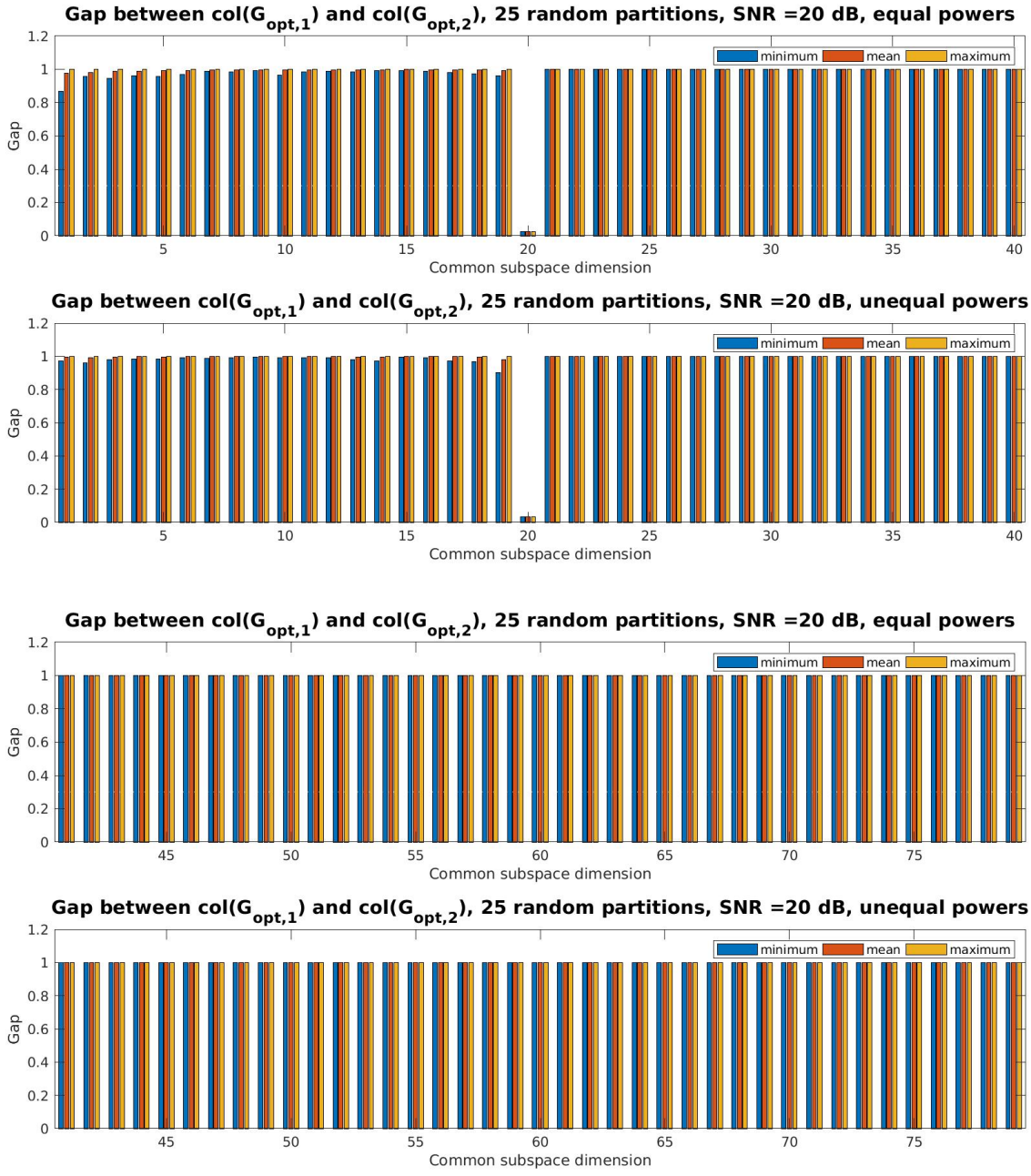


Figure A.14: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=20 dB. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension. **First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

2. SNR=15 dB

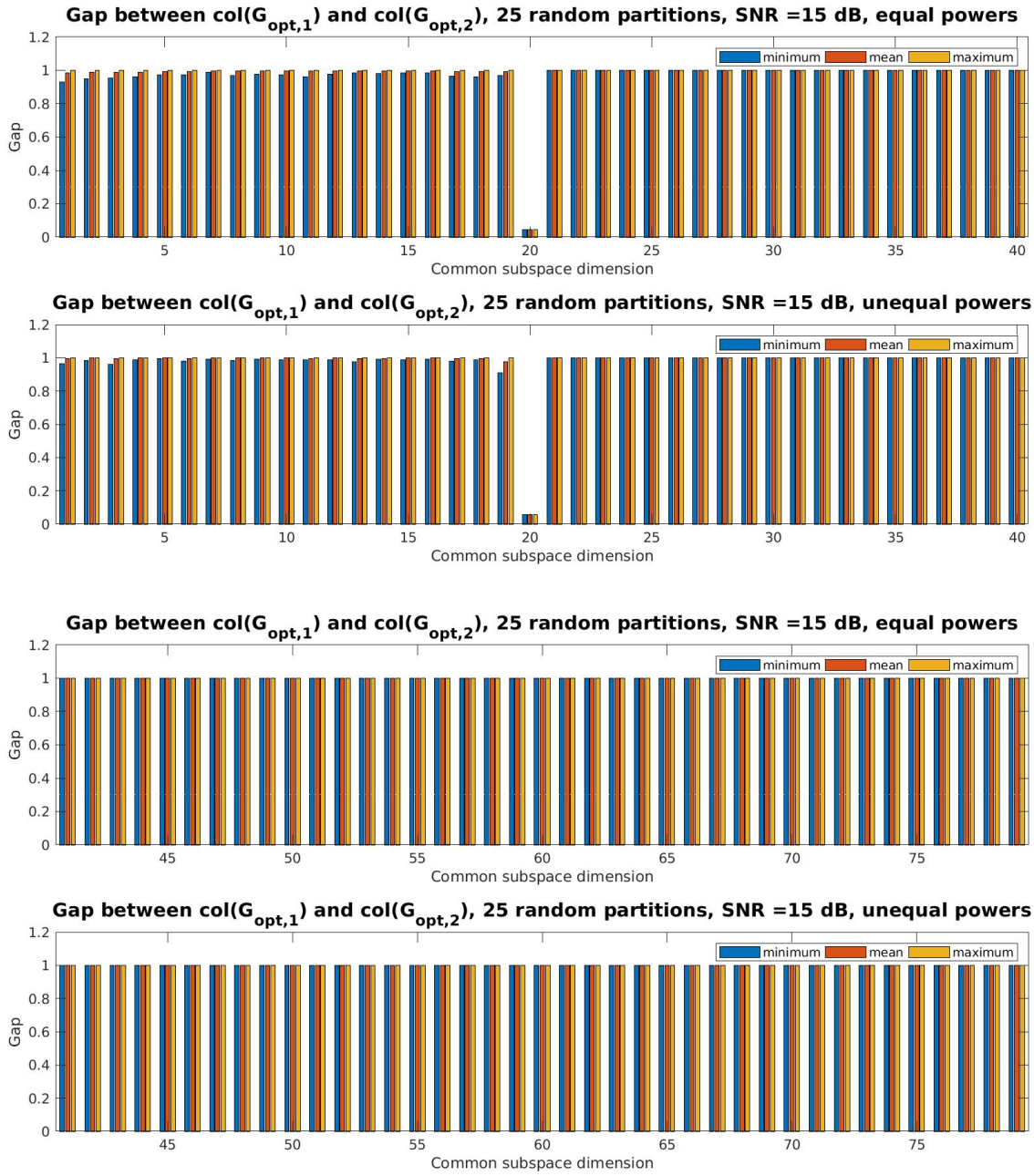


Figure A.15: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension, for SNR=15 dB. **First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

3. SNR=10 dB

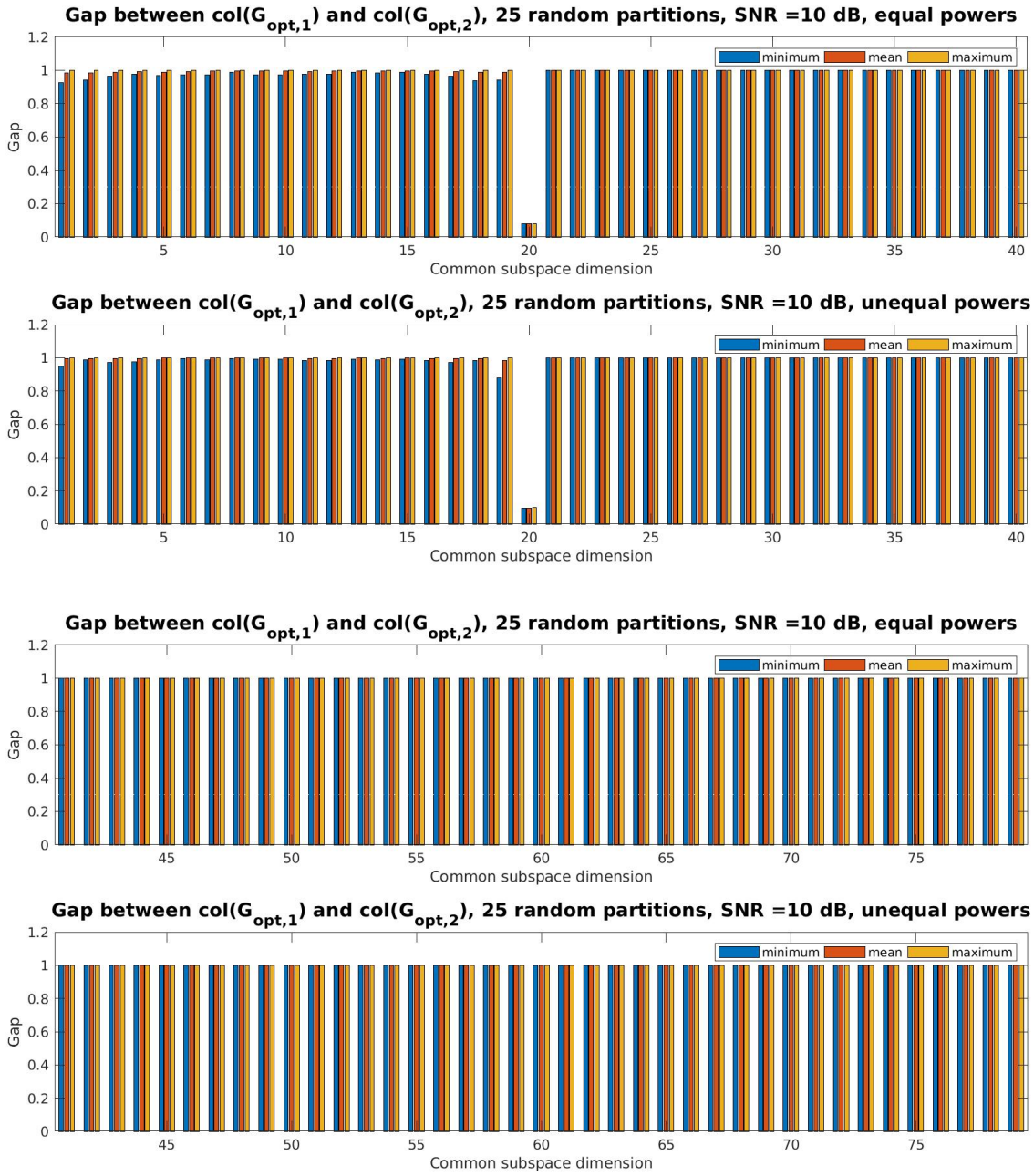


Figure A.16: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=10 dB. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension. **First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

4. SNR=5 dB

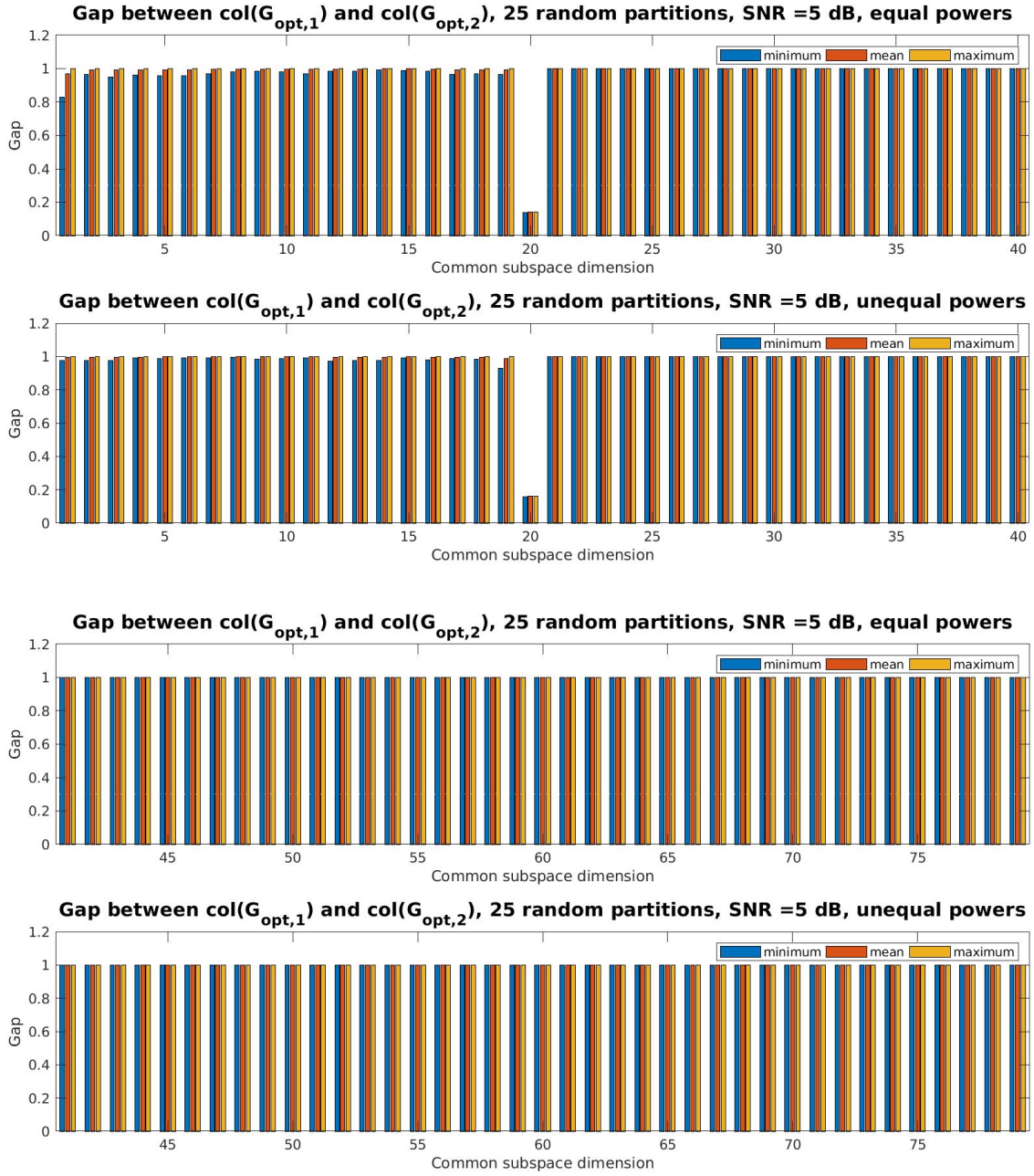


Figure A.17: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=5 dB. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension. **First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

5. SNR=−5 dB

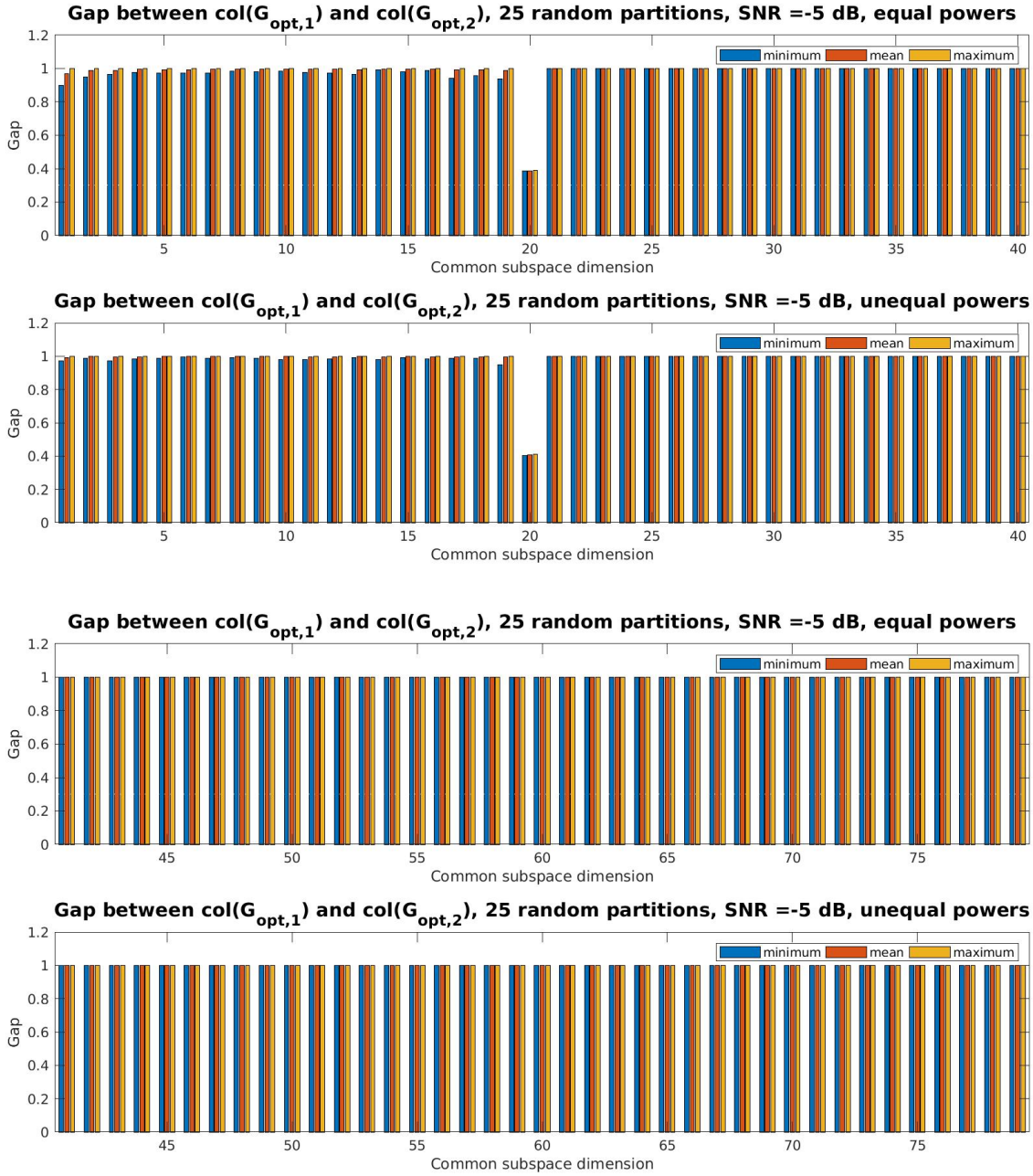


Figure A.18: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=−5 dB. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension.**First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

6. SNR=−10 dB

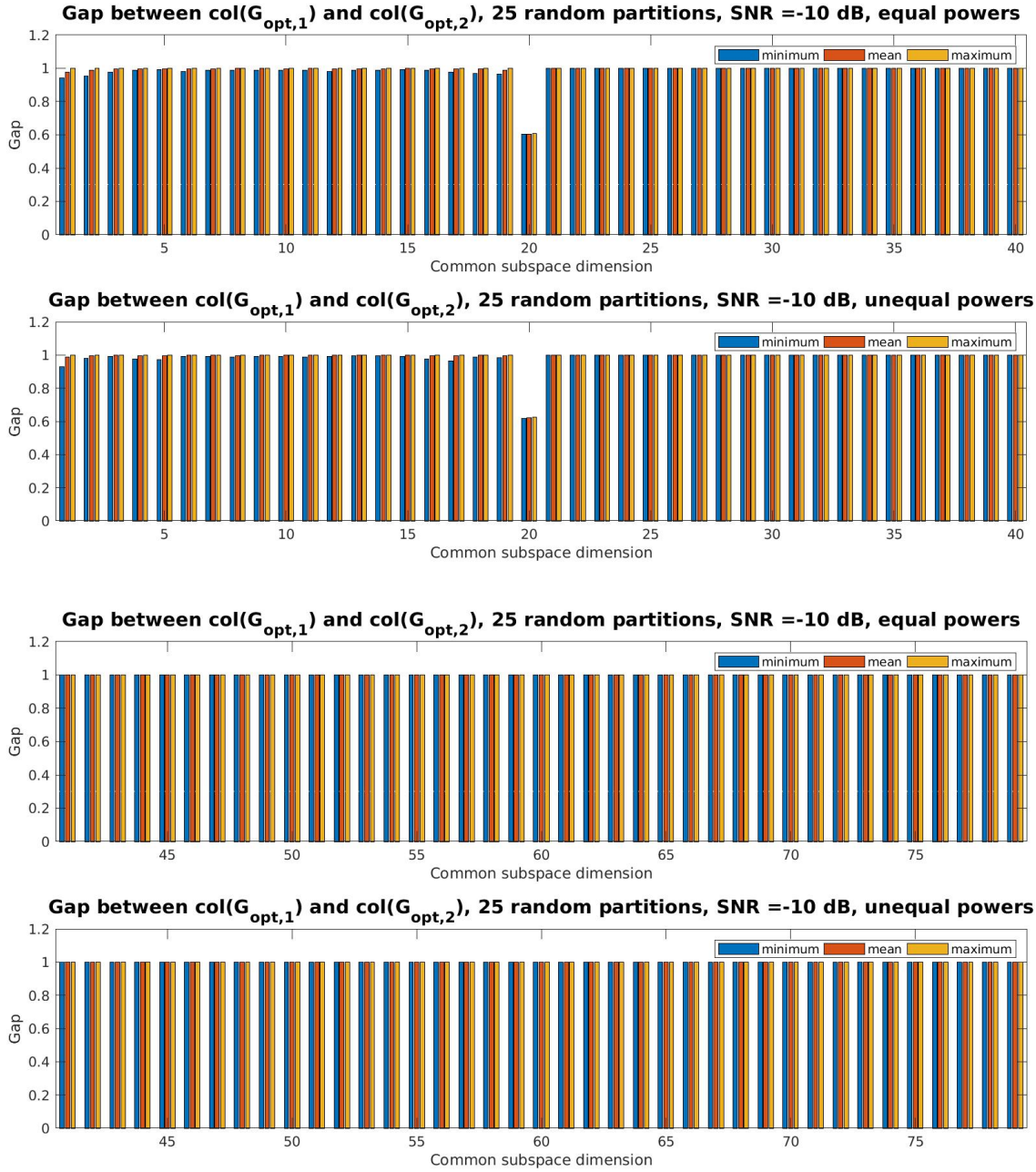


Figure A.19: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=−10 dB. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension.**First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

7. SNR=−15 dB

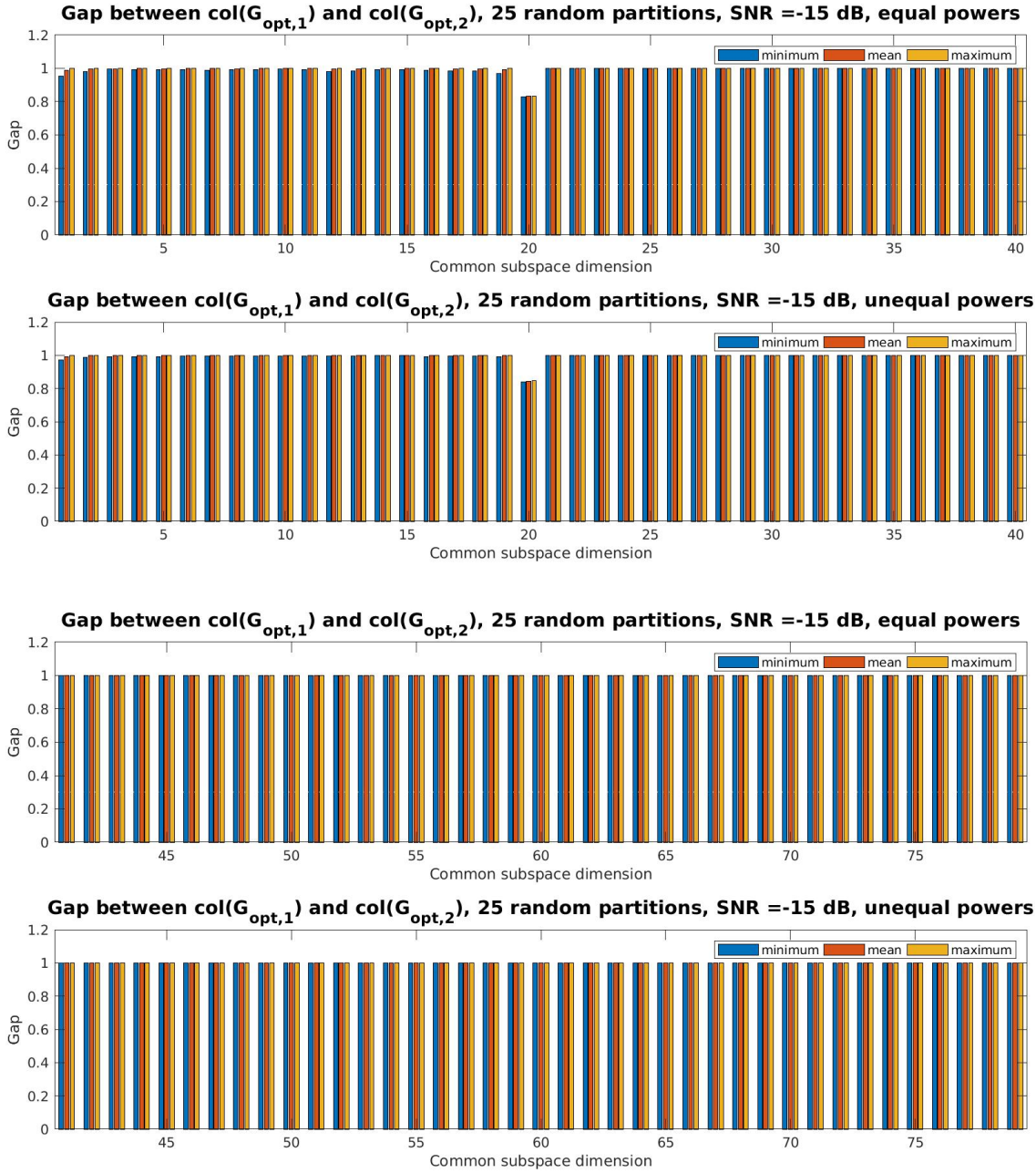


Figure A.20: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=−15 dB. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension. **First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

8. SNR=−20 dB

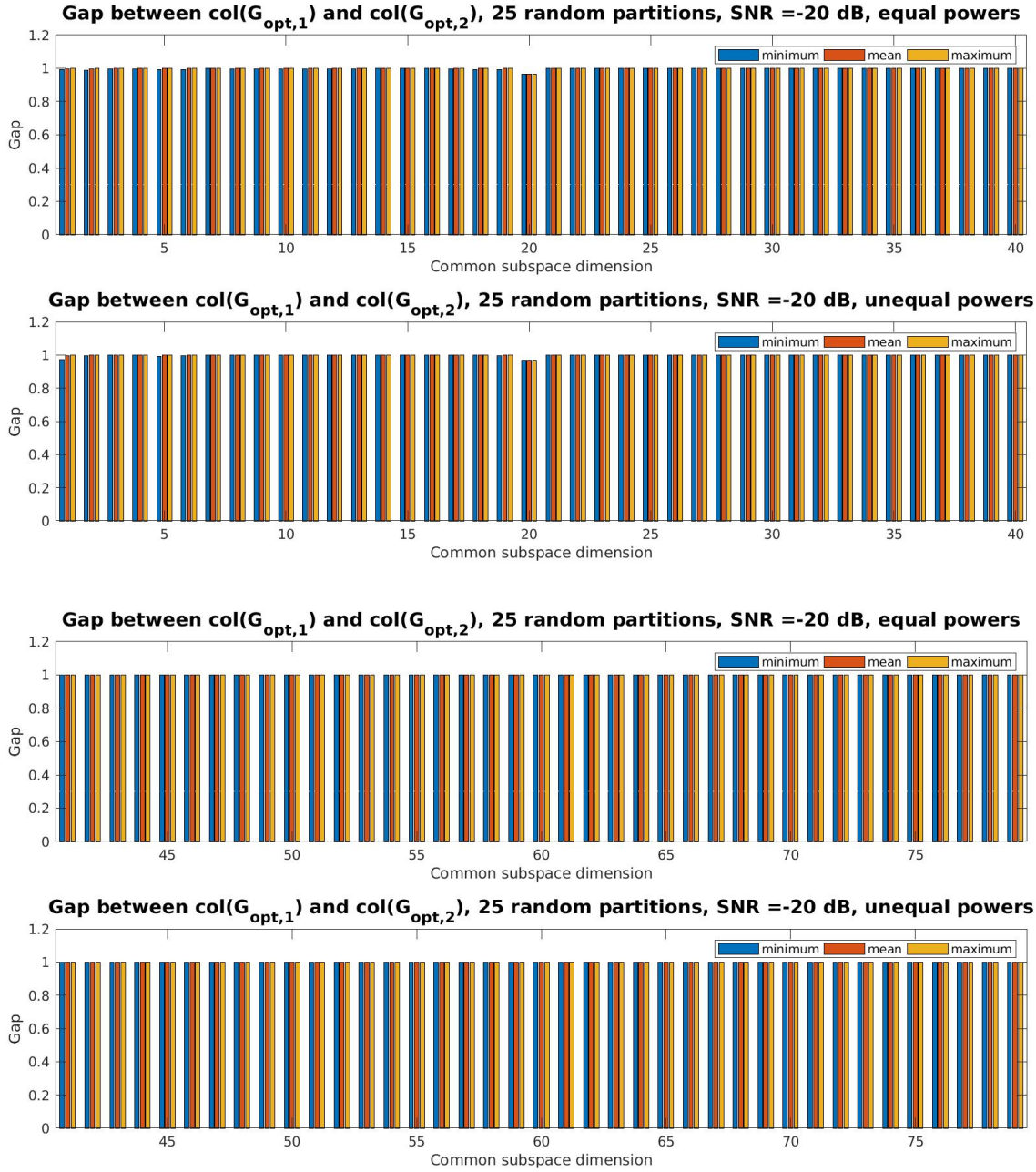


Figure A.21: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=−20 dB. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension. **First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

9. SNR=−25 dB

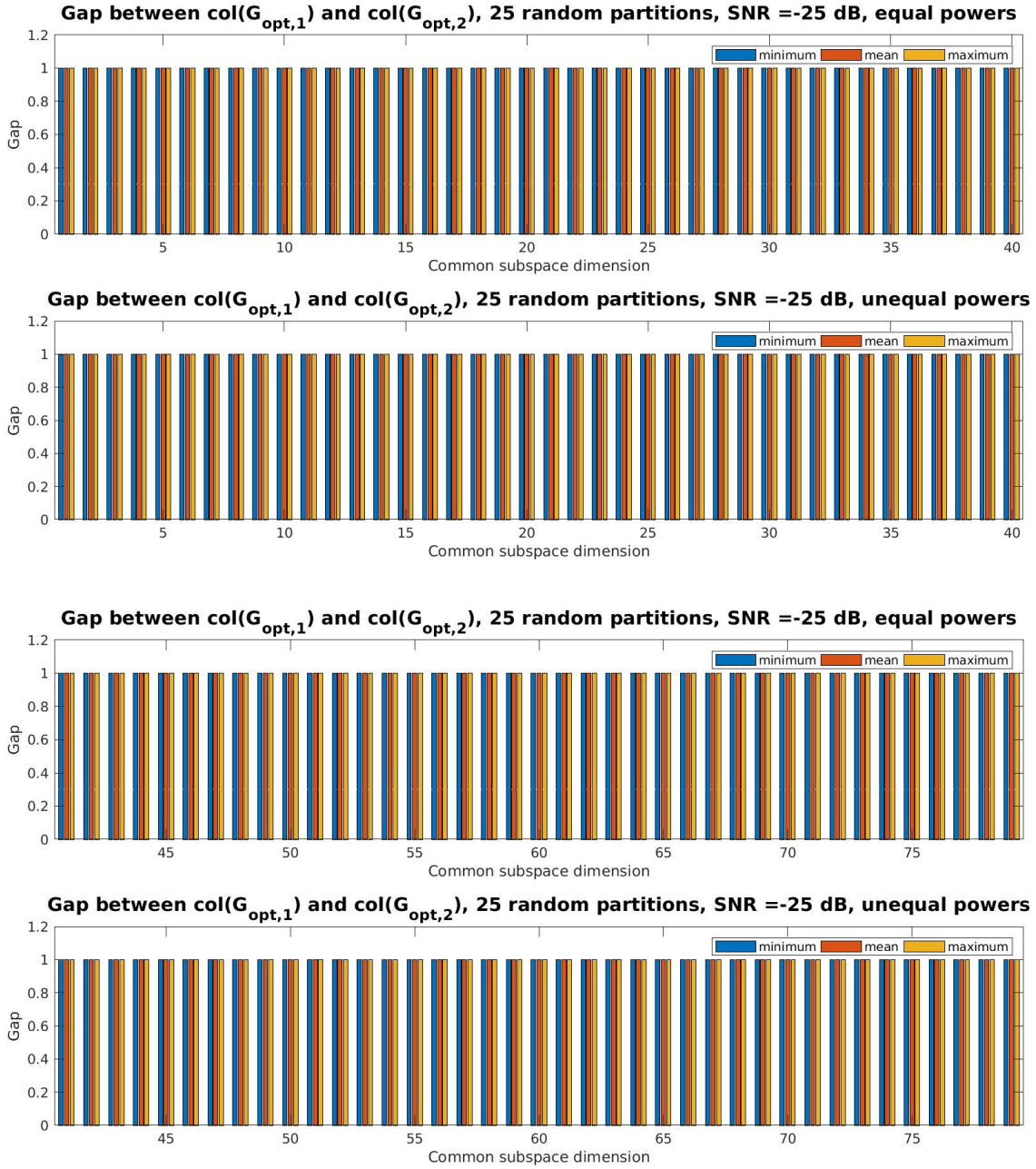


Figure A.22: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=−25 dB. We illustrate the minimum, mean and maximum value of the gaps for each common subspace dimension.**First and Third image**: rank-one terms of equal power. **Second and Fourth image**: rank-one terms of unequal power. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.99931	0.99385	0.99033	0.98621	0.99146	0.99628	0.99462	0.99463	0.99435	0.99513	0.99579
	2	0.99988	0.99911	0.99781	0.99594	0.9983	0.99442	0.996	0.99685	0.99747	0.99339	0.99776
	3	0.99992	0.99944	0.99815	0.99798	0.99563	0.9946	0.99682	0.9957	0.9959	0.99688	0.99592
	4	0.99994	0.9995	0.99818	0.99485	0.99621	0.99511	0.9972	0.99684	0.99779	0.99757	0.99845
	5	0.99994	0.9992	0.99779	0.99644	0.99755	0.99803	0.99762	0.99764	0.99894	0.9974	0.99766
D	**6**	0.99992	0.99946	0.99895	0.998	0.99854	0.99756	0.99843	0.99902	0.99727	0.99768	0.99778
I	**7**	0.99996	0.99964	0.99884	0.99868	0.99793	0.9986	0.99893	0.99773	0.99825	0.99828	0.99661
M	**8**	0.99997	0.99966	0.99863	0.997	0.99823	0.99836	0.99831	0.99883	0.99777	0.99835	0.99772
E	**9**	0.99996	0.9997	0.99888	0.99879	0.99835	0.99826	0.99752	0.99828	0.99799	0.99874	0.99786
N	**10**	0.99998	0.99968	0.99888	0.9985	0.99766	0.99781	0.9983	0.99793	0.99812	0.99812	0.99881
S	**11**	0.99998	0.99963	0.99883	0.99848	0.99705	0.99713	0.99824	0.99715	0.99719	0.99792	0.99907
I	**12**	0.99998	0.99972	0.99919	0.99823	0.99704	0.99738	0.99694	0.99623	0.99762	0.99822	0.99763
O	**13**	0.99997	0.99972	0.99898	0.9981	0.99753	0.99638	0.99651	0.99745	0.99526	0.99685	0.99829
N	**14**	0.99997	0.99977	0.99953	0.99852	0.99686	0.99598	0.99699	0.99684	0.99722	0.99606	0.99895
	15	0.99998	0.9997	0.99941	0.9979	0.99857	0.99732	0.99771	0.99777	0.99804	0.99878	0.99787
	16	0.99997	0.99947	0.99884	0.99627	0.99654	0.9974	0.99715	0.99666	0.99864	0.99851	0.99847
	17	0.99998	0.99959	0.99889	0.99462	0.99708	0.99743	0.99725	0.99527	0.99675	0.99773	0.99768
	18	0.99997	0.99971	0.99926	0.99832	0.99727	0.99697	0.99505	0.99642	0.99682	0.99495	0.99736
	19	0.99999	0.99941	0.99878	0.99633	0.99674	0.99442	0.98754	0.98378	0.97461	0.98034	0.98969
	20	0.99999	0.96789	0.84333	0.62093	0.40623	0.25817	0.16062	0.096647	0.056118	0.031989	0.0003215
	21	1	0.99999	1	1	0.99999	1	0.99999	0.99999	1	1	0.99999

Table A.1: Mean value of the gaps for each common subspace dimension and each SNR, for the case in which rank-one terms are of unequal power. When the common subspace dimension is greater than 21 the mean value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	1	0.99999	0.99997	0.99999	0.99993	0.99999	1	0.99997	0.99999	0.99999	0.99994
	2	1	1	1	0.99998	0.99998	1	0.99994	1	0.99999	0.99991	0.99998
	3	1	1	1	1	1	1	1	0.99998	1	0.99999	1
	4	1	1	1	0.99999	1	1	0.99999	1	1	0.99996	1
	5	1	1	1	1	1	0.99999	0.99999	0.99996	1	1	1
D	**6**	1	1	1	1	1	0.99999	1	1	1	0.99999	0.99999
I	**7**	1	1	1	0.99999	0.99999	1	1	1	0.99999	1	0.99997
M	**8**	1	1	0.99999	1	0.99995	1	0.99998	1	0.99998	1	1
E	**9**	1	1	0.99996	0.99998	1	1	1	0.99999	1	0.99999	1
N	**10**	1	1	1	0.99997	1	0.99999	1	0.99998	1	0.99998	1
S	**11**	1	1	1	1	0.99999	1	0.99999	1	1	0.99998	1
I	**12**	1	1	1	0.99999	1	0.99998	1	0.99994	1	1	1
O	**13**	1	1	1	1	0.99999	0.99999	0.99998	0.99999	0.99999	1	0.99998
N	**14**	1	1	1	1	1	1	0.99998	0.99998	0.99998	0.99999	1
	15	1	1	1	1	1	0.99997	1	0.99997	1	1	0.99994
	16	1	1	1	0.99997	1	0.99998	1	1	1	1	1
	17	1	1	1	0.99999	1	0.99996	1	0.99998	0.99999	1	0.99996
	18	1	1	0.99999	1	0.99997	0.99996	1	0.99999	0.99997	1	0.99999
	19	1	1	1	1	0.99998	1	1	1	0.99999	0.99997	0.99998
	20	1	0.96992	0.84592	0.62467	0.41028	0.26605	0.16187	0.097628	0.056775	0.032242	0.00032589

Table A.2: Maximum value of the gaps for each common subspace dimension and each SNR,for the case in which rank-one terms are of unequal power. When the common subspace dimension is greater than 20 the maximum value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.99713	0.97013	0.97096	0.92882	0.97305	0.98261	0.97478	0.94809	0.96398	0.9727	0.97855
	2	0.99905	0.99608	0.9859	0.97905	0.98656	0.96536	0.97455	0.98693	0.98522	0.96129	0.98879
	3	0.99958	0.9975	0.99026	0.99188	0.97218	0.96349	0.97723	0.97023	0.96	0.98171	0.95639
	4	0.99983	0.99744	0.99152	0.97705	0.98222	0.96656	0.99065	0.9768	0.98623	0.98525	0.99059
	5	0.99978	0.99225	0.99155	0.97058	0.98917	0.99131	0.98755	0.98616	0.99535	0.9836	0.98248
D	**6**	0.99976	0.9968	0.9958	0.98995	0.99407	0.98885	0.99145	0.99691	0.97834	0.99021	0.98941
I	**7**	0.9998	0.99849	0.99394	0.99341	0.98763	0.99379	0.99069	0.98934	0.99188	0.98746	0.98284
M	**8**	0.99987	0.99753	0.99571	0.98762	0.9899	0.98997	0.99401	0.99585	0.98347	0.99276	0.98697
E	**9**	0.99978	0.99894	0.99557	0.99324	0.9893	0.99384	0.98554	0.9906	0.98993	0.99435	0.98879
N	**10**	0.99986	0.99882	0.99579	0.98971	0.97794	0.99179	0.98714	0.99047	0.98683	0.99122	0.99469
S	**11**	0.99991	0.99776	0.9956	0.98621	0.98	0.98836	0.99254	0.98501	0.98723	0.99055	0.99493
I	**12**	0.99988	0.99931	0.99723	0.99088	0.98527	0.98555	0.9736	0.98339	0.98758	0.99106	0.98816
O	**13**	0.99992	0.99913	0.99407	0.99413	0.99055	0.98504	0.97491	0.9911	0.97673	0.9797	0.99245
N	**14**	0.99986	0.99882	0.99781	0.99448	0.97964	0.98304	0.97611	0.98888	0.9914	0.97136	0.99388
	15	0.99989	0.99792	0.99792	0.99258	0.99034	0.99261	0.99148	0.99345	0.98832	0.99541	0.98431
	16	0.99989	0.99817	0.99007	0.97681	0.98467	0.99126	0.98125	0.98212	0.99314	0.9929	0.99498
	17	0.99987	0.99798	0.99465	0.96506	0.98608	0.98399	0.9866	0.97391	0.9812	0.97226	0.99263
	18	0.99986	0.99891	0.99716	0.98763	0.9859	0.97326	0.98488	0.98471	0.98912	0.96985	0.98481
	19	0.99995	0.99691	0.99145	0.98199	0.94916	0.97364	0.92763	0.87964	0.90791	0.90207	0.92759
	20	0.99996	0.96647	0.84074	0.61827	0.40345	0.25506	0.15943	0.095923	0.055503	0.031696	0.00031821
	21	0.99999	0.99998	0.99999	0.99997	0.99998	0.99998	0.99995	0.99993	0.99999	0.99998	0.99996
	22	0.99999	0.99999	0.99999	0.99999	0.99998	0.99998	0.99999	0.99999	0.99999	0.99999	0.99999
	23	0.99998	0.99999	0.99999	1	0.99999	0.99999	0.99999	0.99999	1	0.99999	0.99999
	24	0.99999	0.99999	1	1	0.99999	0.99999	1	1	0.99999	1	0.99999
	25	1	1	0.99999	1	1	1	1	0.99999	1	0.99999	1
	26	1	1	1	1	1	0.99999	1	1	1	1	1
	27	1	1	1	1	1	0.99999	1	1	1	1	1
	28	1	1	1	1	1	1	0.99999	1	1	1	1

Table A.3: Minimum value of the gaps for each common subspace dimension and each SNR, for the case in which rank-one terms are of unequal power. When the common subspace dimension is greater than 28 the minimum value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.99954	0.99775	0.98761	0.976	0.96811	0.96827	0.96968	0.98447	0.98487	0.97577	0.97492
	2	0.9998	0.99737	0.99486	0.98949	0.99003	0.98961	0.99033	0.9847	0.98812	0.98233	0.98812
	3	0.99988	0.99903	0.99799	0.99422	0.98947	0.9874	0.99057	0.98913	0.98909	0.9875	0.98515
	4	0.99991	0.99923	0.99816	0.99692	0.99513	0.99238	0.99132	0.99415	0.98775	0.98782	0.99083
	5	0.99993	0.99919	0.99766	0.99686	0.99404	0.99478	0.99214	0.9902	0.99349	0.99082	0.99197
D	**6**	0.99993	0.99921	0.99809	0.99548	0.99288	0.99395	0.99096	0.99317	0.99268	0.99175	0.99464
I	**7**	0.99995	0.99958	0.99819	0.99663	0.99616	0.99594	0.99519	0.9956	0.99676	0.99684	0.99685
M	**8**	0.99997	0.99923	0.99766	0.99824	0.9967	0.99705	0.99757	0.99804	0.99511	0.99662	0.99793
E	**9**	0.99995	0.99925	0.99826	0.99823	0.99725	0.99708	0.99574	0.99663	0.99678	0.99737	0.99866
N	**10**	0.99998	0.99953	0.99911	0.99748	0.99739	0.99792	0.99558	0.99451	0.99527	0.99738	0.99676
S	**11**	0.99998	0.99956	0.9987	0.99823	0.99659	0.99777	0.99504	0.99397	0.99464	0.99702	0.99544
I	**12**	0.99998	0.99953	0.99614	0.99557	0.99525	0.99474	0.99643	0.99433	0.99537	0.99702	0.99685
O	**13**	0.99998	0.99953	0.9969	0.99573	0.99364	0.99618	0.99721	0.99641	0.99685	0.99762	0.99687
N	**14**	0.99997	0.9996	0.9988	0.99764	0.99664	0.99775	0.99839	0.99662	0.99653	0.99784	0.99702
	15	0.99996	0.9995	0.99906	0.99851	0.99728	0.99755	0.99807	0.997	0.99663	0.9982	0.99767
	16	0.99997	0.99959	0.99817	0.99736	0.99741	0.99694	0.99611	0.99513	0.9968	0.99713	0.99559
	17	0.99998	0.99944	0.99656	0.99433	0.99318	0.99161	0.99325	0.99195	0.99282	0.99516	0.99481
	18	0.99993	0.99866	0.99599	0.99468	0.99273	0.9879	0.99301	0.98823	0.99246	0.99488	0.99202
	19	0.99994	0.99832	0.99152	0.99028	0.98847	0.98963	0.99278	0.98971	0.9906	0.99085	0.99273
	20	0.99995	0.96406	0.83194	0.60414	0.3875	0.23669	0.14005	0.08065	0.045768	0.025823	0.00025857
	21	1	1	0.99999	1	0.99999	1	1	0.99999	0.99999	1	0.99999

Table A.4: Mean value of the gaps for each common subspace dimension and each SNR, for the case in which rank-one terms are of equal power. When the common subspace dimension is greater than 21 the mean value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	1	1	0.99995	0.99984	0.99944	1	0.99969	0.99998	0.99998	1	0.99969
	2	1	0.99999	0.99999	0.99999	0.99983	0.99999	0.99999	0.99999	1	0.99997	1
	3	1	1	1	0.99997	0.9999	1	0.99999	0.99987	0.99993	0.99999	0.99998
	4	1	1	0.99999	0.99998	0.99999	1	1	0.99999	0.99993	1	0.9996
	5	1	1	0.99999	1	1	1	0.99993	0.99994	0.99999	0.99996	0.99998
D	**6**	1	1	1	0.99998	1	0.99998	1	0.99998	0.99997	1	0.99999
I	**7**	1	1	1	1	1	0.99999	1	1	1	0.99999	1
M	**8**	1	1	1	0.99999	0.99999	1	1	1	1	0.99999	1
E	**9**	1	1	0.99998	1	0.99999	1	0.99996	1	1	1	0.99999
N	**10**	1	1	1	1	1	1	1	1	1	1	1
S	**11**	1	1	0.99999	1	0.99995	0.99994	0.99997	1	1	0.99995	0.9997
I	**12**	1	1	0.99999	0.99997	0.99996	0.99992	1	1	0.99993	0.99998	1
O	**13**	1	1	1	0.99999	0.99988	0.99999	1	0.99999	0.99999	1	1
N	**14**	1	0.99999	0.99999	1	1	1	1	1	1	1	1
	15	1	1	1	1	0.99996	1	1	1	1	1	1
	16	1	1	1	0.99997	0.99999	0.99999	0.99985	1	0.99995	1	0.99997
	17	1	1	0.99999	1	1	0.99999	0.99979	0.99987	0.99995	0.99999	0.99998
	18	1	0.99999	0.99999	1	0.99998	0.99998	1	0.99999	1	1	1
	19	1	1	0.99998	1	0.99999	0.99993	1	1	0.99992	0.99995	1
	20	1	0.96513	0.83417	0.60756	0.38939	0.23824	0.14048	0.080887	0.046005	0.026028	0.00025973

Table A.5: Maximum value of the gaps for each common subspace dimension and each SNR,for the case in which rank-one terms are of equal power. When the common subspace dimension is greater than 20 the maximum value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.99818	0.99128	0.95285	0.93991	0.89891	0.83447	0.82796	0.92685	0.92866	0.86616	0.93484
	2	0.99927	0.98915	0.97985	0.95202	0.94777	0.92494	0.96436	0.93987	0.9499	0.95686	0.93675
	3	0.99948	0.99558	0.99435	0.97673	0.96551	0.95246	0.95063	0.96339	0.95422	0.94527	0.92746
	4	0.99952	0.9951	0.9932	0.9896	0.97636	0.94361	0.96269	0.97505	0.96138	0.96056	0.95424
	5	0.99953	0.99415	0.99267	0.99113	0.97415	0.98532	0.95672	0.9695	0.97131	0.95908	0.97231
D	**6**	0.99944	0.99088	0.99048	0.97943	0.97347	0.95752	0.95786	0.97126	0.97285	0.96734	0.96895
I	**7**	0.99967	0.99839	0.98749	0.98829	0.97473	0.9552	0.96804	0.97166	0.98719	0.98792	0.98472
M	**8**	0.99982	0.99633	0.99255	0.98941	0.98305	0.97618	0.98012	0.98725	0.96862	0.98475	0.98658
E	**9**	0.9997	0.9943	0.99168	0.98977	0.98083	0.98448	0.98426	0.97098	0.97733	0.9917	0.99467
N	**10**	0.99991	0.99734	0.99703	0.98746	0.98577	0.99291	0.98169	0.97292	0.97352	0.96387	0.97749
S	**11**	0.99988	0.99768	0.9939	0.98899	0.97572	0.98839	0.96796	0.97659	0.9596	0.98602	0.97927
I	**12**	0.99987	0.99633	0.97964	0.98117	0.9715	0.97165	0.98611	0.97722	0.97727	0.98795	0.98073
O	**13**	0.99988	0.99696	0.98471	0.98647	0.9641	0.98203	0.9858	0.98733	0.98566	0.98322	0.98709
N	**14**	0.99984	0.99829	0.99395	0.98985	0.99044	0.99198	0.99062	0.98561	0.98156	0.99117	0.9824
	15	0.99982	0.99839	0.99125	0.99119	0.98129	0.99195	0.98976	0.98725	0.98623	0.99193	0.9862
	16	0.99986	0.99861	0.98999	0.98783	0.98833	0.98181	0.9864	0.97652	0.9827	0.98918	0.96928
	17	0.99993	0.99665	0.98571	0.97768	0.94261	0.96594	0.96485	0.96639	0.96609	0.98069	0.98096
	18	0.99935	0.99255	0.98595	0.96834	0.95815	0.95504	0.9678	0.93764	0.96116	0.9726	0.97933
	19	0.9997	0.99215	0.96714	0.96317	0.93814	0.95827	0.96502	0.94249	0.96945	0.95998	0.95985
	20	0.99976	0.96345	0.83063	0.60226	0.38582	0.23559	0.13957	0.080361	0.045532	0.02571	0.00025755
	21	0.99997	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99997	0.99998	0.99997	0.99995
	22	0.99999	0.99998	0.99999	0.99998	0.99998	0.99999	0.99999	0.99999	0.99998	0.99999	0.99999
	23	0.99999	0.99999	0.99999	1	0.99999	0.99999	0.99999	0.99999	1	1	1
	24	0.99999	1	0.99999	1	1	0.99999	1	1	1	0.99999	1
	25	1	1	1	1	1	1	1	1	1	0.99999	1
	26	1	1	0.99999	1	1	1	0.99999	1	0.99999	0.99999	1
	27	1	1	1	1	1	0.99999	1	1	1	1	1
	28	1	1	1	1	1	1	0.99999	1	1	1	1

Table A.6: Minimum value of the gaps for each common subspace dimension and each SNR, for the case in which rank-one terms are of equal power. When the common subspace dimension is greater than 28 the minimum value of gaps is equal to 1.

A.2.2 Case 2: Common orthonormal matrix A with each subject having correlated temporal components S_k.

1. SNR=20 dB, SNR=15 dB,

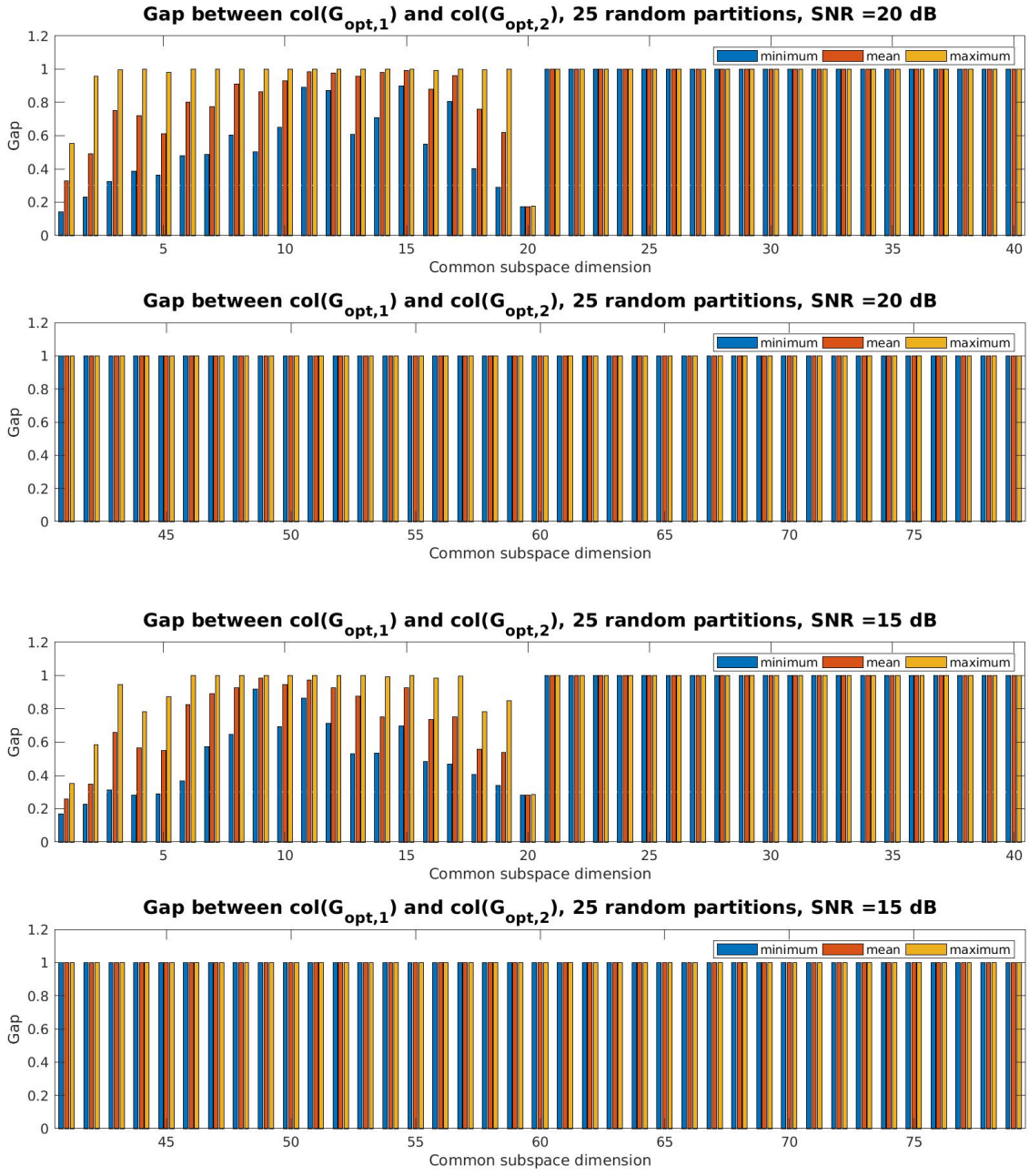


Figure A.23: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First, Second image** SNR=20 dB. **Third, Fourth image**: SNR=15 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

2. SNR=10 dB, SNR=5 dB

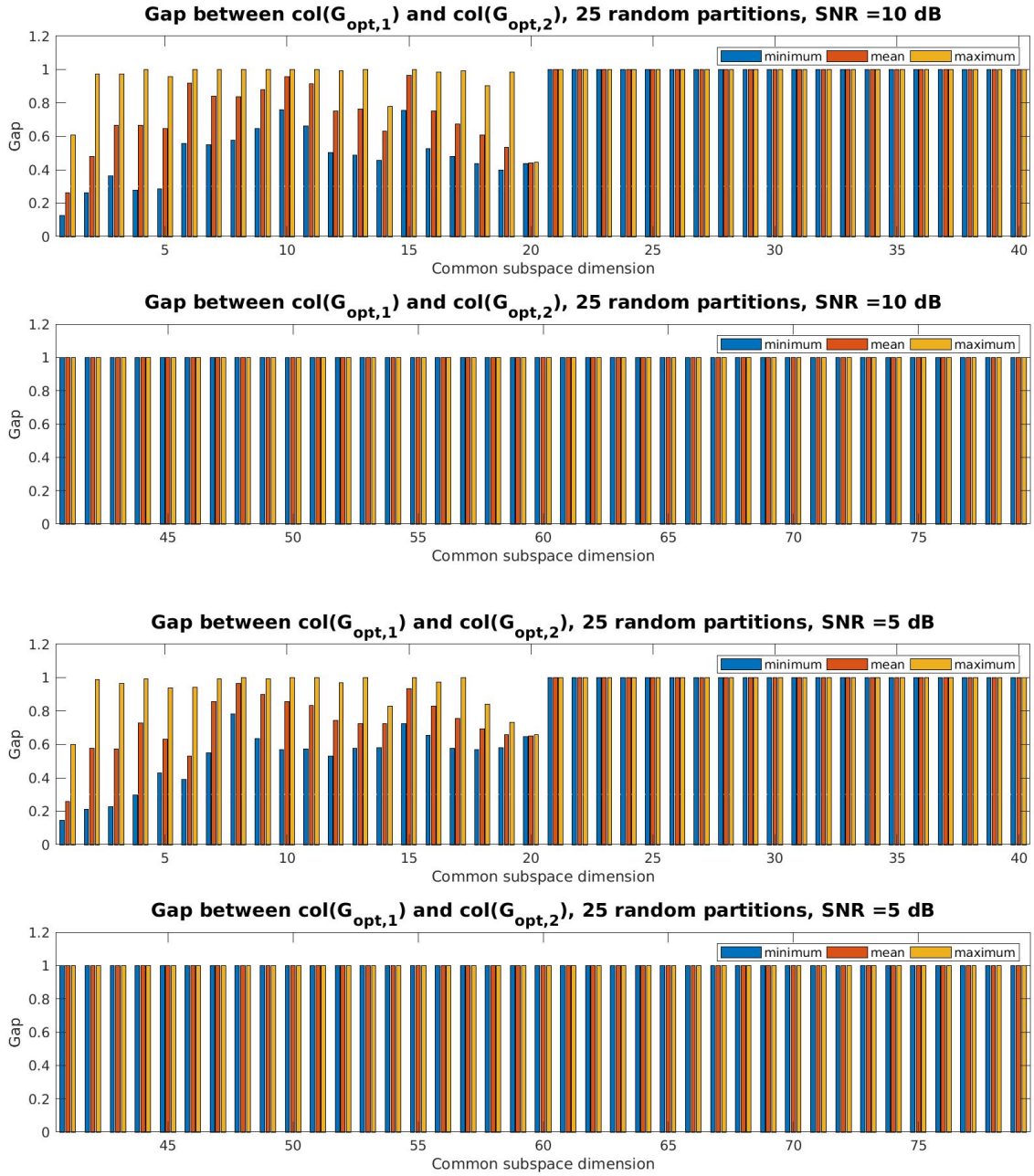


Figure A.24: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First, Second image**: SNR=10 dB. **Third, Fourth image**: SNR=5 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

3. SNR=−5 dB, SNR=−10 dB

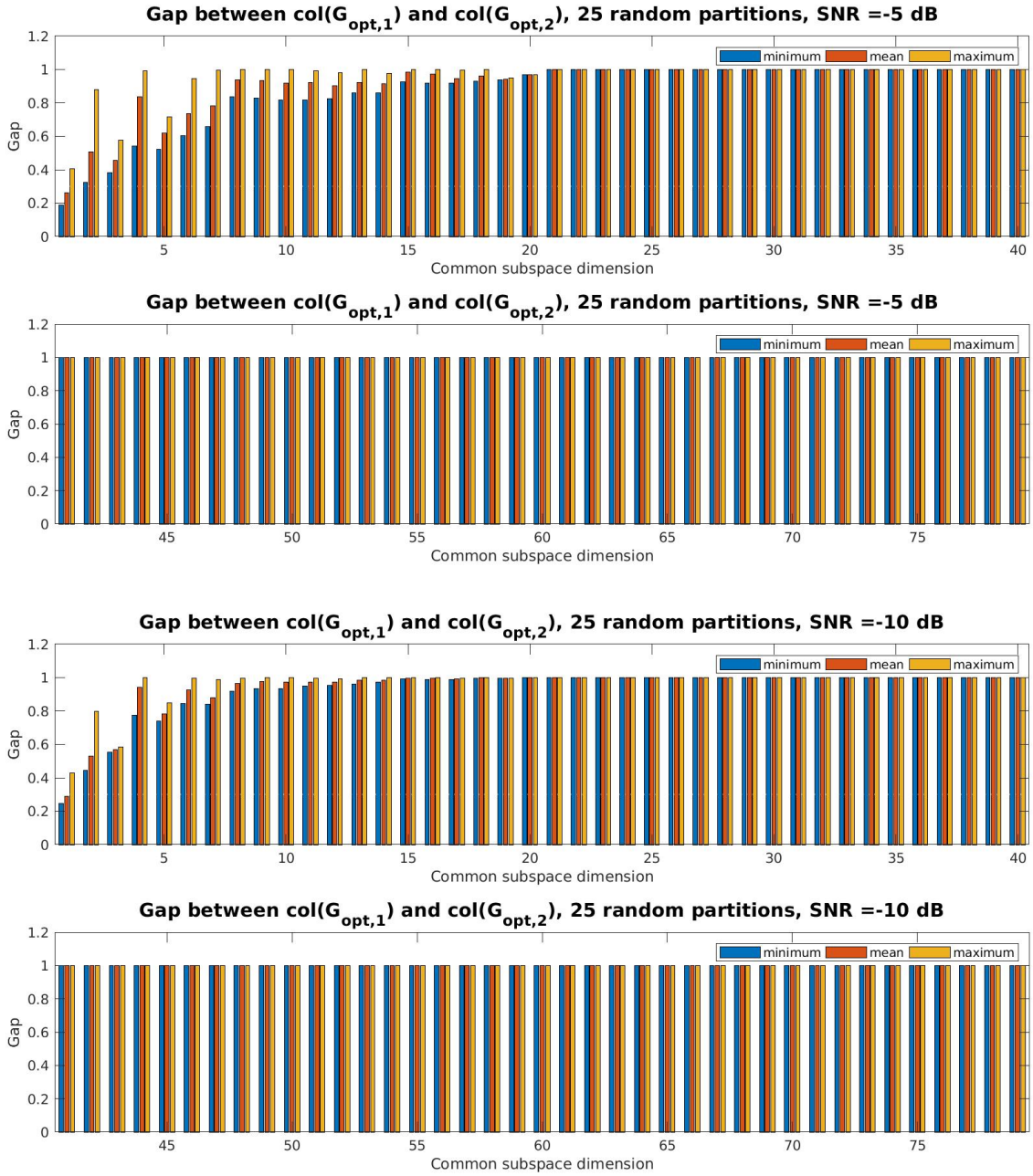


Figure A.25: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First, Second image**: SNR=−5 dB. **Third, Fourth image**: SNR=−10 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

4. SNR=−15 dB, SNR=−20 dB

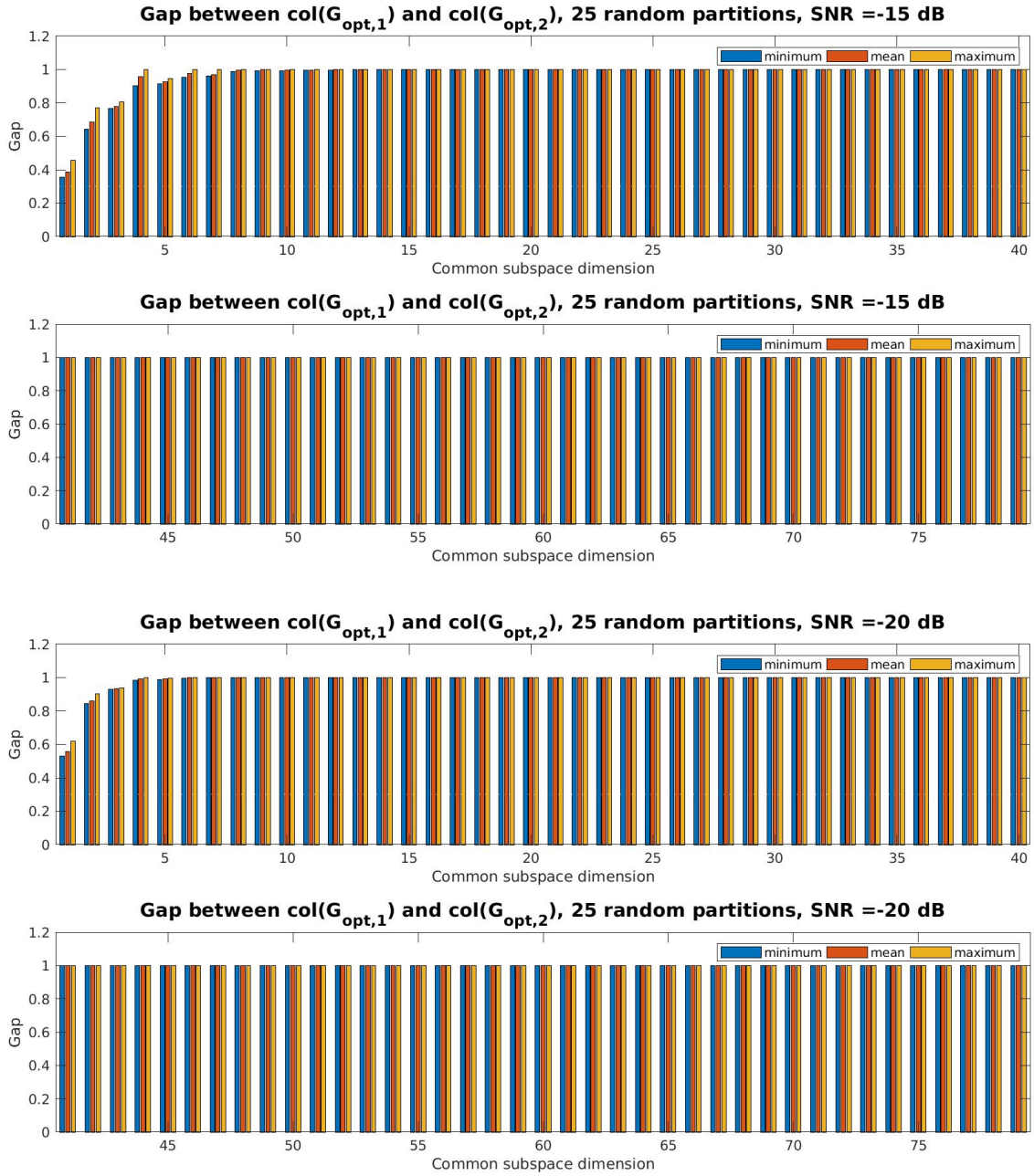


Figure A.26: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First, Second image**: SNR=−15 dB. **Third, Fourth image**: SNR=0 dB: SNR=−20 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value. Upper image: SNR=−15 dB., Lower image: SNR=−20 dB.

5. SNR=−25 dB

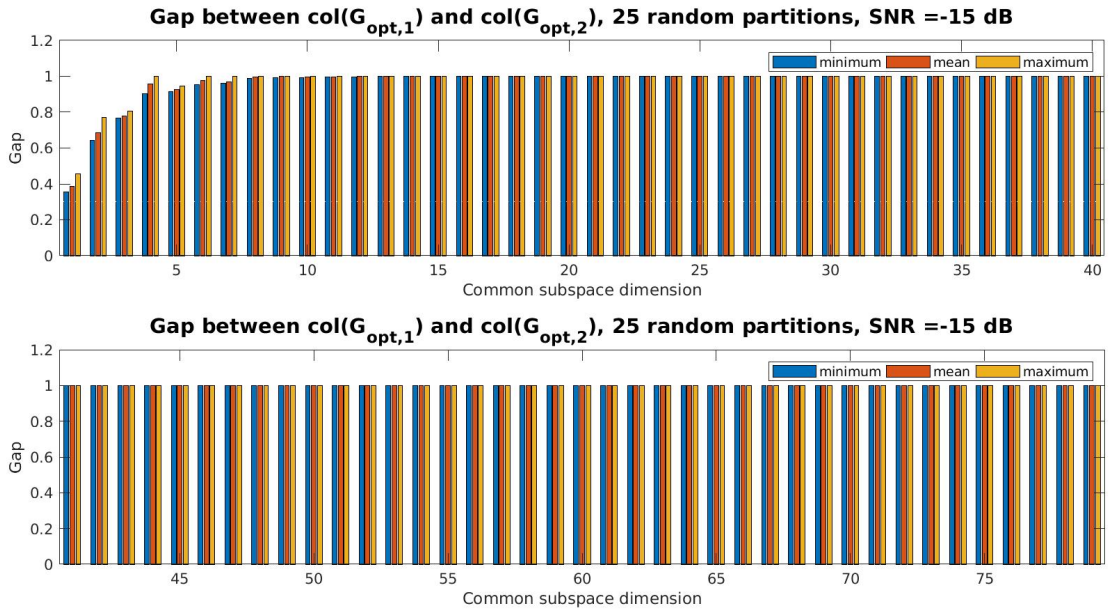


Figure A.27: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=−25 dB. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.77118	0.55653	0.3858	0.28778	0.26124	0.25434	0.25767	0.26119	0.25821	0.32716	0.3695
	2	0.97378	0.86144	0.68401	0.53162	0.50642	0.69811	0.57502	0.47918	0.3491	0.49034	0.55686
	3	0.99477	0.93422	0.77799	0.56912	0.45623	0.42269	0.57452	0.66761	0.65764	0.75119	0.85379
	4	0.99999	0.9939	0.95722	0.94004	0.8386	0.79234	0.72689	0.66736	0.5653	0.7191	0.75876
	5	1	0.99167	0.92568	0.78066	0.61936	0.52947	0.63032	0.64607	0.54795	0.60995	0.65656
D	**6**	1	0.99915	0.97743	0.92716	0.73769	0.56327	0.53041	0.91968	0.82397	0.80151	0.83053
I	**7**	1	0.99966	0.97004	0.88105	0.78178	0.79459	0.85529	0.83937	0.89016	0.77365	0.80784
M	**8**	1	0.99999	0.99527	0.96526	0.93838	0.91243	0.96656	0.83631	0.92558	0.91224	0.96118
E	**9**	1	1	0.99832	0.97644	0.93398	0.87936	0.90007	0.87784	0.98466	0.86497	0.89268
N	**10**	1	1	0.99666	0.97333	0.91838	0.85869	0.85597	0.95911	0.94505	0.9285	0.92432
S	**11**	1	1	0.99801	0.97118	0.92243	0.84423	0.83365	0.9136	0.97231	0.98346	0.97984
I	**12**	1	1	0.99855	0.97274	0.90391	0.81184	0.74338	0.75071	0.92735	0.9767	0.93786
O	**13**	1	1	0.99986	0.98494	0.92091	0.79862	0.72539	0.76479	0.8752	0.95627	0.97837
N	**14**	1	1	0.99994	0.98291	0.91503	0.79573	0.72367	0.63183	0.74957	0.97986	0.91912
	15	1	1	0.99999	0.99799	0.98337	0.96868	0.93531	0.96359	0.92576	0.99346	0.81255
	16	1	1	1	0.99736	0.97471	0.9256	0.82976	0.75334	0.73618	0.87777	0.91968
	17	1	1	1	0.9923	0.94508	0.83475	0.75663	0.67285	0.75258	0.96068	0.71876
	18	1	1	1	0.99823	0.96238	0.8592	0.69133	0.6074	0.55854	0.75753	0.77495
	19	1	1	1	0.99589	0.94319	0.80385	0.65869	0.53325	0.53666	0.61784	0.97725
	20	1	1	1	0.99982	0.96858	0.85253	0.64881	0.4409	0.28323	0.17465	0.0020529
	21	1	1	1	0.99999	1	1	0.99999	1	1	1	0.99999

Table A.7: Mean value of the gaps for each common subspace dimension and each SNR. When the common subspace dimension is greater than 21 the mean value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.79207	0.61948	0.45705	0.43105	0.40639	0.51728	0.59966	0.60891	0.35217	0.55389	0.55298
	2	0.98169	0.9016	0.76911	0.79856	0.87855	1	0.98669	0.97425	0.58576	0.95747	0.96627
	3	0.99721	0.93941	0.80696	0.58447	0.5756	0.72404	0.96651	0.97359	0.94669	0.99609	0.99962
	4	1	0.99996	0.99901	0.99965	0.99355	0.99012	0.99061	0.99816	0.78158	0.99945	1
	5	1	0.99491	0.94647	0.8497	0.71482	0.73902	0.93718	0.95539	0.87095	0.97908	0.98377
D	**6**	1	1	0.99994	0.9976	0.94615	0.72884	0.9433	0.99985	0.99955	0.99998	0.99999
I	**7**	1	0.99999	0.99834	0.98845	0.99593	0.99084	0.99072	1	1	0.99985	0.99999
M	**8**	1	1	0.99999	0.99693	0.99986	0.99977	0.99993	1	0.99998	0.99993	1
E	**9**	1	1	1	0.99983	0.99936	0.99996	0.99113	0.99997	0.99999	0.99997	1
N	**10**	1	1	1	0.99892	0.99999	0.99992	0.99983	0.99991	0.99998	0.99999	0.99998
S	**11**	1	1	0.99997	0.99765	0.99223	0.96701	0.99996	0.99998	0.9997	0.99998	0.99999
I	**12**	1	1	0.99996	0.99222	0.98172	0.96311	0.9677	0.99068	1	0.99994	0.99996
O	**13**	1	1	1	0.99969	0.99864	0.96909	0.99955	0.99911	0.99997	1	0.99999
N	**14**	1	1	1	0.99964	0.97759	0.87847	0.83012	0.77783	0.99244	0.99922	0.99952
	15	1	1	1	0.99999	0.99981	0.99985	0.99976	0.99996	0.99913	0.99999	1
	16	1	1	1	0.99998	0.99998	0.99978	0.97258	0.98541	0.98626	0.99234	0.99982
	17	1	1	1	0.99798	0.99564	0.99825	0.99848	0.99207	0.99771	0.99956	0.99821
	18	1	1	1	1	0.99994	0.99969	0.83922	0.90153	0.78207	0.99601	1
	19	1	1	1	0.99754	0.95042	0.84812	0.7324	0.9861	0.84718	0.99909	1
	20	1	1	1	1	0.97048	0.85896	0.65731	0.44347	0.28644	0.17695	0.002116

Table A.8: Maximum value of the gaps for each common subspace dimension and each SNR. When the common subspace dimension is greater than 20 the maximum value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.75737	0.5306	0.3538	0.24733	0.18803	0.14989	0.14446	0.12753	0.16714	0.14312	0.18252
	2	0.96791	0.8449	0.64286	0.44519	0.32268	0.24038	0.21215	0.26113	0.22839	0.2301	0.23041
	3	0.99322	0.93085	0.76872	0.55256	0.38318	0.28348	0.22903	0.36415	0.3108	0.3247	0.37332
	4	0.99997	0.98567	0.90278	0.7734	0.54036	0.41133	0.29573	0.27611	0.28094	0.38584	0.37829
	5	0.99999	0.98969	0.91396	0.73934	0.52332	0.37836	0.42975	0.28686	0.28907	0.36396	0.37341
D	**6**	0.99999	0.99745	0.95344	0.84428	0.60487	0.43087	0.3907	0.55558	0.36823	0.47865	0.48852
I	**7**	1	0.99899	0.96073	0.84041	0.65633	0.52905	0.54885	0.55102	0.57241	0.48708	0.35474
M	**8**	1	0.99996	0.98668	0.91816	0.83761	0.66248	0.78189	0.57683	0.64628	0.60256	0.75446
E	**9**	1	0.99999	0.9929	0.93497	0.82842	0.65869	0.63598	0.64812	0.91731	0.50171	0.64125
N	**10**	1	0.99999	0.99224	0.93527	0.81556	0.63078	0.57064	0.7593	0.69398	0.6487	0.59963
S	**11**	1	0.99999	0.99499	0.94842	0.81683	0.59246	0.57228	0.66323	0.86274	0.89235	0.78887
I	**12**	1	0.99999	0.9973	0.95294	0.82682	0.65783	0.53015	0.50106	0.71296	0.87322	0.70959
O	**13**	1	1	0.99954	0.96271	0.86087	0.68556	0.57755	0.48602	0.52831	0.60889	0.88807
N	**14**	1	1	0.99976	0.97172	0.86048	0.73435	0.57902	0.45575	0.53511	0.71065	0.68031
	15	1	1	0.99996	0.99288	0.92792	0.87585	0.72542	0.75357	0.69551	0.9001	0.57093
	16	1	1	0.99999	0.98735	0.91976	0.8145	0.65471	0.52585	0.48538	0.54767	0.61359
	17	1	1	0.99999	0.98785	0.91667	0.75148	0.57692	0.47768	0.46857	0.8054	0.47
	18	1	1	0.99999	0.99487	0.93028	0.77177	0.56876	0.43625	0.4066	0.40169	0.4689
	19	1	1	0.99999	0.99494	0.93865	0.7879	0.58157	0.39745	0.33844	0.28778	0.83701
	20	1	1	1	0.99955	0.96714	0.84864	0.64475	0.4376	0.28034	0.17374	0.0019691
	21	1	1	0.99999	0.99996	0.99998	0.99998	0.99996	0.99998	0.99998	0.99997	0.99998
	22	1	1	1	0.99999	0.99999	0.99999	0.99999	0.99999	0.99998	0.99999	0.99999
	23	1	1	1	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
	24	1	1	1	0.99999	1	1	0.99999	0.99999	0.99999	1	0.99999
	25	1	1	1	0.99999	1	1	0.99999	1	1	0.99999	1
	26	1	1	1	1	1	1	1	1	1	1	0.99999
	27	1	1	1	1	1	1	0.99999	1	1	1	1

Table A.9: Minimum value of the gaps for each common subspace dimension and each SNR. When the common subspace dimension is greater than 27 the minimum value of gaps is equal to 1.

A.2.3 Case 3: Corrupted common orthonormal matrix A with each subject having correlated temporal components S_k.

1. SNR=20 dB, SNR=15 dB,

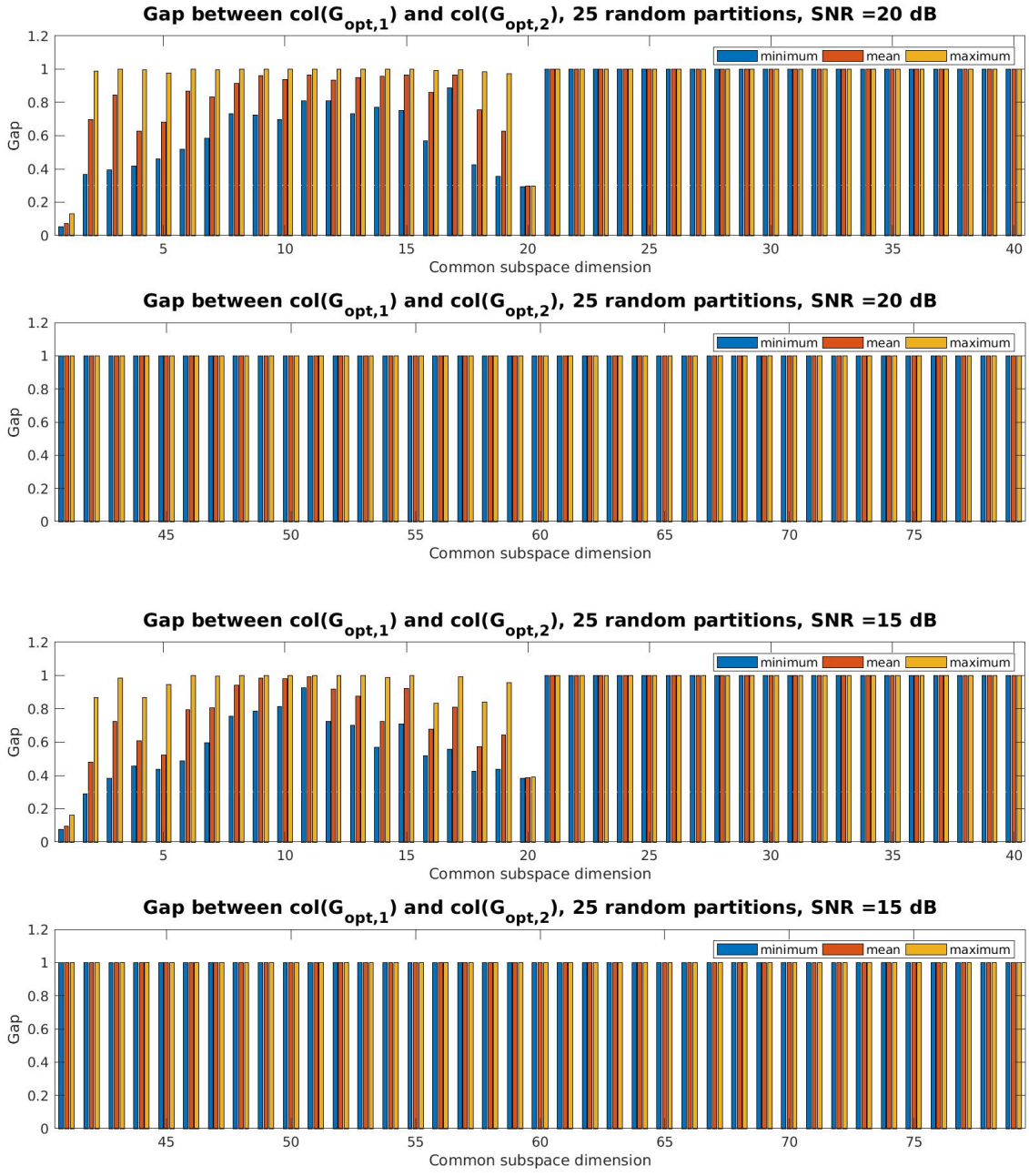


Figure A.28: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First, Second image**: SNR=20 dB. **Third, Fourth image**: SNR=15 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

2. SNR=10 dB, SNR=5 dB

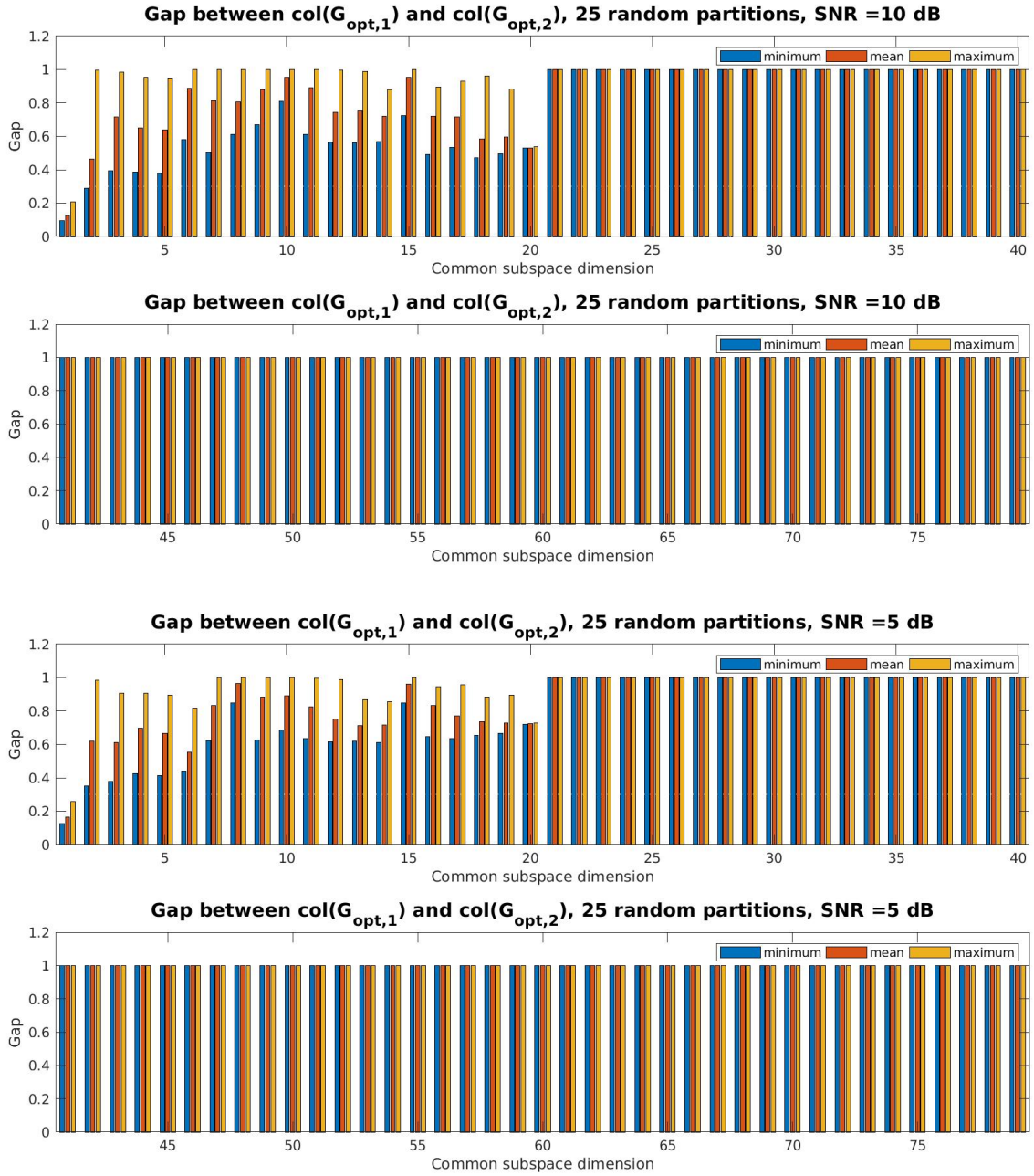


Figure A.29: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First, Second image**: SNR=10 dB. **Third, Fourth image**: SNR=5 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

3. SNR=−5 dB, SNR=−10 dB

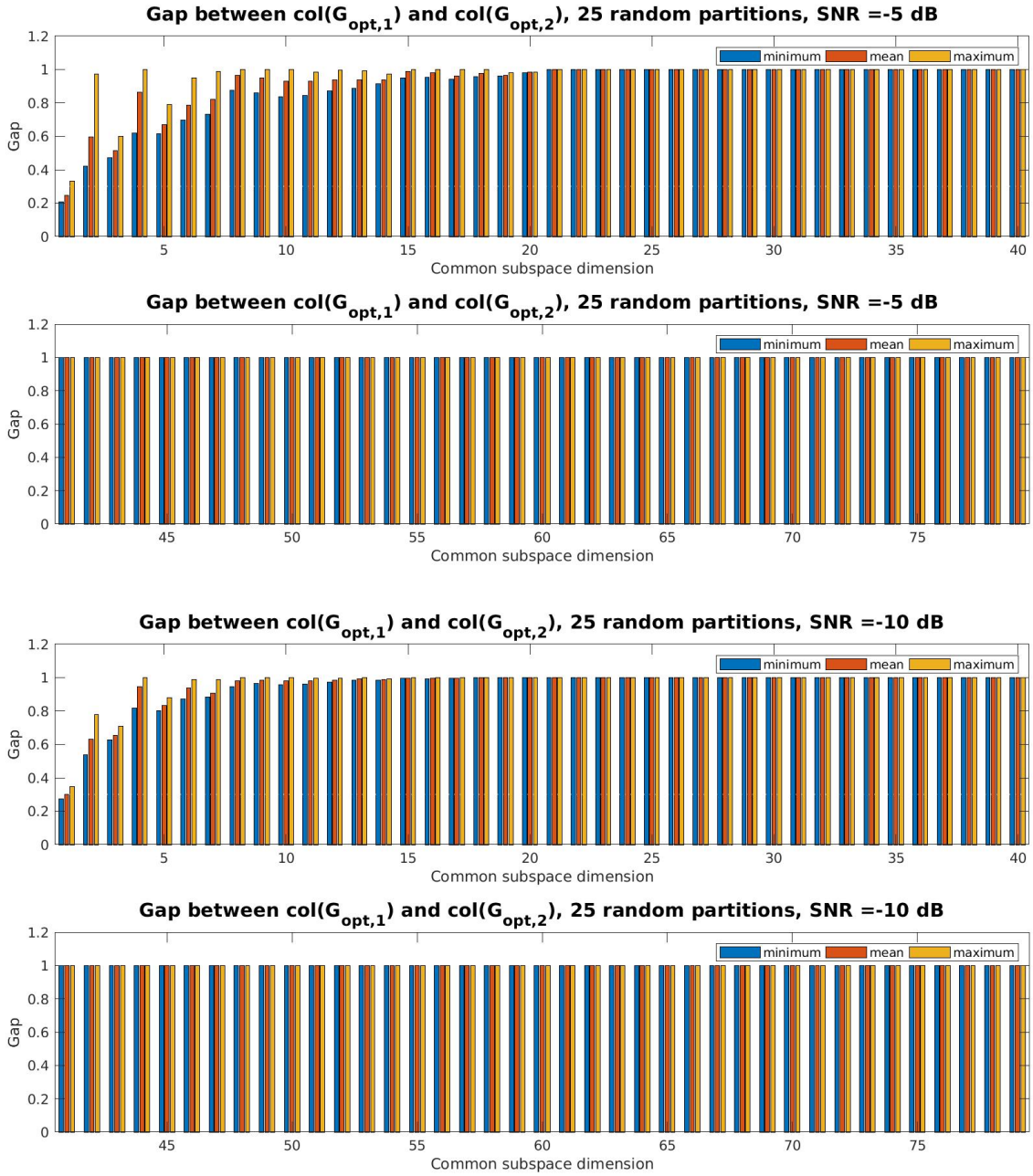


Figure A.30: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First, Second image**: SNR=−5 dB. **Third, Fourth image**: SNR=−10 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

4. SNR=−15 dB, SNR=−20 dB

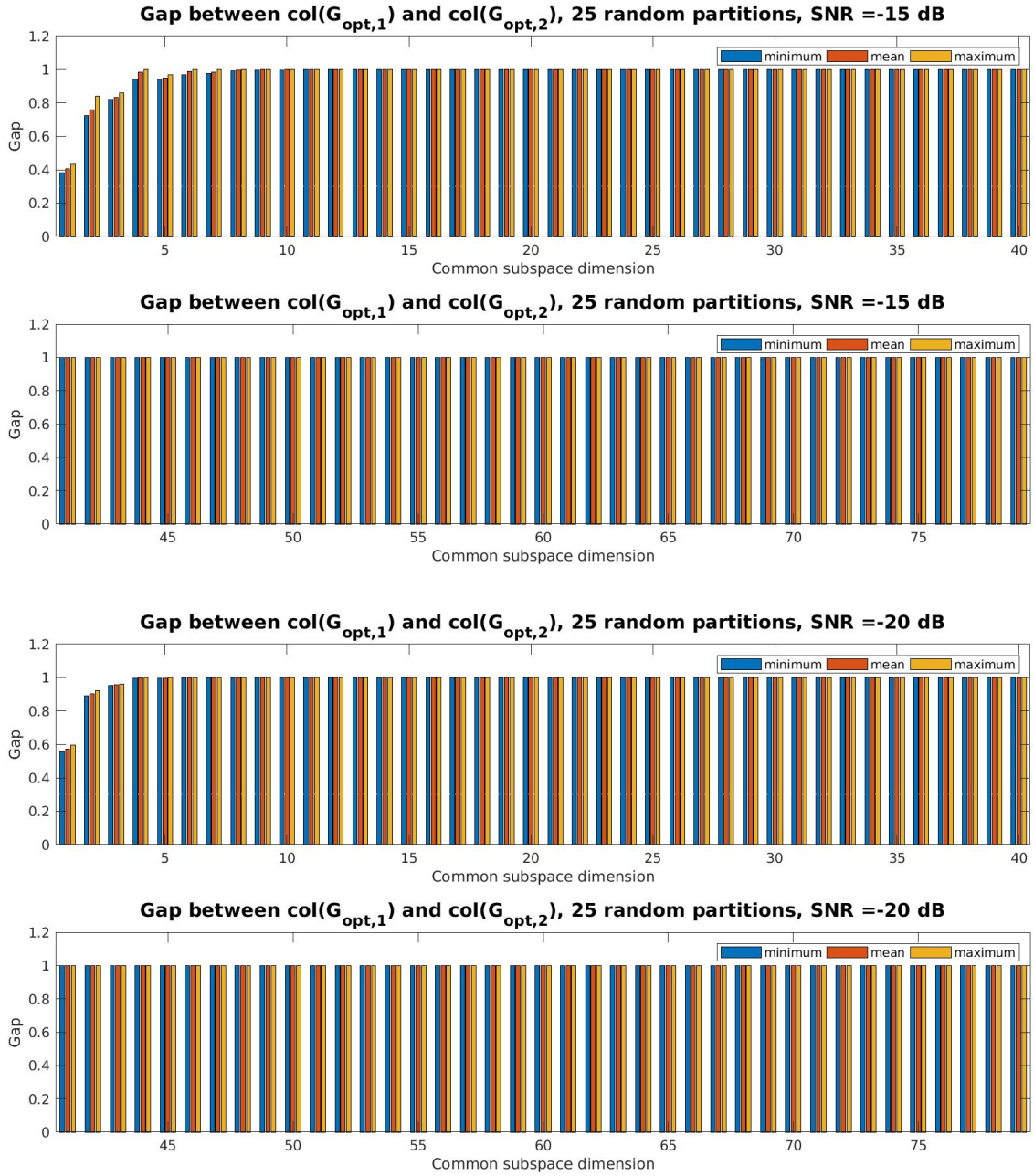


Figure A.31: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. **First, Second image**: SNR=−15 dB. **Third, Fourth image**: SNR=−20 dB. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

5. SNR=−25 dB

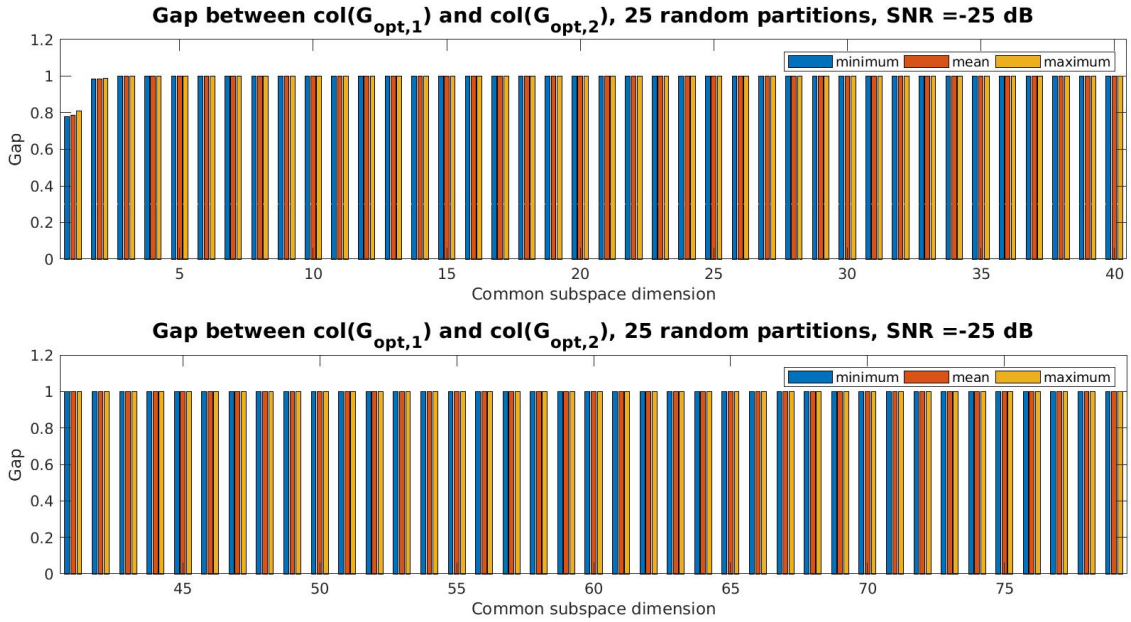


Figure A.32: We consider 25 random partitions of data. For each and every of the partitions we estimate the $\mathbf{G}_{opt,1}$ and $\mathbf{G}_{opt,2}$ and compute the gap for each common subspace dimension, for SNR=−25 dB. We illustrate the minimum, mean and maximum value of the gap for each common subspace dimension. Blue color: minimum value, Red color: mean value, Yellow color: maximum value.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.78744	0.57209	0.40434	0.30034	0.2461	0.20629	0.16605	0.12658	0.093297	0.073296	0.00069647
	2	0.98541	0.90263	0.75877	0.62965	0.59615	0.72371	0.62091	0.4636	0.47989	0.69718	0.9979
	3	0.99924	0.9583	0.83379	0.65415	0.51444	0.48126	0.61235	0.71498	0.72506	0.84301	0.99826
	4	0.99999	0.99828	0.98461	0.94569	0.86369	0.72336	0.69571	0.64953	0.60704	0.62859	0.99855
	5	1	0.99664	0.95114	0.83192	0.66954	0.58827	0.66715	0.63761	0.52277	0.68268	0.99895
D	**6**	1	0.99997	0.98653	0.93645	0.78735	0.64029	0.55432	0.88749	0.79343	0.86831	0.99897
I	**7**	1	0.99999	0.98257	0.90764	0.82144	0.86822	0.83132	0.81432	0.8068	0.83264	0.99798
M	**8**	1	1	0.998	0.97986	0.96464	0.95326	0.96663	0.80706	0.94163	0.91542	0.99862
E	**9**	1	1	0.99948	0.98545	0.94957	0.87991	0.88539	0.87972	0.98442	0.96	0.99847
N	**10**	1	1	0.99907	0.98123	0.93095	0.86918	0.8897	0.9553	0.97941	0.93745	0.99828
S	**11**	1	1	0.99955	0.98106	0.93134	0.86177	0.82627	0.89288	0.99206	0.9669	0.9929
I	**12**	1	1	0.99997	0.98574	0.93918	0.86711	0.75242	0.74306	0.92022	0.93383	0.99559
O	**13**	1	1	0.99999	0.99083	0.93704	0.84683	0.71296	0.74985	0.87391	0.95071	0.9967
N	**14**	1	1	1	0.98923	0.93863	0.84452	0.71779	0.7193	0.72403	0.95824	0.99725
	15	1	1	1	0.99803	0.98777	0.96349	0.95978	0.9525	0.92082	0.9654	0.99553
	16	1	1	1	0.99739	0.98215	0.93445	0.83441	0.71899	0.6772	0.85905	0.99562
	17	1	1	1	0.99729	0.95945	0.88296	0.77273	0.71734	0.80933	0.96462	0.99225
	18	1	1	1	0.99953	0.97563	0.8828	0.73411	0.58506	0.57358	0.7556	0.98238
	19	1	1	1	0.99982	0.96686	0.8666	0.72627	0.59774	0.64316	0.629	0.96073
	20	1	1	1	0.99999	0.98318	0.89657	0.72343	0.53158	0.38541	0.29644	0.2309
	21	1	1	1	1	1	1	1	1	0.99999	0.99999	1

Table A.10: Mean value of the gaps for each common subspace dimension and each SNR. When the common subspace dimension is greater than 21 the mean value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.80864	0.59519	0.43115	0.34915	0.33362	0.3104	0.25721	0.20936	0.16243	0.12819	0.00071997
	2	0.98906	0.92225	0.84184	0.77898	0.97383	0.98023	0.9837	0.99637	0.86664	0.98774	1
	3	1	0.96245	0.85933	0.71008	0.59906	0.83954	0.90831	0.98488	0.98601	0.99934	0.99998
	4	1	0.99999	0.99999	0.99992	0.99968	0.9792	0.90849	0.952	0.86961	0.99769	1
	5	1	0.99986	0.96791	0.87891	0.78952	0.81432	0.8954	0.95013	0.94725	0.9769	0.99999
D	**6**	1	1	0.99868	0.98707	0.94821	0.903	0.81922	1	0.99952	0.9997	0.99999
I	**7**	1	1	0.999	0.98907	0.98641	0.99992	0.99989	0.99996	0.99651	0.99546	1
M	**8**	1	1	0.99999	0.99958	1	0.99994	0.99986	0.99913	0.99999	1	1
E	**9**	1	1	1	0.99999	1	0.98871	0.99926	0.99944	1	1	1
N	**10**	1	1	1	1	0.99943	0.99524	0.99999	0.99951	1	0.99972	1
S	**11**	1	1	0.99999	0.99745	0.98624	0.98398	0.99701	0.99983	0.99997	0.99998	0.99992
I	**12**	1	1	1	0.99686	0.99715	0.98367	0.98782	0.99423	0.99995	0.99976	0.99999
O	**13**	1	1	1	0.99999	0.99313	0.99178	0.86971	0.98983	0.99985	0.99982	0.99996
N	**14**	1	1	1	0.99324	0.9742	0.9197	0.85589	0.88033	0.98826	0.99971	1
	15	1	1	1	1	0.99955	0.99898	1	0.99984	0.99966	1	1
	16	1	1	1	0.99997	0.99998	0.99947	0.94524	0.89322	0.83292	0.99417	1
	17	1	1	1	0.99948	0.99994	1	0.95819	0.92971	0.99407	0.99805	1
	18	1	1	1	1	0.99997	0.96712	0.8816	0.96282	0.84053	0.98506	1
	19	1	1	1	1	0.97936	0.89369	0.89371	0.88264	0.9555	0.972	1
	20	1	1	1	1	0.9854	0.90065	0.72723	0.53657	0.38974	0.29834	0.23304

Table A.11: Maximum value of the gaps for each common subspace dimension and each SNR. When the common subspace dimension is greater than 20 the maximum value of gaps is equal to 1.

		SNR (dB)										
		-25	**-20**	**-15**	**-10**	**-5**	**0**	**5**	**10**	**15**	**20**	**noiseless**
	1	0.77769	0.55713	0.38147	0.2751	0.20665	0.16842	0.12606	0.096409	0.074671	0.051413	0.00067436
	2	0.98261	0.89127	0.72246	0.53704	0.42244	0.41453	0.35143	0.29096	0.29074	0.36576	0.98436
	3	0.99832	0.95531	0.82199	0.62681	0.47054	0.36666	0.37793	0.39406	0.38085	0.39253	0.98777
	4	0.99998	0.9942	0.94003	0.81743	0.6206	0.50673	0.42662	0.38798	0.45644	0.41766	0.99138
	5	0.99999	0.99548	0.94249	0.80019	0.61675	0.46913	0.41511	0.37985	0.43726	0.45856	0.99562
D	**6**	0.99999	0.99982	0.96865	0.8708	0.69899	0.53995	0.43985	0.5816	0.48888	0.51984	0.99619
I	**7**	0.99999	0.99997	0.97737	0.88219	0.73354	0.60909	0.62349	0.50474	0.59474	0.58446	0.98522
M	**8**	1	0.99999	0.99415	0.94518	0.87741	0.74644	0.84917	0.61147	0.75689	0.73276	0.98928
E	**9**	1	0.99999	0.99767	0.96316	0.86061	0.73911	0.62726	0.67162	0.78683	0.72429	0.98821
N	**10**	1	1	0.99711	0.95909	0.83501	0.71208	0.68546	0.80796	0.81348	0.69643	0.98883
S	**11**	1	1	0.99898	0.96256	0.8458	0.69872	0.63513	0.60975	0.92761	0.81157	0.96312
I	**12**	1	0.99999	0.99975	0.97188	0.87302	0.73812	0.61456	0.56353	0.72286	0.8083	0.97926
O	**13**	1	1	0.99995	0.98253	0.88889	0.74831	0.61932	0.56254	0.70047	0.73394	0.97566
N	**14**	1	1	0.99999	0.98587	0.91298	0.75602	0.61077	0.56809	0.57013	0.77	0.98747
	15	1	1	0.99999	0.99434	0.94791	0.87343	0.84896	0.72469	0.70711	0.7506	0.97712
	16	1	1	0.99999	0.9937	0.95312	0.83173	0.64655	0.48951	0.51857	0.56732	0.97132
	17	1	1	1	0.99532	0.94177	0.80838	0.63377	0.53273	0.55616	0.88571	0.93916
	18	1	1	0.99999	0.99888	0.95871	0.83809	0.65477	0.4703	0.42629	0.42658	0.80221
	19	1	1	0.99999	0.99944	0.96305	0.84801	0.66536	0.49538	0.43851	0.35472	0.68341
	20	1	1	1	0.99996	0.98176	0.89374	0.72021	0.52876	0.38278	0.29479	0.22879
	21	1	1	1	1	0.99999	0.99997	0.99999	0.99999	0.99996	0.99995	0.99997
	22	1	1	1	1	0.99999	0.99999	1	0.99999	0.99999	0.99999	1
	23	1	1	1	0.99999	0.99999	0.99999	1	0.99999	0.99999	0.99999	0.99999
	24	1	1	1	1	0.99999	1	1	0.99999	0.99999	1	1
	25	1	1	1	1	1	1	1	0.99999	1	1	0.99999
	26	1	1	1	0.99999	1	1	0.99999	1	1	1	1
	27	1	1	1	1	1	1	1	1	1	1	1
	28	1	1	1	1	1	1	1	0.99999	1	1	0.99999
	29	1	1	1	1	1	1	1	0.99999	1	1	1

Table A.12: Minimum value of the gaps for each common subspace dimension and each SNR. When the common subspace dimension is greater than 29 the minimum value of gaps is equal to 1.

A.3 Figures for experimental evaluation of the estimation of the common spatial factor A

A.3.1 Case 1: Common orthonormal matrix A with each subject having independent temporal components S_k.

Penalty method

· **Comparison of Penalty method solutions**

· **Noiseless case**

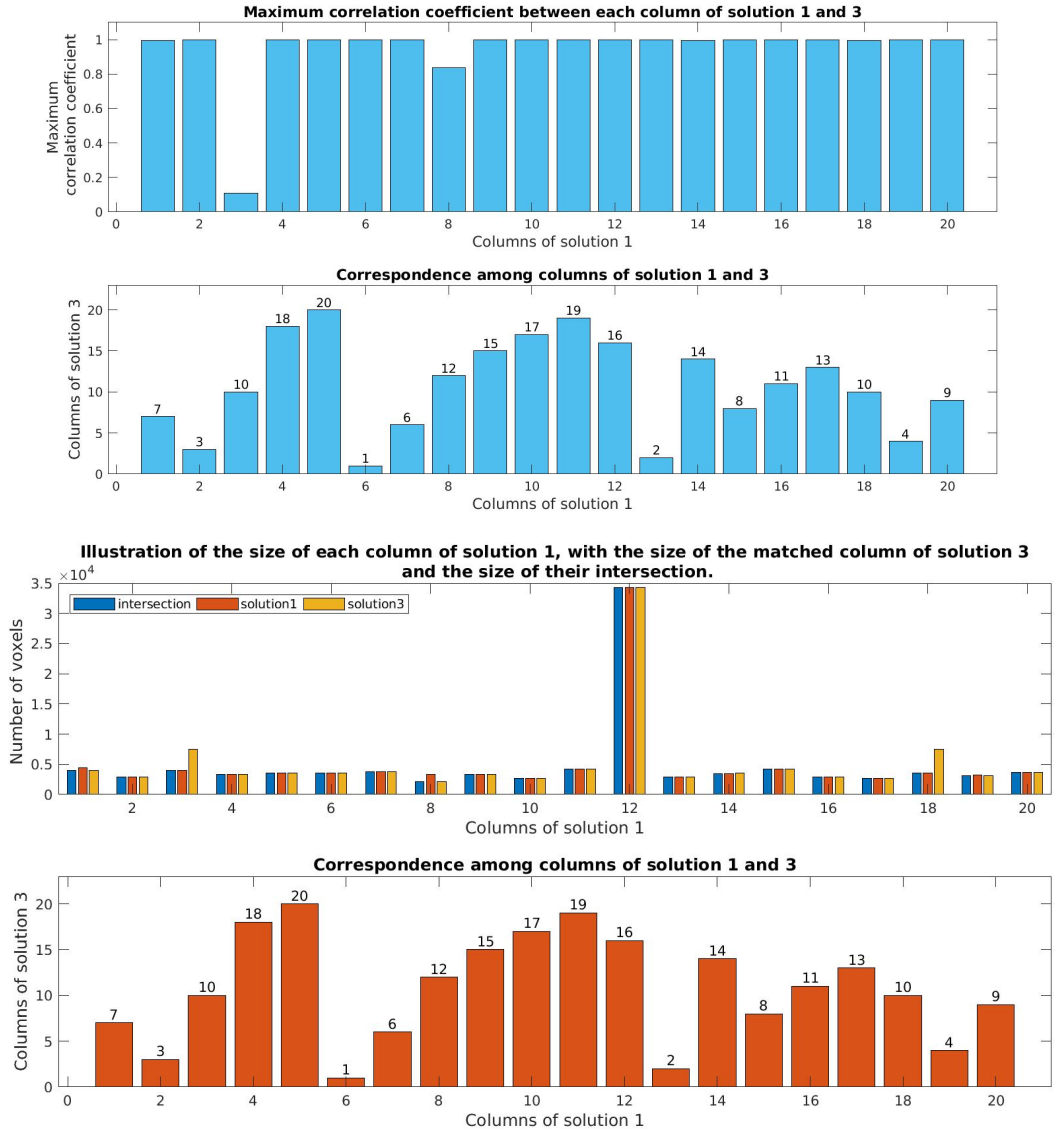


Figure A.33: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on intersection. Solutions 1, 3 correspond to penalty method and noiseless data.

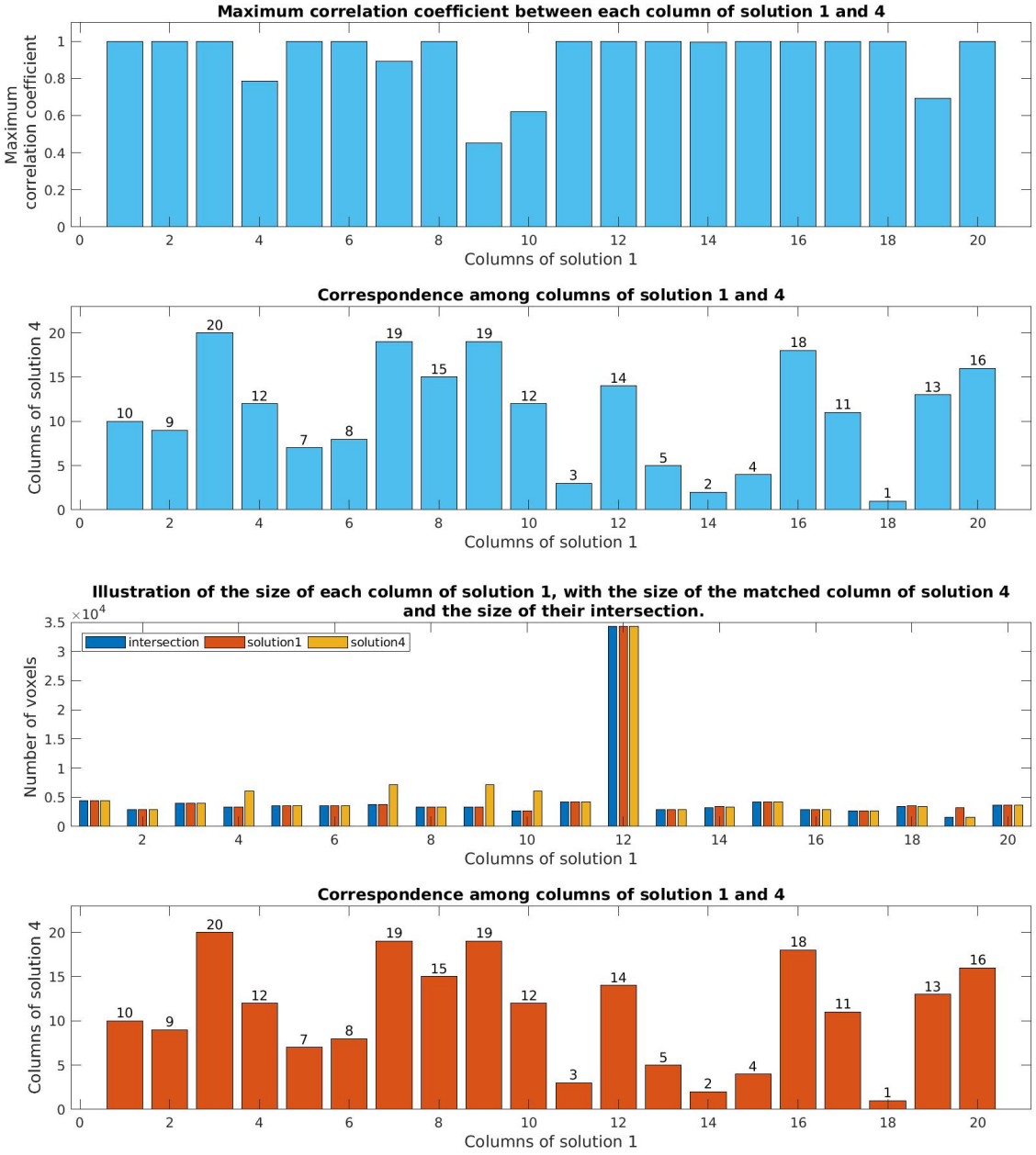


Figure A.34: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on intersection. Solutions 1, 4 correspond to penalty method and noiseless data.

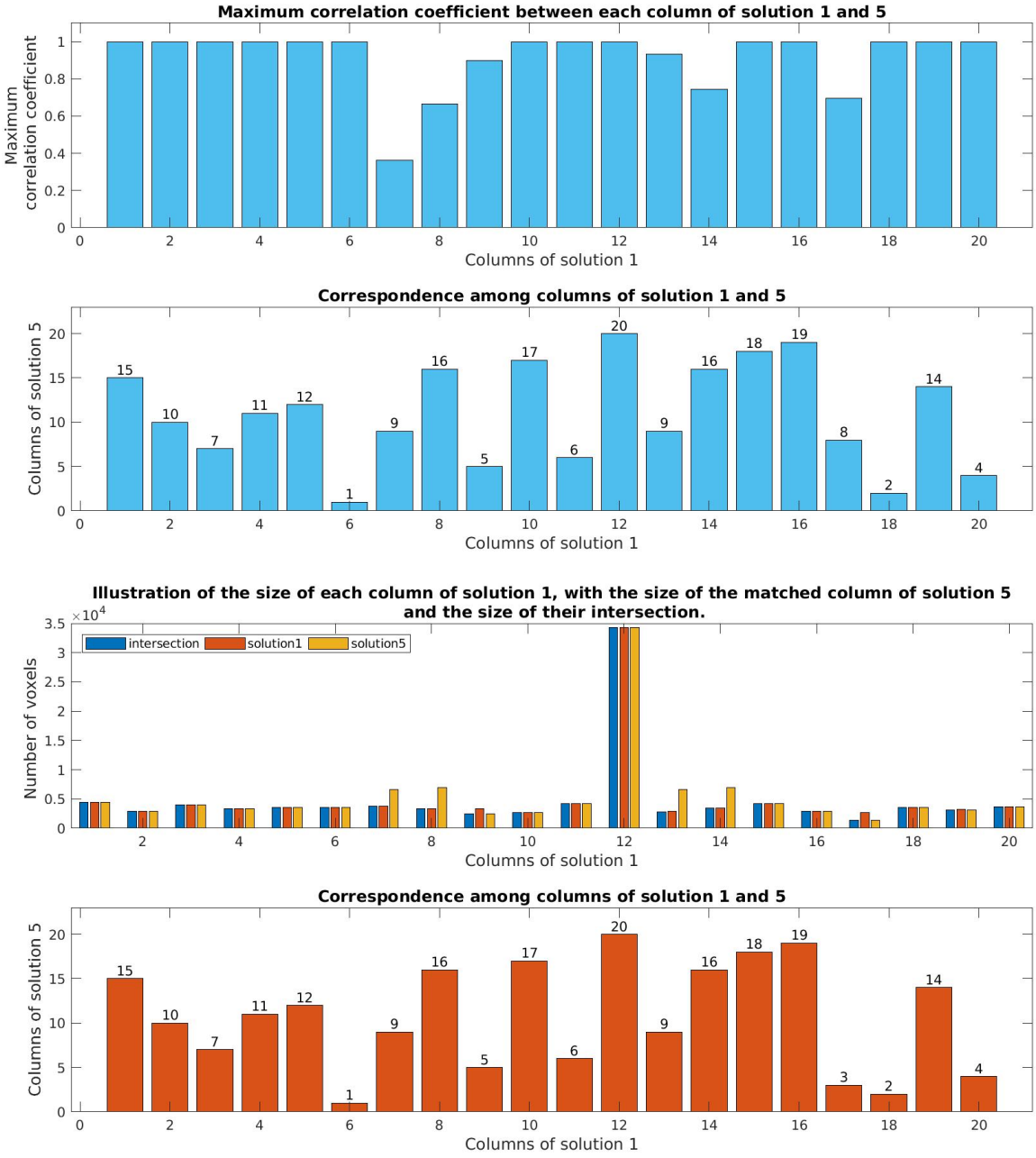


Figure A.35: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 5. We associate the column r of the solution 1 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on intersection. Solutions 1, 5 correspond to penalty method and noiseless data.

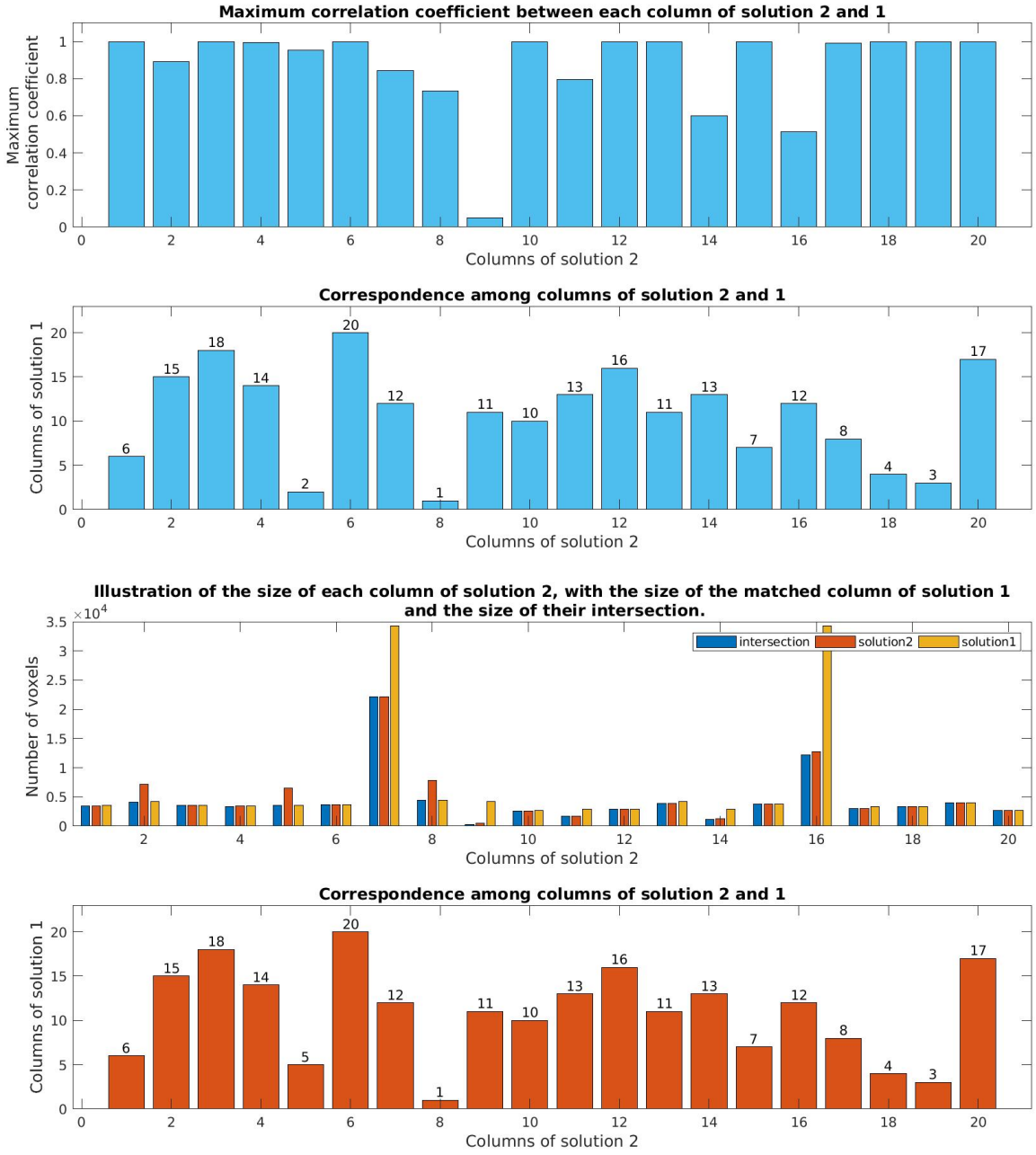


Figure A.36: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on intersection. Solutions 2, 1 correspond to penalty method and noiseless data.

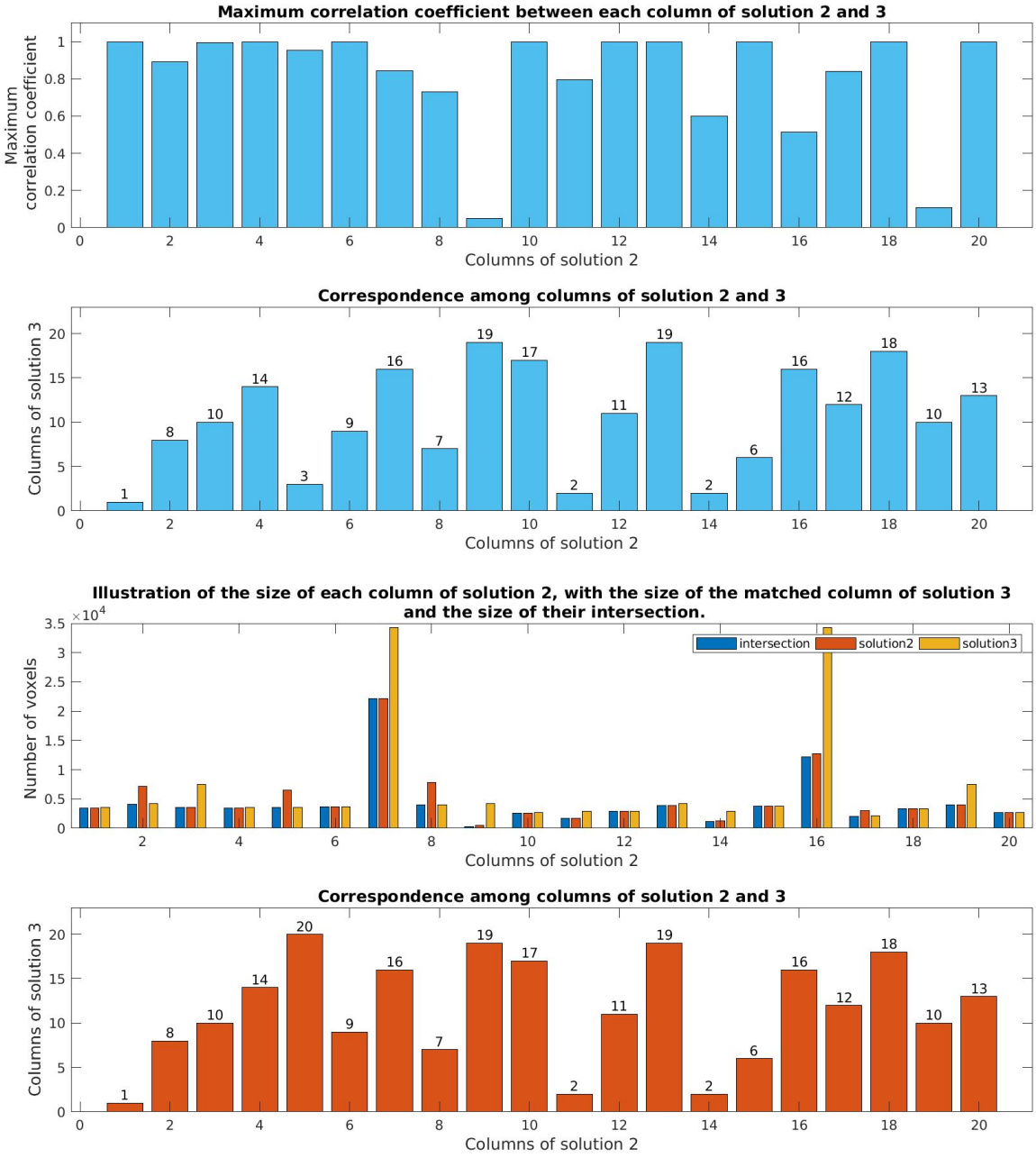


Figure A.37: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on intersection. Solutions 2, 3 correspond to penalty method and noiseless data.

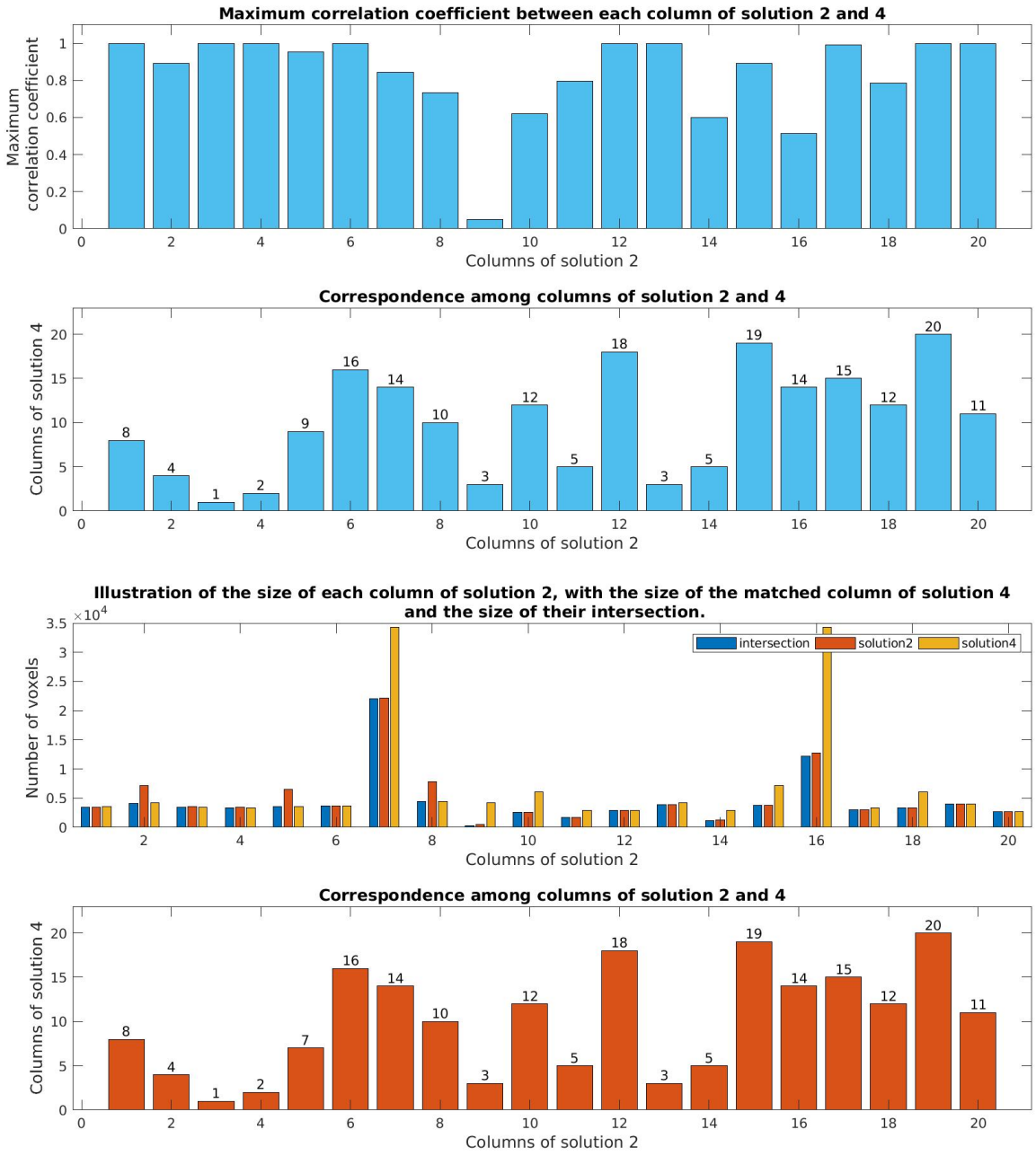


Figure A.38: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on intersection. Solutions 2, 4 correspond to penalty method and noiseless data.

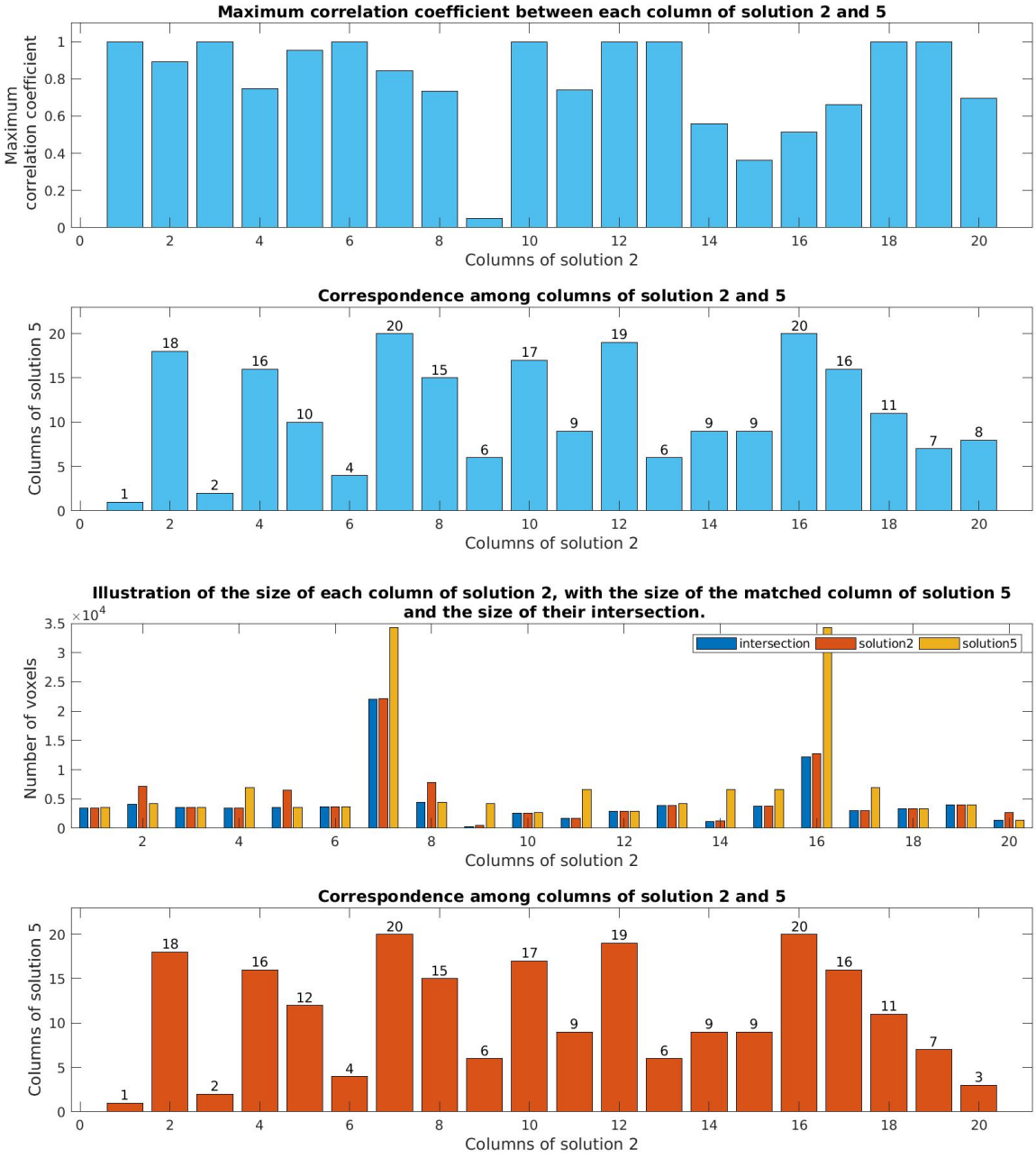


Figure A.39: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 5. We associate the column r of the solution 2 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on intersection. Solutions 2, 5 correspond to penalty method and noiseless data.

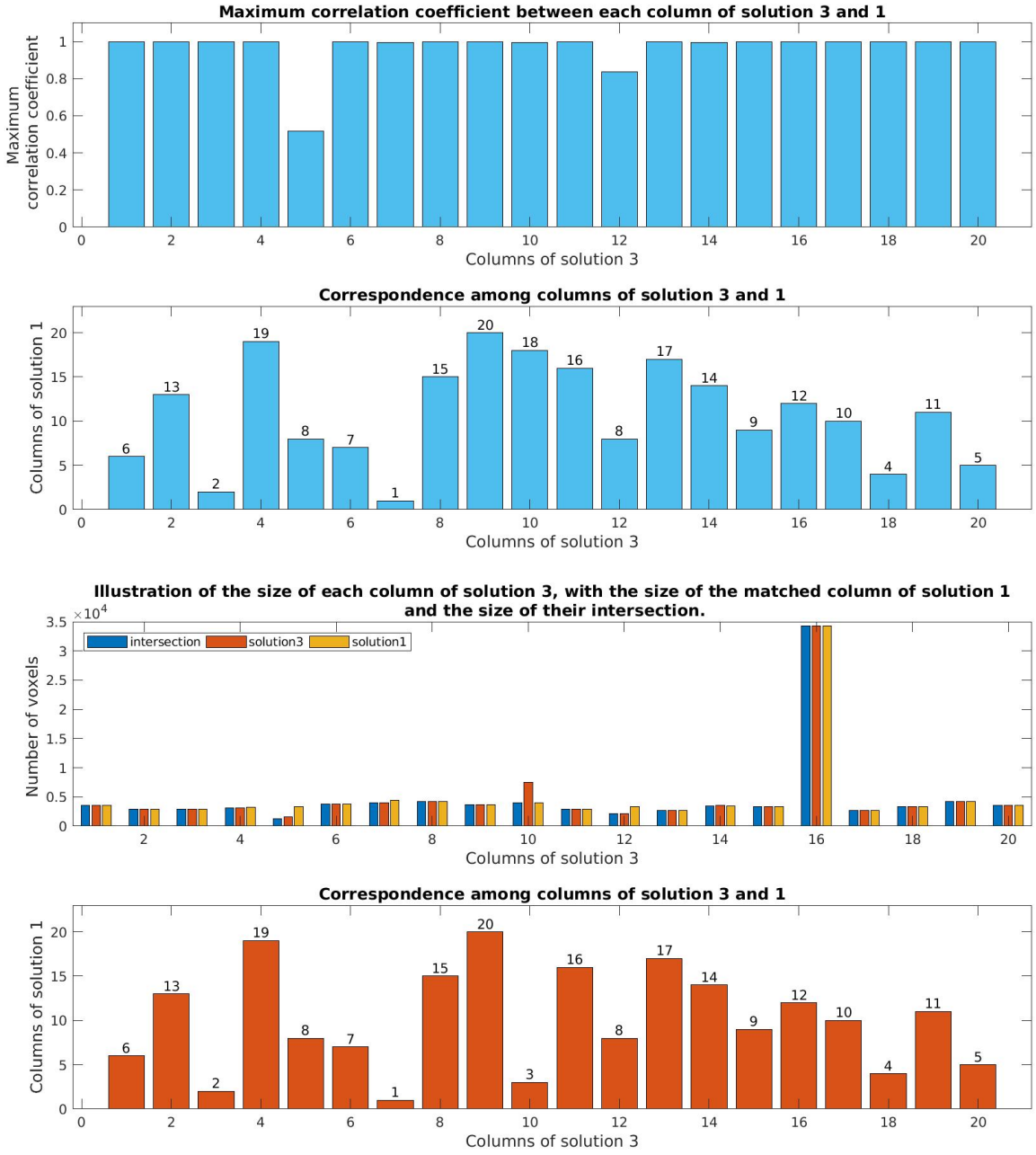


Figure A.40: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on intersection. Solutions 3, 1 correspond to penalty method and noiseless data.

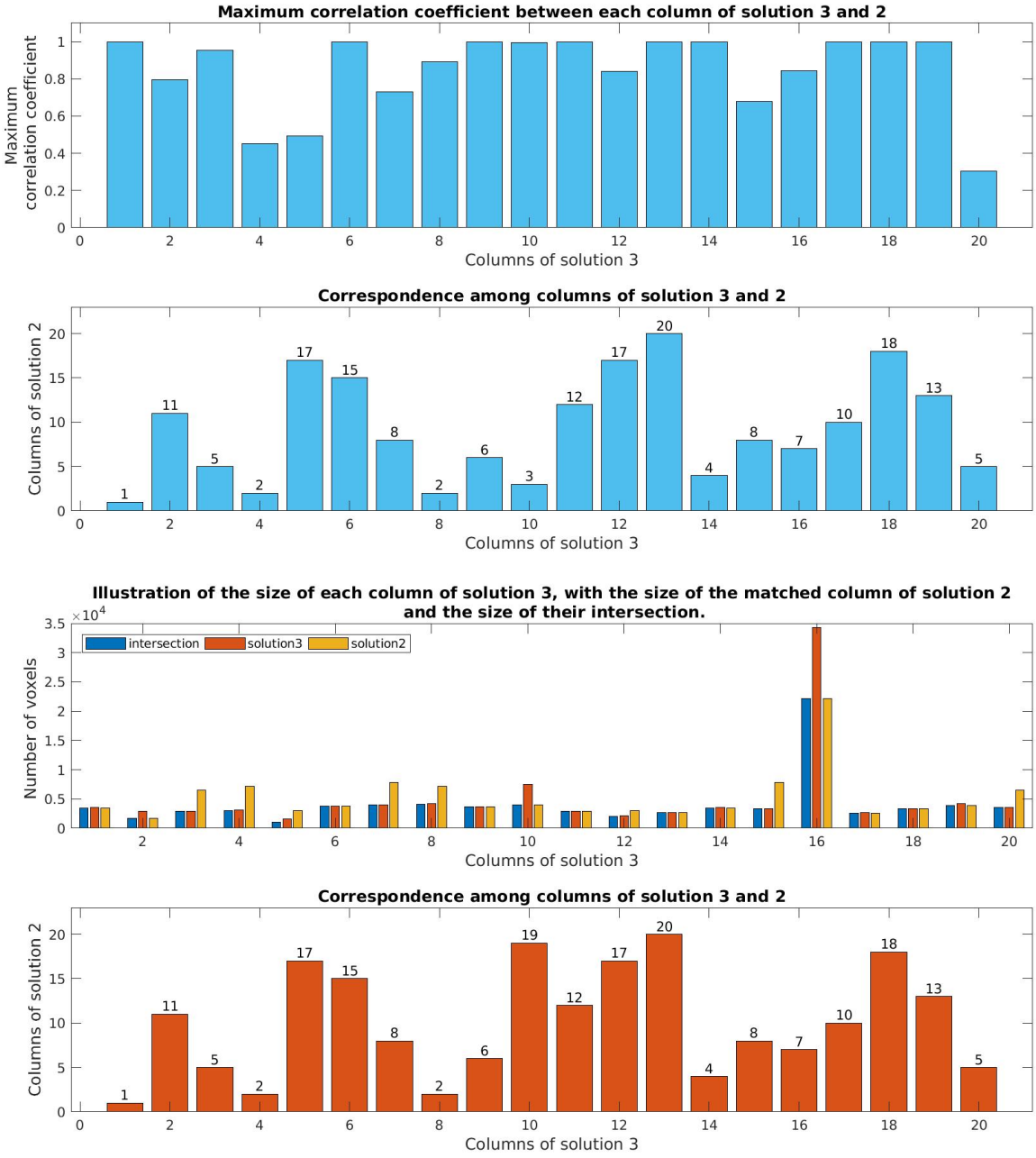


Figure A.41: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on intersection. Solutions 3, 2 correspond to penalty method and noiseless data.

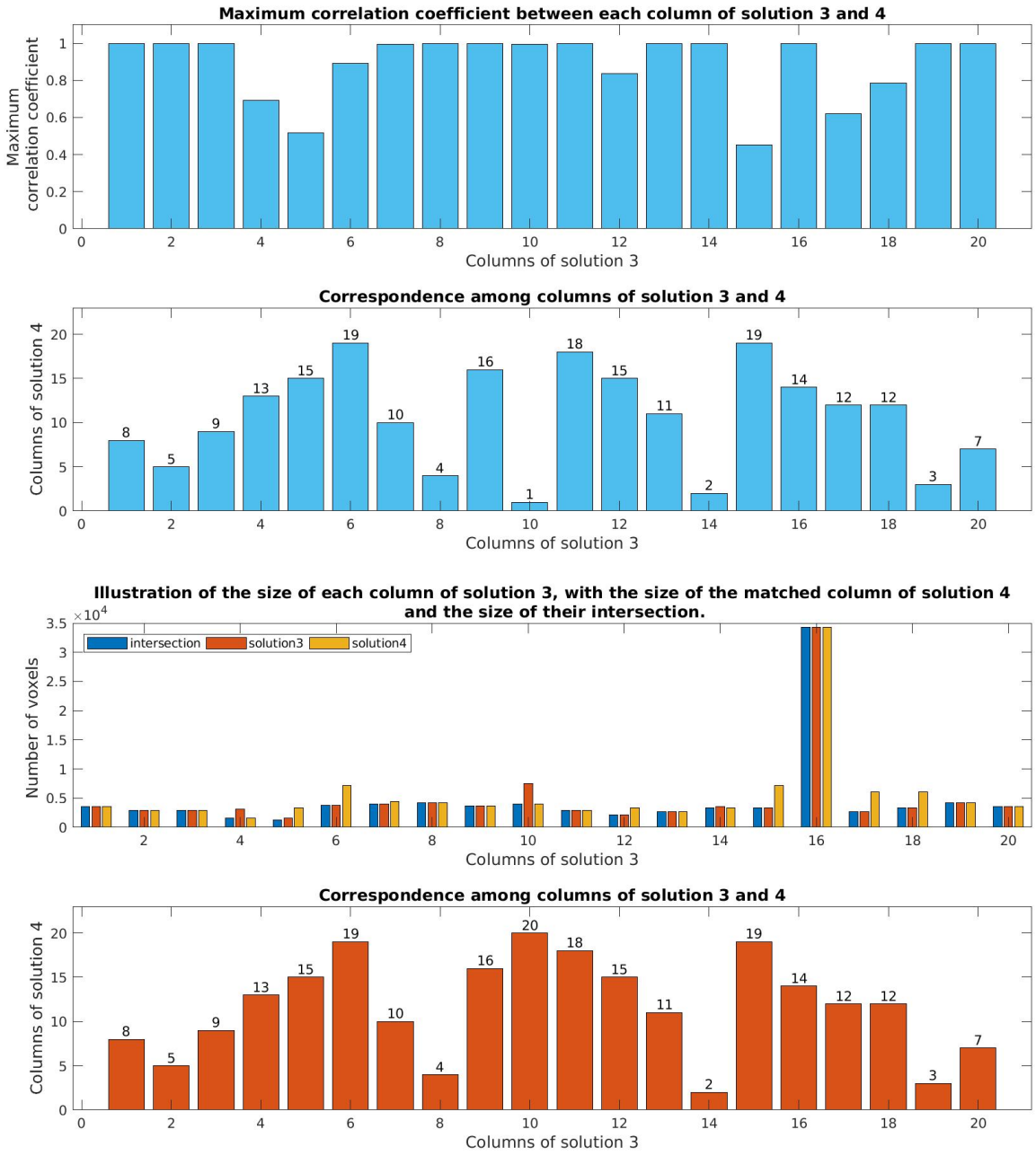


Figure A.42: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on intersection. Solutions 3, 4 correspond to penalty method and noiseless data.

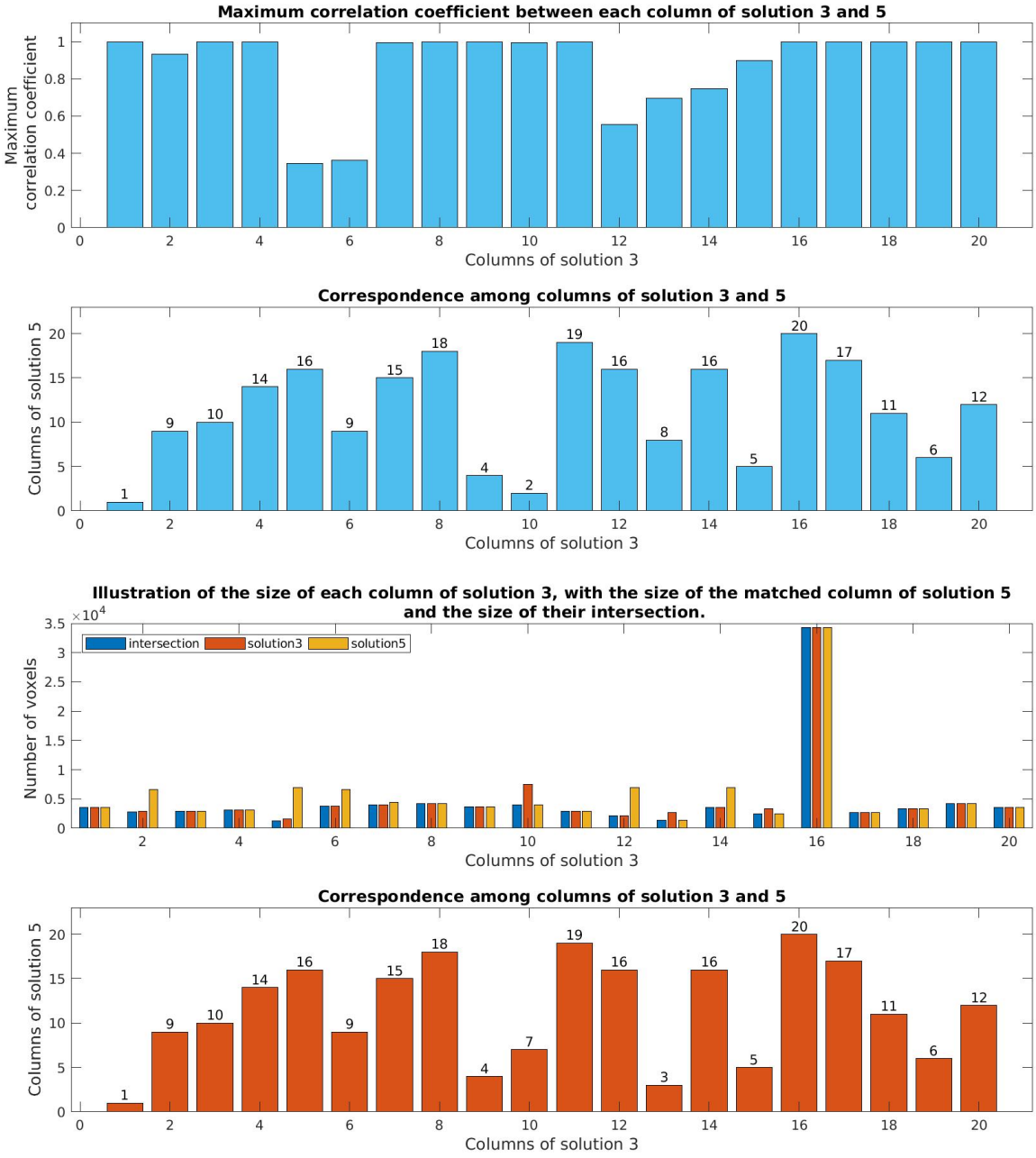


Figure A.43: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 5. We associate the column r of the solution 3 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on intersection. Solutions 3, 5 correspond to penalty method and noiseless data.

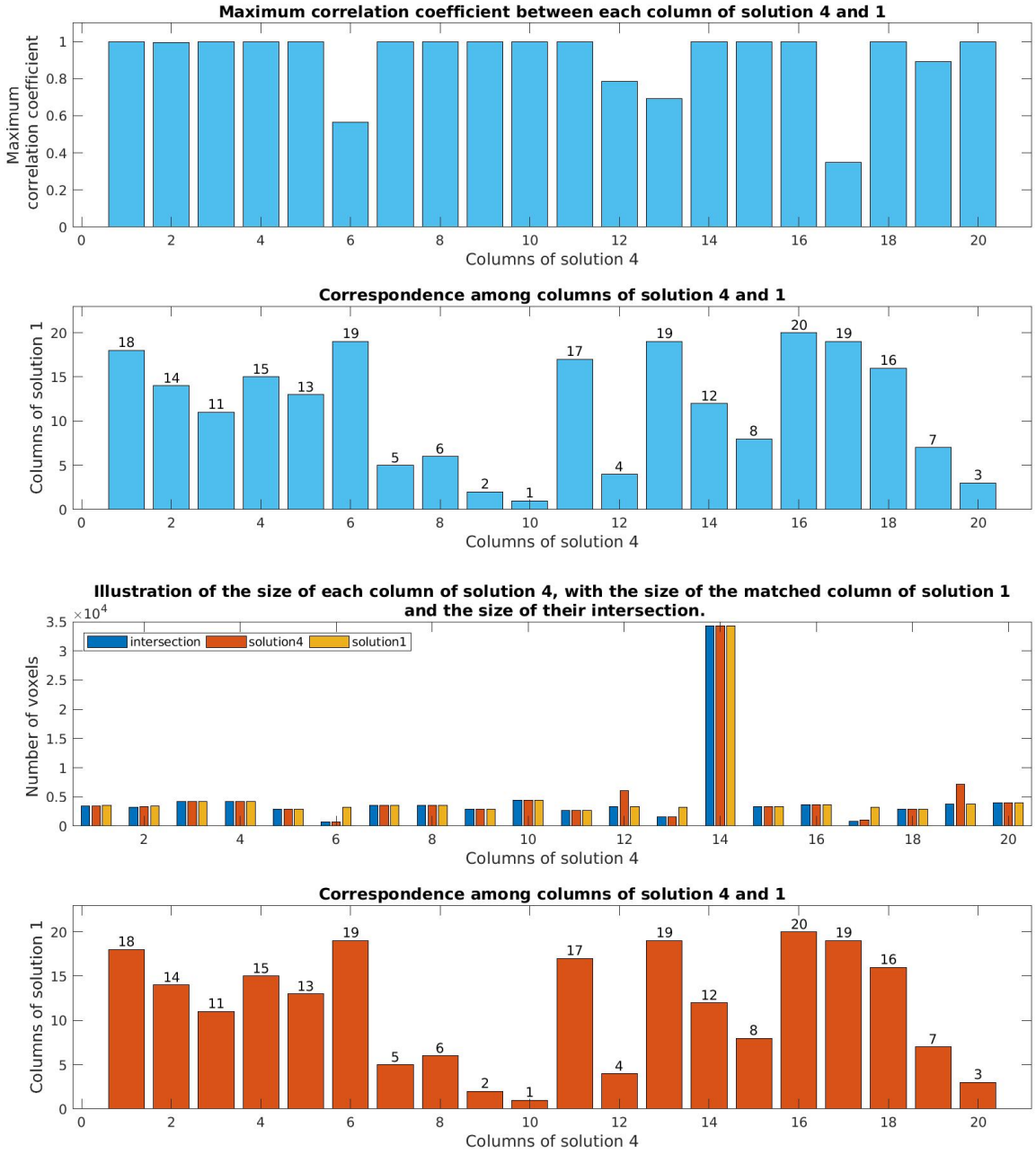


Figure A.44: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on intersection. Solutions 4, 1 correspond to penalty method and noiseless data.

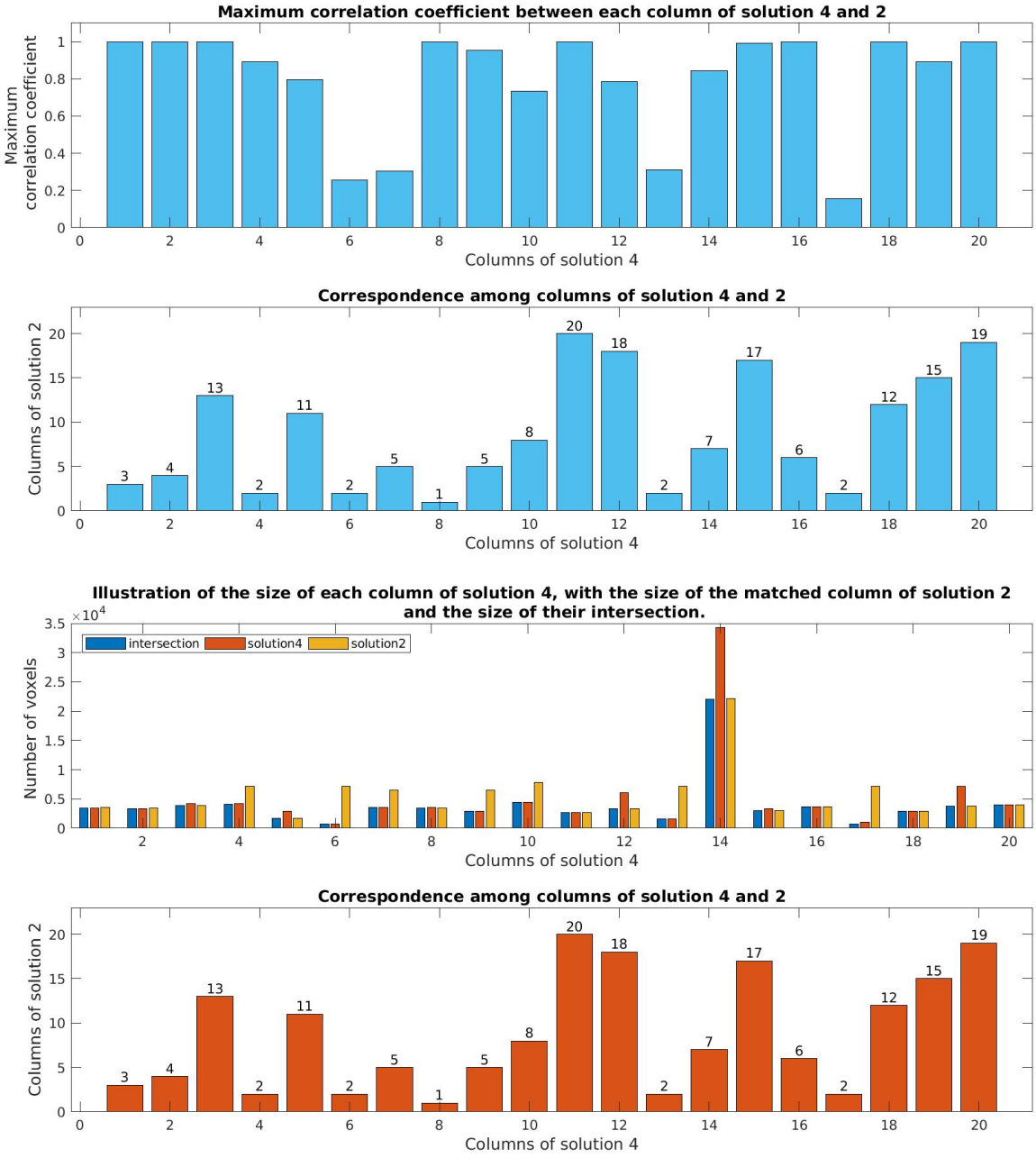


Figure A.45: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on intersection. Solutions 4, 2 correspond to penalty method and noiseless data.

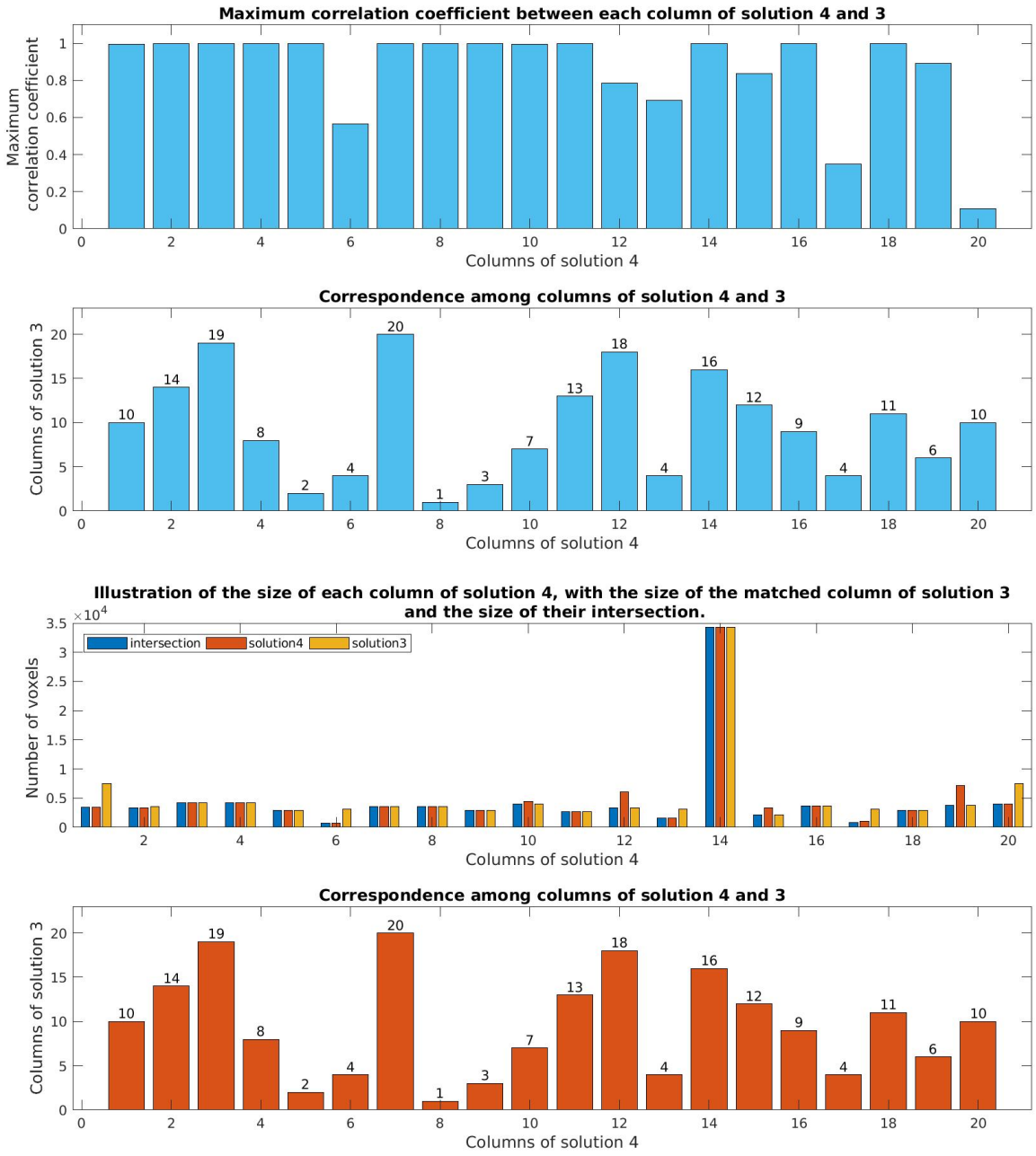


Figure A.46: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on intersection. Solutions 4, 3 correspond to penalty method and noiseless data.

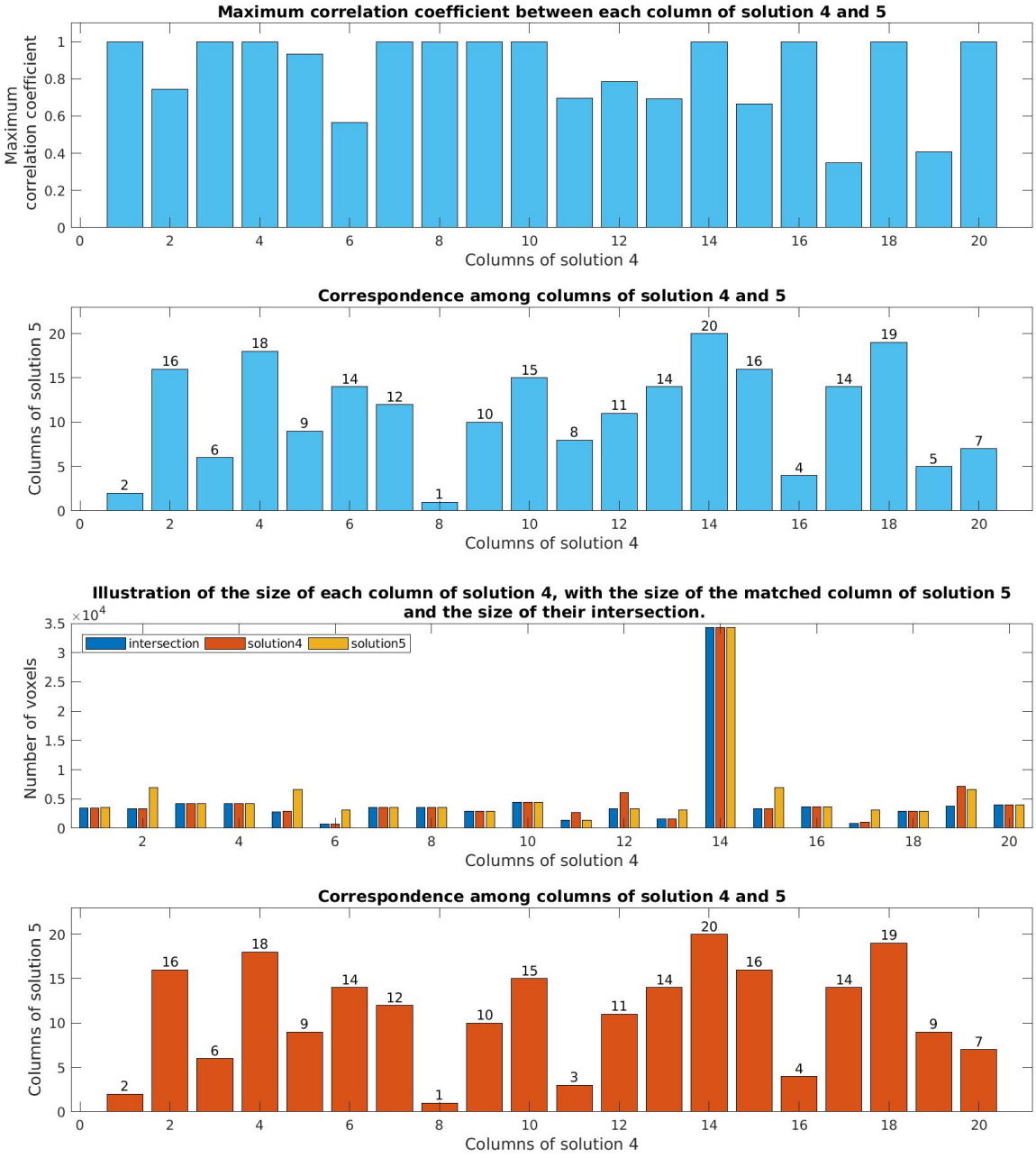


Figure A.47: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 5. We associate the column r of the solution 4 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on intersection. Solutions 4, 5 correspond to penalty method and noiseless data.

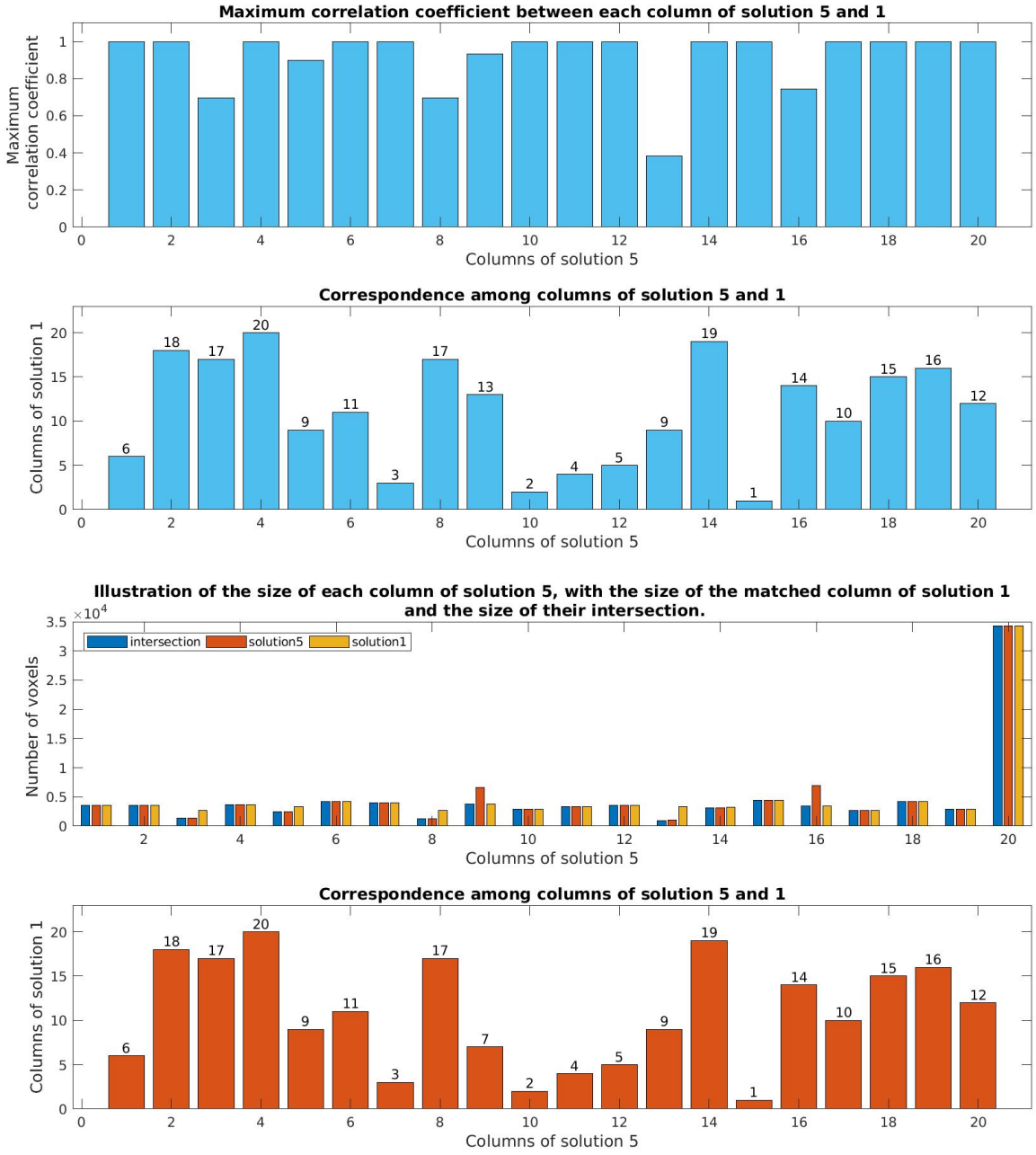


Figure A.48: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 1. We associate the column r of the solution 5 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on intersection. Solutions 5, 1 correspond to penalty method and noiseless data.

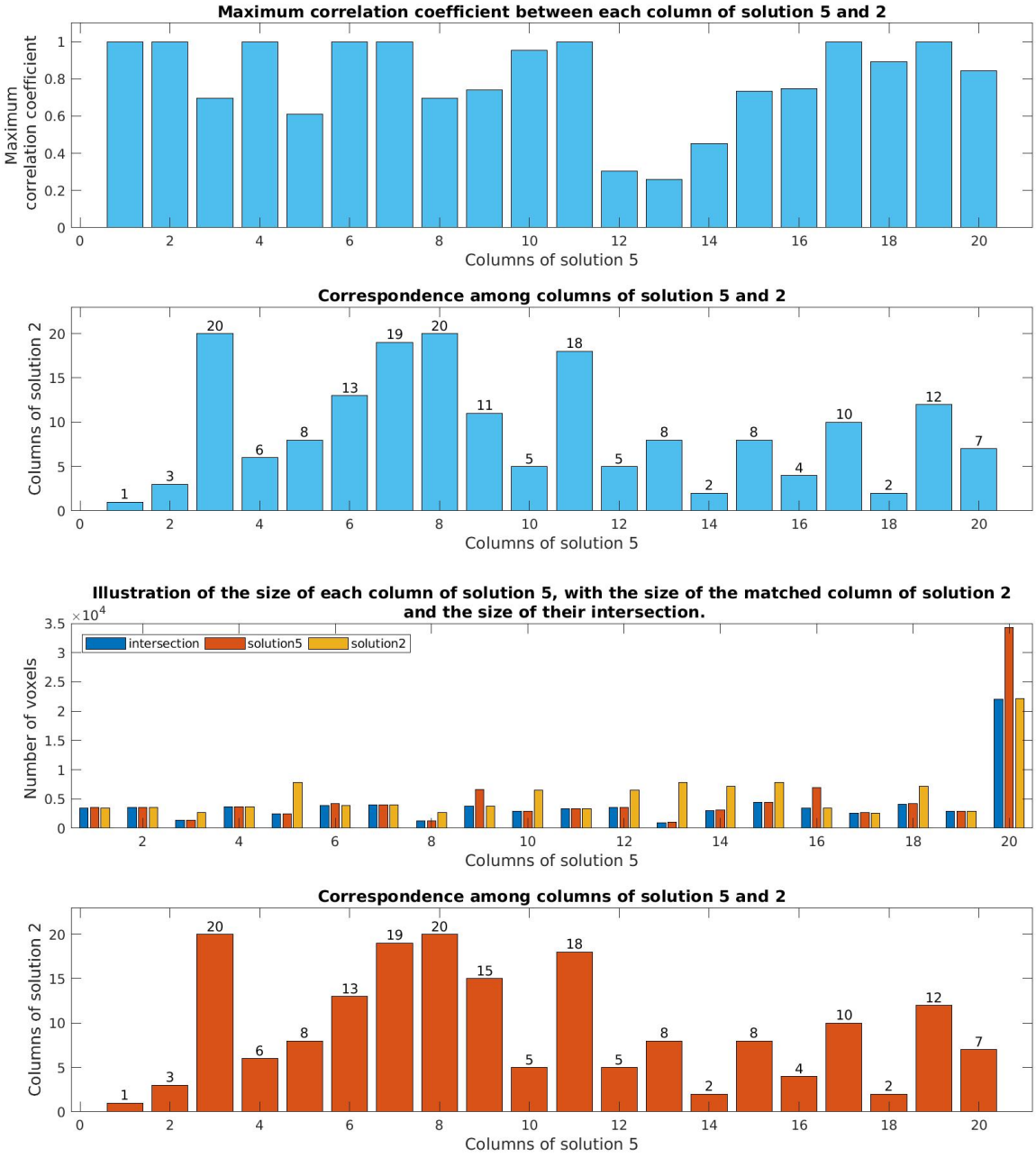


Figure A.49: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 2. We associate the column r of the solution 5 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on intersection. Solutions 5, 2 correspond to penalty method and noiseless data.

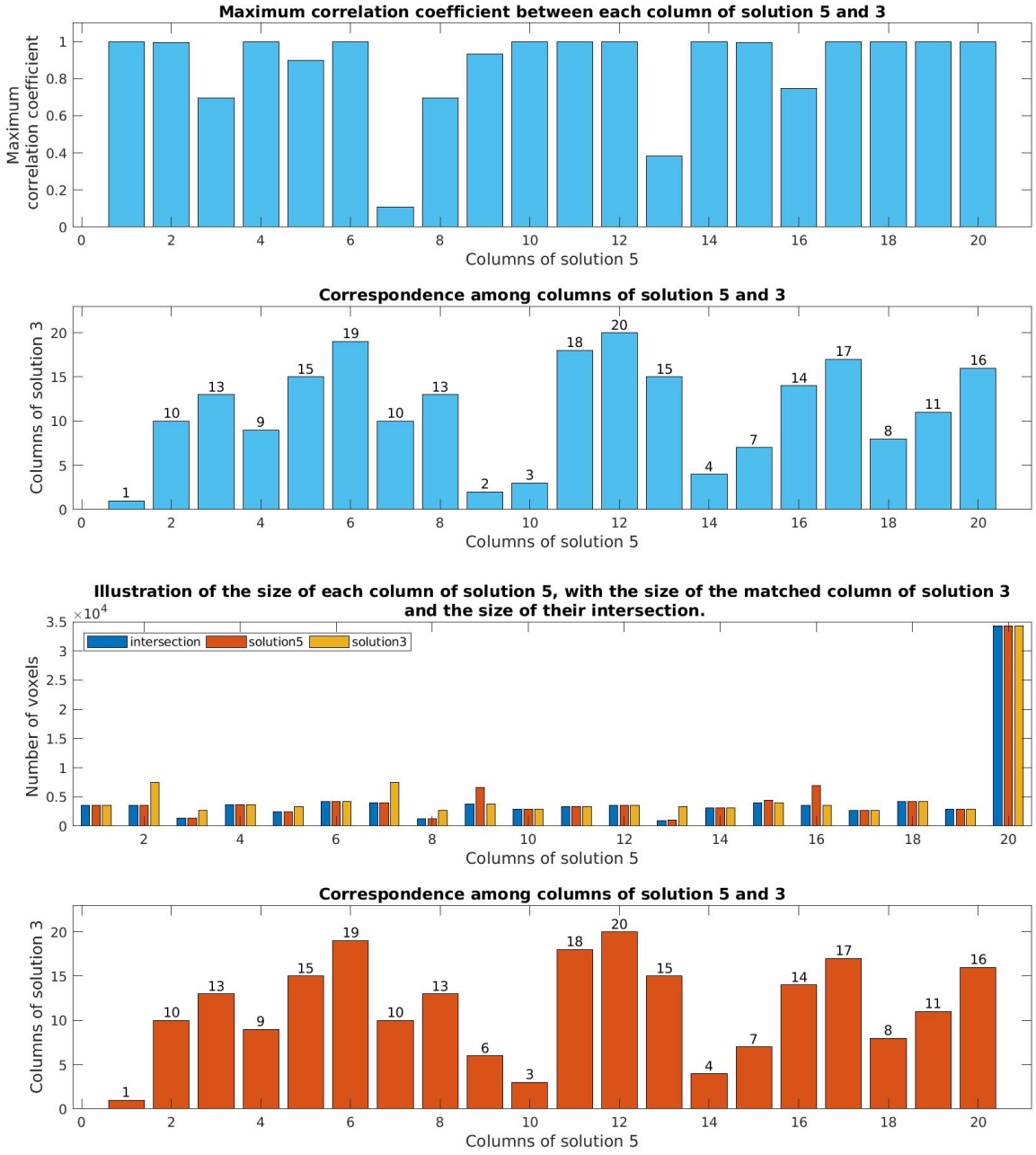


Figure A.50: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 3. We associate the column r of the solution 5 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on intersection. Solutions 5, 3 correspond to penalty method and noiseless data.

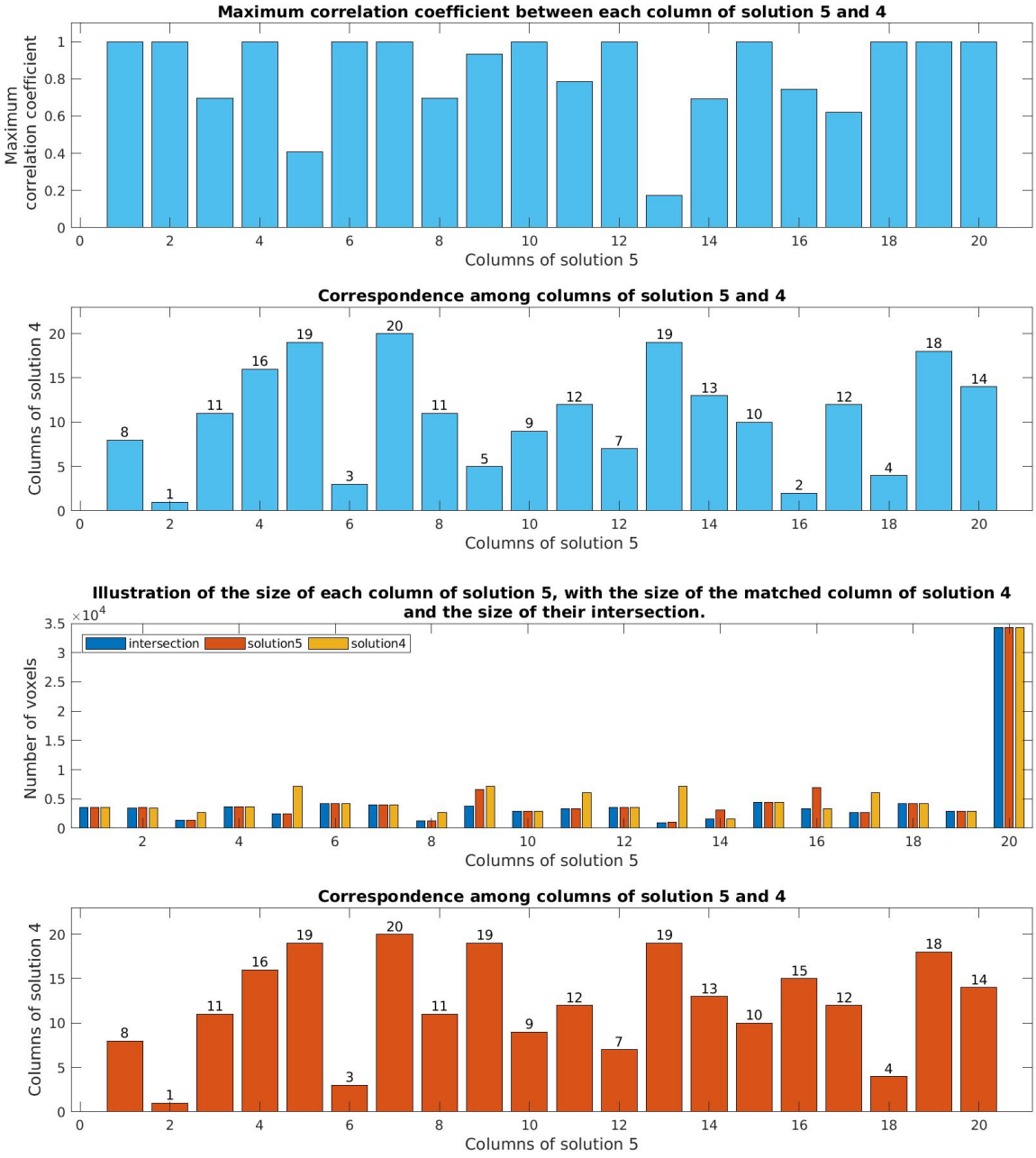


Figure A.51: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 4. We associate the column r of the solution 5 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on intersection. Solutions 5, 4 correspond to penalty method and noiseless data.

· **SNR=0 dB**

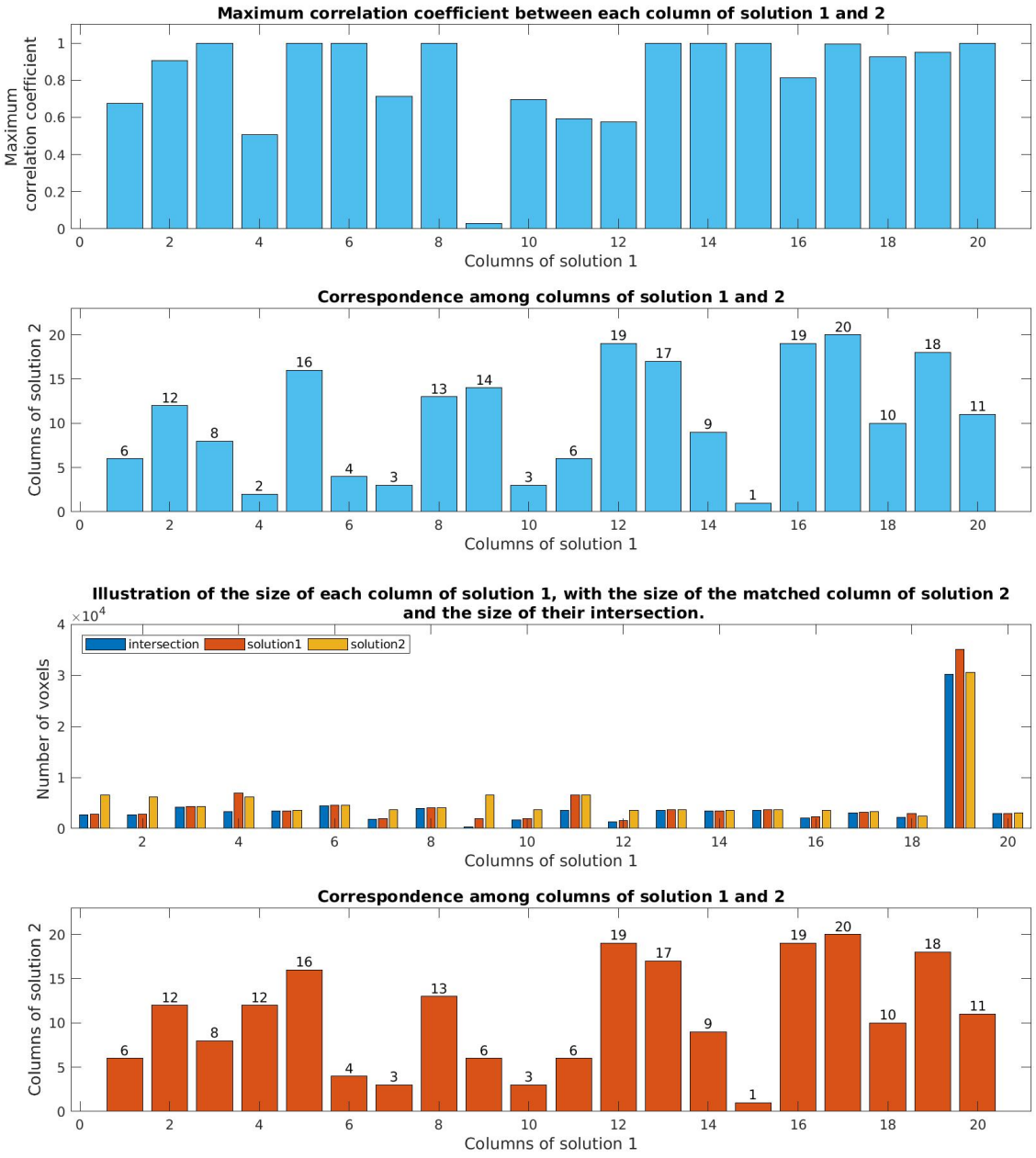


Figure A.52: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated, based on intersection. Solutions 1, 2 correspond to penalty method and SNR=0 dB.

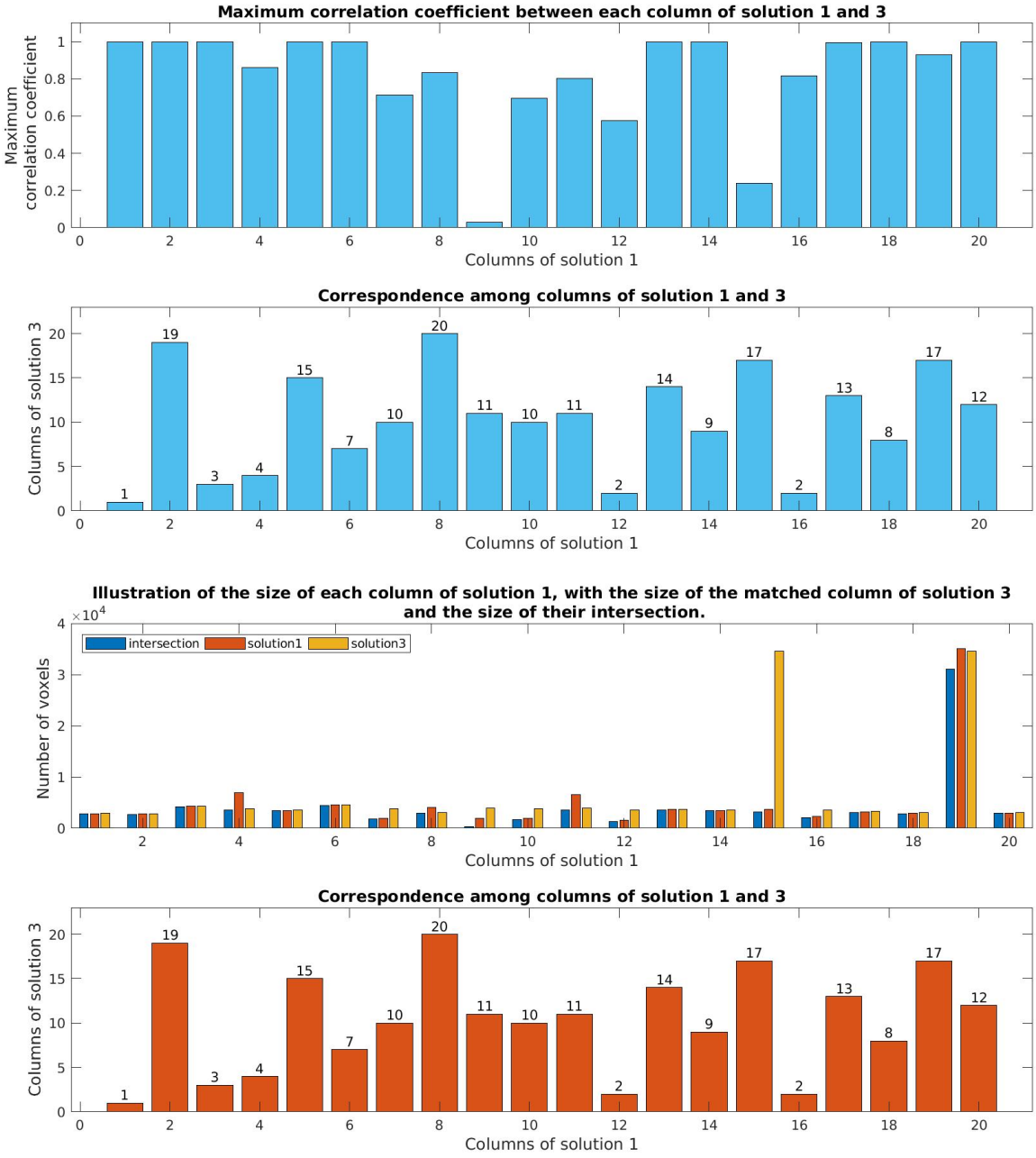


Figure A.53: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on intersection. Solutions 1, 3 correspond to penalty method and SNR=0 dB.

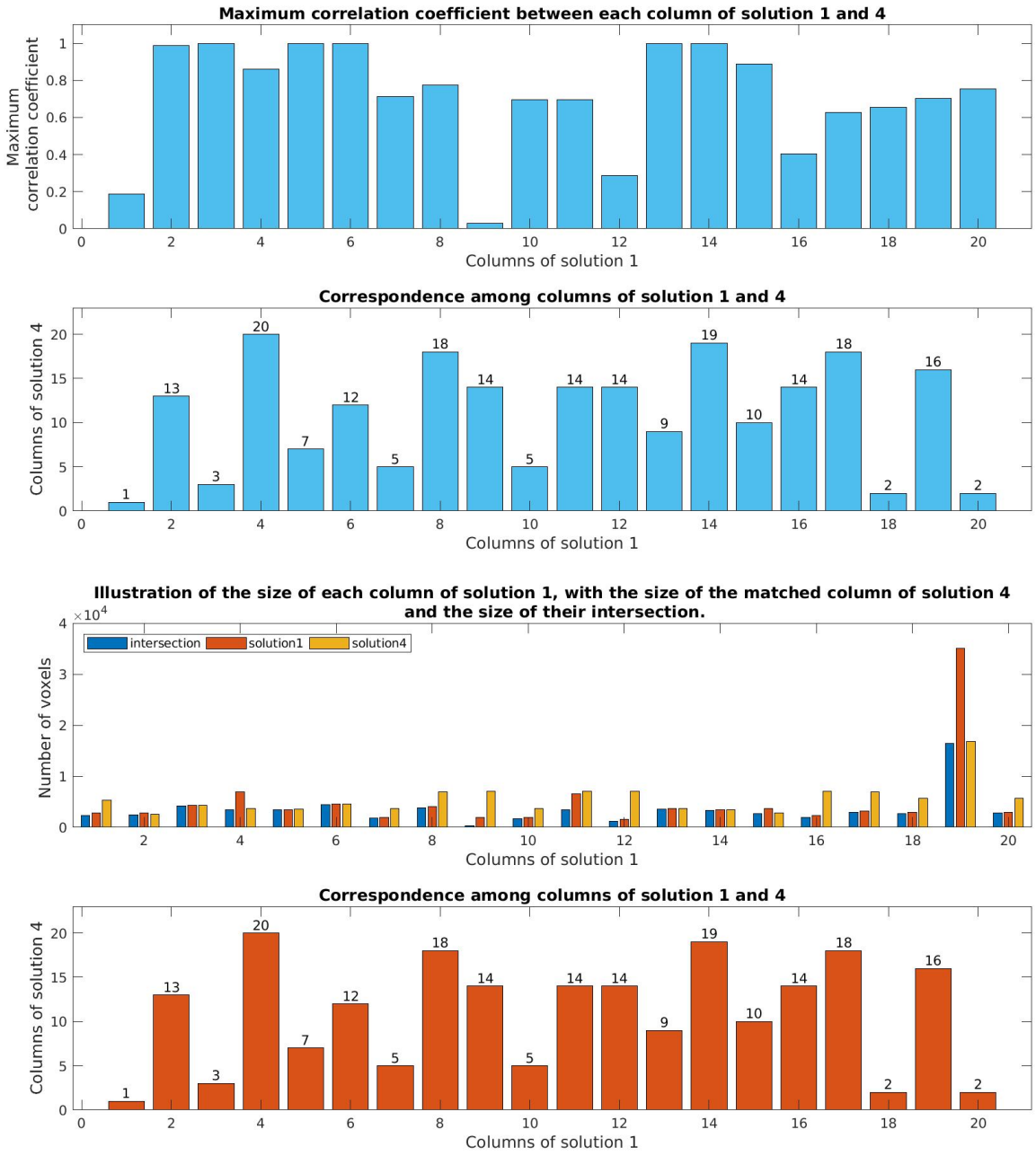


Figure A.54: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on intersection. Solutions 1, 4 correspond to penalty method and SNR=0 dB.

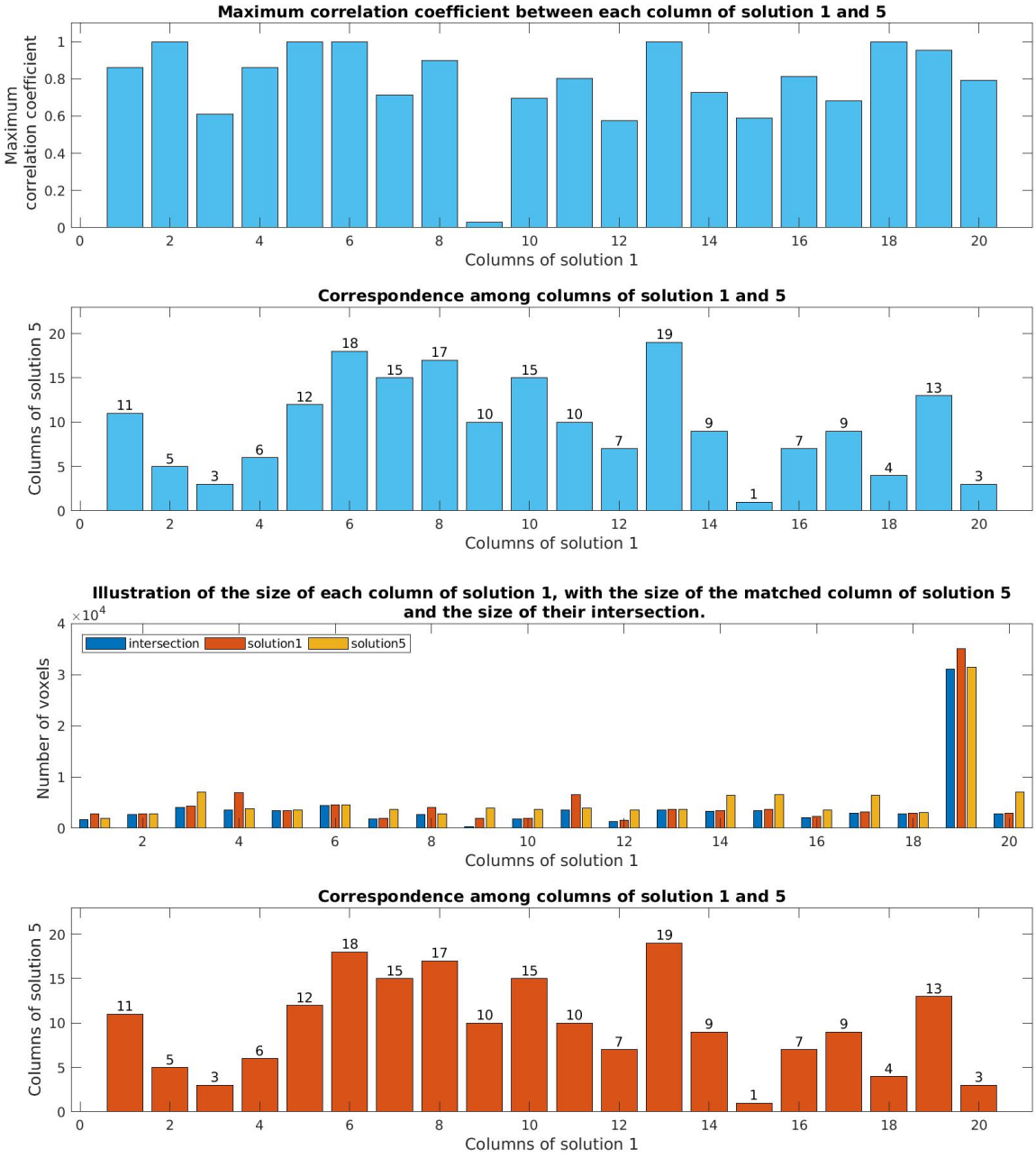


Figure A.55: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 5. We associate the column r of the solution 1 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on intersection. Solutions 1, 5 correspond to penalty method and SNR=0 dB.

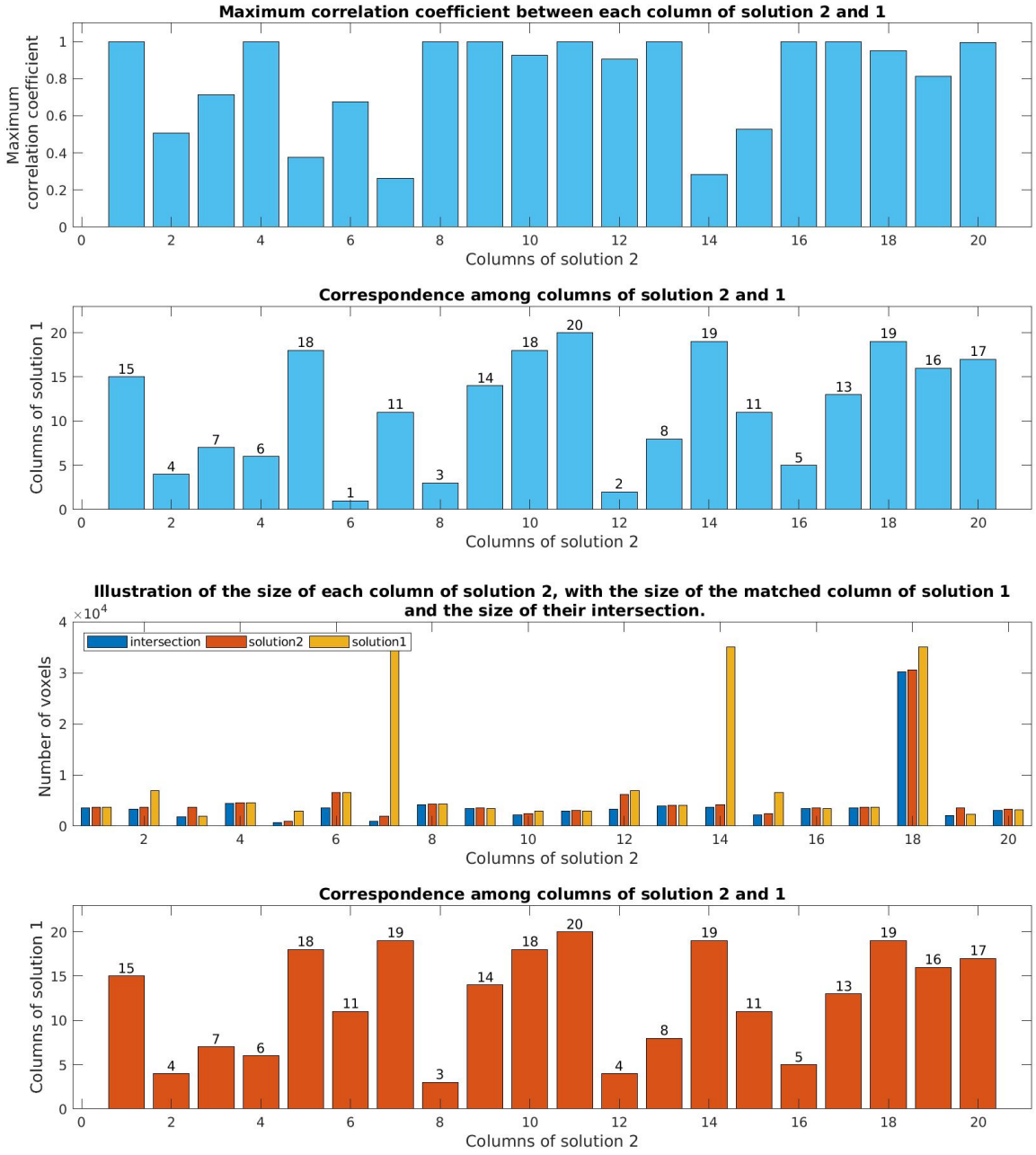


Figure A.56: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on intersection. Solutions 2, 1 correspond to penalty method and SNR=0 dB.

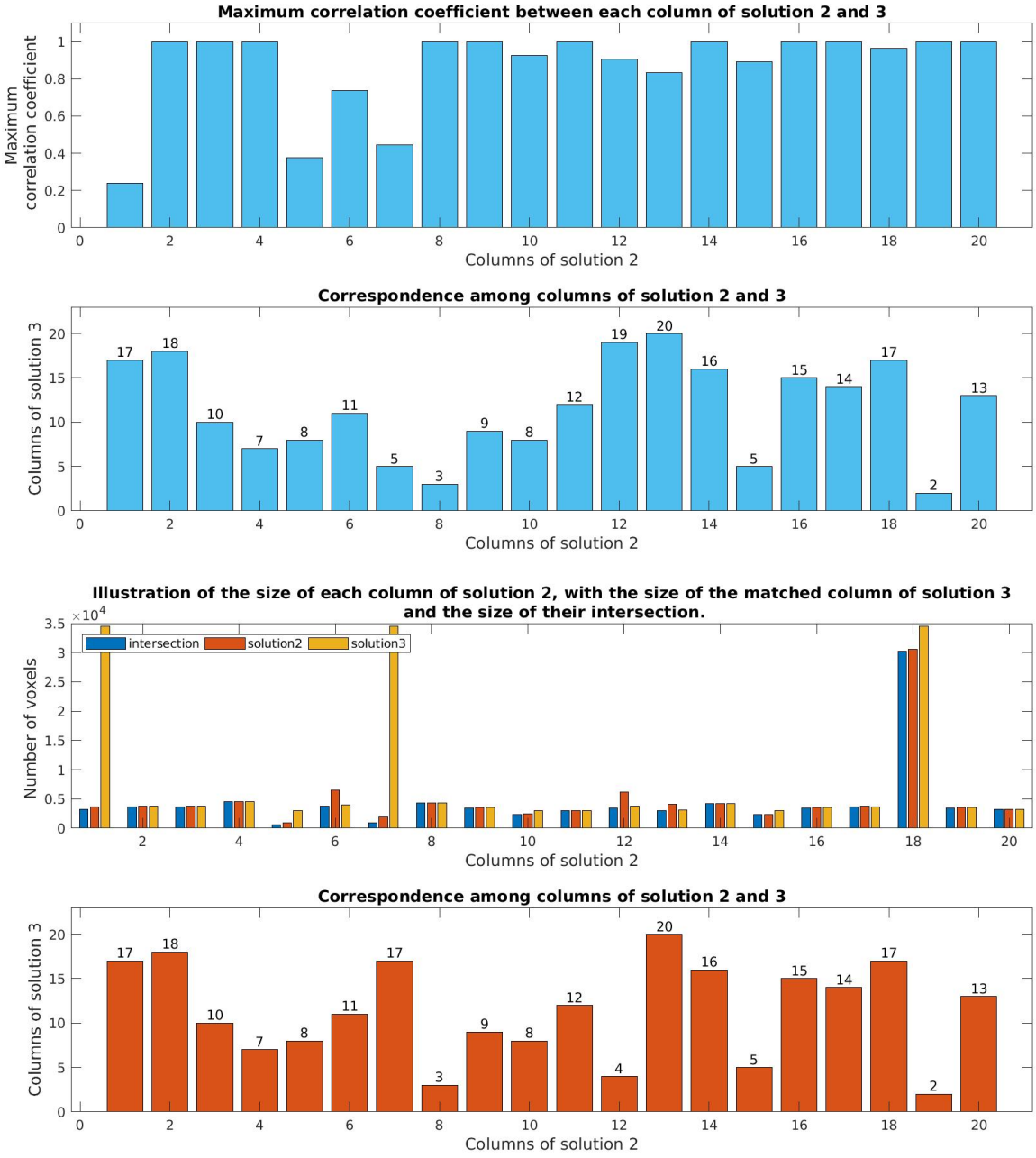


Figure A.57: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on intersection. Solutions 2, 3 correspond to penalty method and SNR=0 dB.

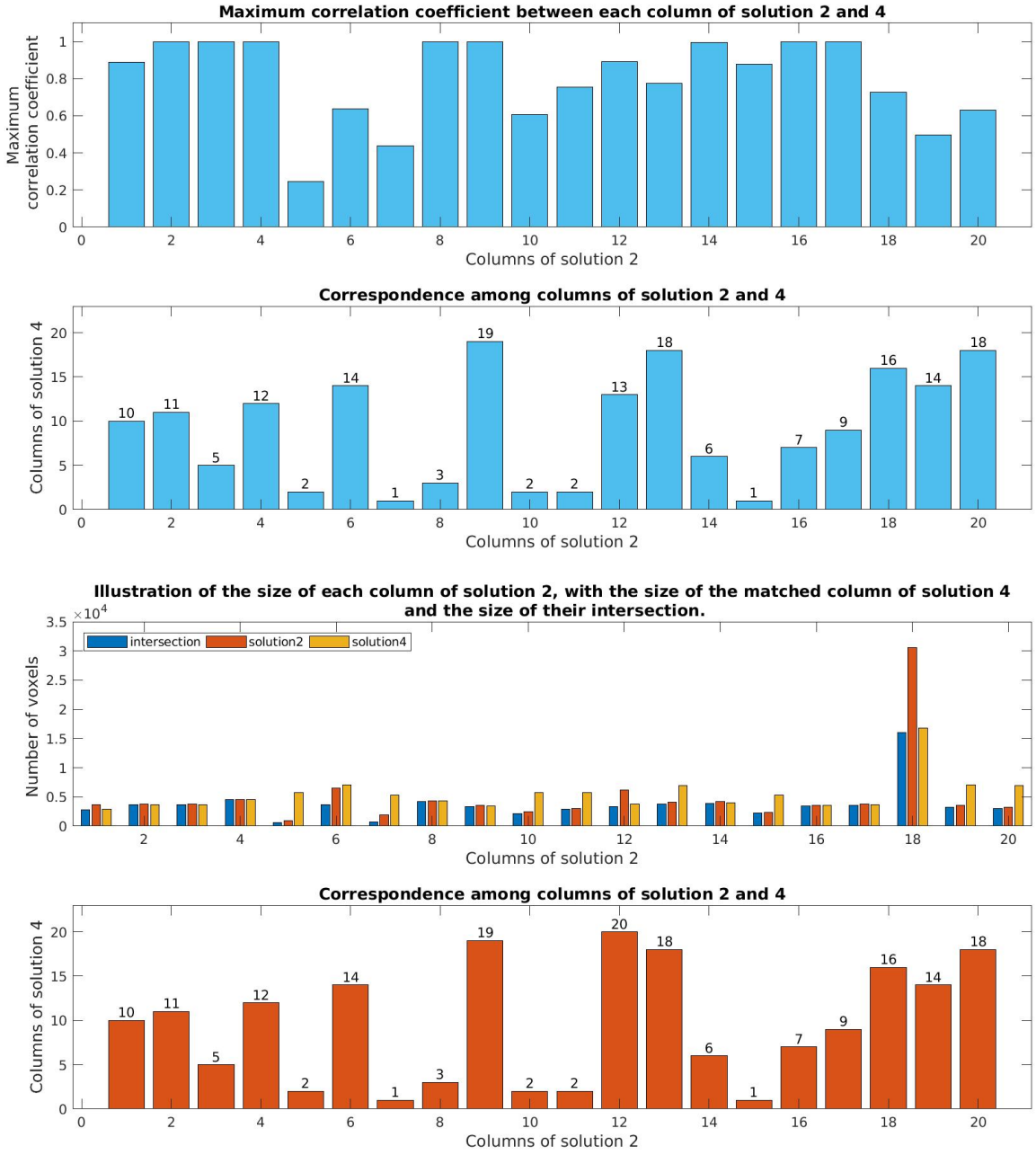


Figure A.58: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on intersection. Solutions 2, 4 correspond to penalty method and SNR=0 dB.

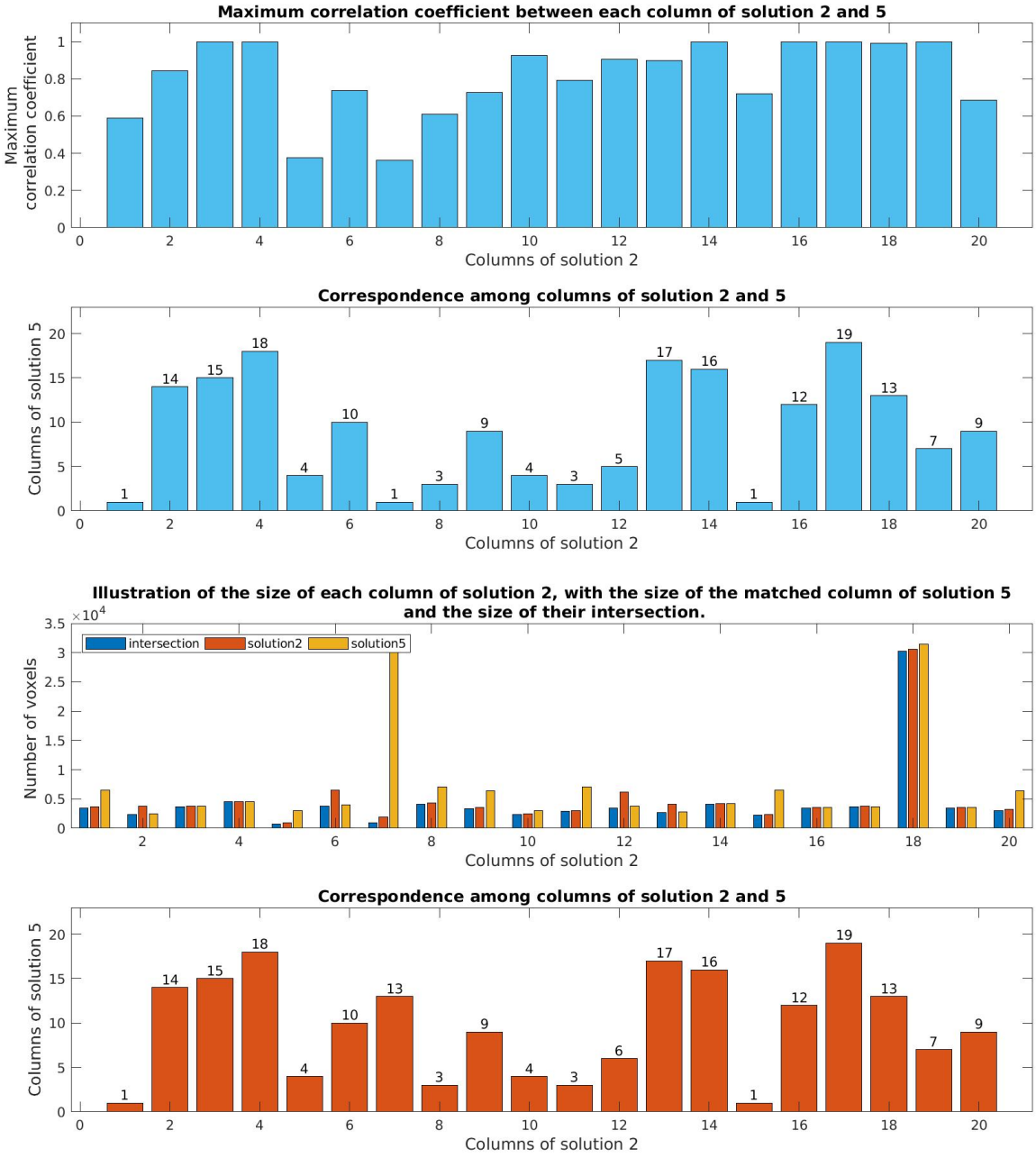


Figure A.59: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 5. We associate the column r of the solution 2 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on intersection. Solutions 2, 5 correspond to penalty method and SNR=0 dB.

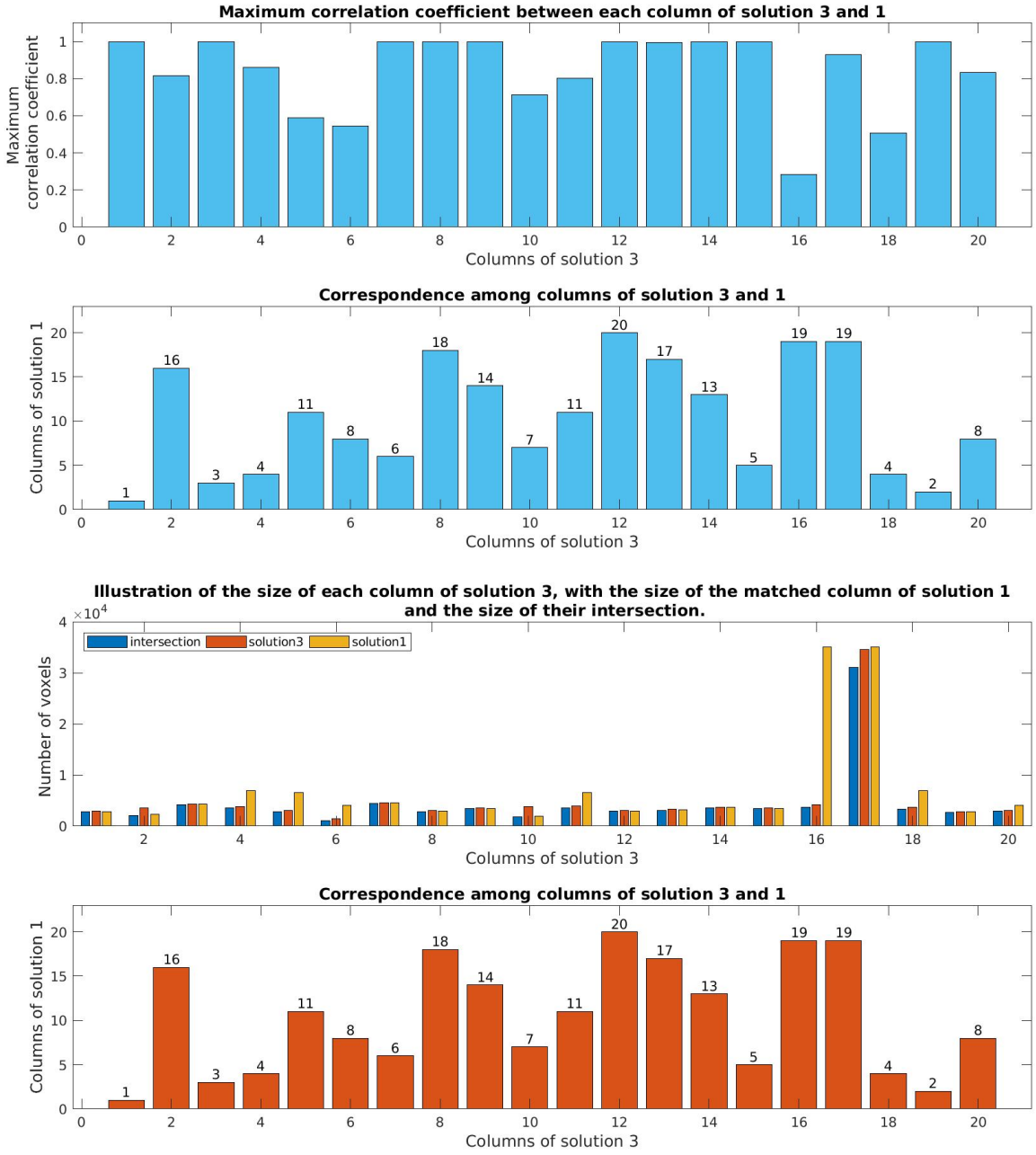


Figure A.60: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on intersection. Solutions 3, 1 correspond to penalty method and SNR=0 dB.

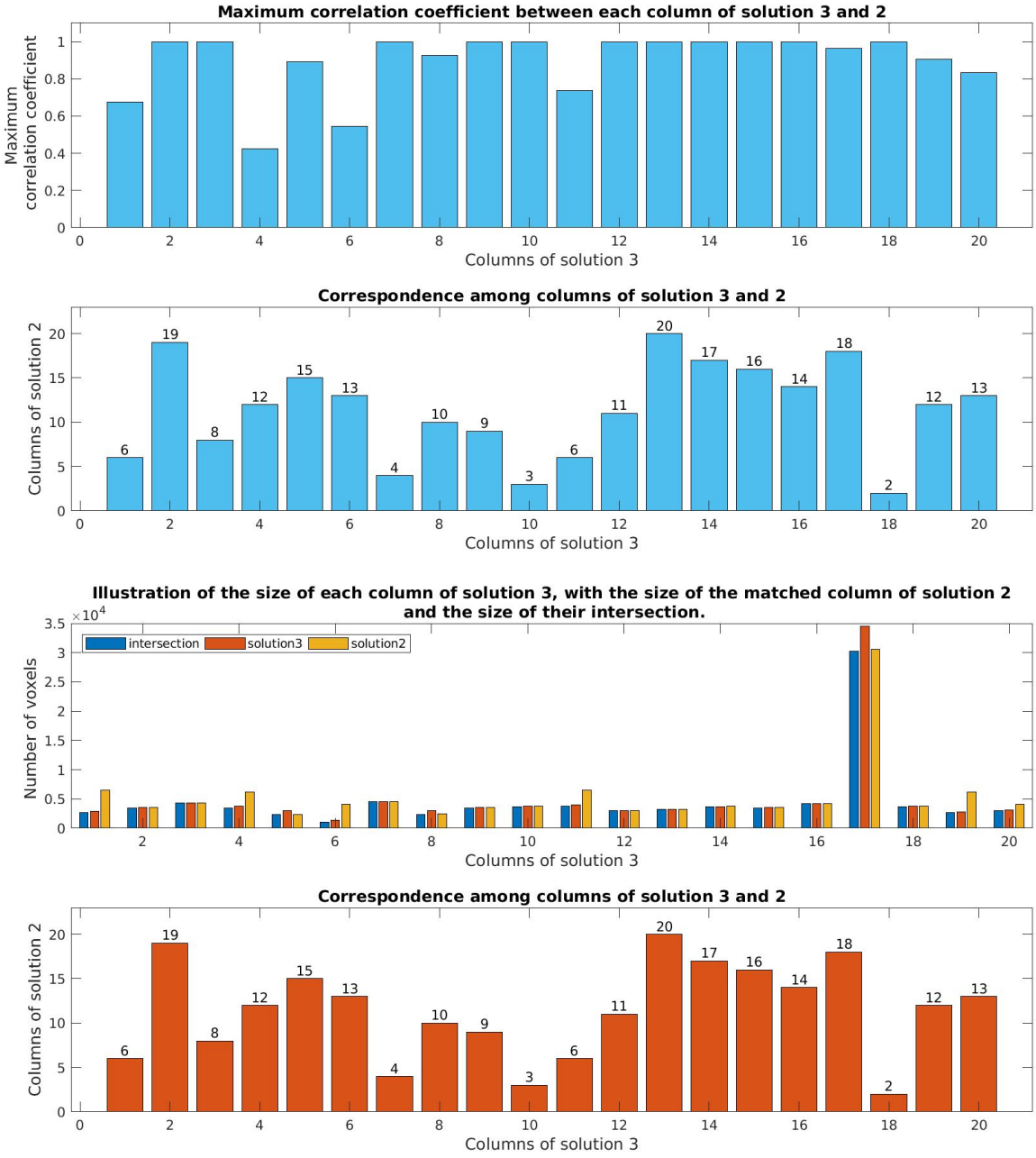


Figure A.61: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on intersection. Solutions 3, 2 correspond to penalty method and SNR=0 dB.

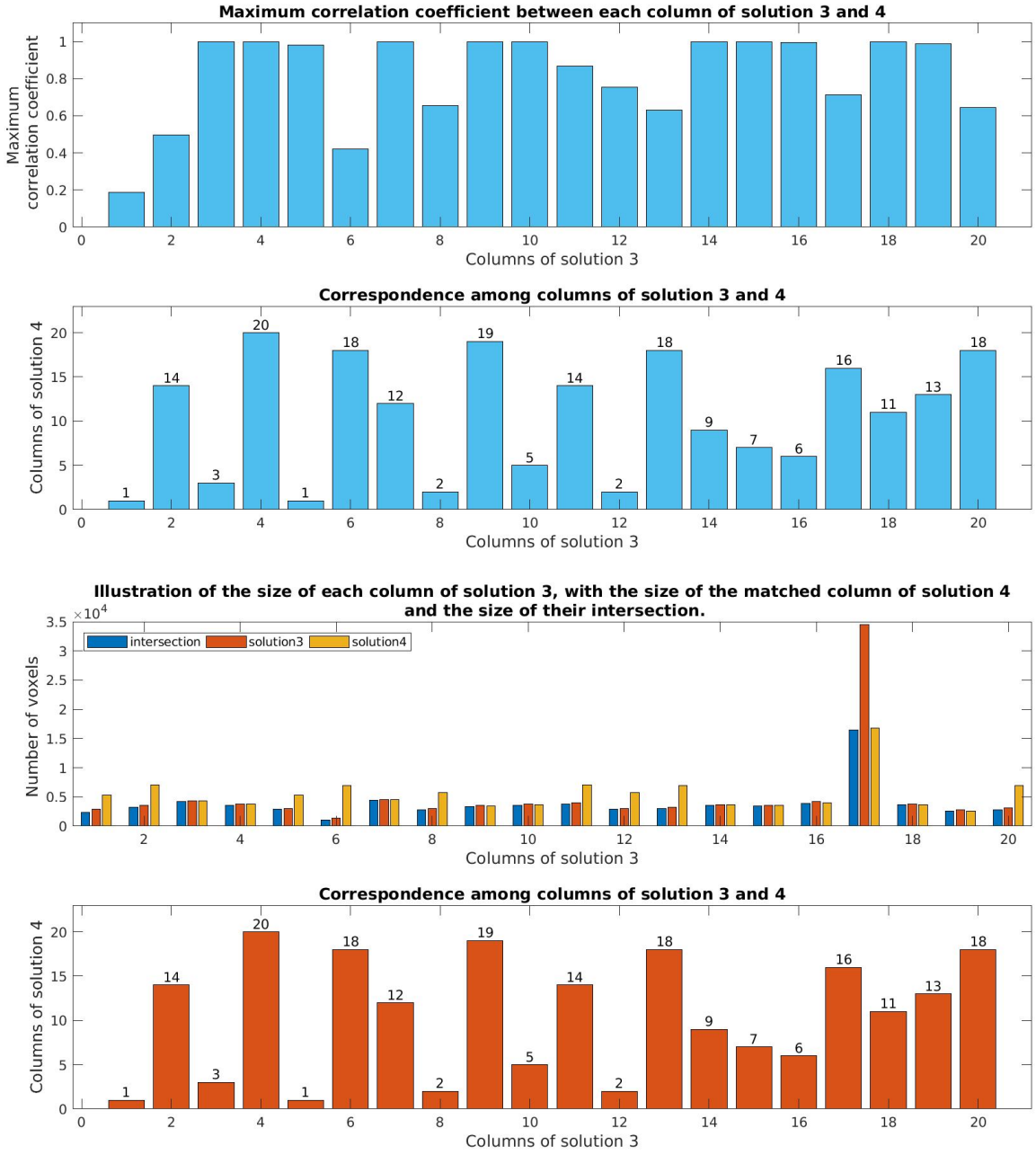


Figure A.62: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on intersection. Solutions 3, 4 correspond to penalty method and SNR=0 dB.

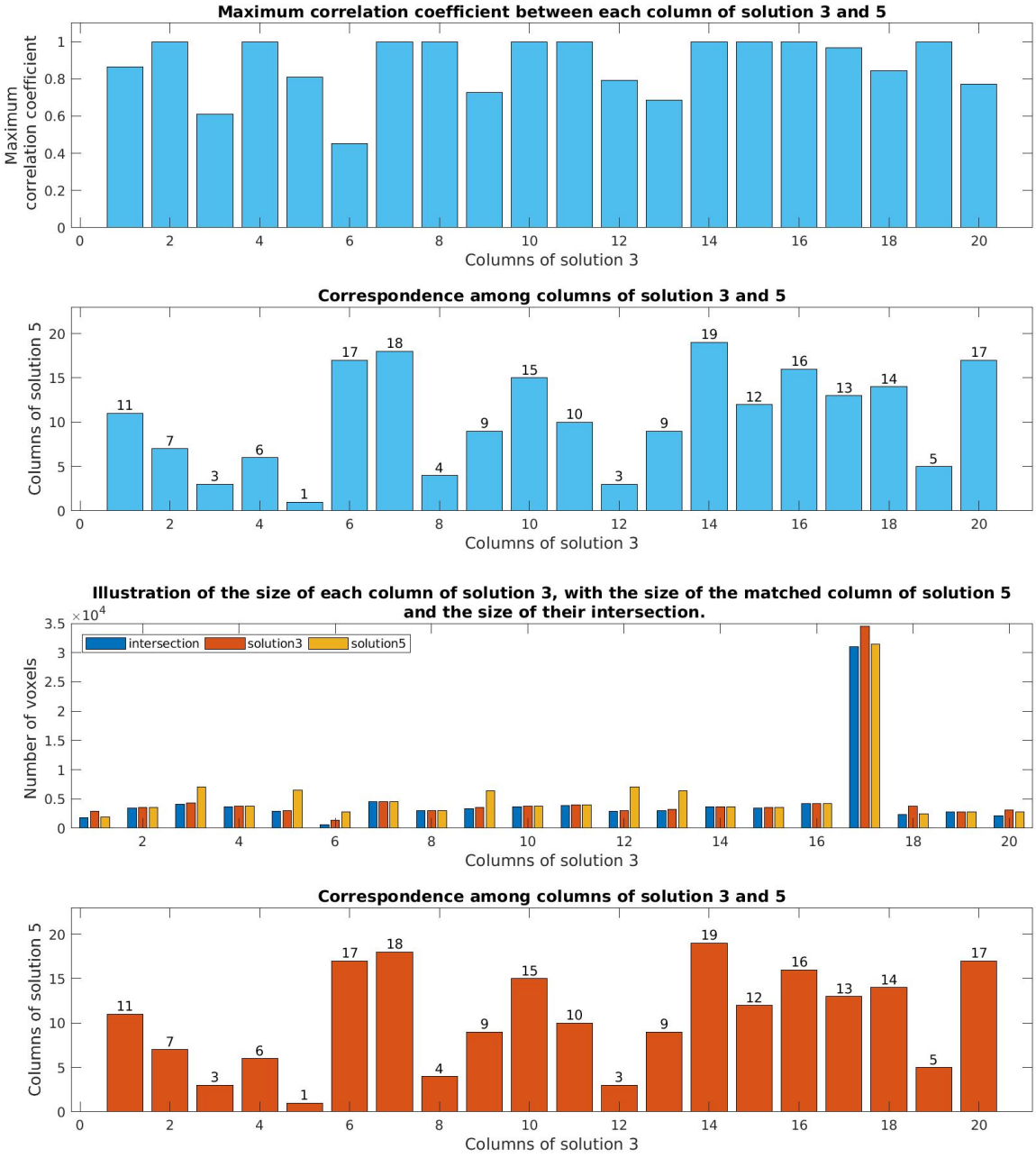


Figure A.63: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 5. We associate the column r of the solution 3 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on intersection. Solutions 3, 5 correspond to penalty method and SNR=0 dB.

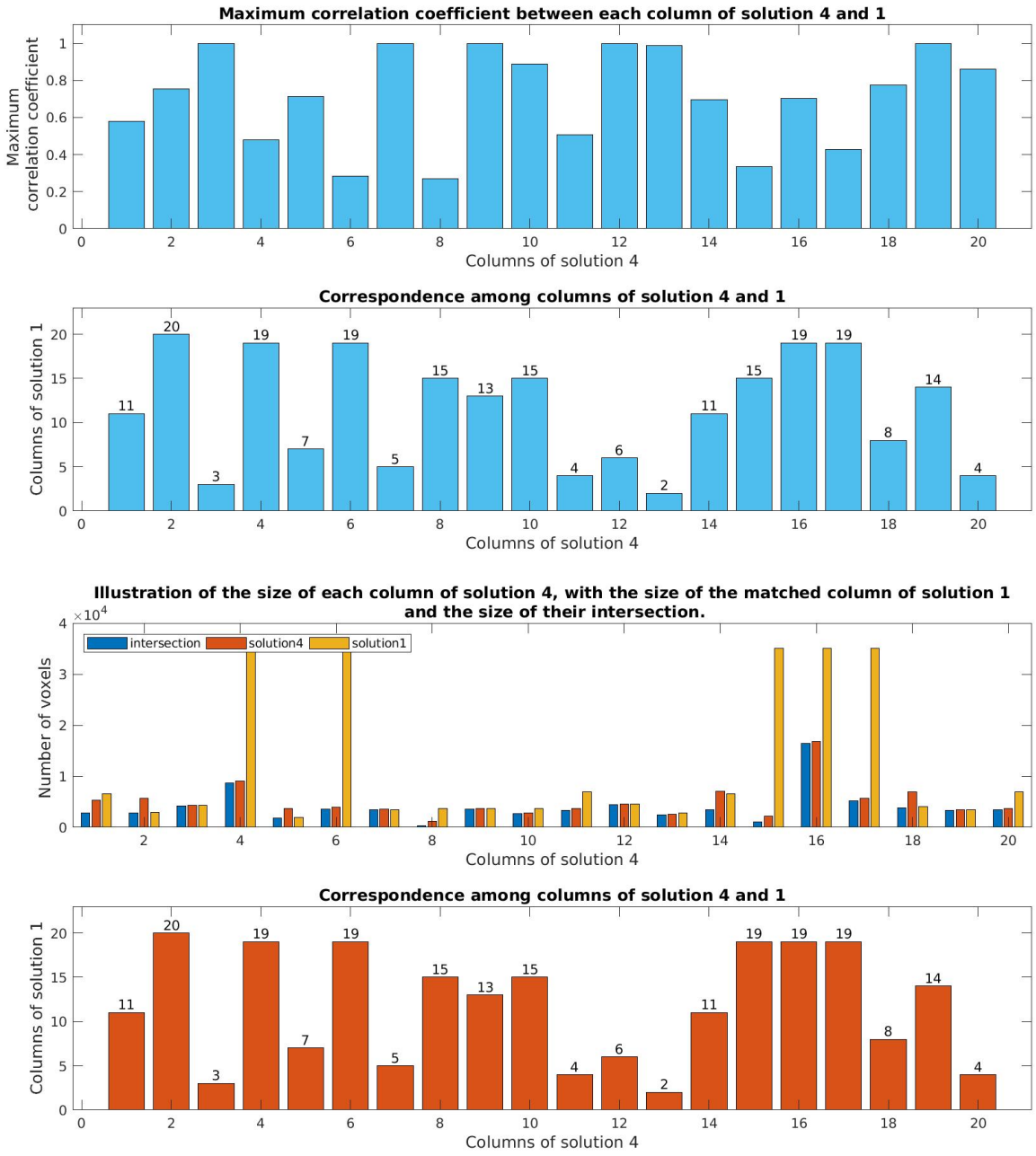


Figure A.64: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on intersection. Solutions 4, 1 correspond to penalty method and SNR=0 dB.

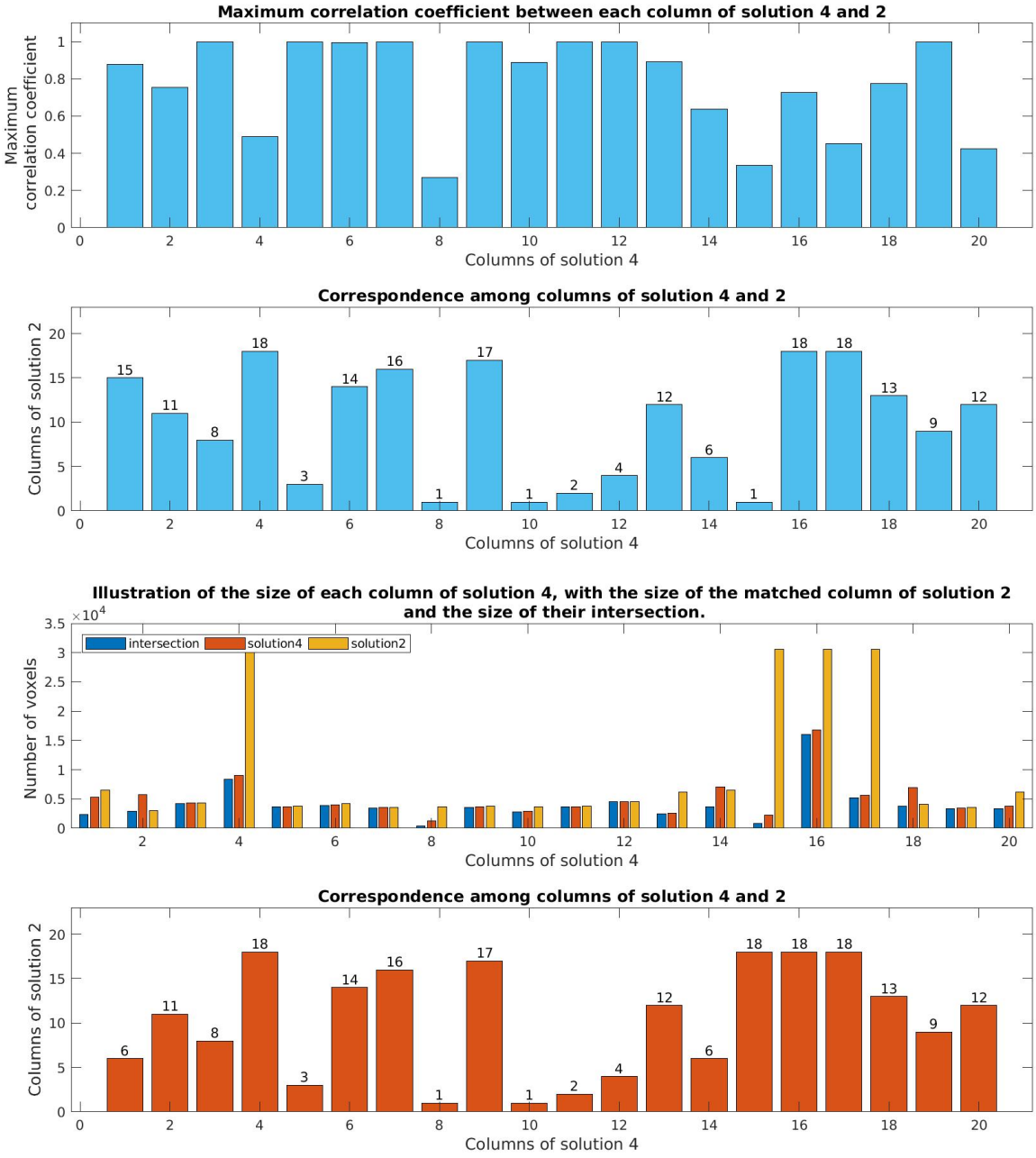


Figure A.65: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on intersection. Solutions 4, 2 correspond to penalty method and SNR=0 dB.

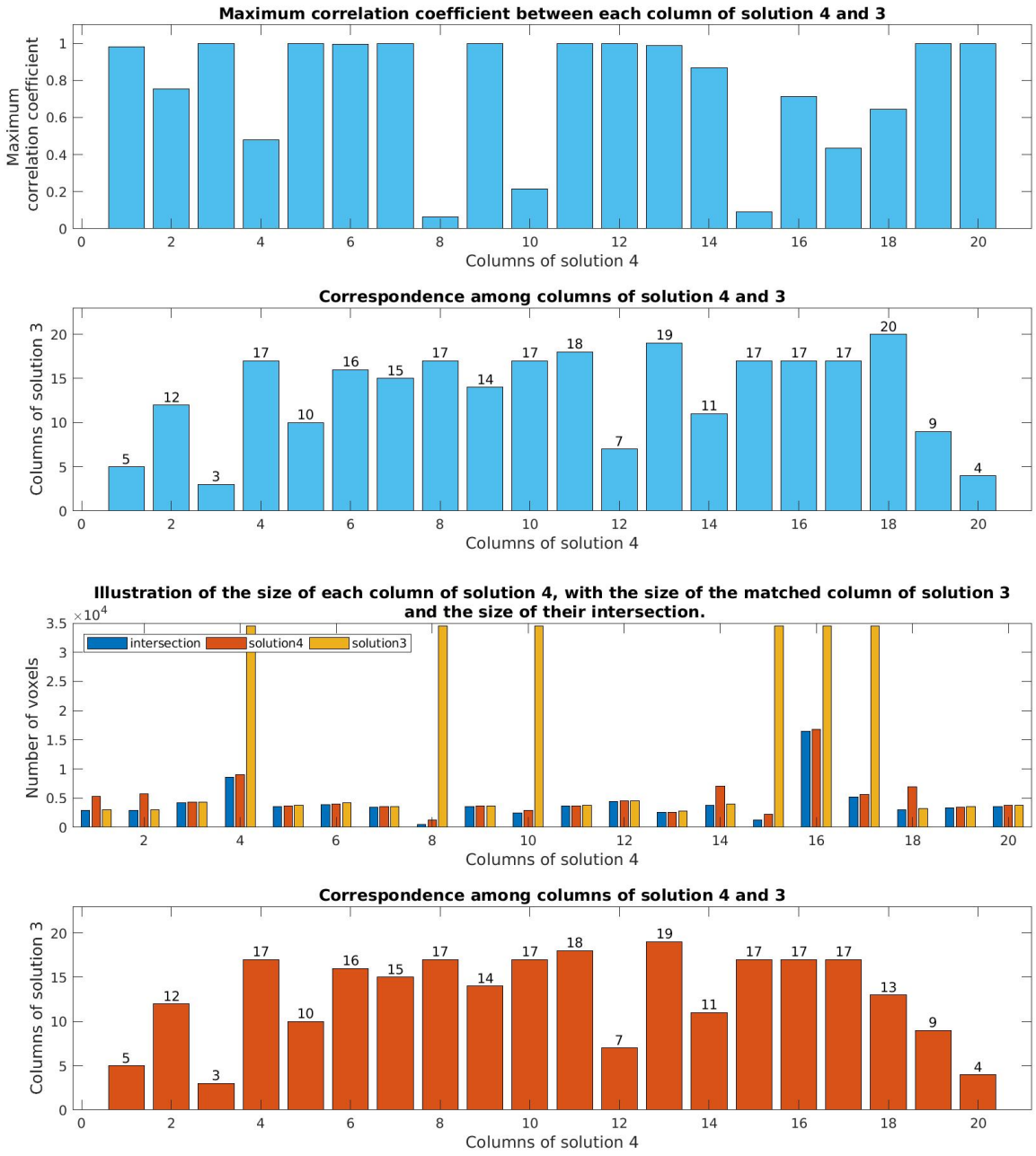


Figure A.66: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on intersection. Solutions 4, 3 correspond to penalty method and SNR=0 dB.

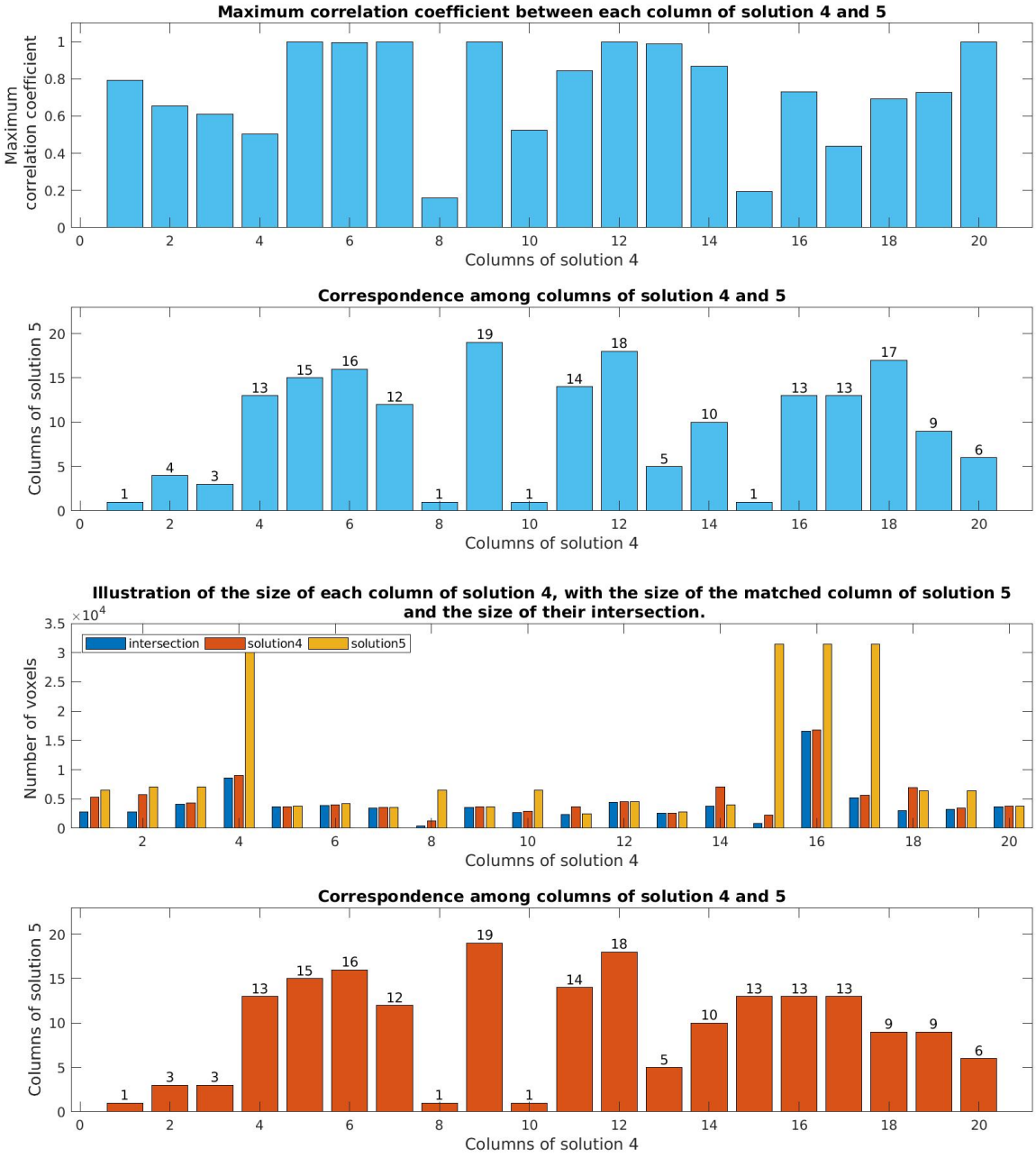


Figure A.67: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 5. We associate the column r of the solution 4 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on intersection. Solutions 4, 5 correspond to penalty method and SNR=0 dB.

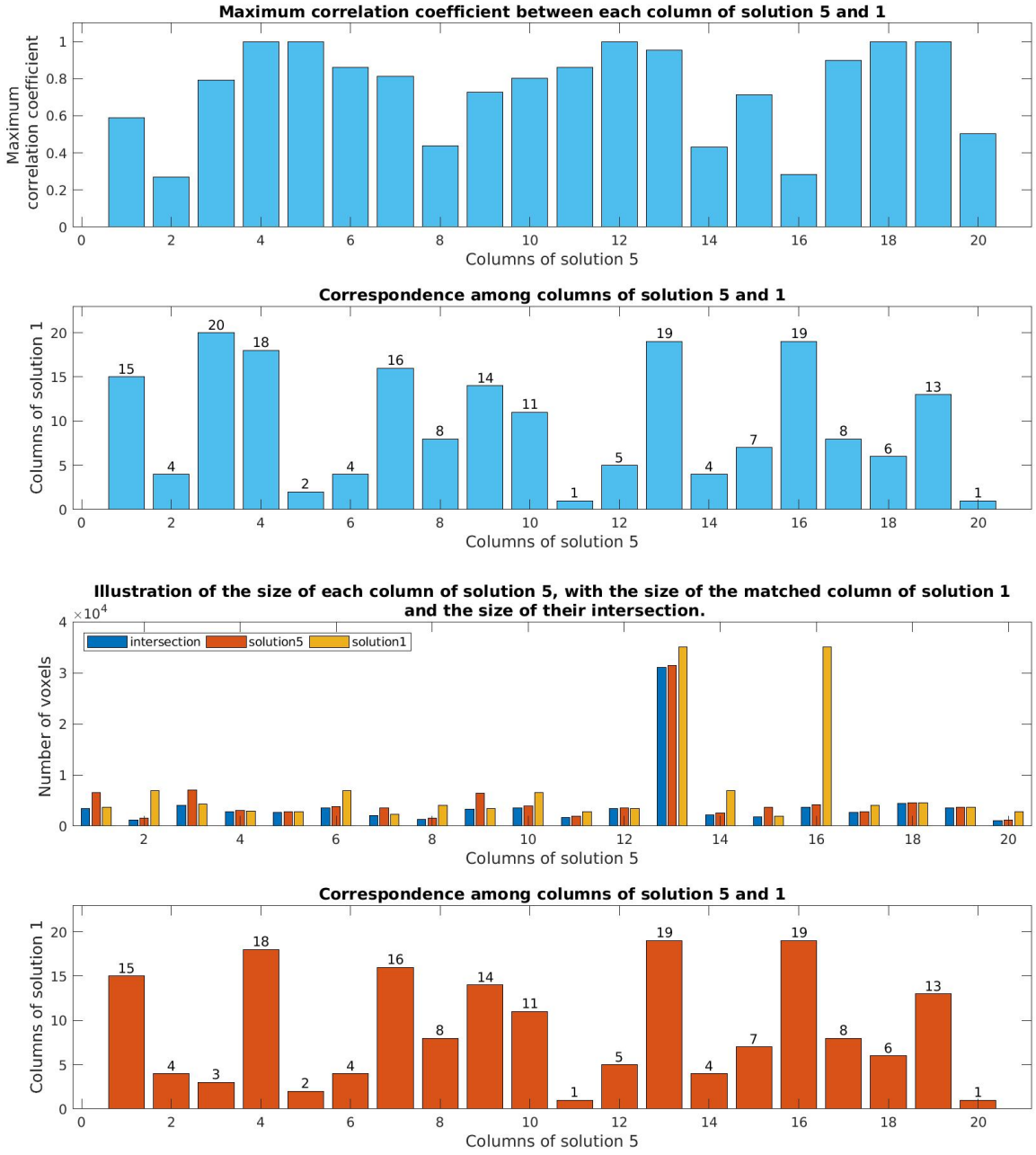


Figure A.68: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 1. We associate the column r of the solution 5 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on intersection. Solutions 5, 1 correspond to penalty method and SNR=0 dB.

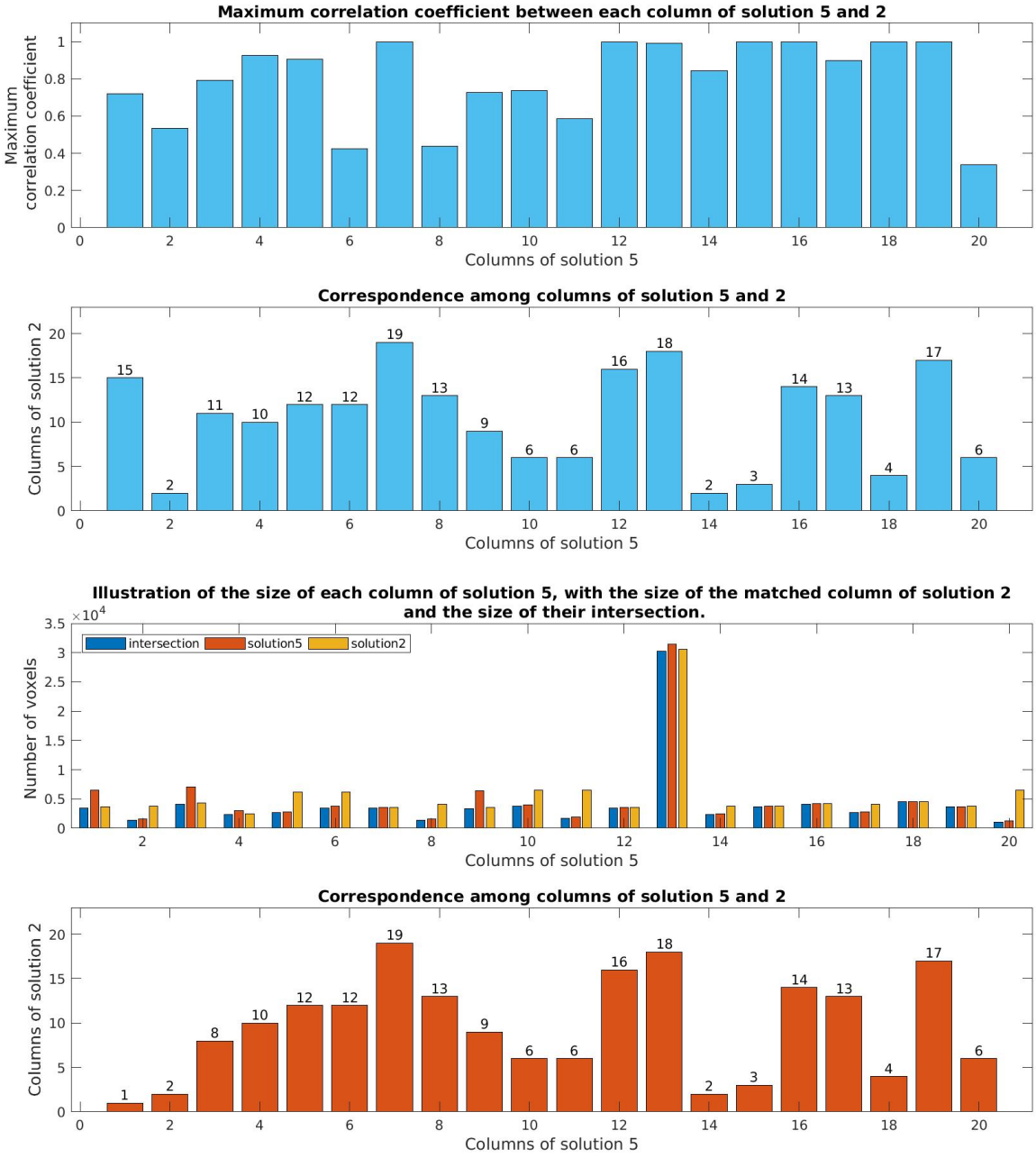


Figure A.69: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 2. We associate the column r of the solution 5 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on intersection. Solutions 5, 2 correspond to penalty method and SNR=0 dB.

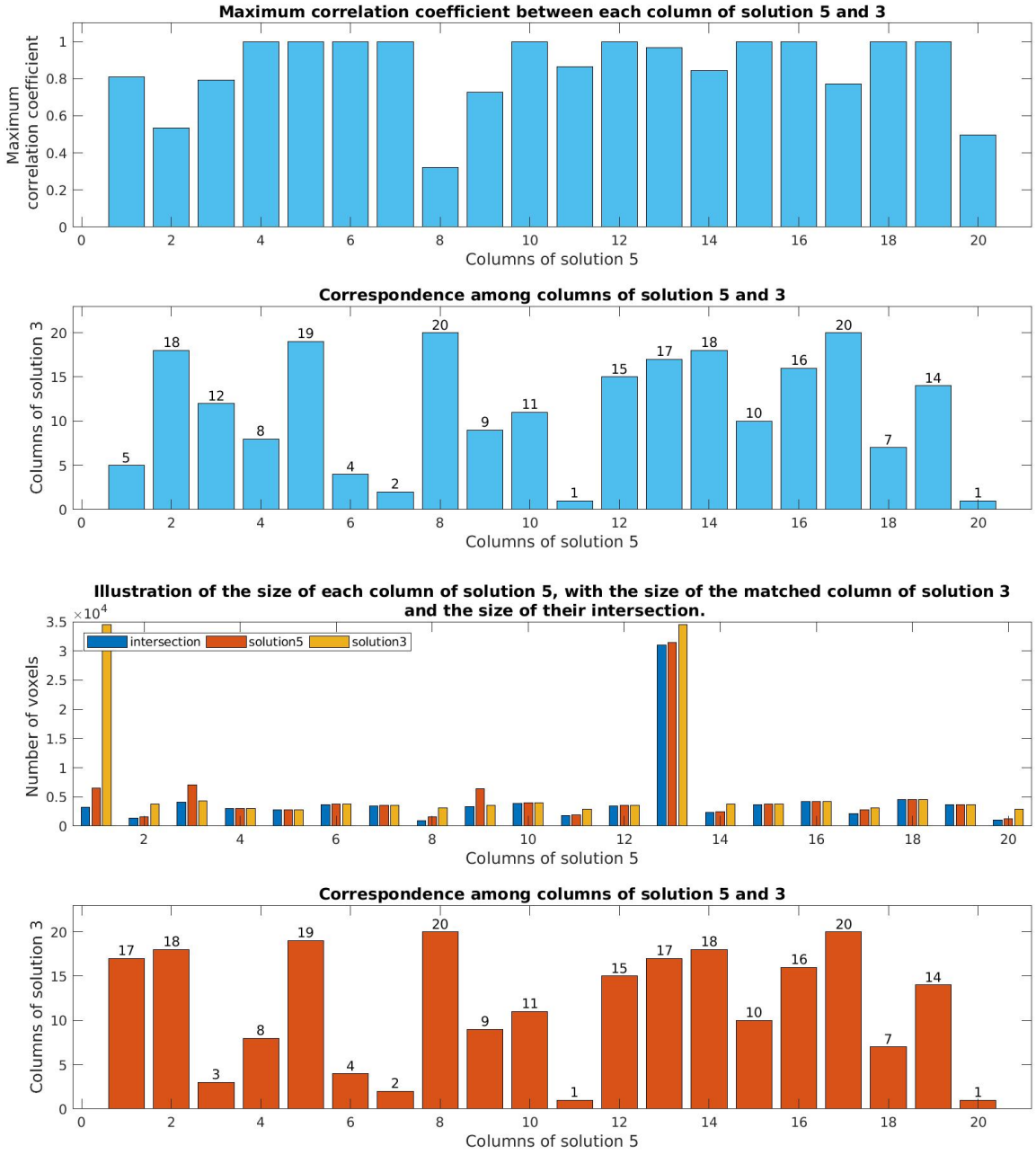


Figure A.70: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 3. We associate the column r of the solution 5 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on intersection. Solutions 5, 3 correspond to penalty method and SNR=0 dB.

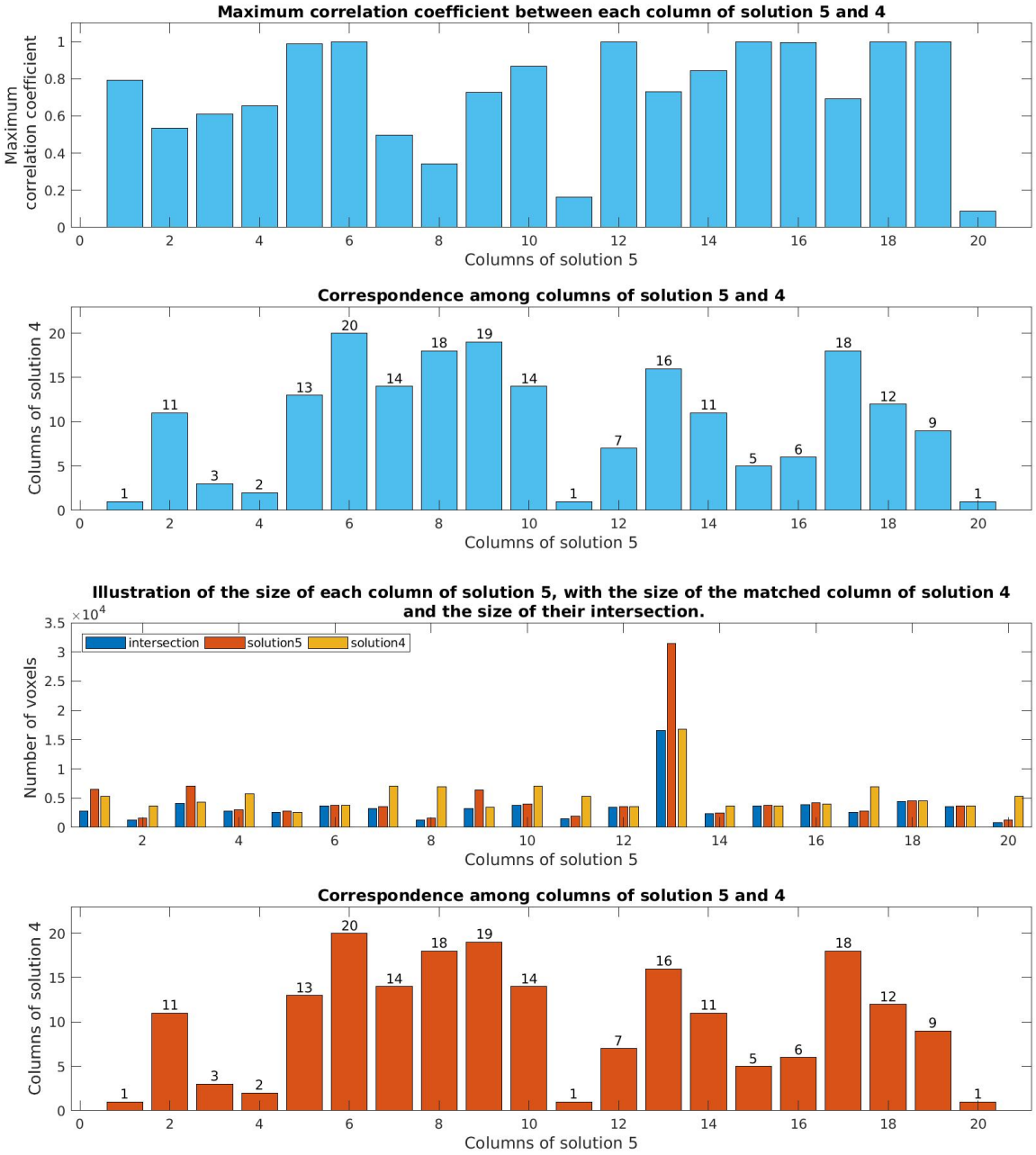


Figure A.71: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 4. We associate the column r of the solution 5 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on intersection. Solutions 5, 4 correspond to penalty method and SNR=0 dB.

· **Comparison of Penalty method solutions to real factor A.**

· **Noiseless case**

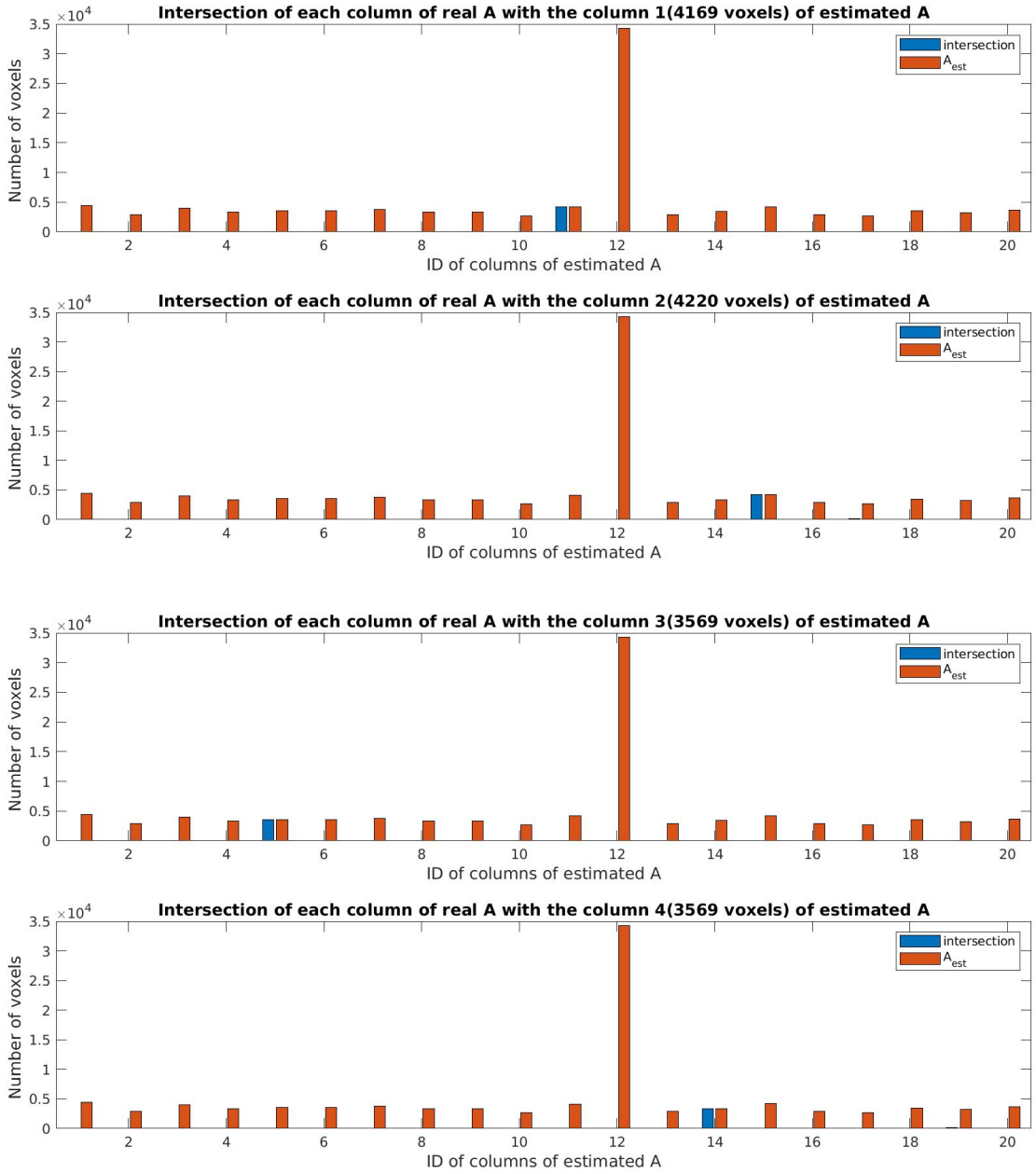


Figure A.72: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to penalty method and noiseless case.

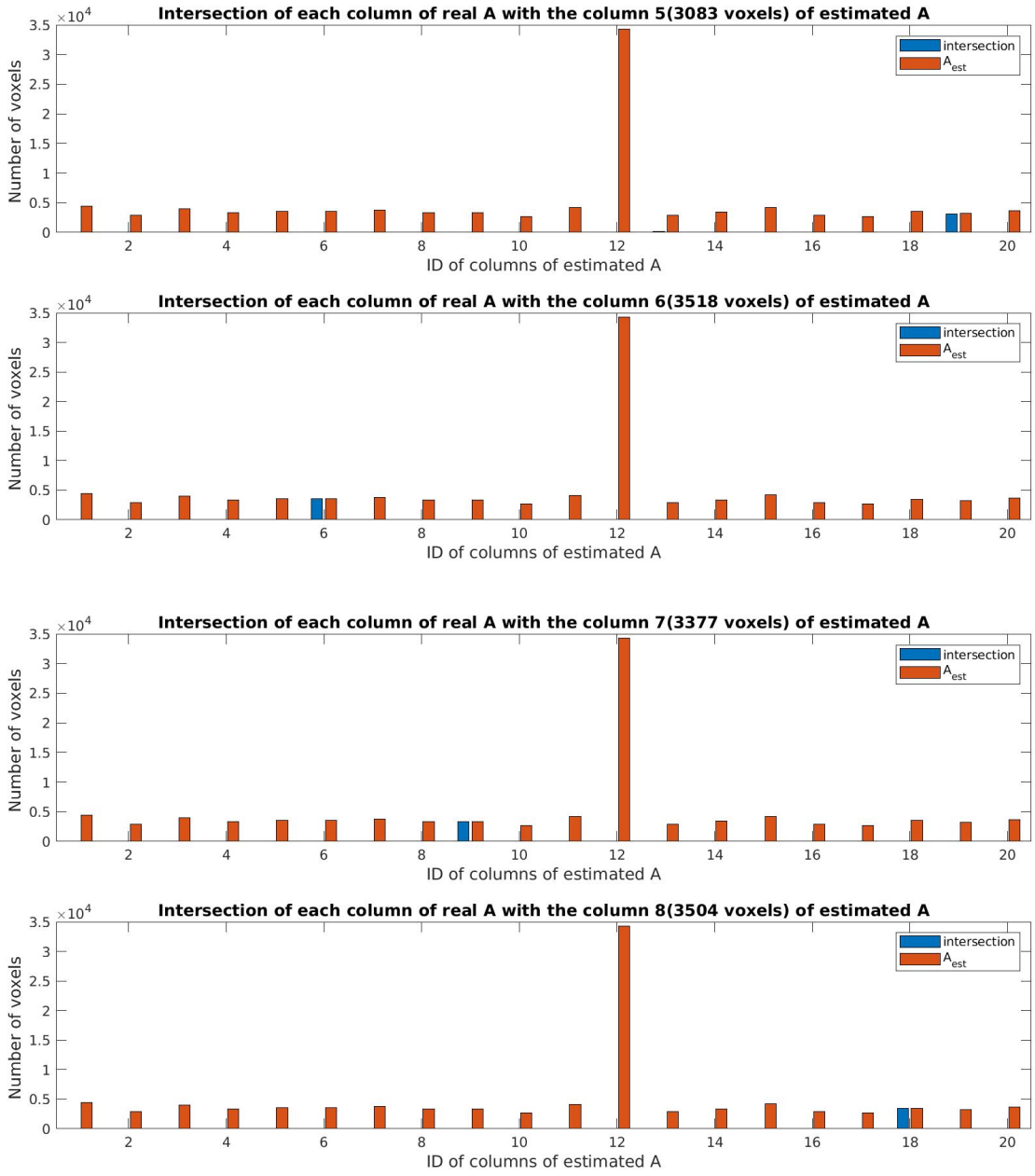


Figure A.73: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to penalty method and noiseless case.

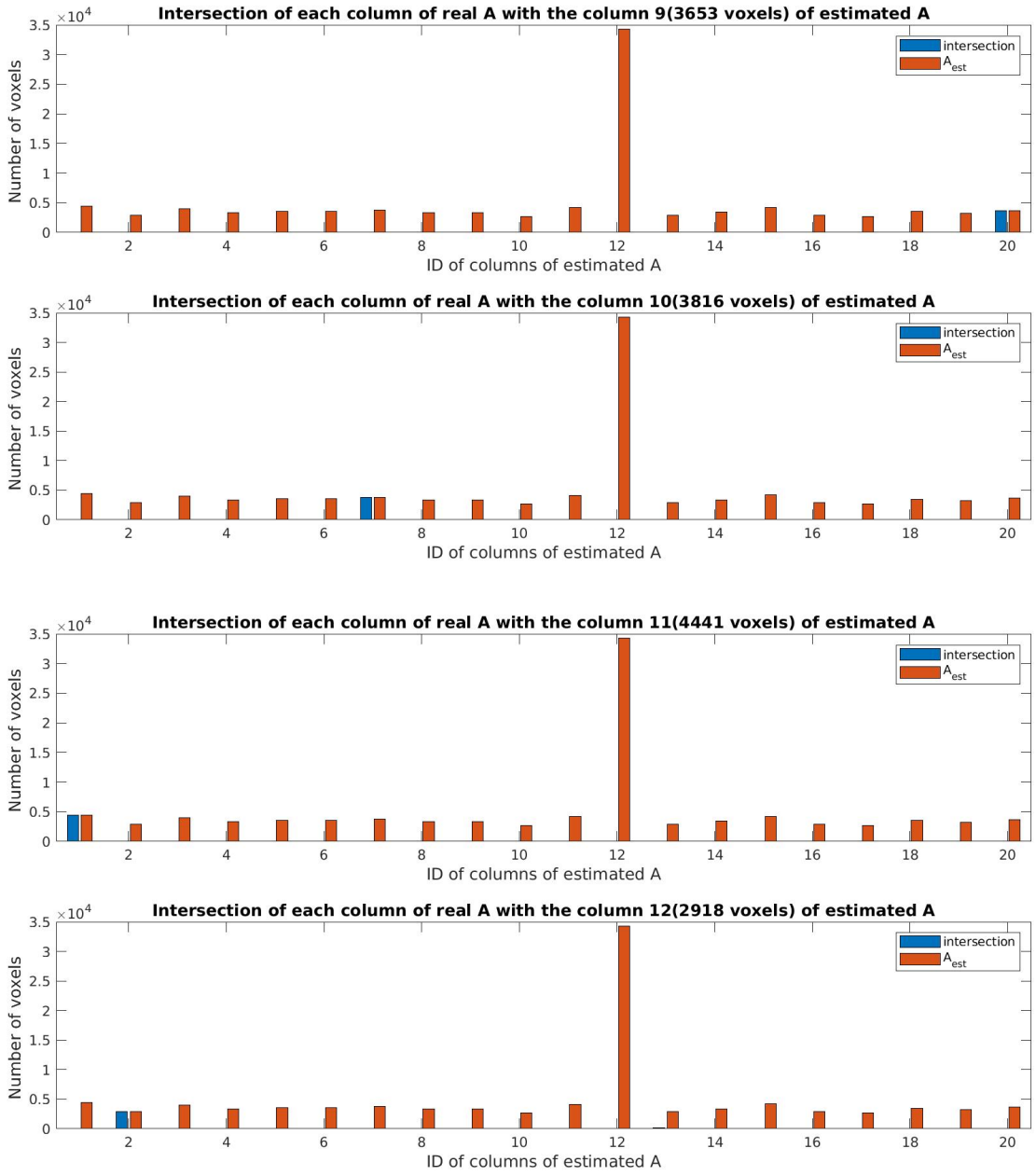


Figure A.74: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to penalty method and noiseless case.

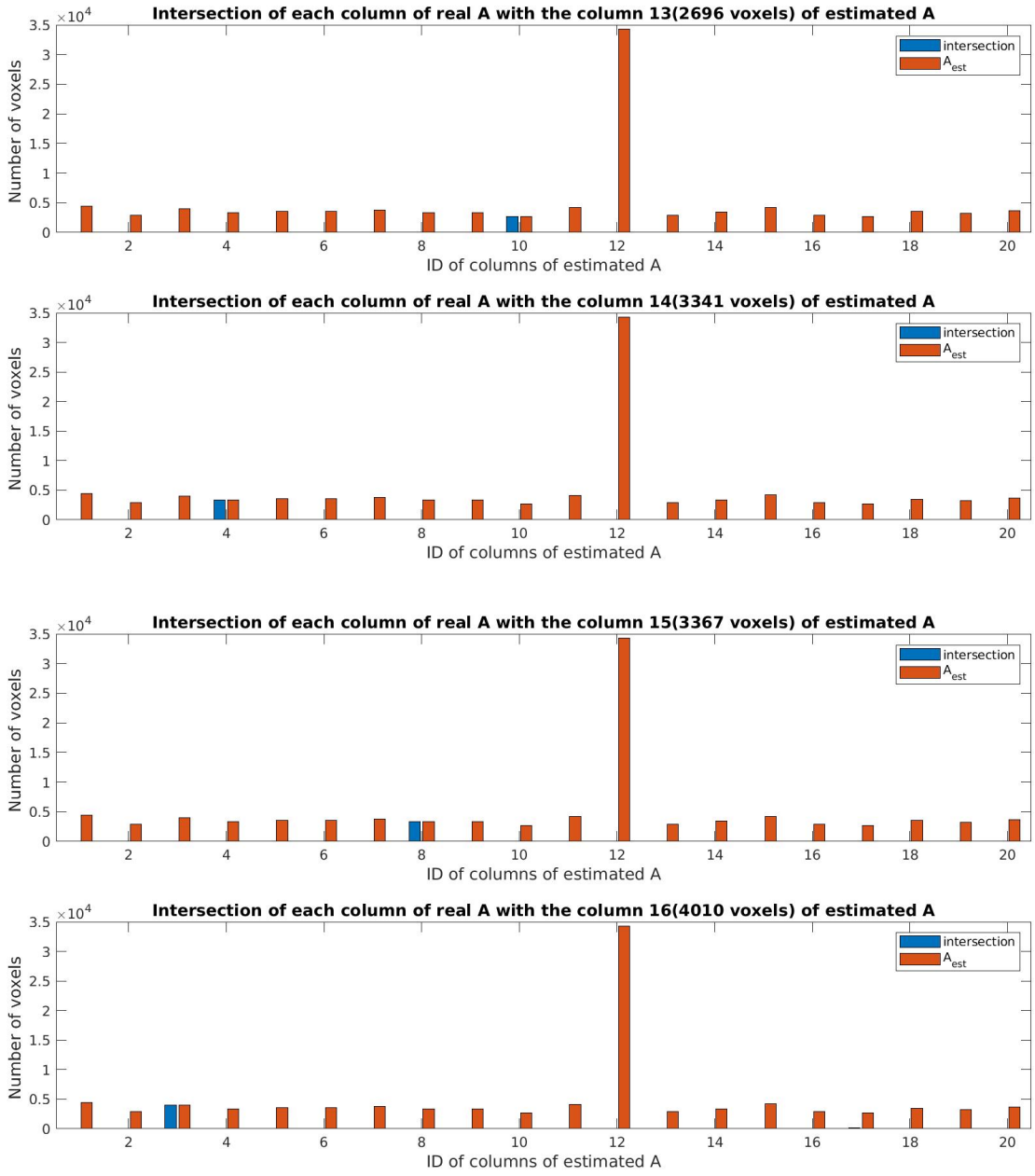


Figure A.75: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to penalty method and noiseless case.

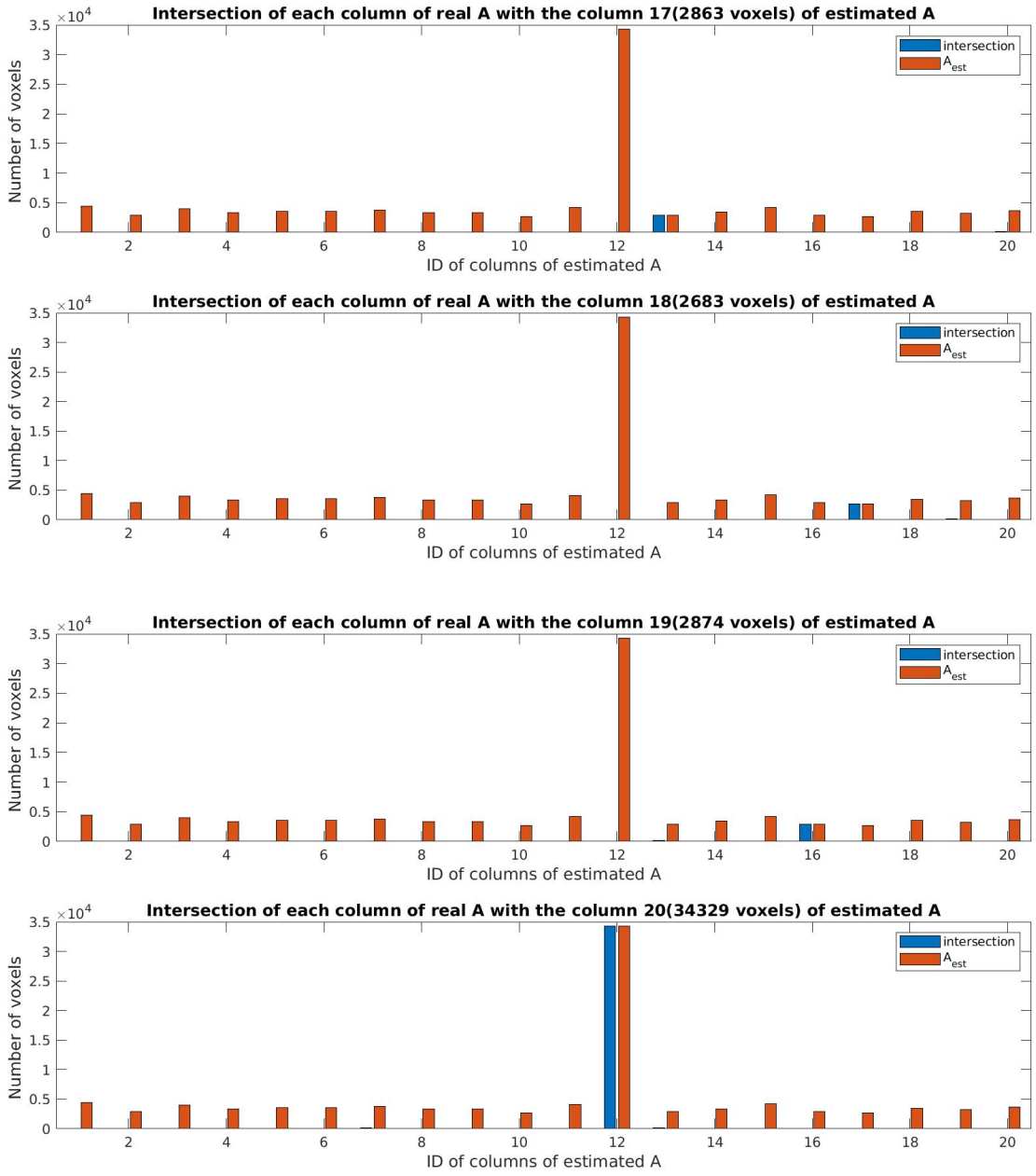


Figure A.76: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to penalty method and noiseless case.

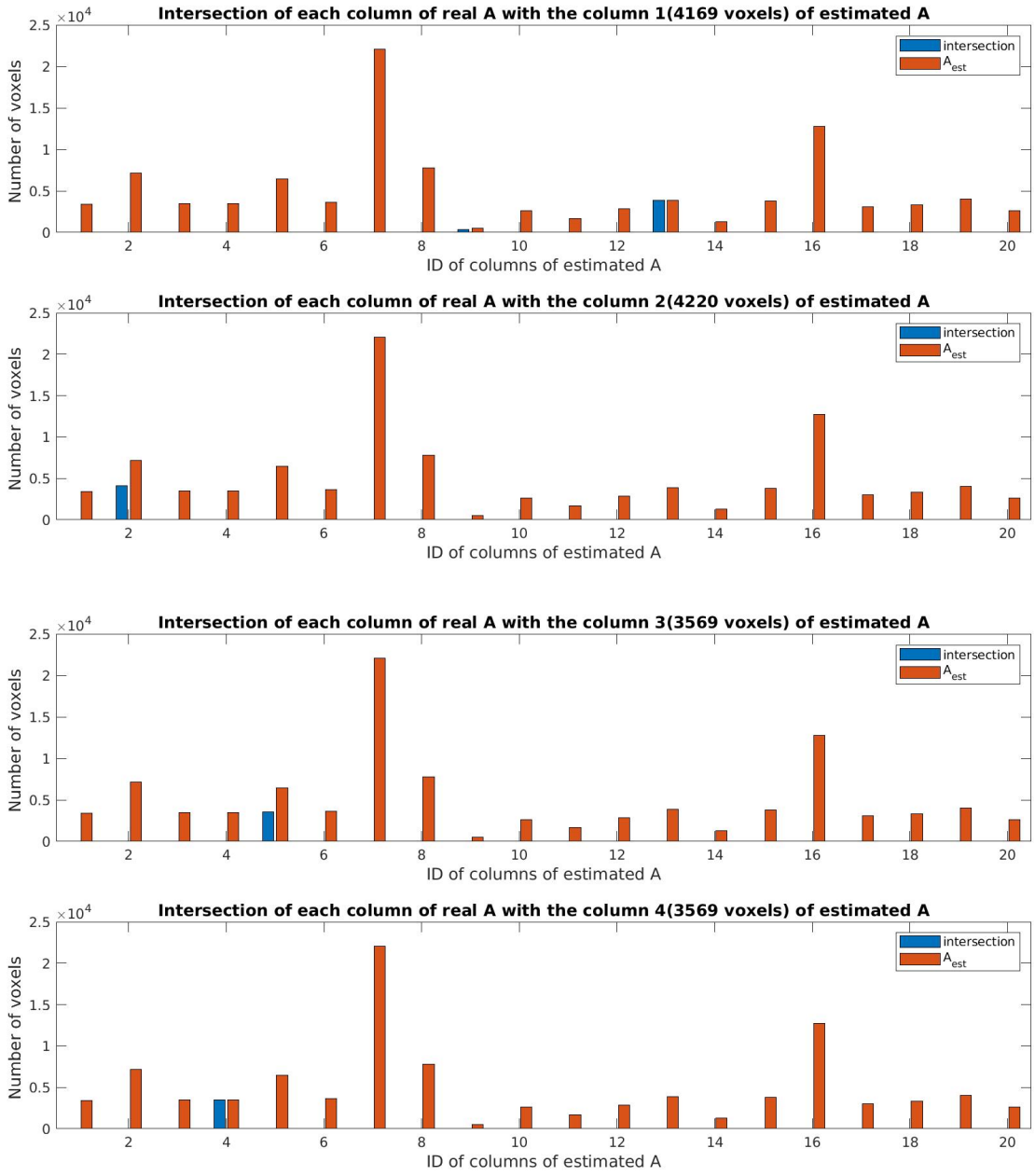


Figure A.77: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 2 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 2 (red bar). The solution 2 corresponds to penalty method and noiseless case.

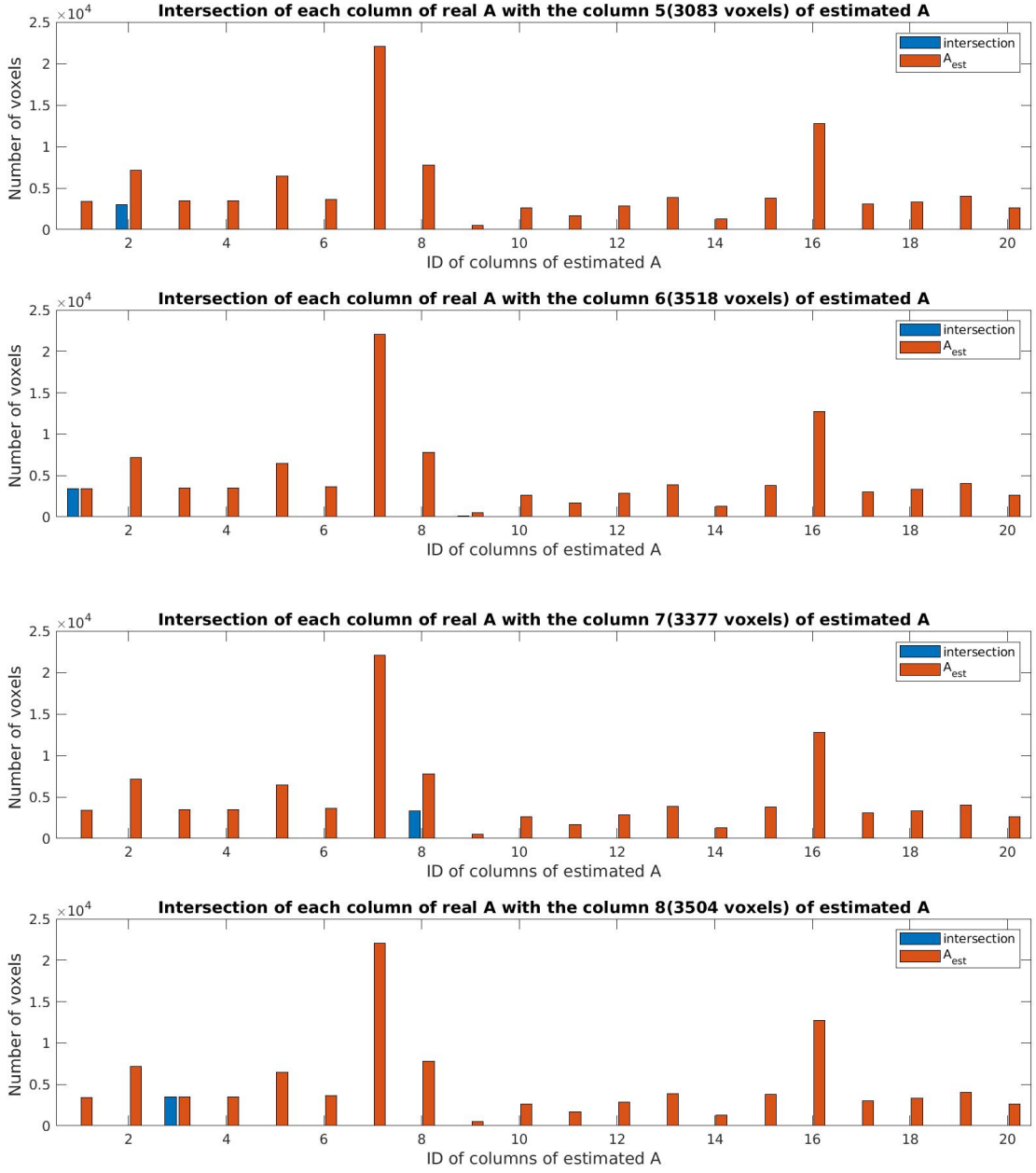


Figure A.78: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 2 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 2 (red bar). The solution 2 corresponds to penalty method and noiseless case.

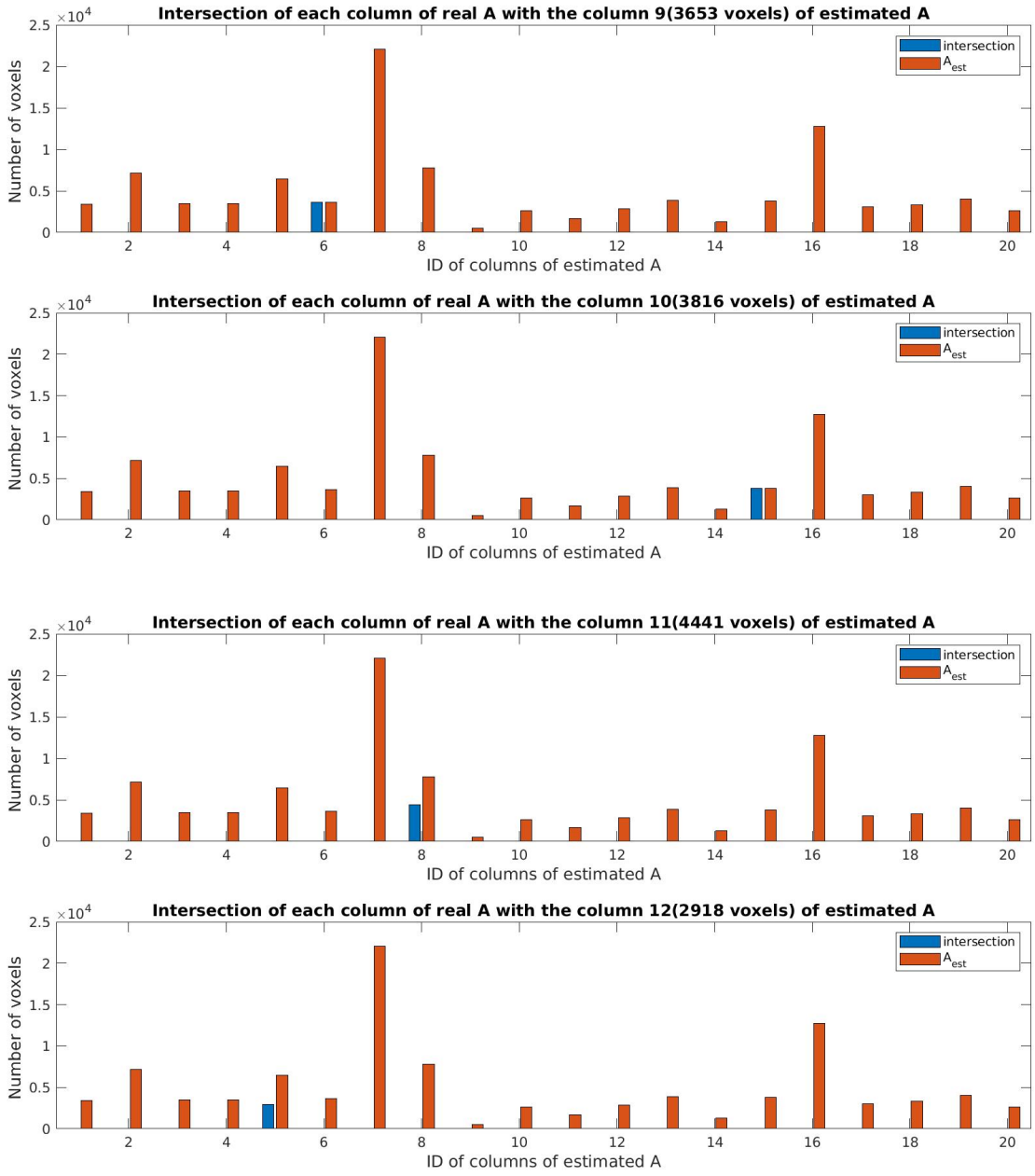


Figure A.79: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 2 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 2 (red bar). The solution 2 corresponds to penalty method and noiseless case.

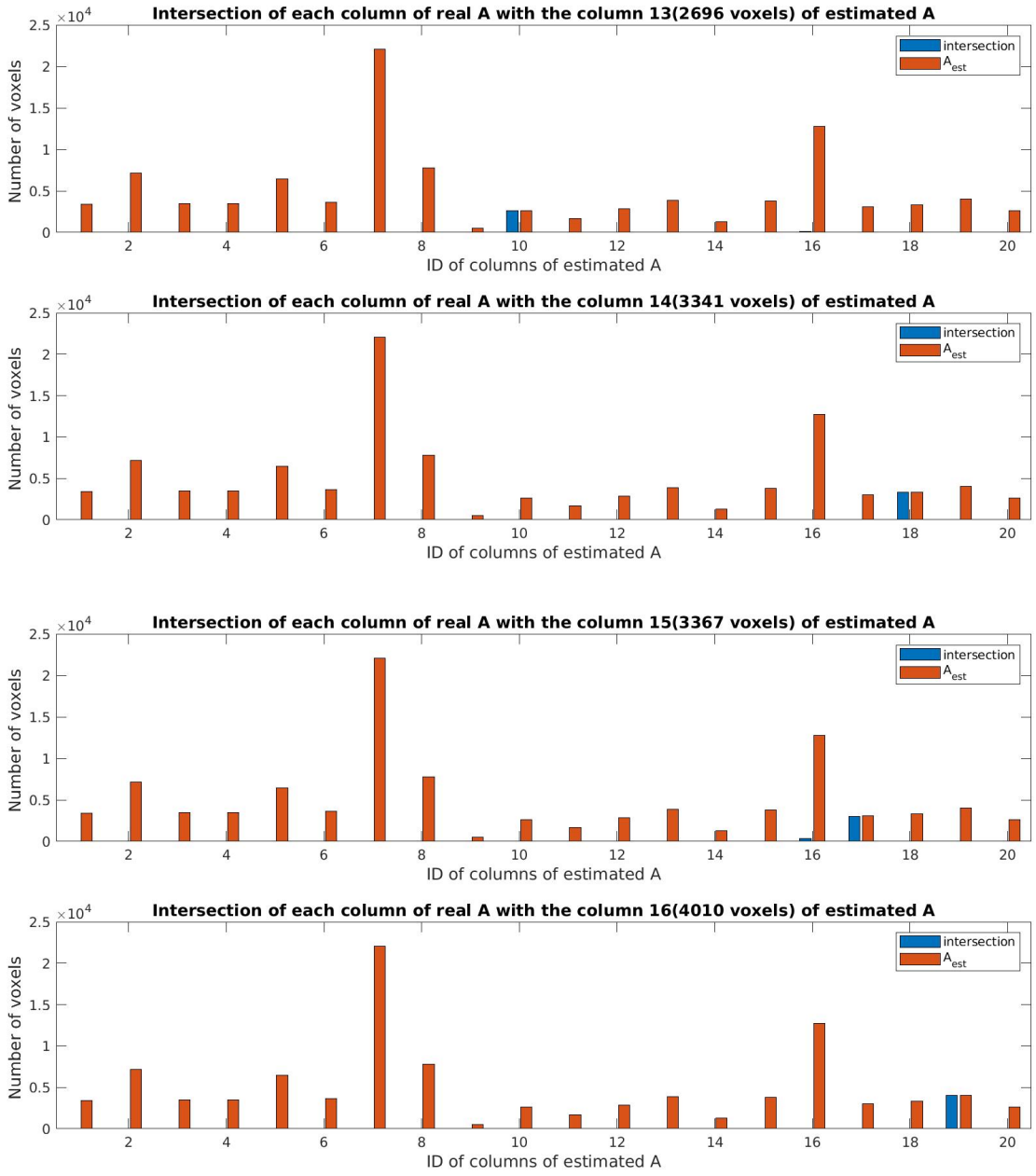


Figure A.80: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 2 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 2 (red bar). The solution 2 corresponds to penalty method and noiseless case.

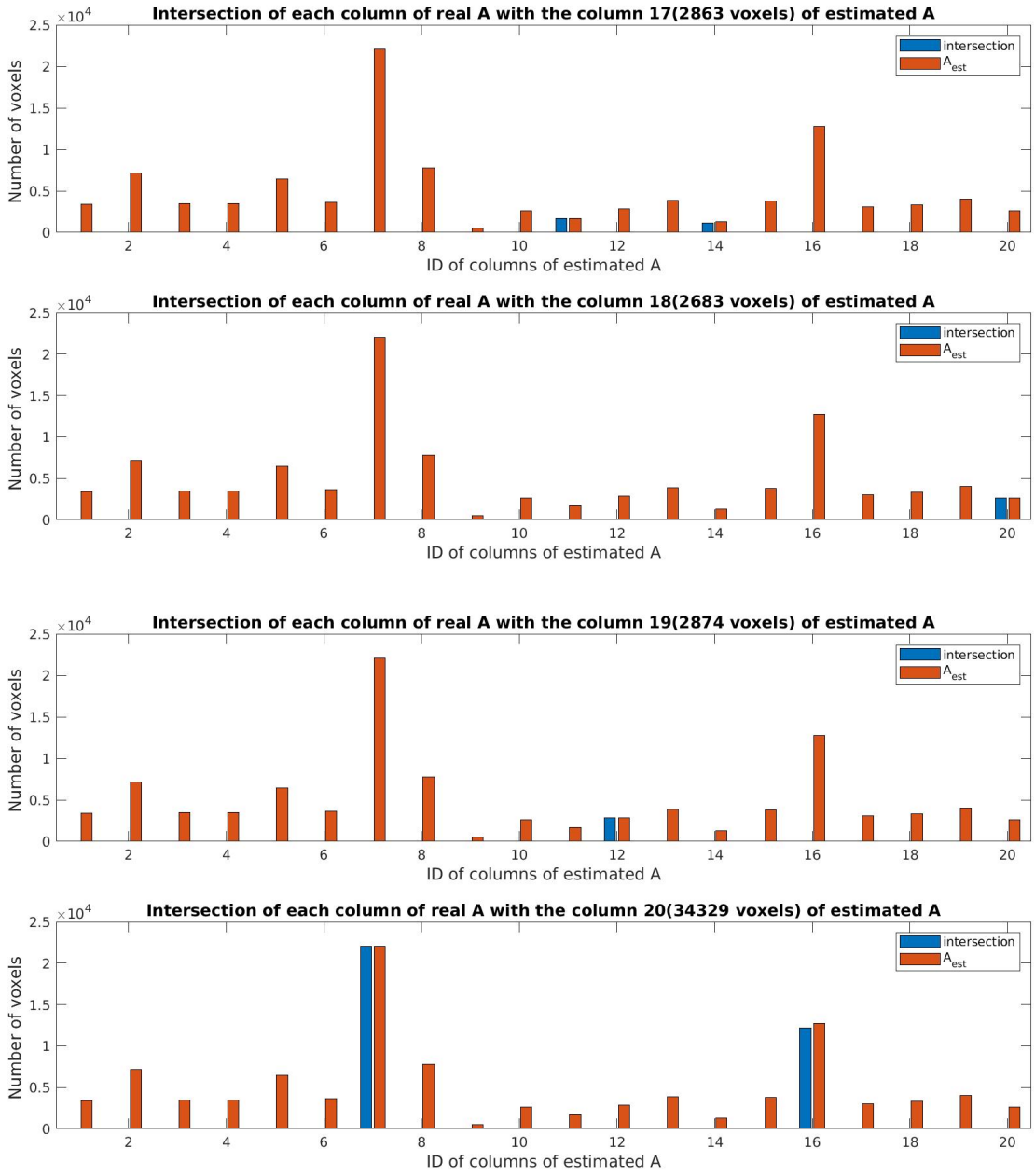


Figure A.81: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 2 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 2 (red bar). The solution 2 corresponds to penalty method and noiseless case.

· **SNR=0 dB**

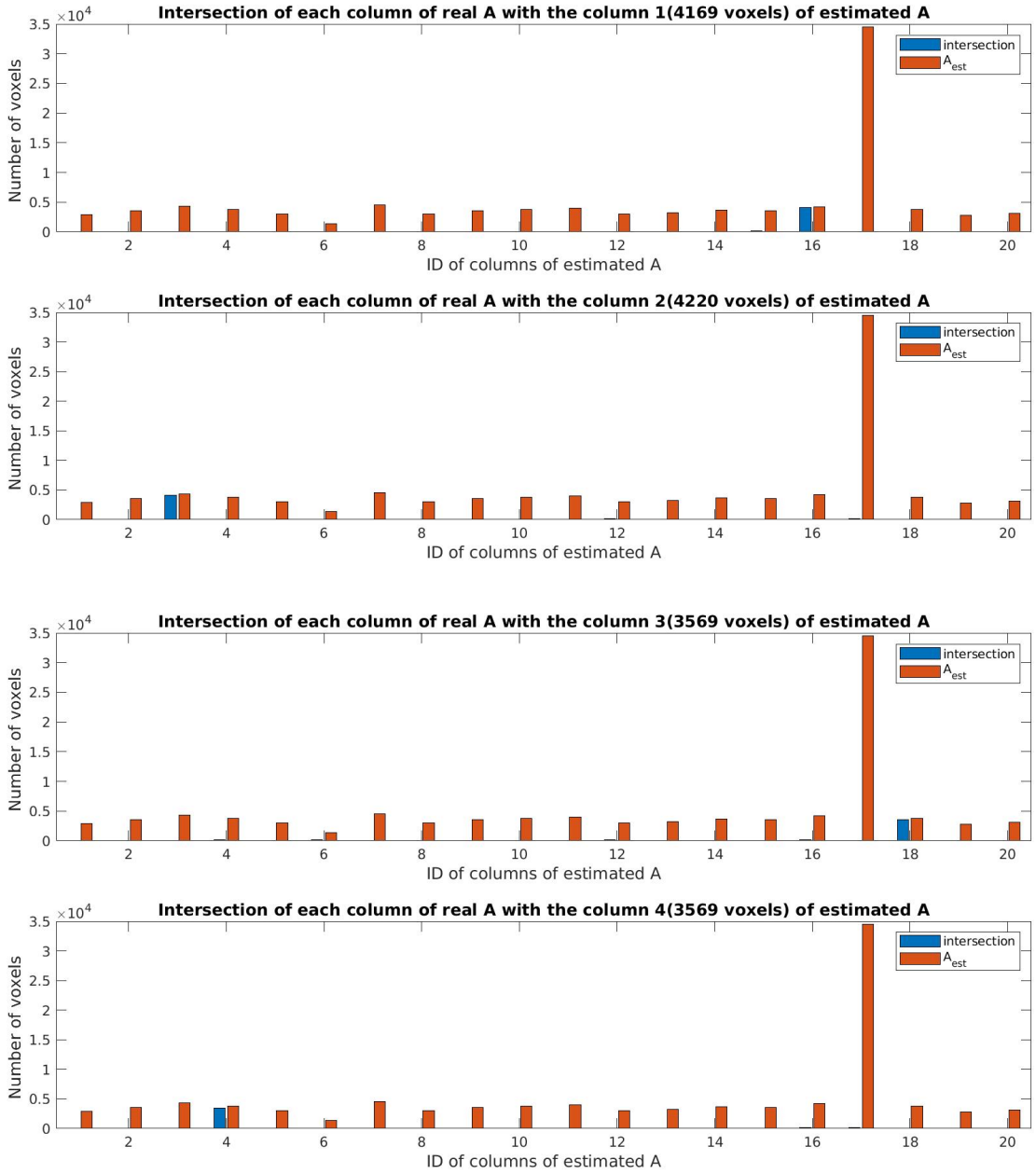


Figure A.82: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 3 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 3 (red bar). The solution 3 corresponds to penalty method and SNR=0 dB.

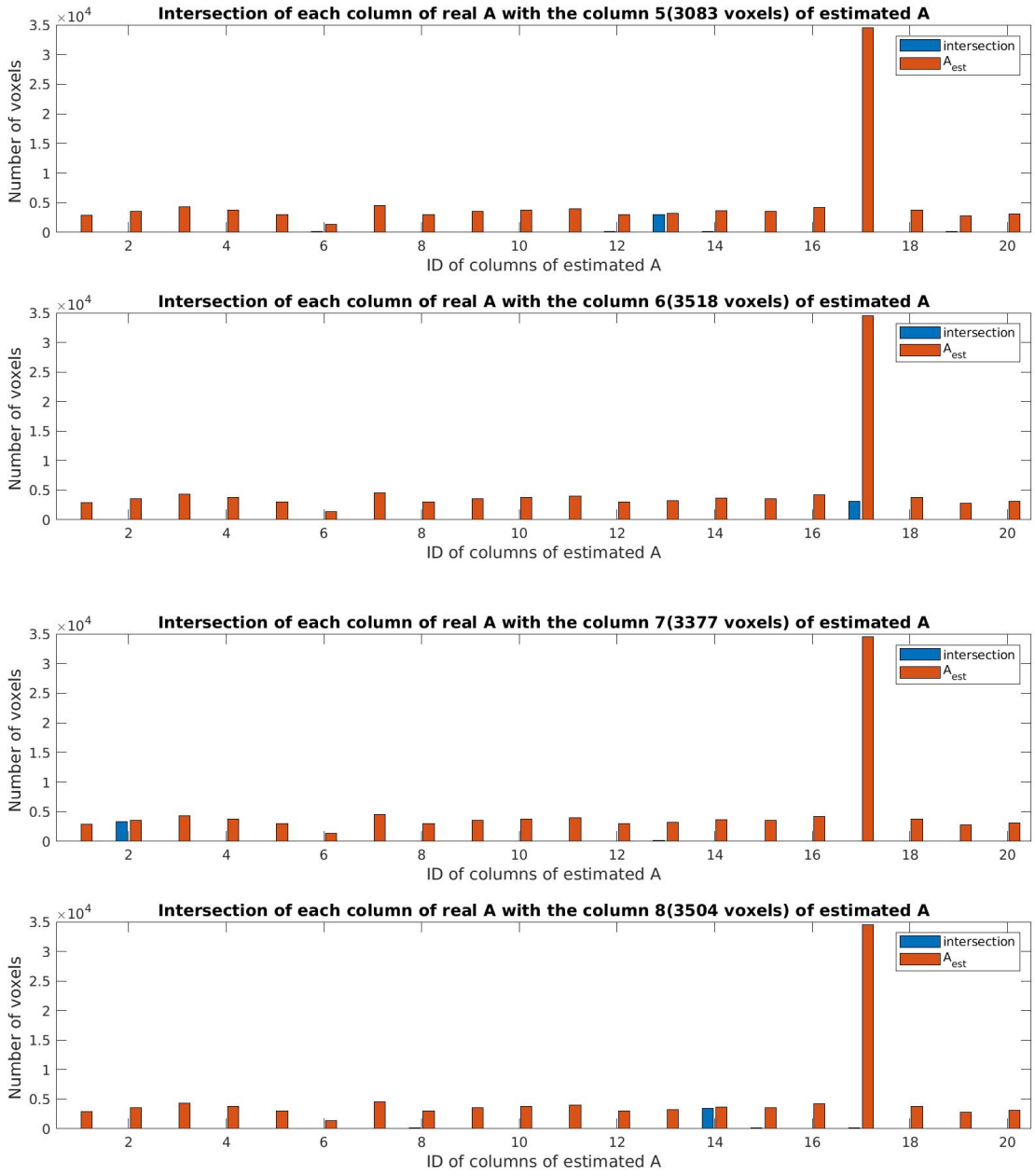


Figure A.83: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 3 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 3 (red bar). The solution 3 corresponds to penalty method and SNR=0 dB.

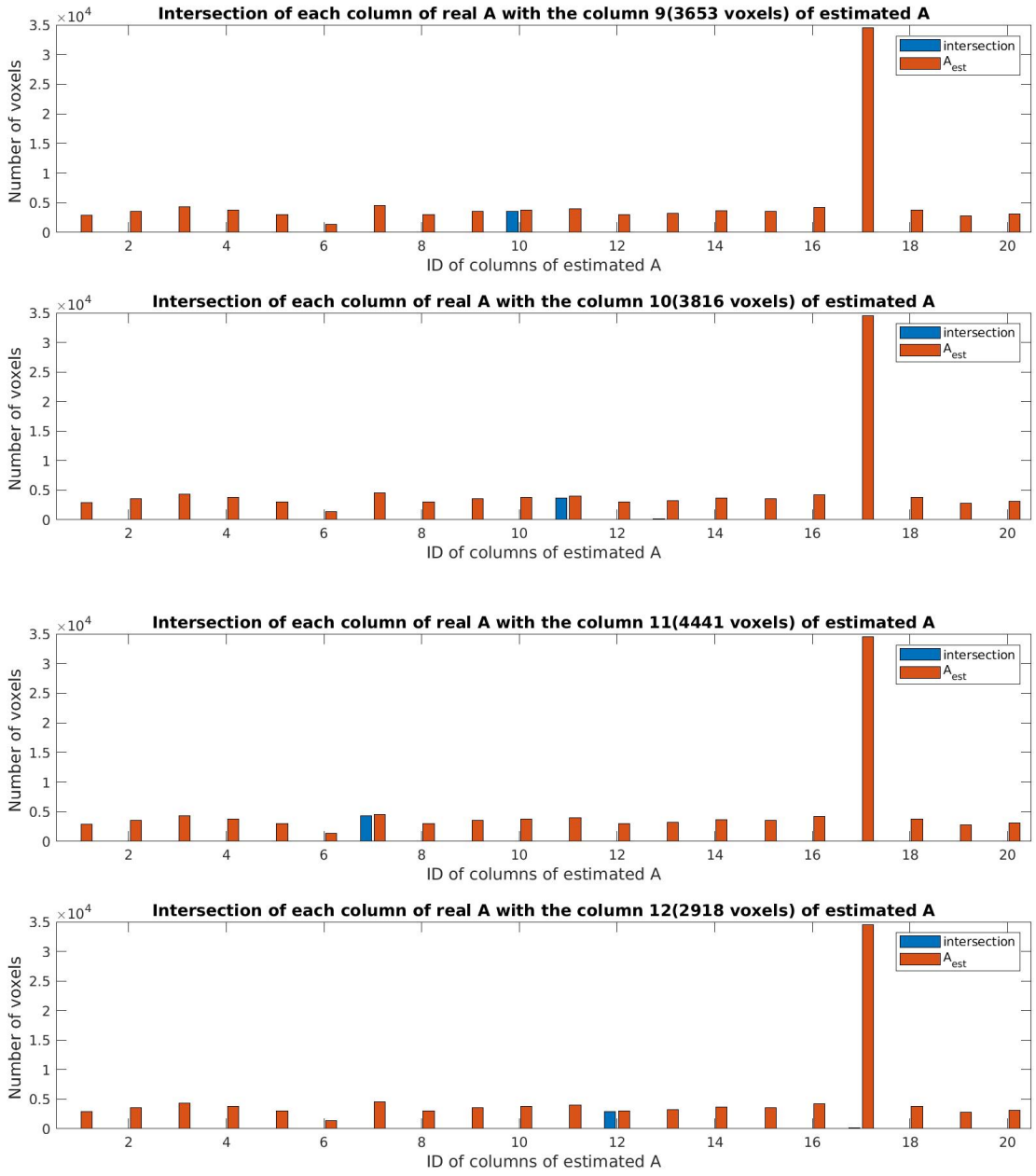


Figure A.84: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 3 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 3 (red bar). The solution 3 corresponds to penalty method and SNR=0 dB.

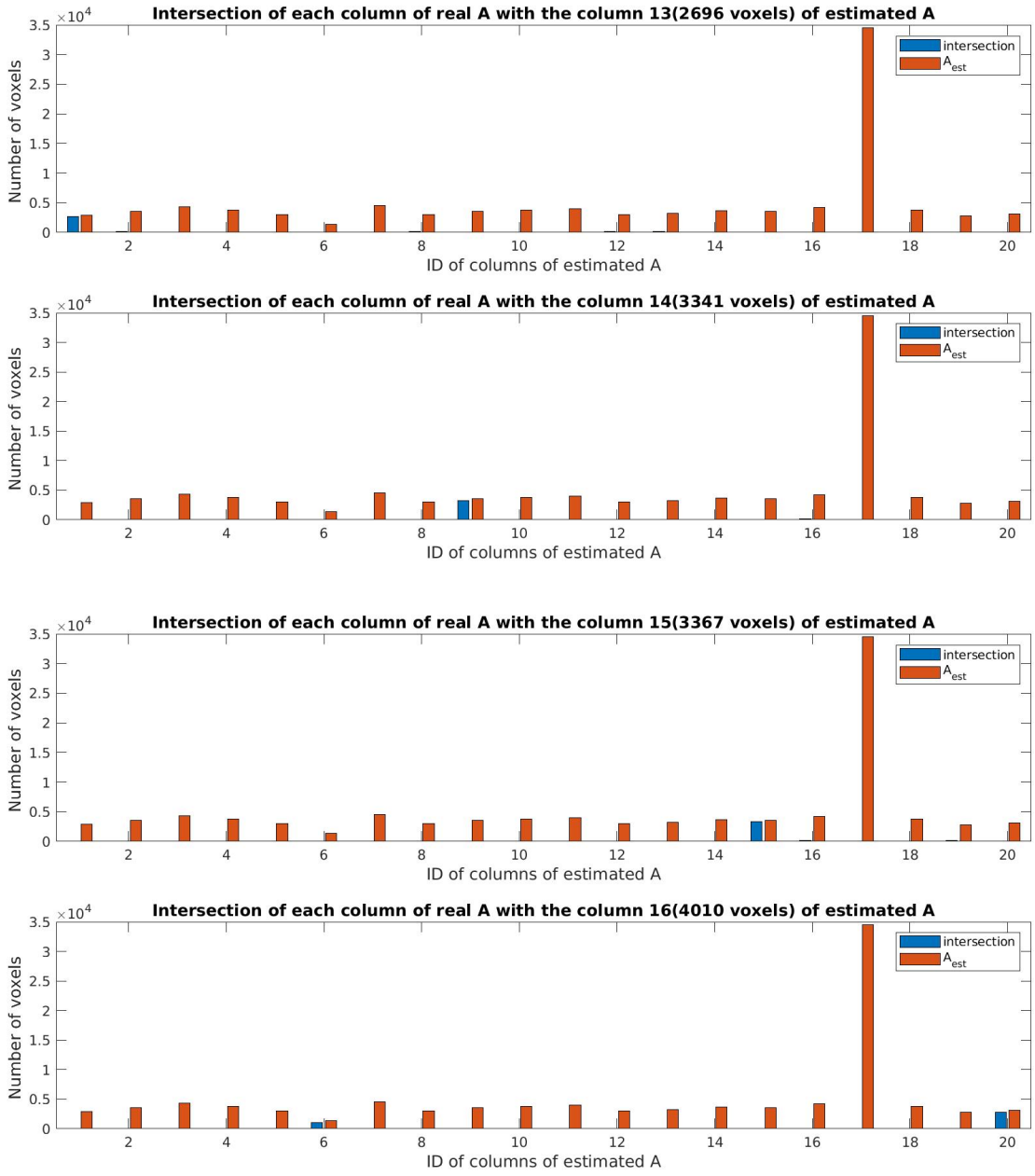


Figure A.85: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 3 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 3 (red bar). The solution 3 corresponds to penalty method and SNR=0 dB.

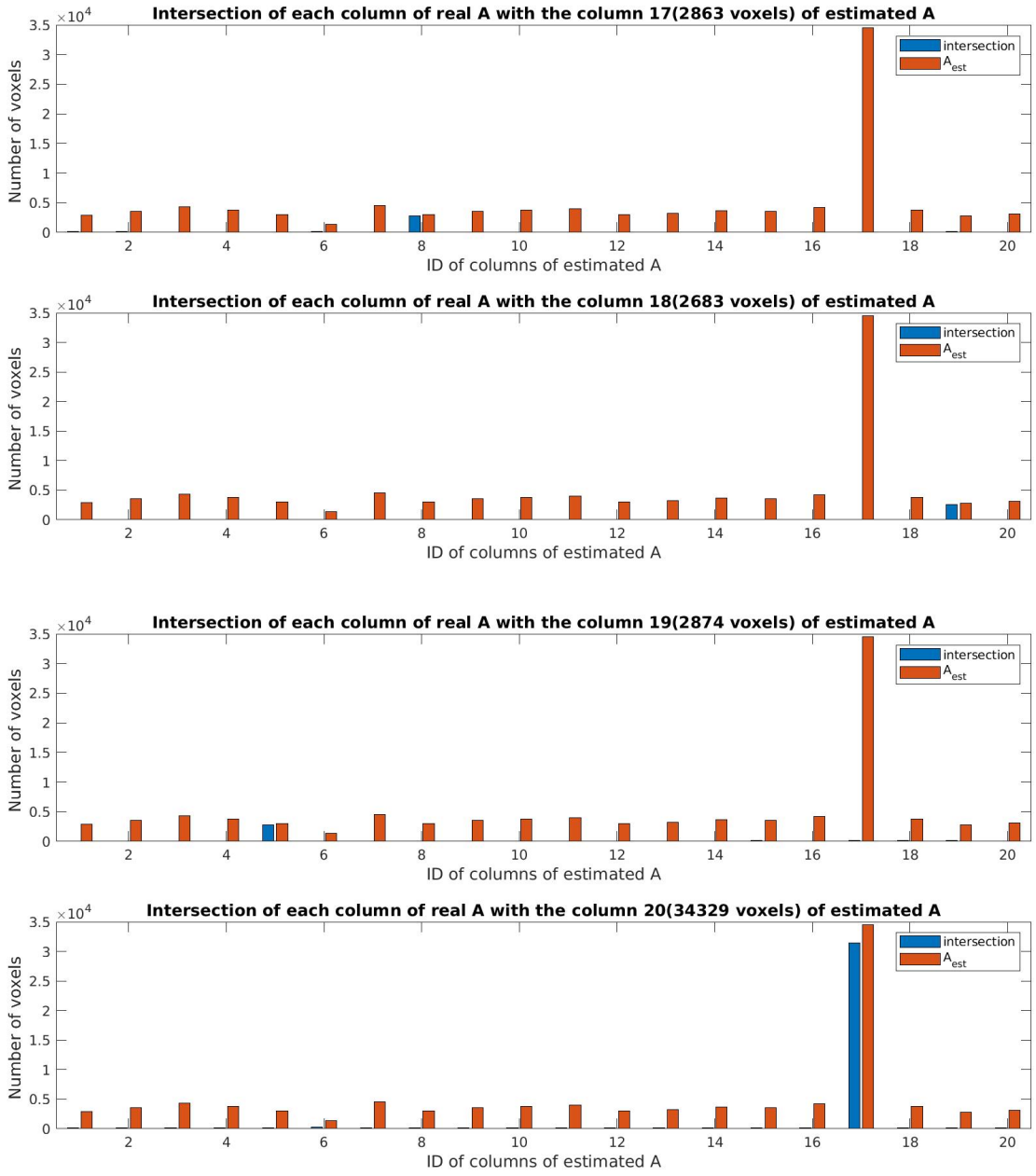


Figure A.86: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 3 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 3 (red bar). The solution 3 corresponds to penalty method and SNR=0 dB.

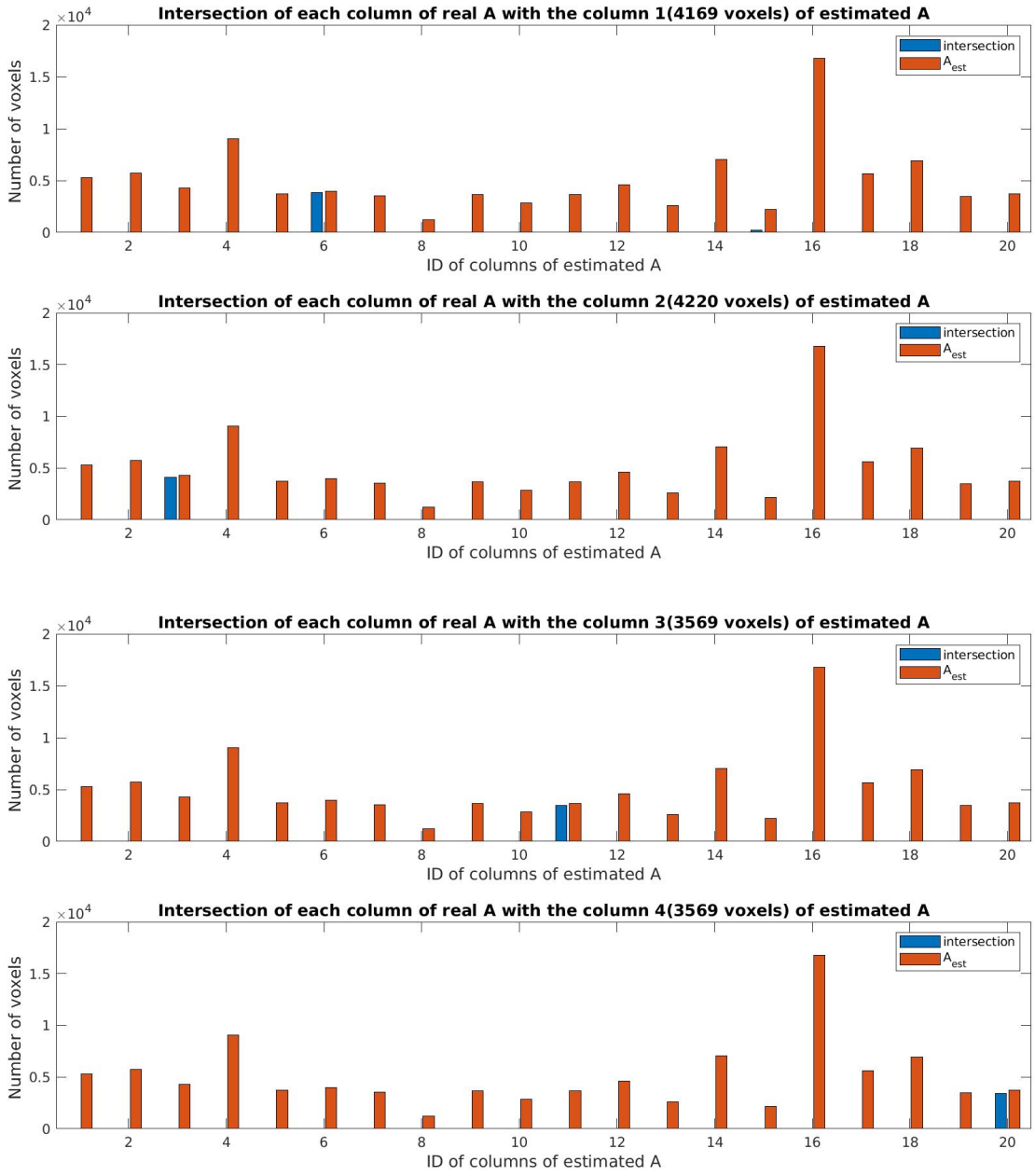


Figure A.87: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 4 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 4 (red bar). The solution 4 corresponds to penalty method and SNR=0 dB.

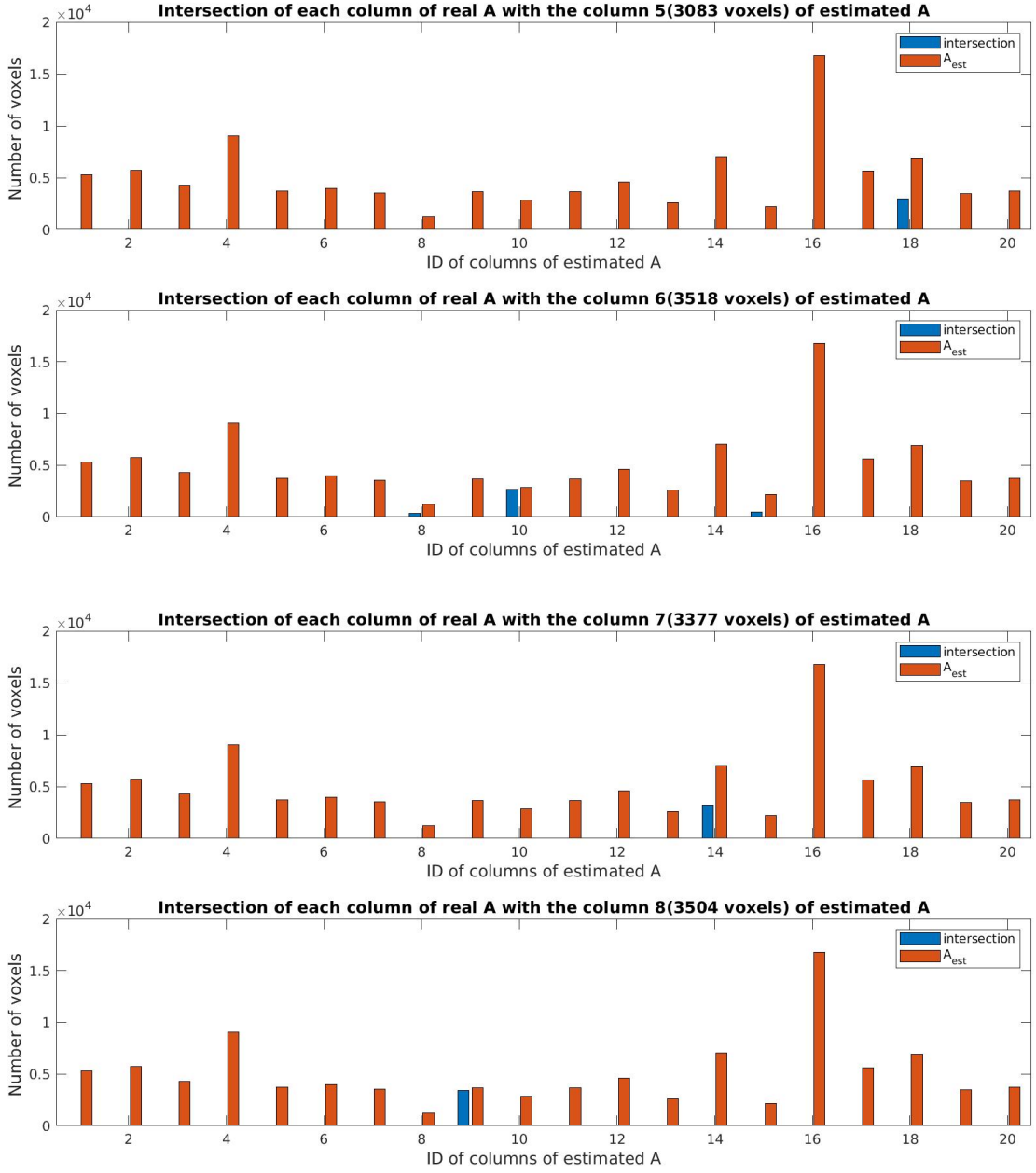


Figure A.88: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 4 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 4 (red bar). The solution 4 corresponds to penalty method and SNR=0 dB.

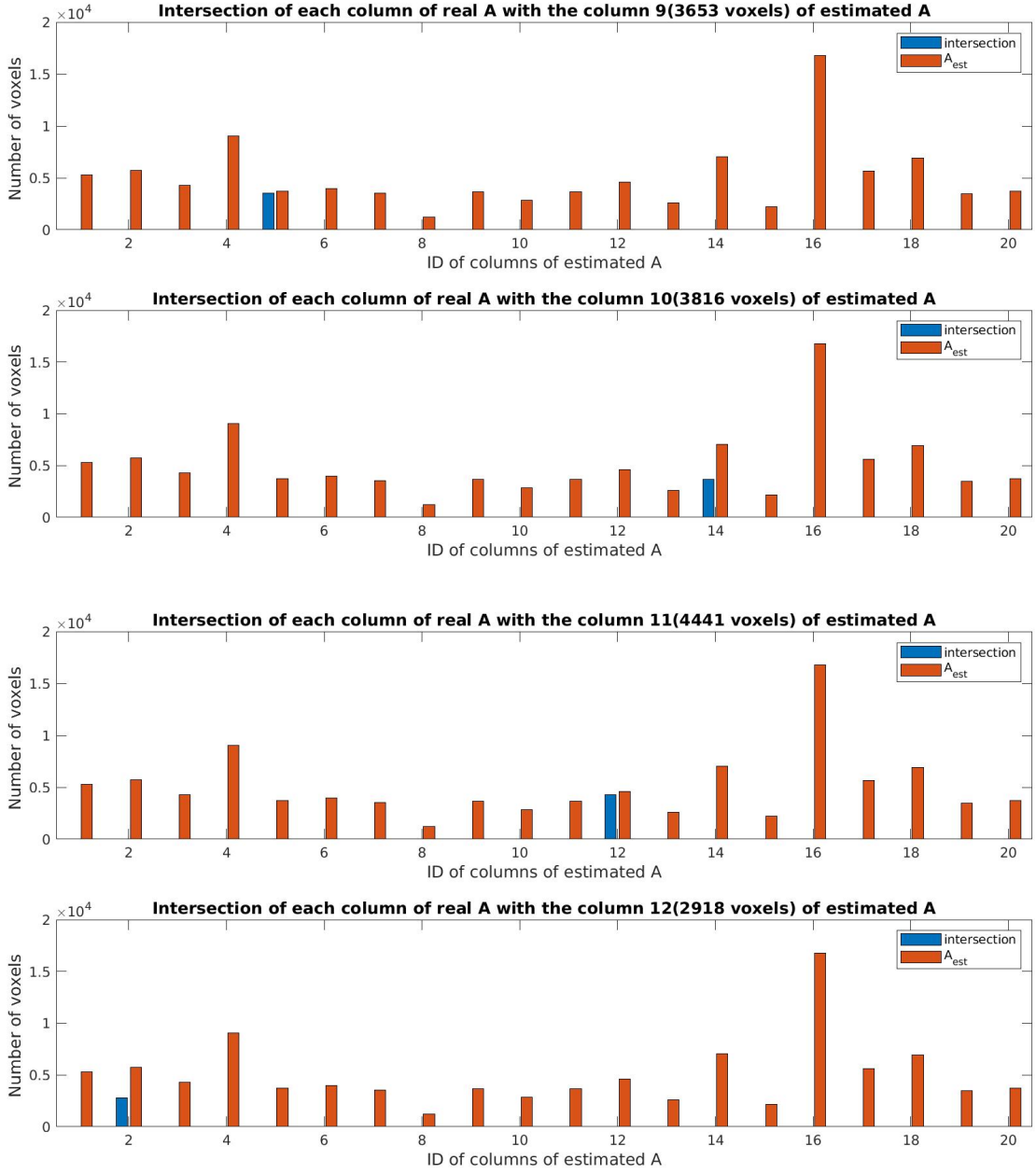


Figure A.89: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 4 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 4 (red bar). The solution 4 corresponds to penalty method and SNR=0 dB.

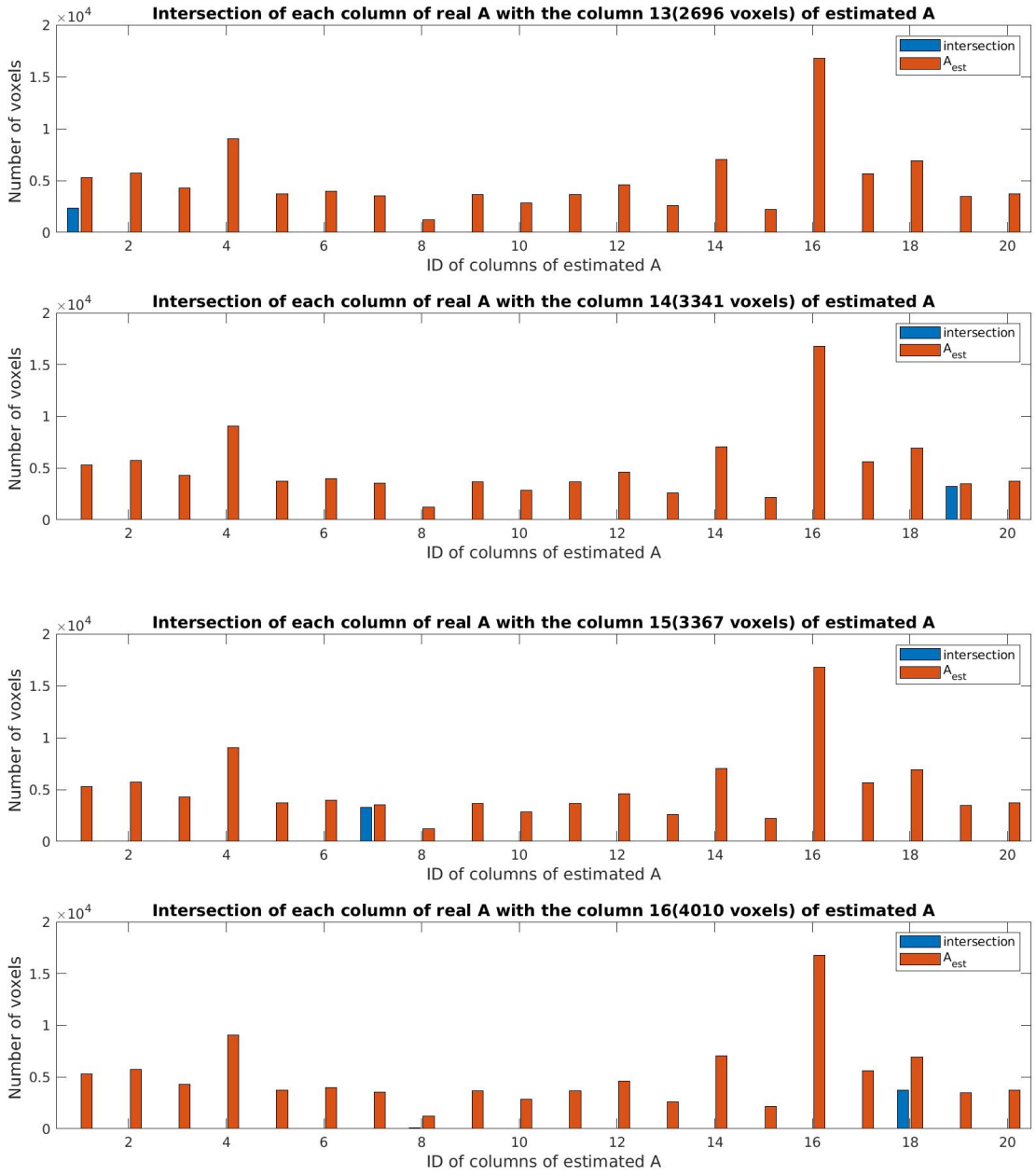


Figure A.90: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 4 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 4 (red bar). The solution 4 corresponds to penalty method and SNR=0 dB.

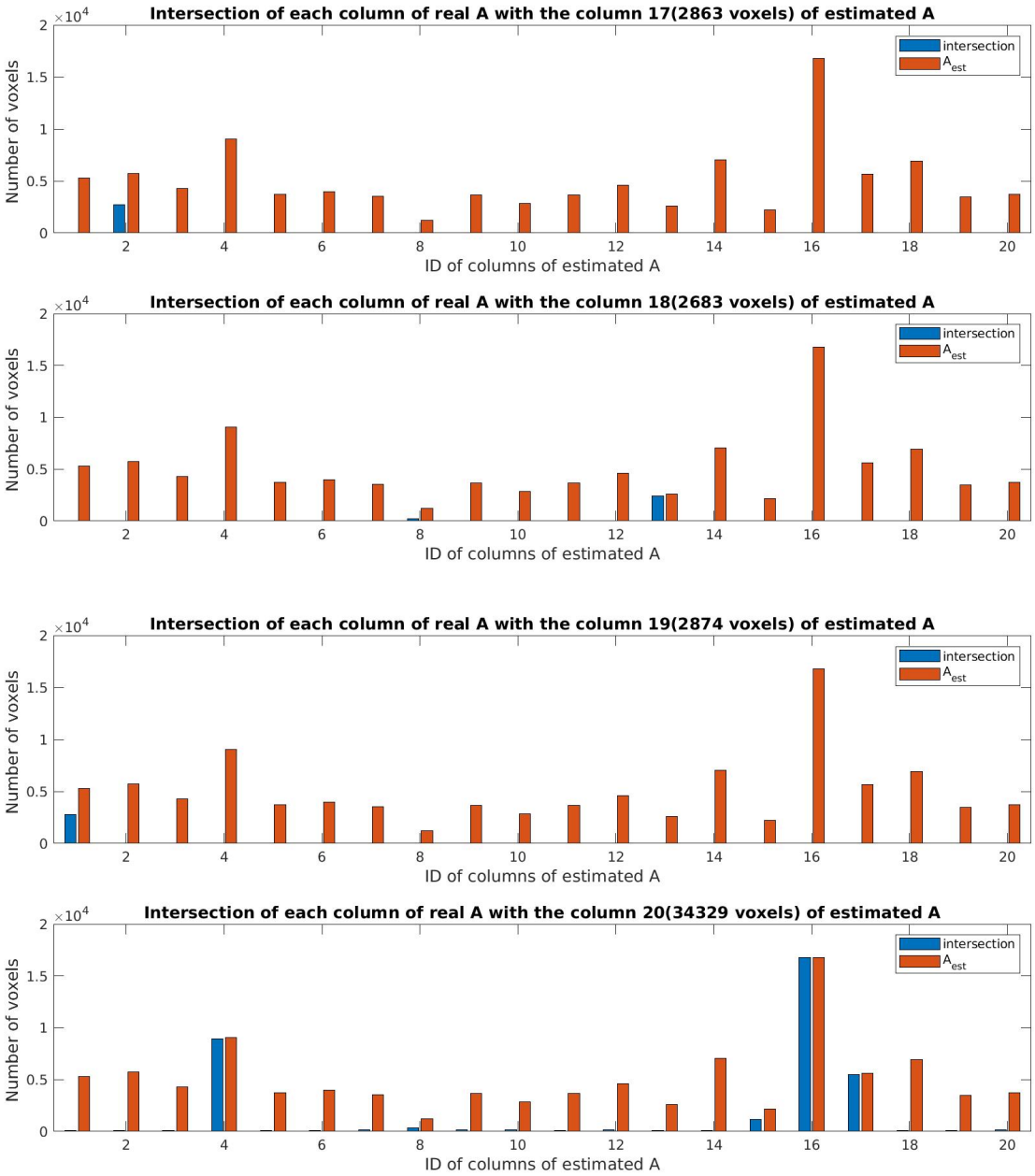


Figure A.91: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 4 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 4 (red bar). The solution 4 corresponds to penalty method and SNR=0 dB.

Accelerated penalty method

· Comparison of Accelerated Penalty method solutions

· Noiseless case

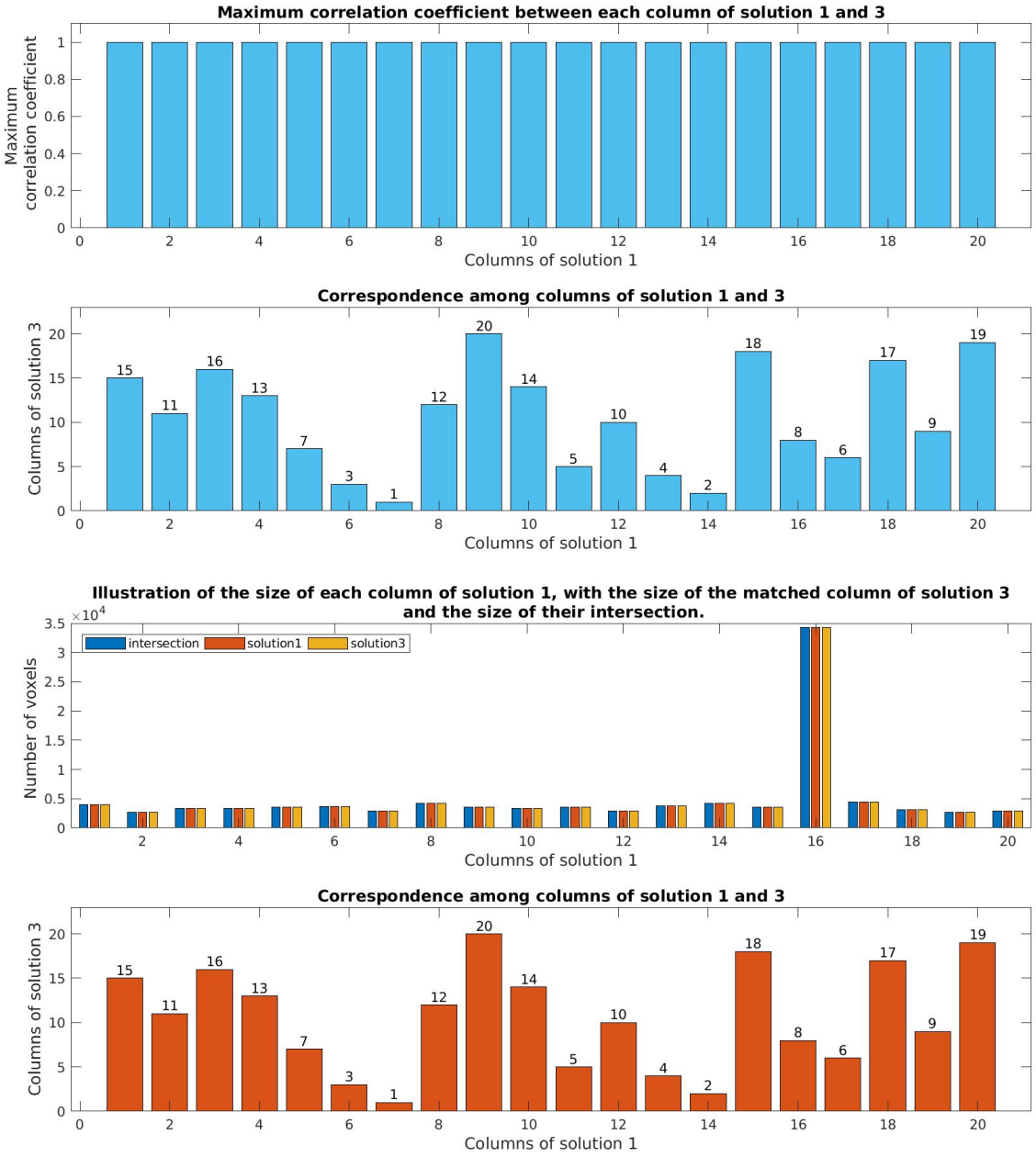


Figure A.92: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on intersection. Solutions 1, 3 correspond to accelerated penalty method and noiseless data.

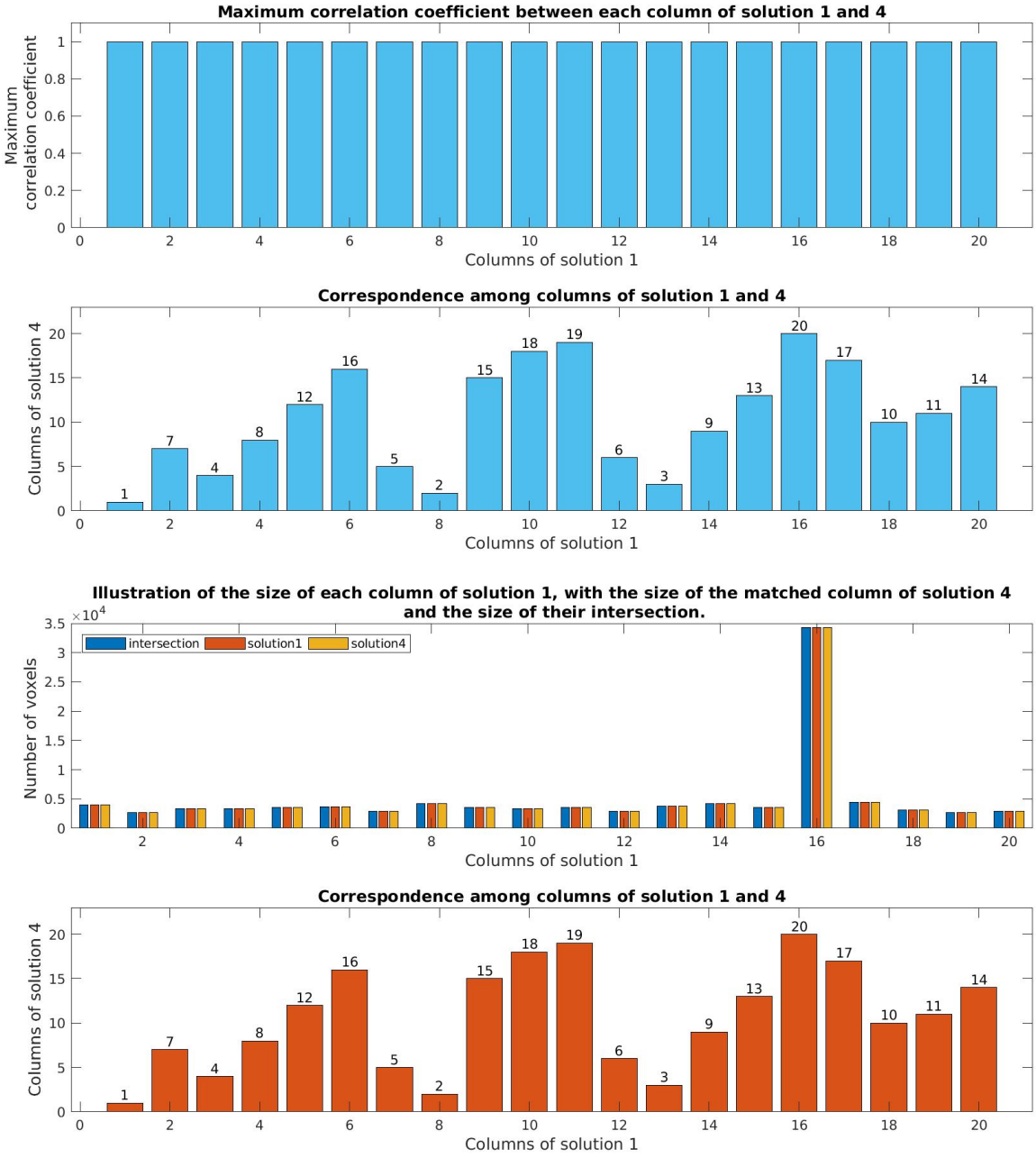


Figure A.93: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on intersection. Solutions 1, 4 correspond to accelerated penalty method and noiseless data.

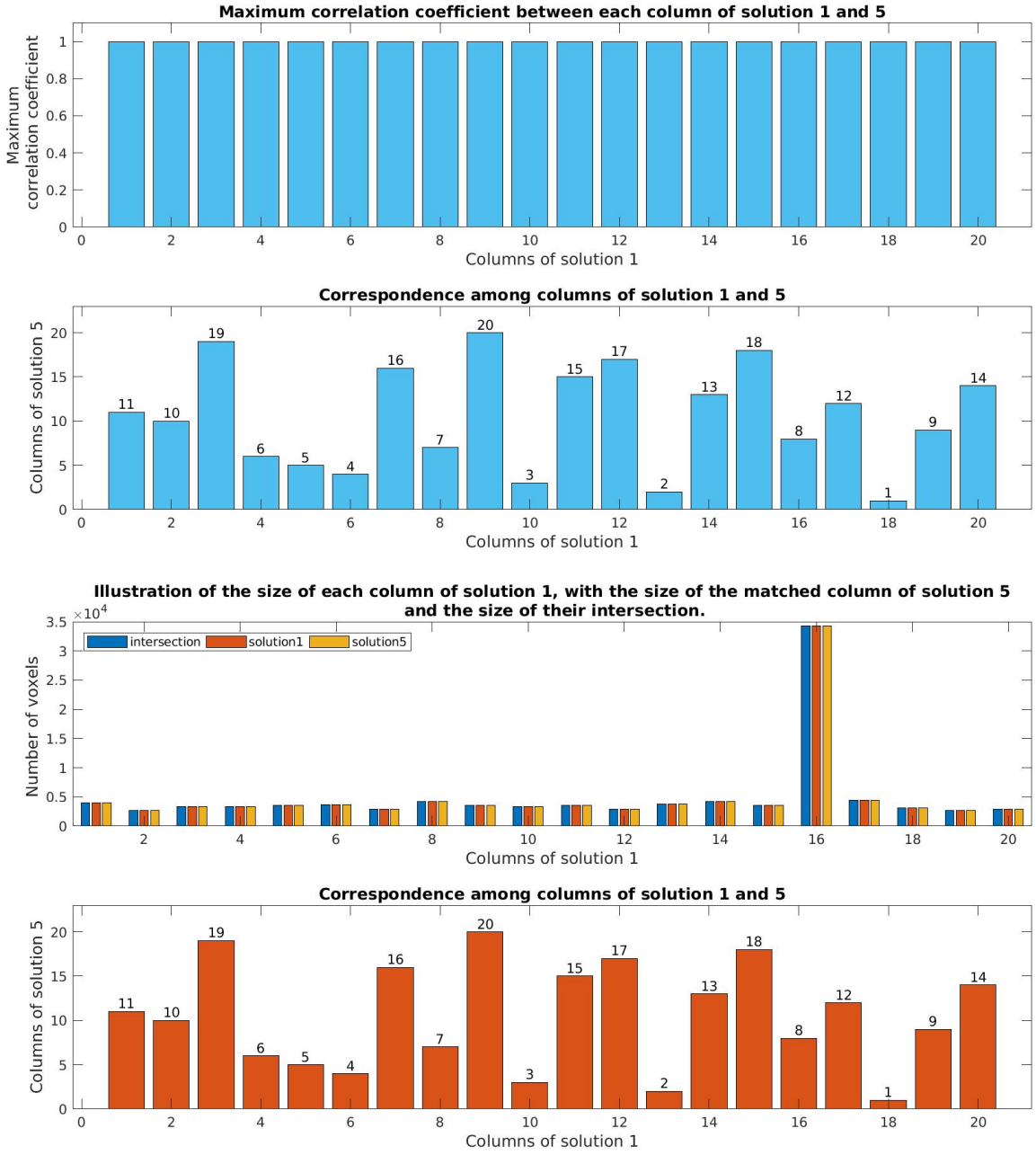


Figure A.94: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 5. We associate the column r of the solution 1 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on intersection. Solutions 1, 5 correspond to accelerated penalty method and noiseless data.

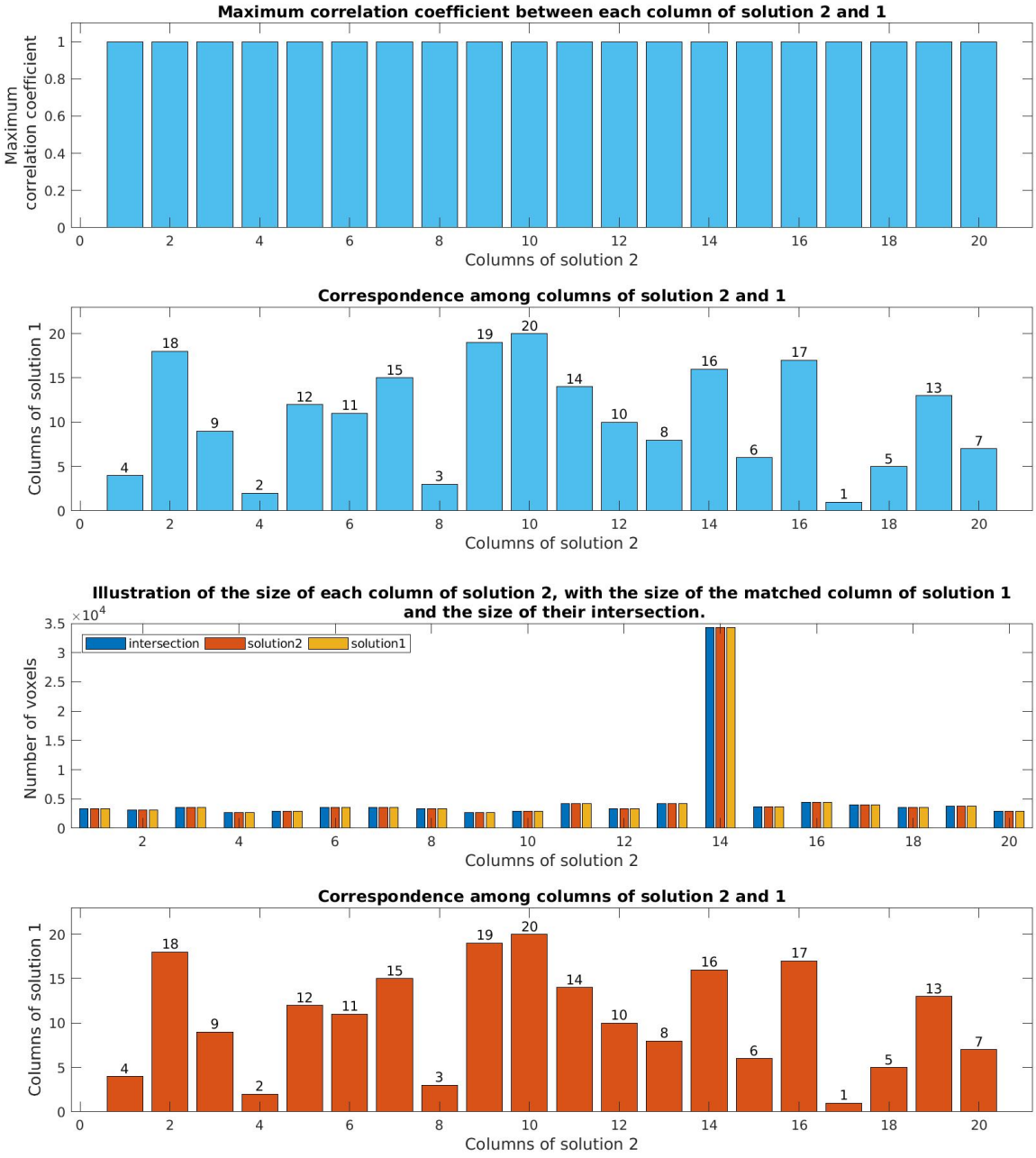


Figure A.95: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on intersection. Solutions 2, 1 correspond to accelerated penalty method and noiseless data.

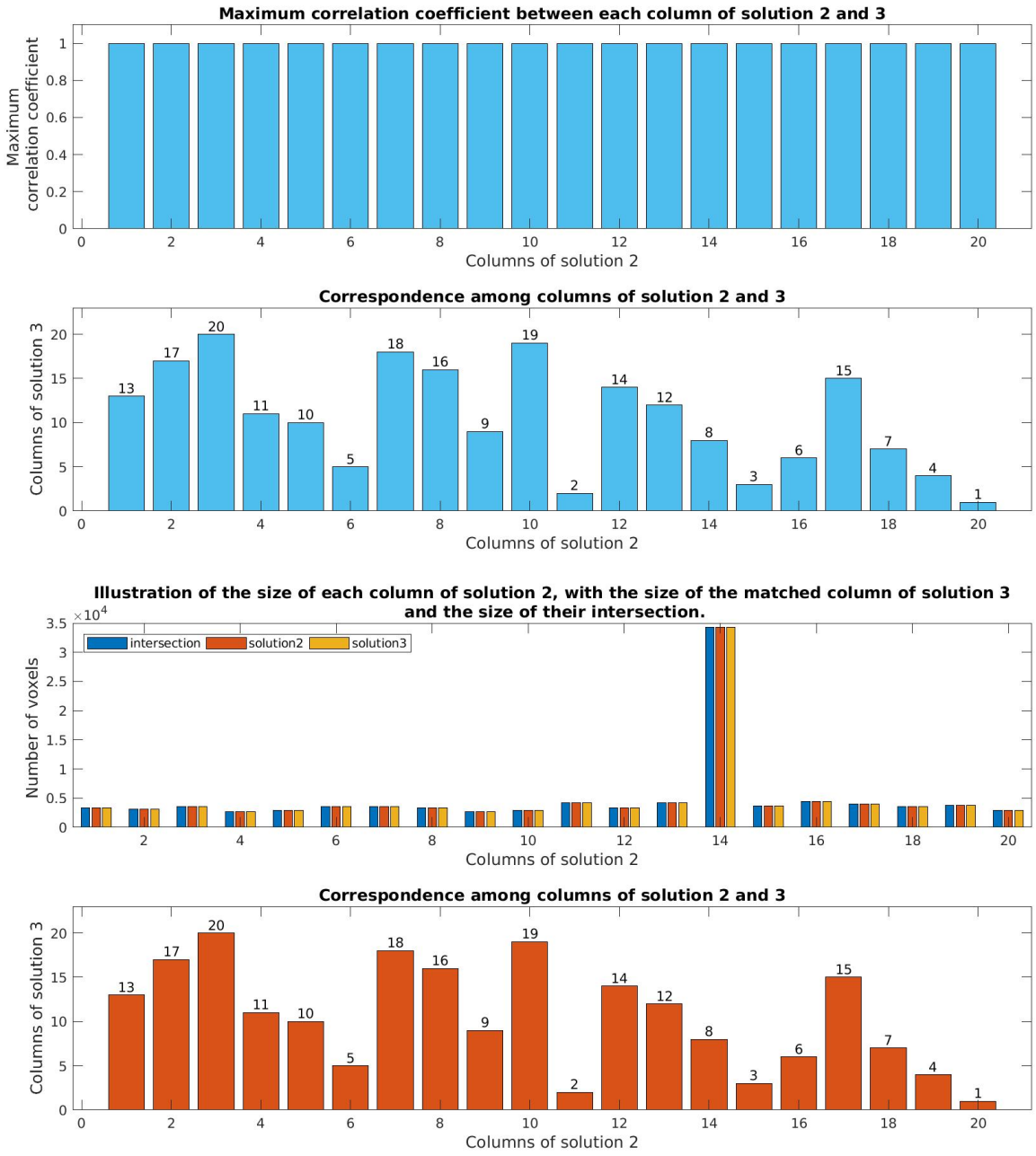


Figure A.96: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on intersection. Solutions 2, 3 correspond to accelerated penalty method and noiseless data.

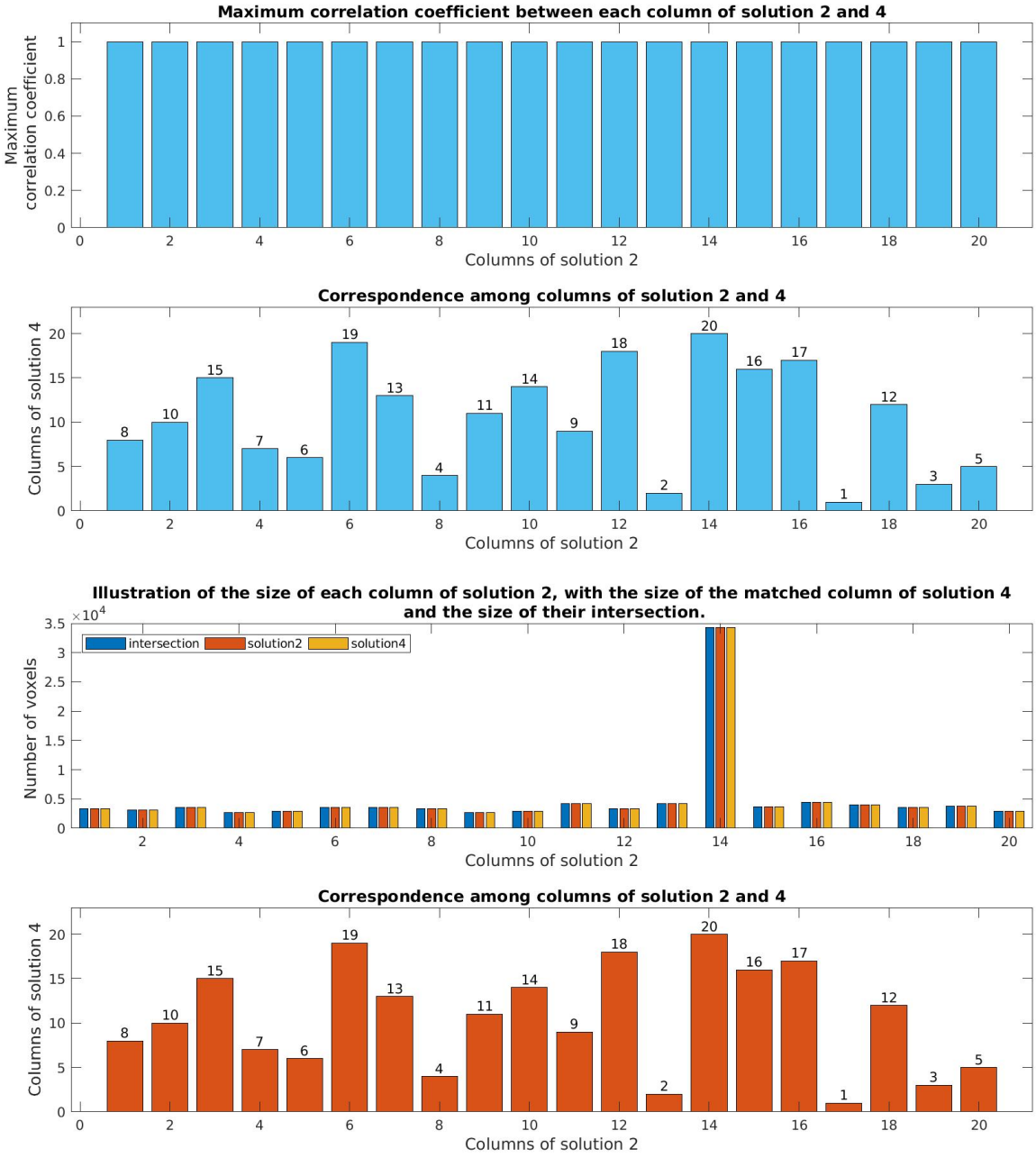


Figure A.97: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on intersection. Solutions 2, 4 correspond to accelerated penalty method and noiseless data.

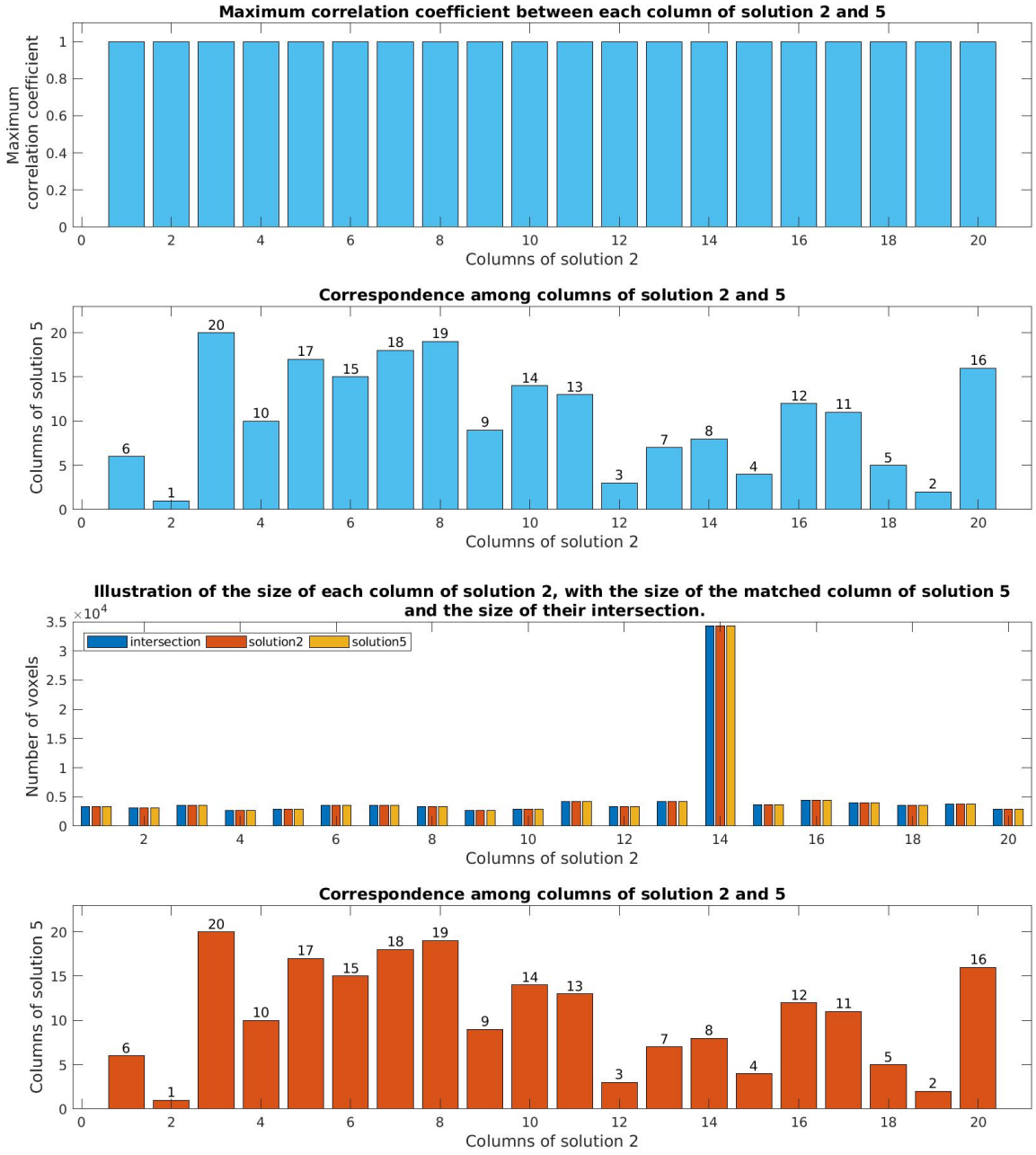


Figure A.98: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 5. We associate the column r of the solution 2 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on intersection. Solutions 2, 5 correspond to accelerated penalty method and noiseless data.

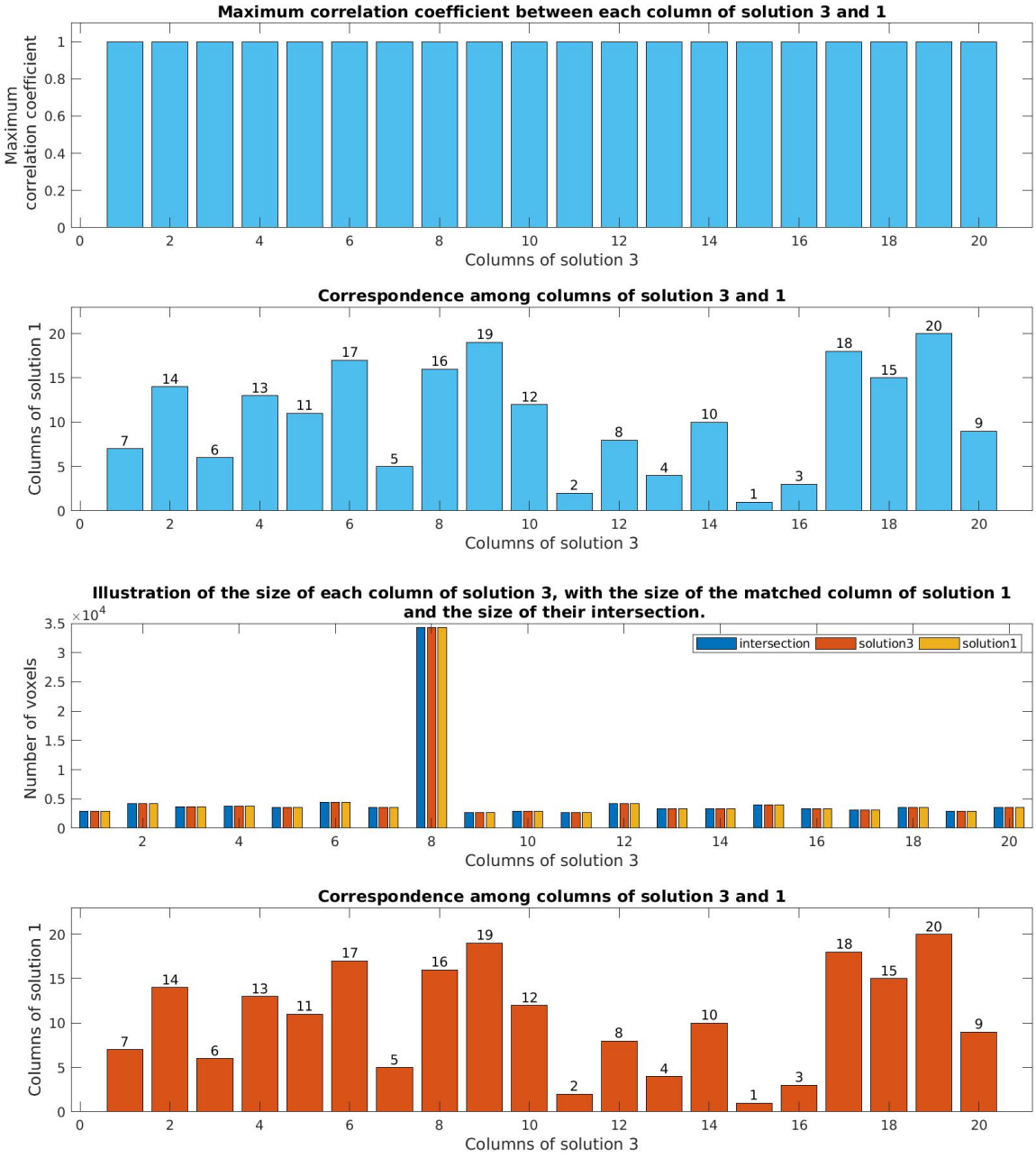


Figure A.99: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on intersection. Solutions 3, 1 correspond to accelerated penalty method and noiseless data.

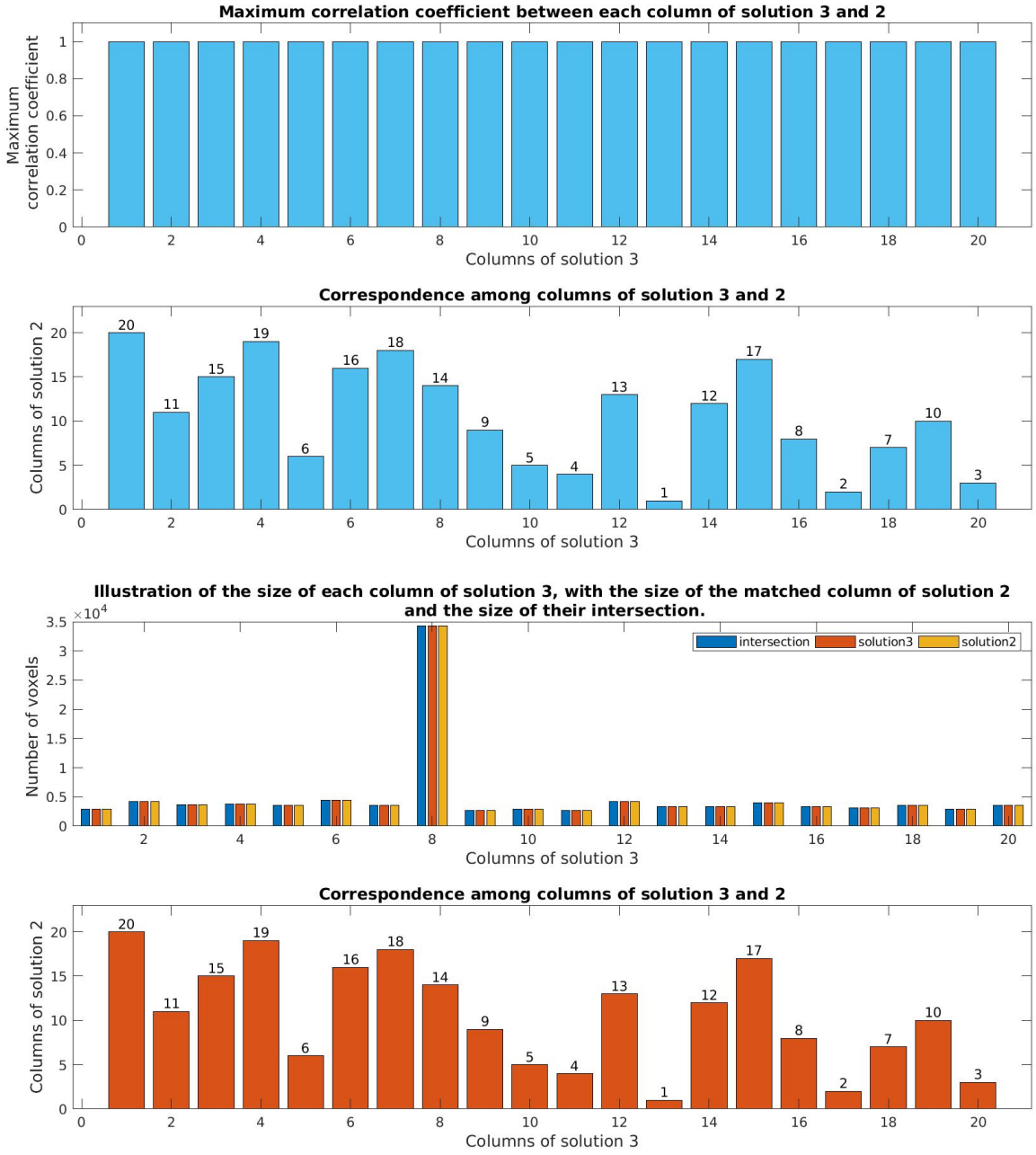


Figure A.100: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on intersection. Solutions 3, 2 correspond to accelerated penalty method and noiseless data.

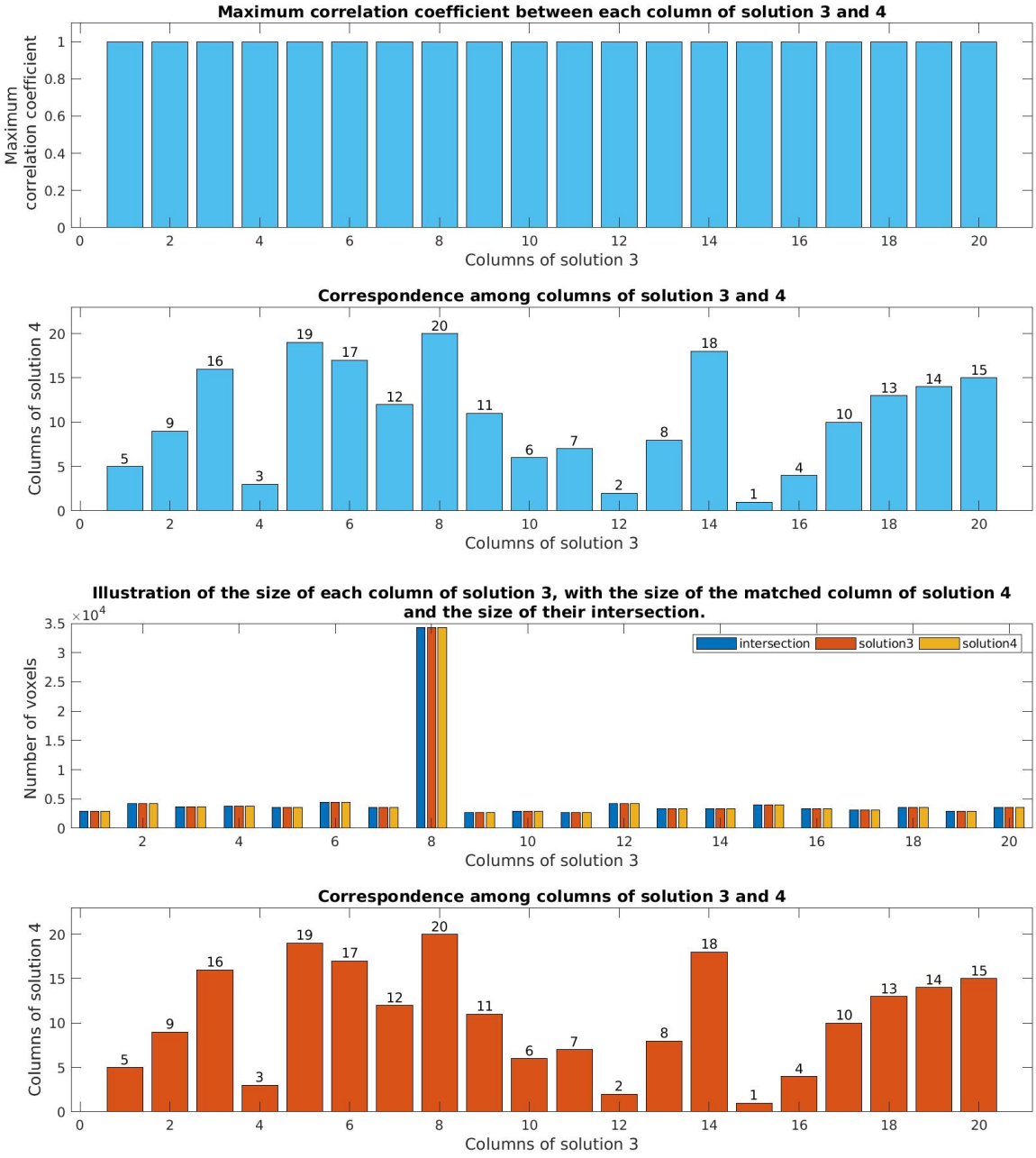


Figure A.101: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on intersection. Solutions 3, 4 correspond to accelerated penalty method and noiseless data.

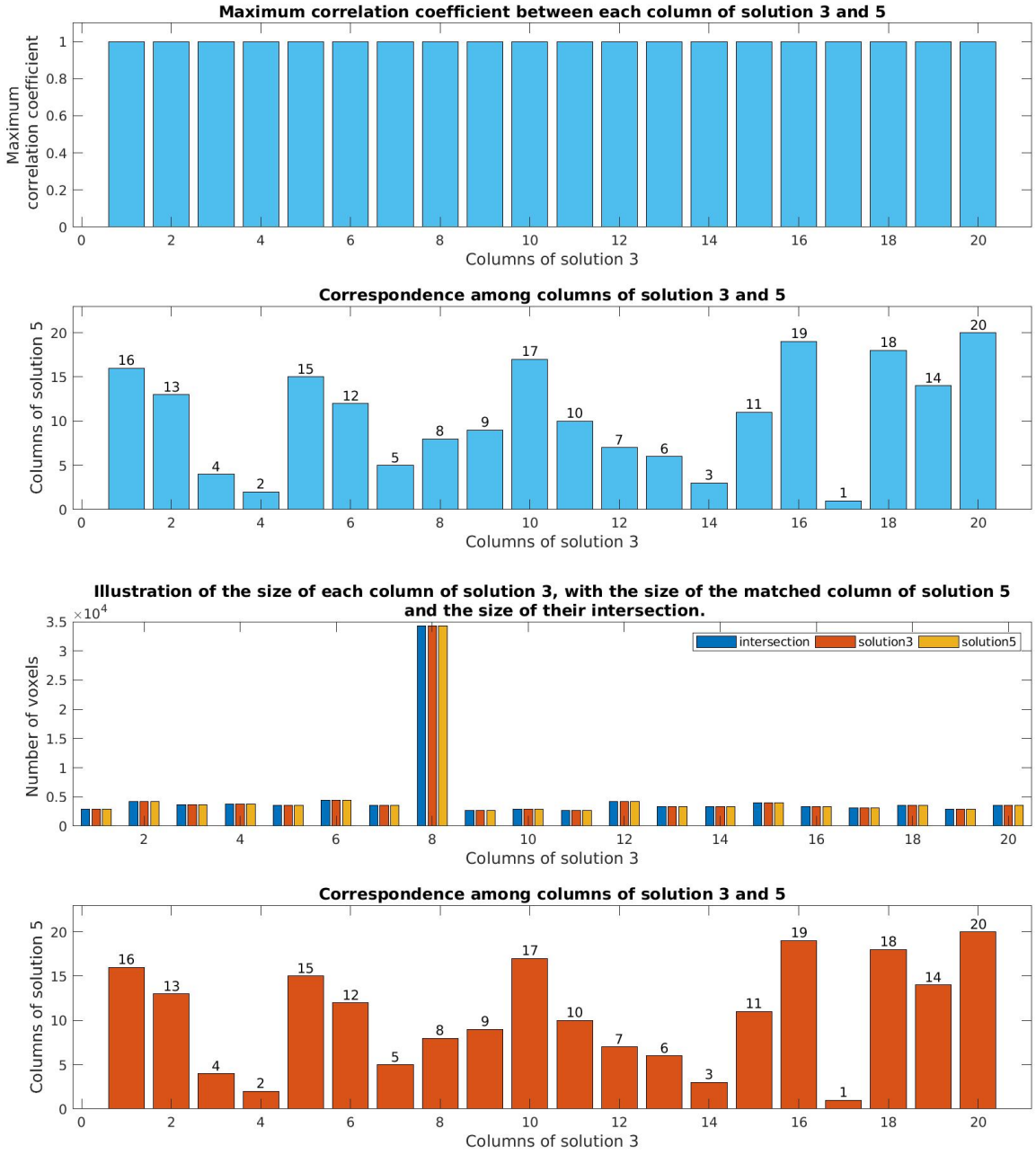


Figure A.102: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 5. We associate the column r of the solution 3 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on intersection. Solutions 3, 5 correspond to accelerated penalty method and noiseless data.

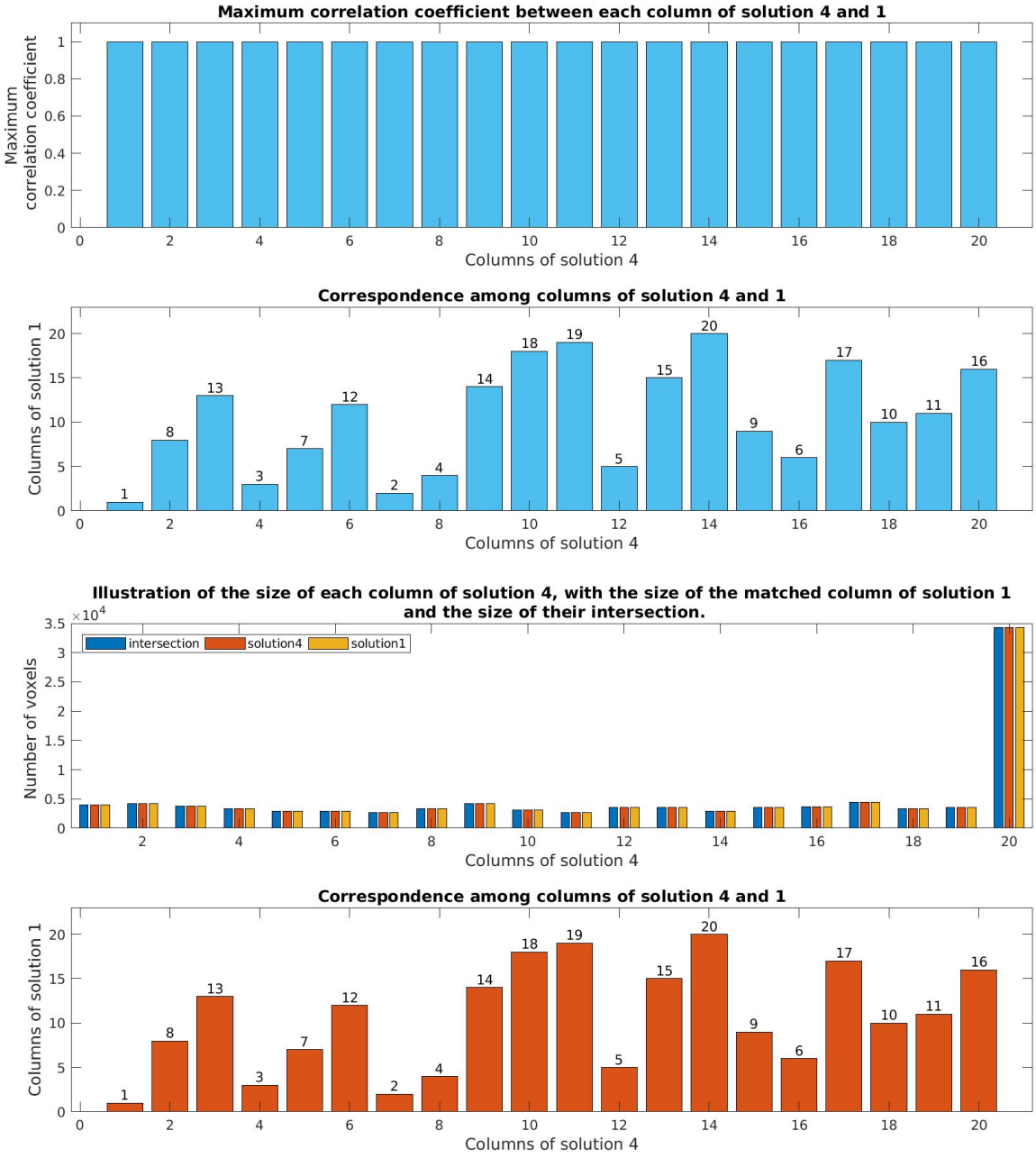


Figure A.103: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on intersection. Solutions 4, 1 correspond to accelerated penalty method and noiseless data.

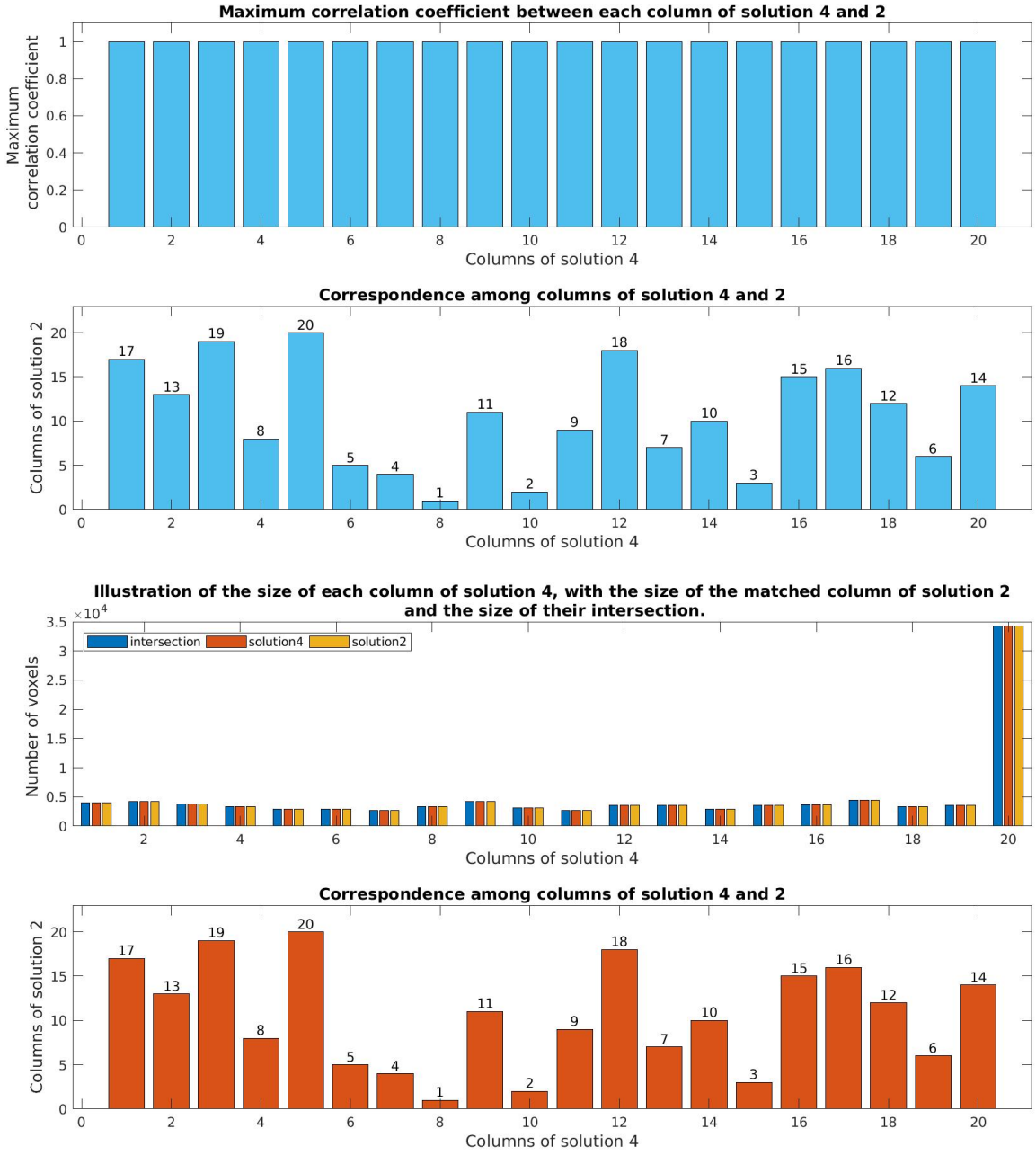


Figure A.104: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on intersection. Solutions 4, 2 correspond to accelerated penalty method and noiseless data.

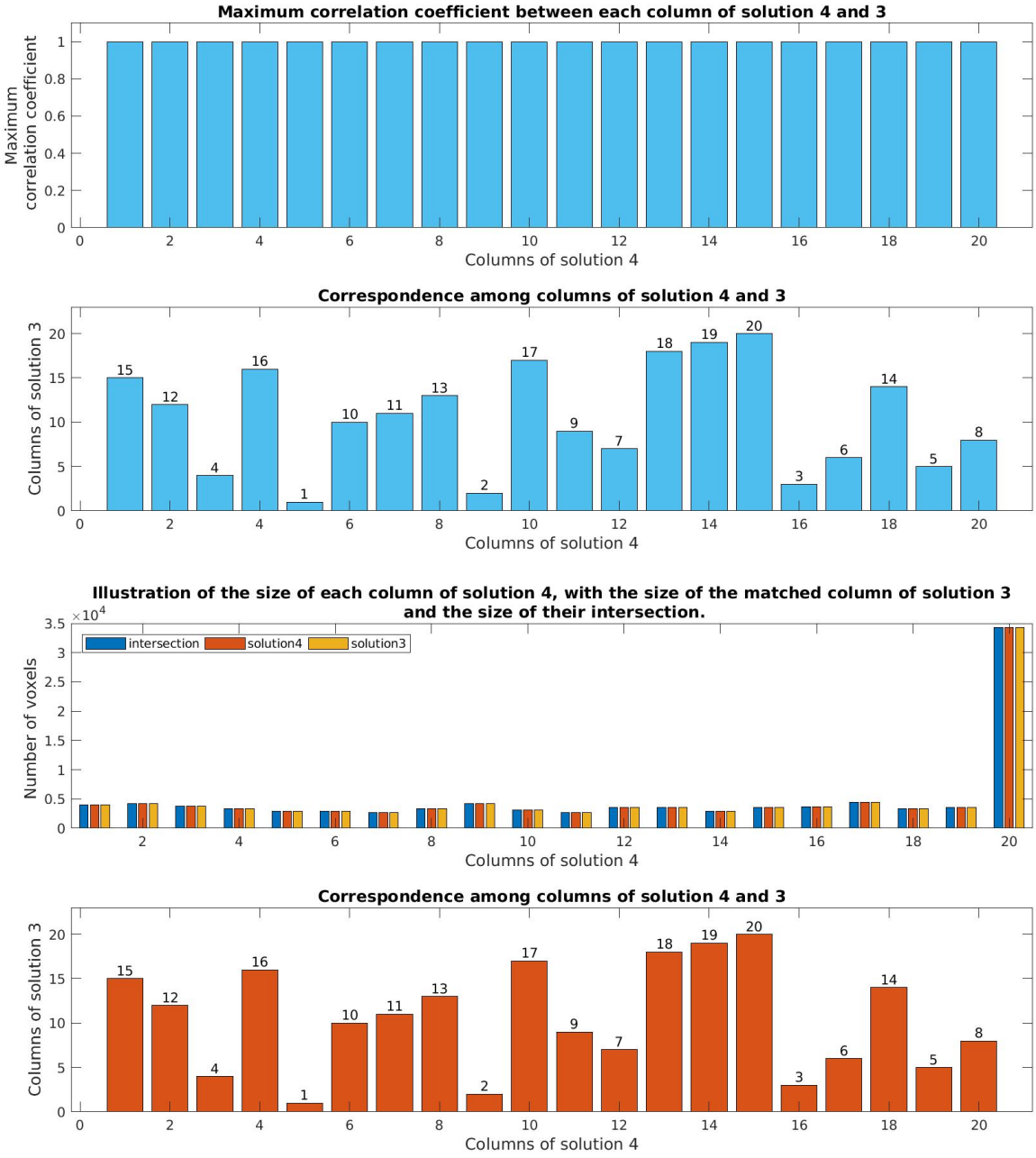


Figure A.105: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on intersection. Solutions 4, 3 correspond to accelerated penalty method and noiseless data.

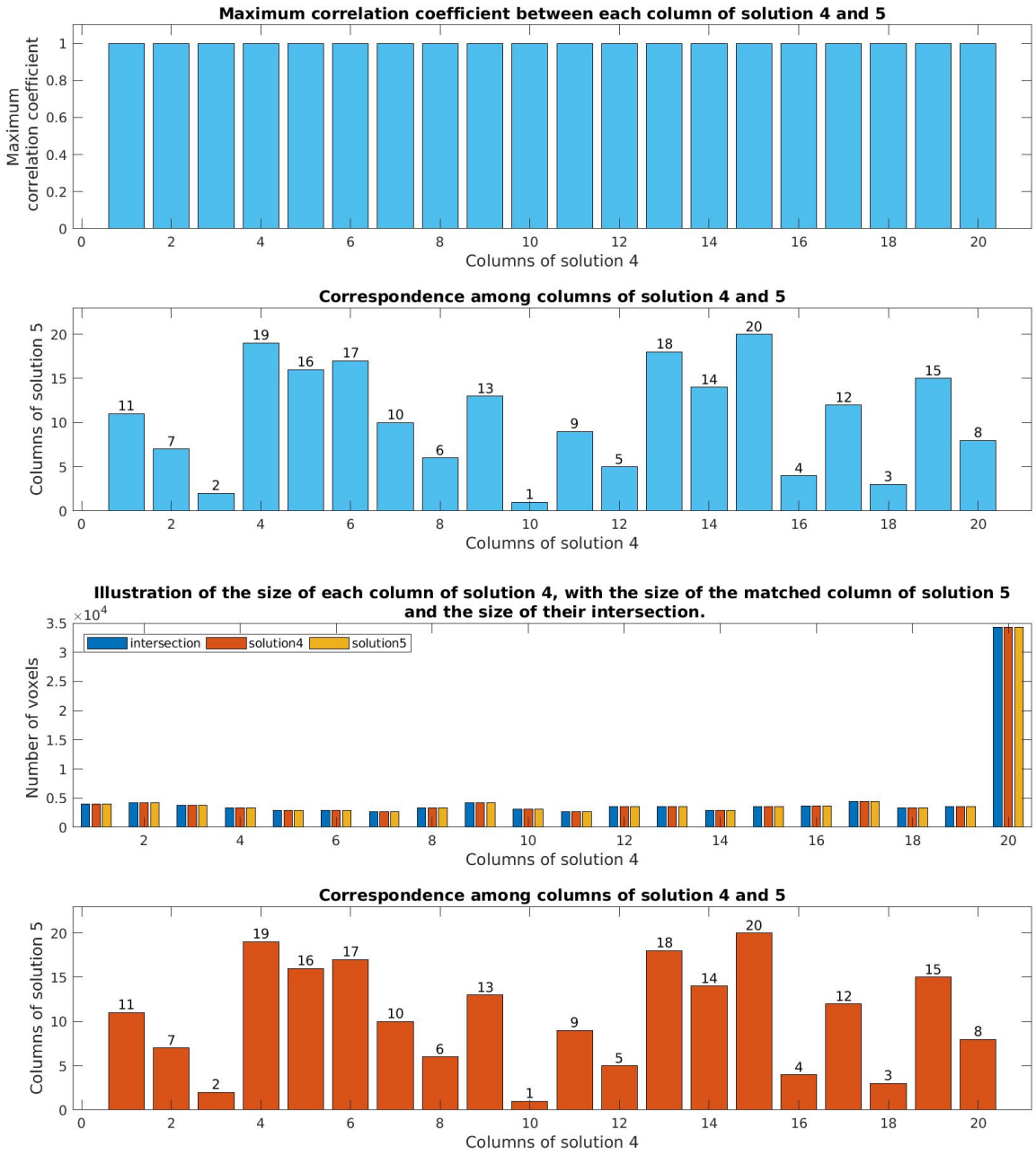


Figure A.106: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 5. We associate the column r of the solution 4 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on intersection. Solutions 4, 5 correspond to accelerated penalty method and noiseless data.

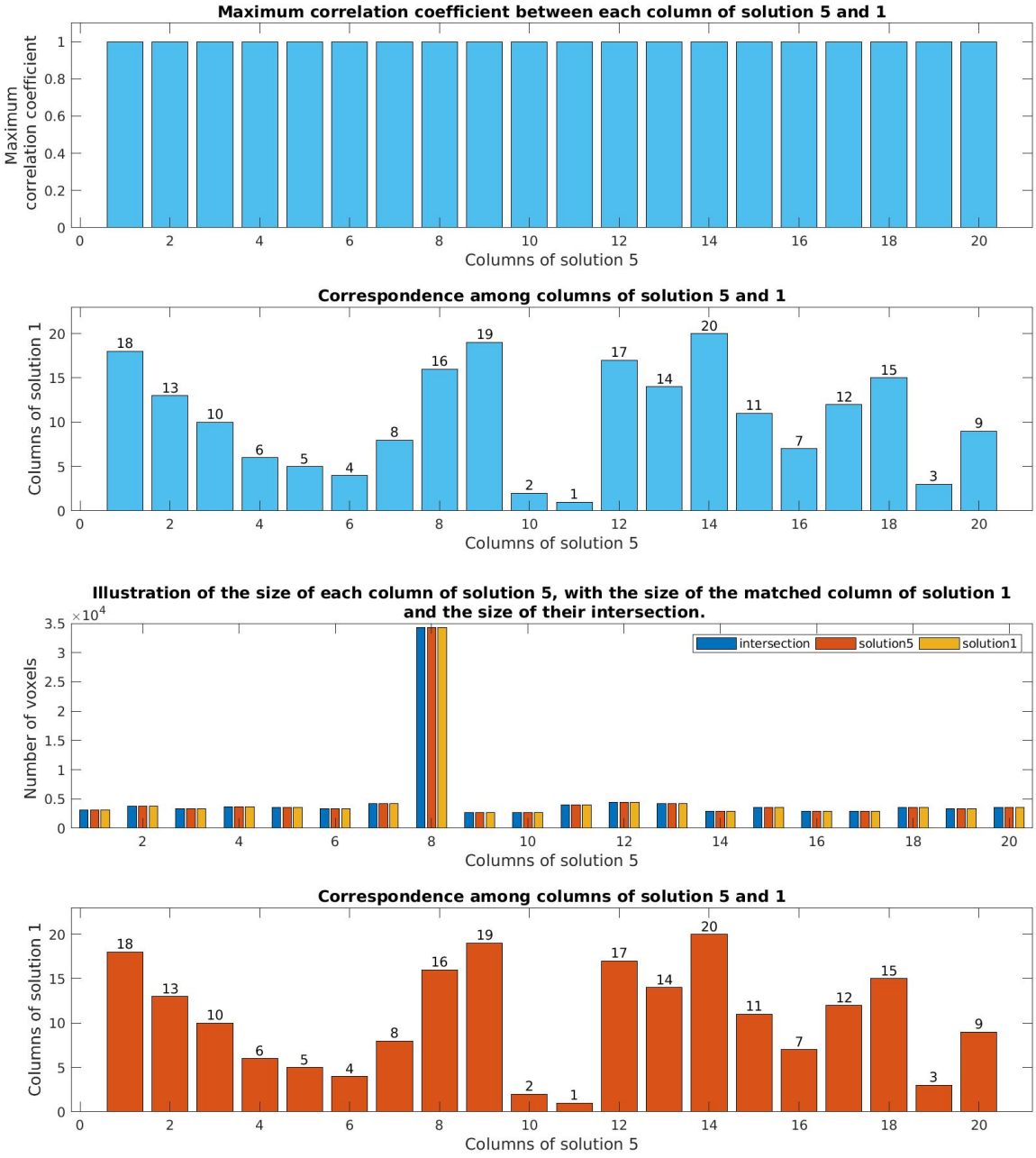


Figure A.107: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 1. We associate the column r of the solution 5 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on intersection. Solutions 5, 1 correspond to accelerated penalty method and noiseless data.

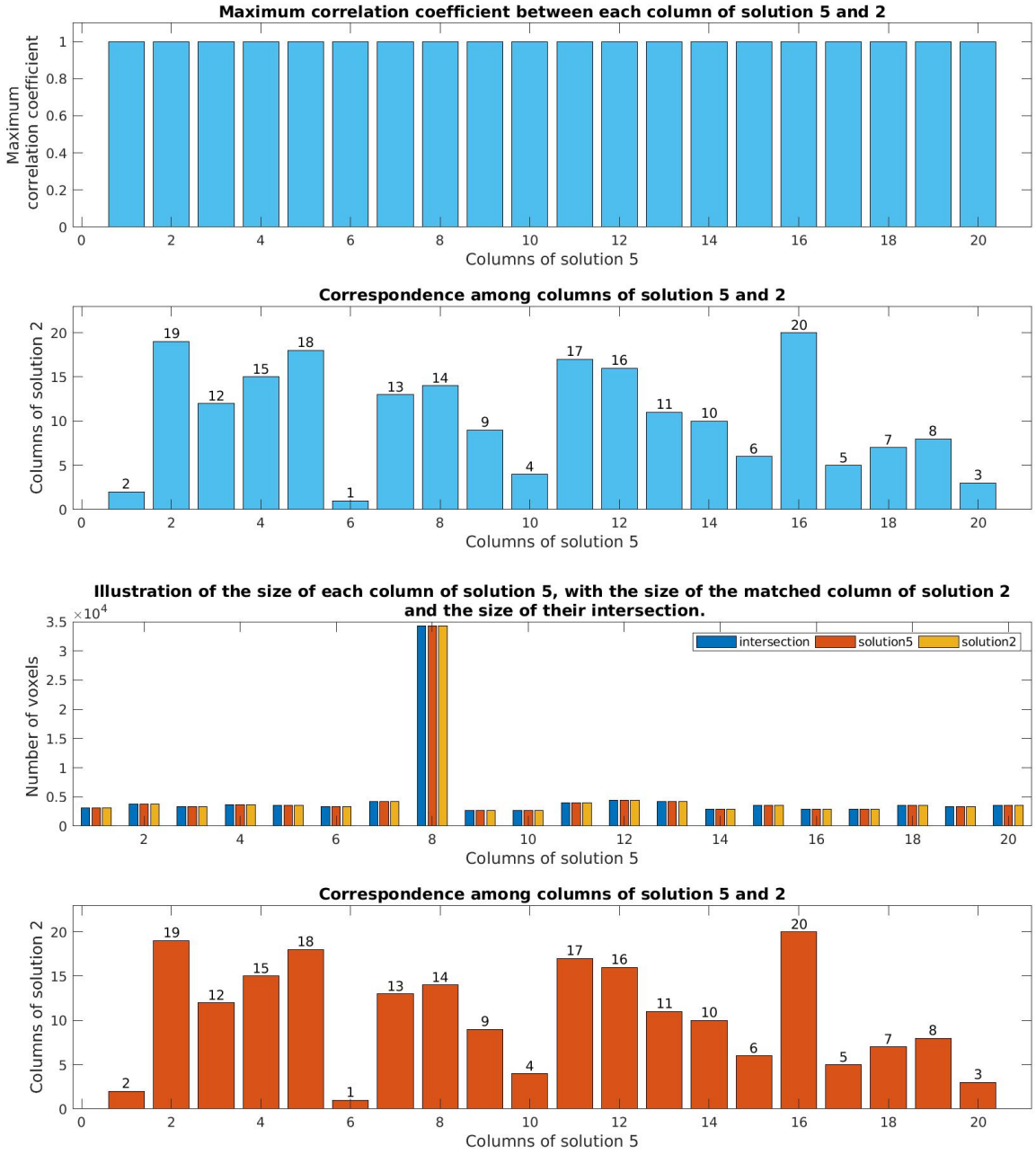


Figure A.108: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 2. We associate the column r of the solution 5 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on intersection. Solutions 5, 2 correspond to accelerated penalty method and noiseless data.

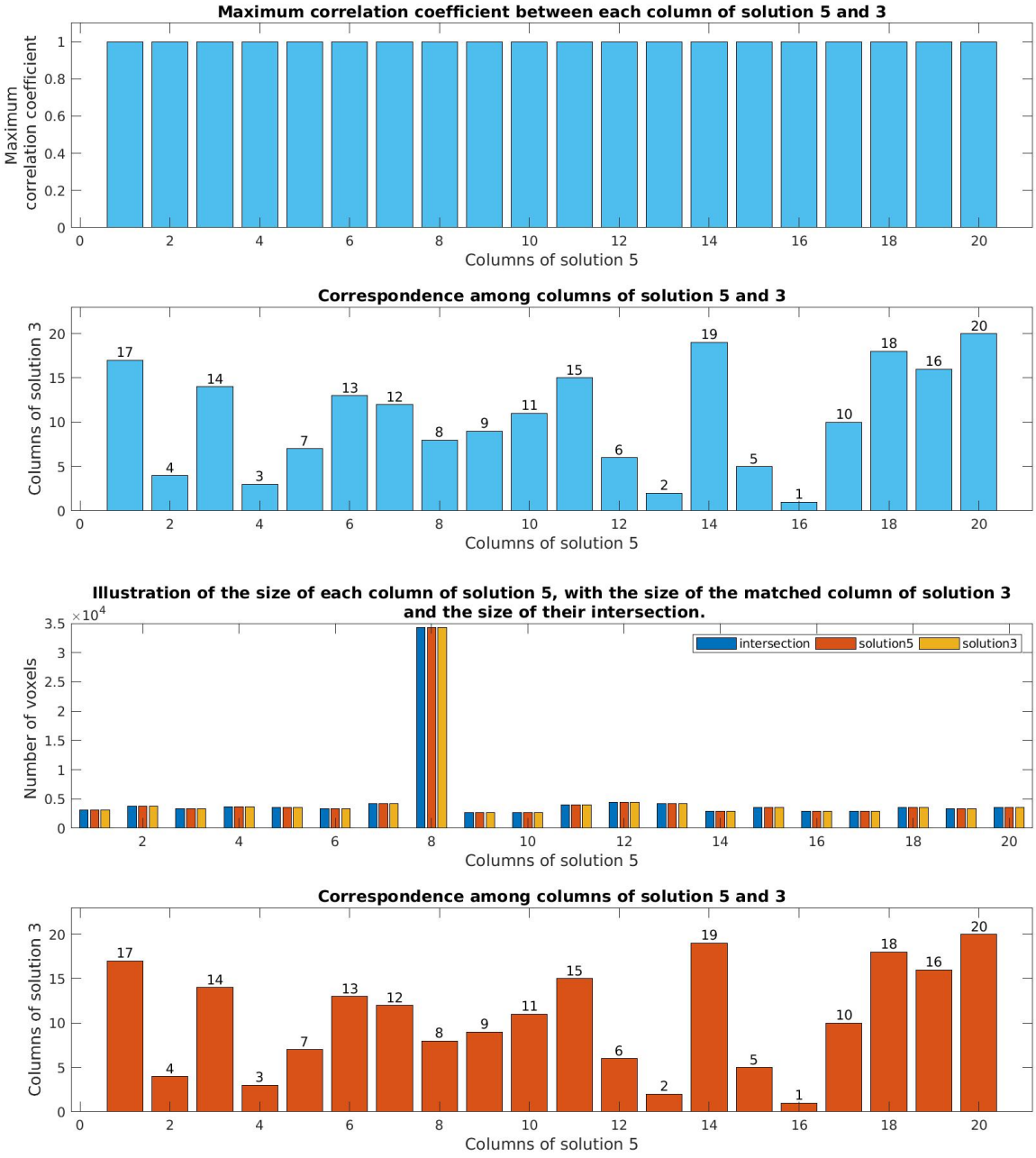


Figure A.109: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 3. We associate the column r of the solution 5 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on intersection. Solutions 5, 3 correspond to accelerated penalty method and noiseless data.

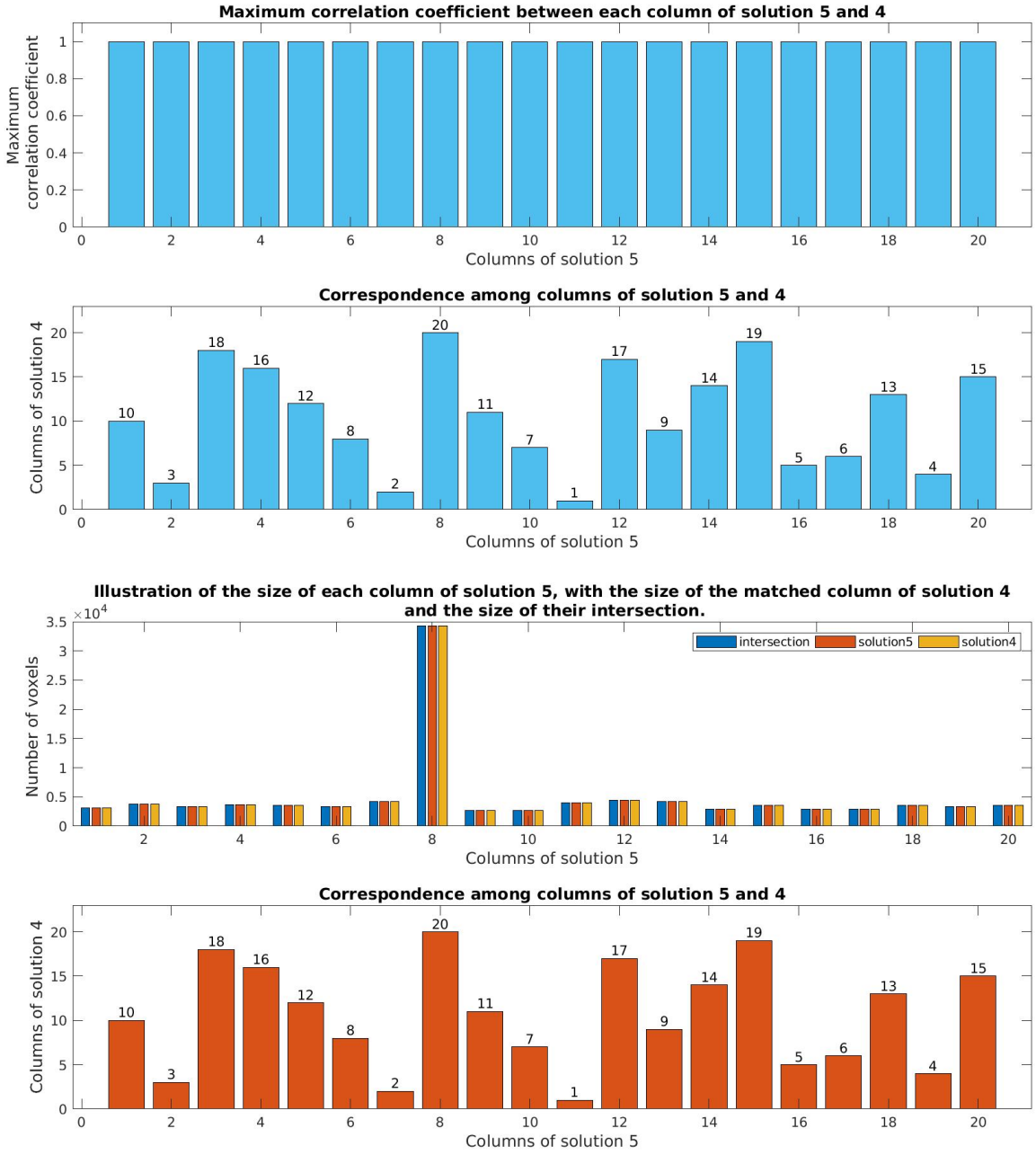


Figure A.110: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 4. We associate the column r of the solution 5 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on intersection. Solutions 5, 4 correspond to accelerated penalty method and noiseless data.

· **SNR=0 dB**

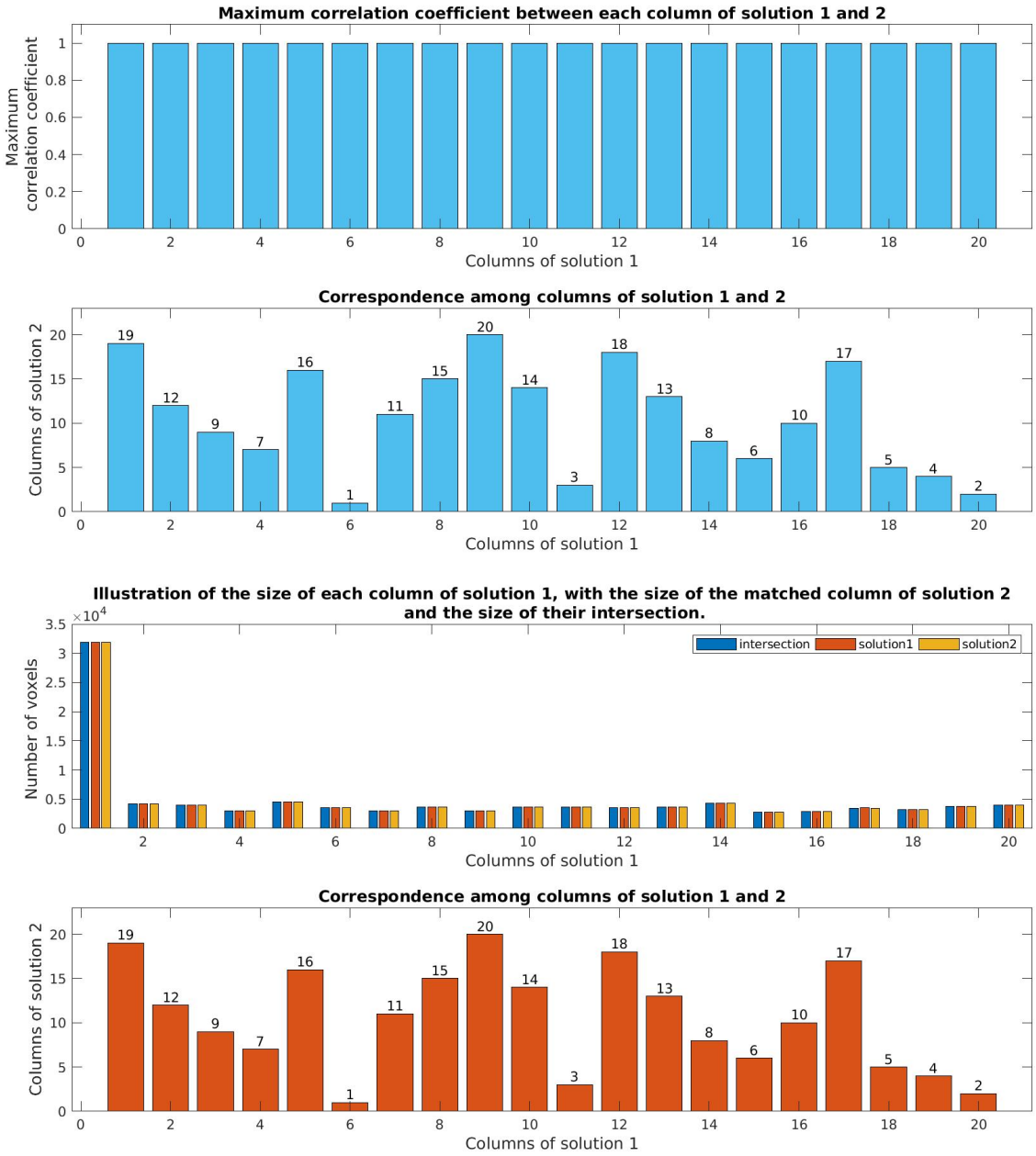


Figure A.111: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated, based on intersection. Solutions 1, 2 correspond to accelerated penalty method and SNR=0 dB.

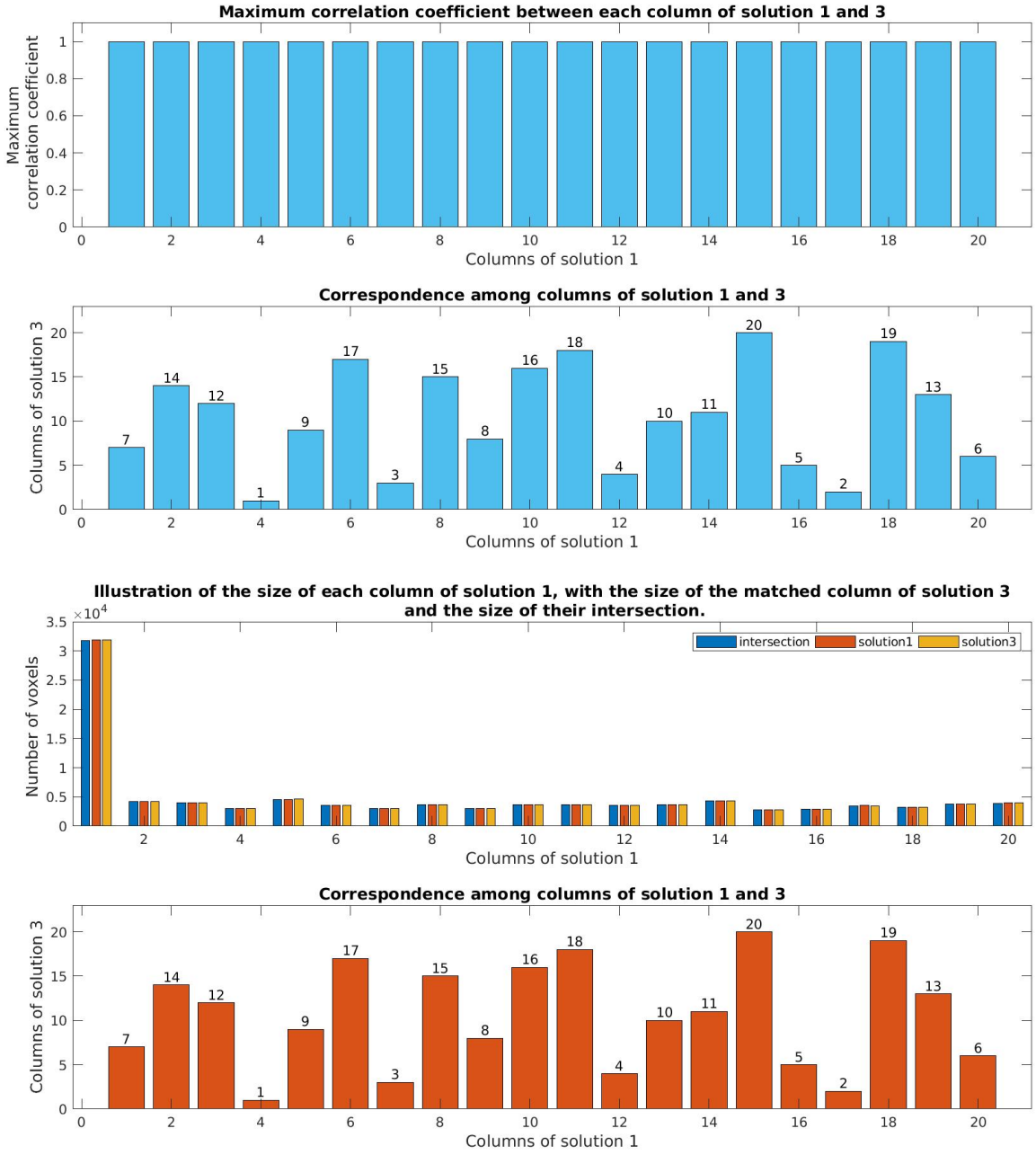


Figure A.112: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on intersection. Solutions 1, 3 correspond to accelerated penalty method and SNR=0 dB.

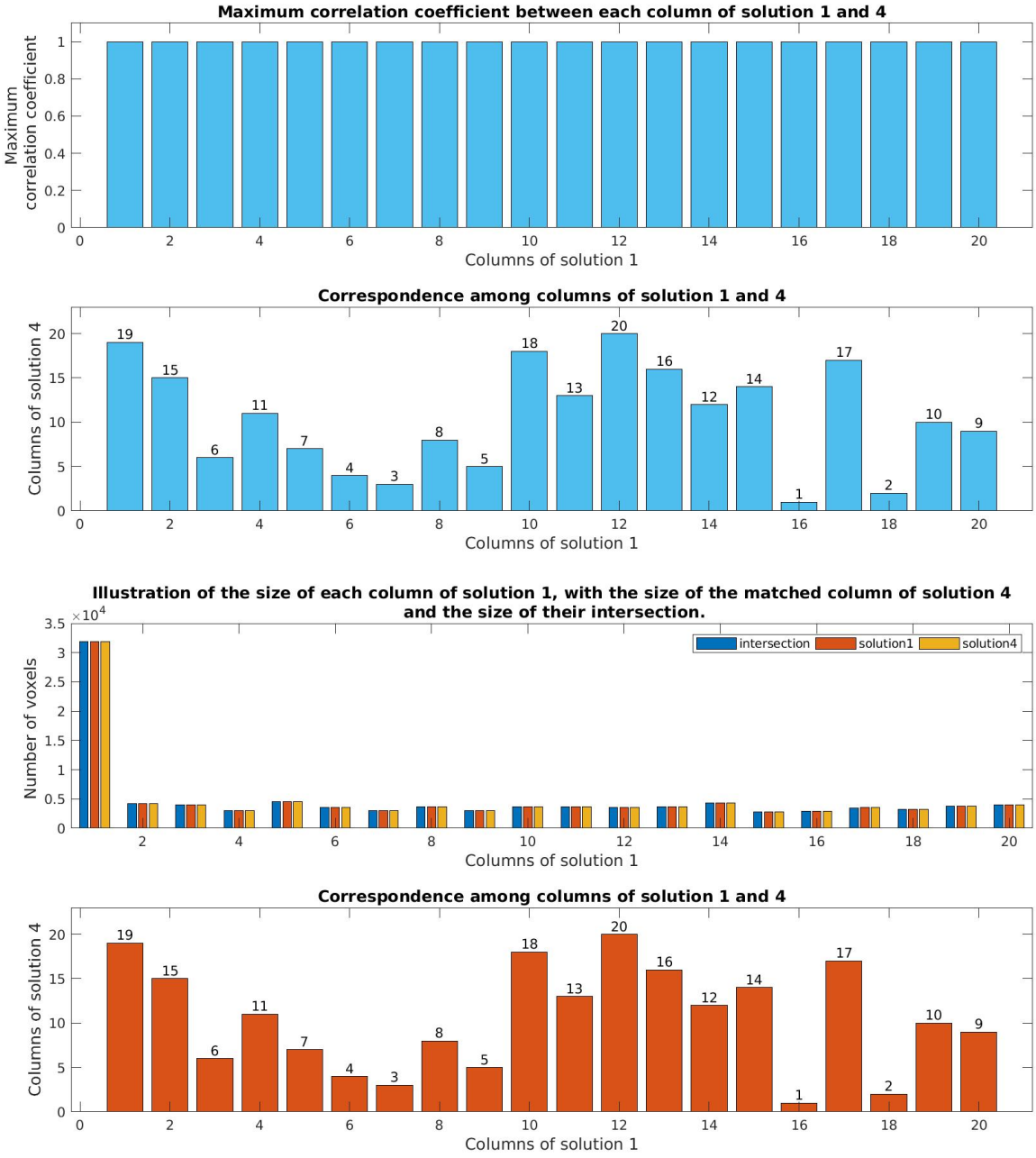


Figure A.113: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on intersection. Solutions 1, 4 correspond to accelerated penalty method and SNR=0 dB.

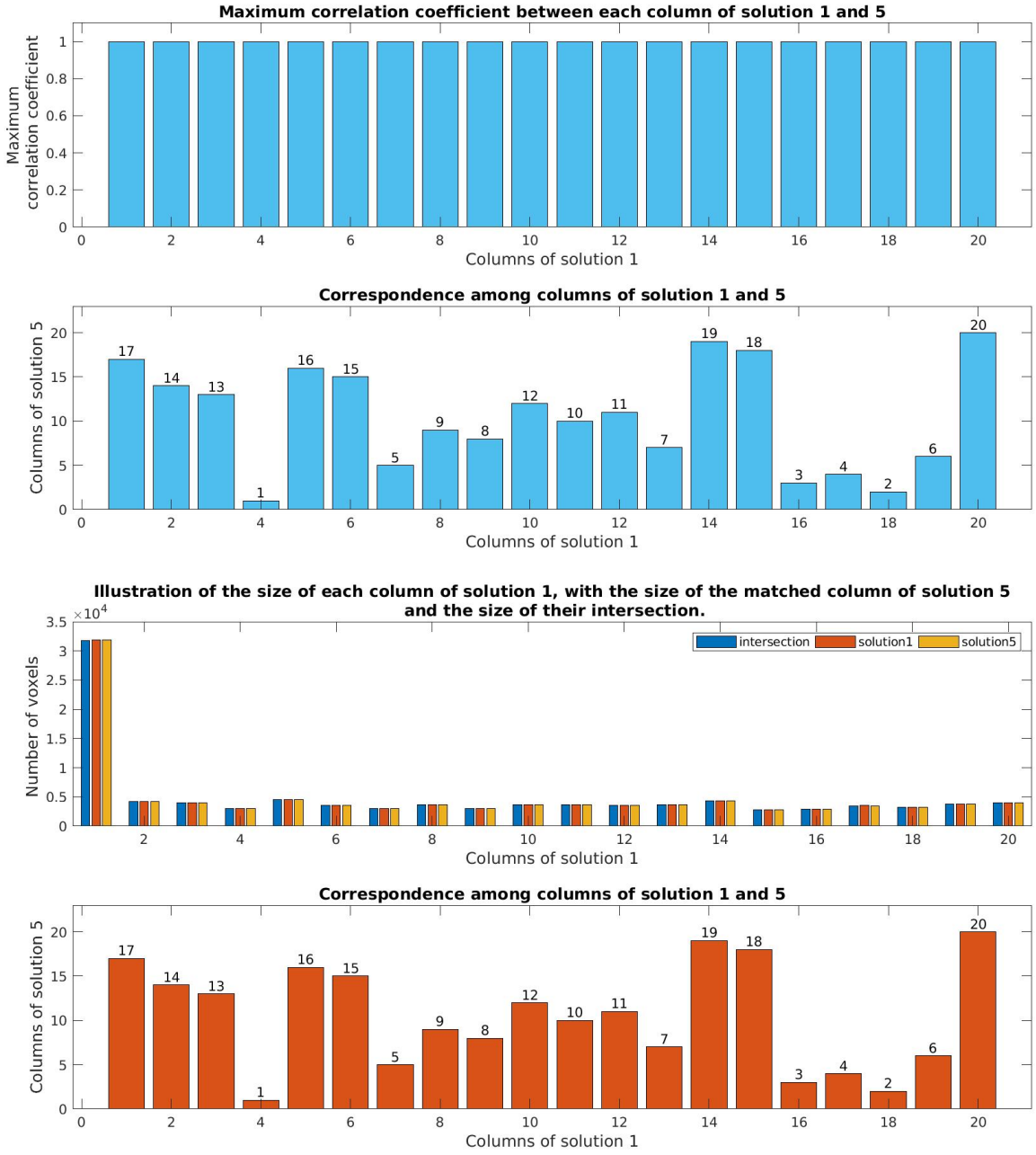


Figure A.114: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 5. We associate the column r of the solution 1 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on intersection. Solutions 1, 5 correspond to accelerated penalty method and SNR=0 dB.

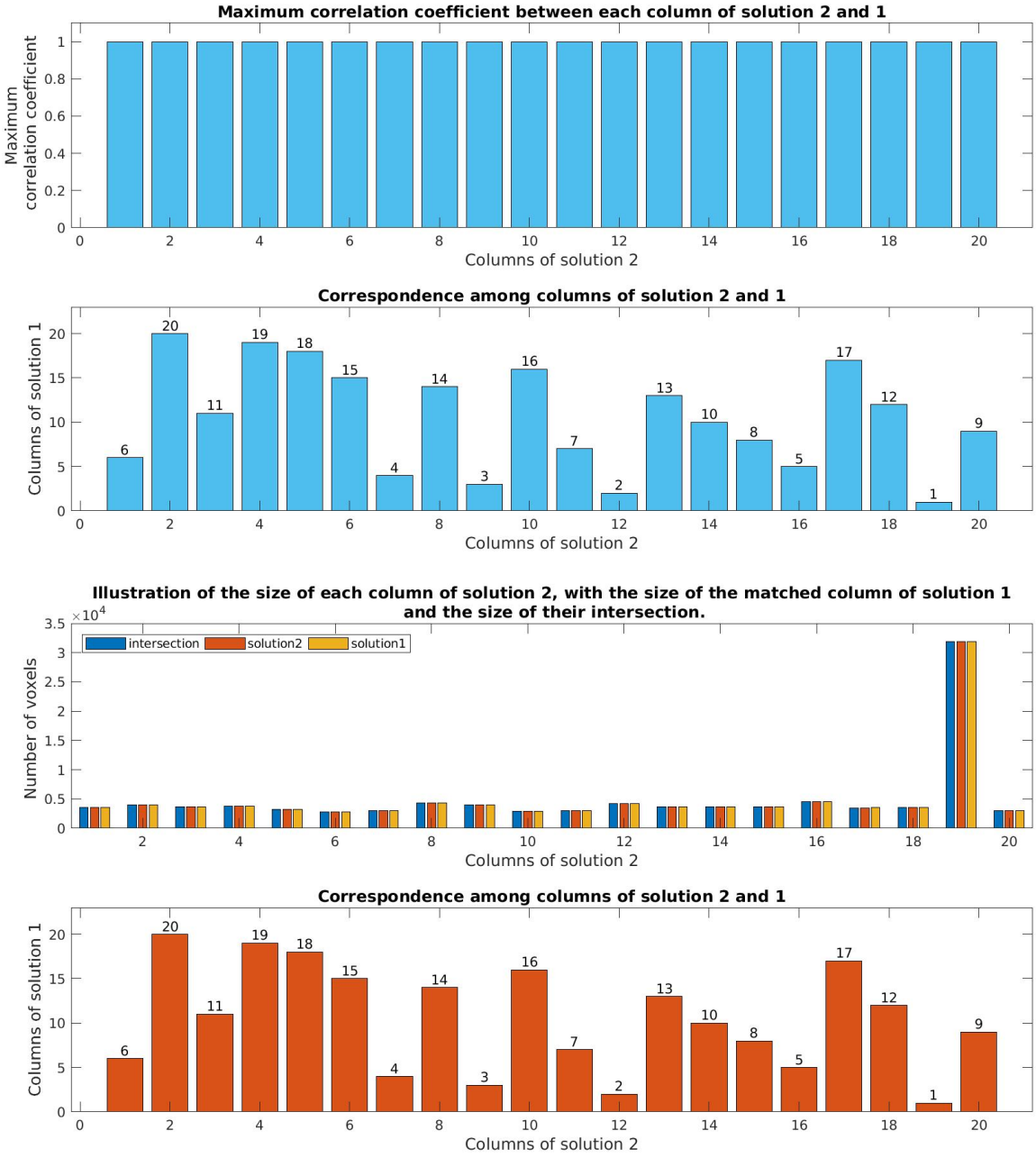


Figure A.115: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on intersection. Solutions 2, 1 correspond to accelerated penalty method and SNR=0 dB.

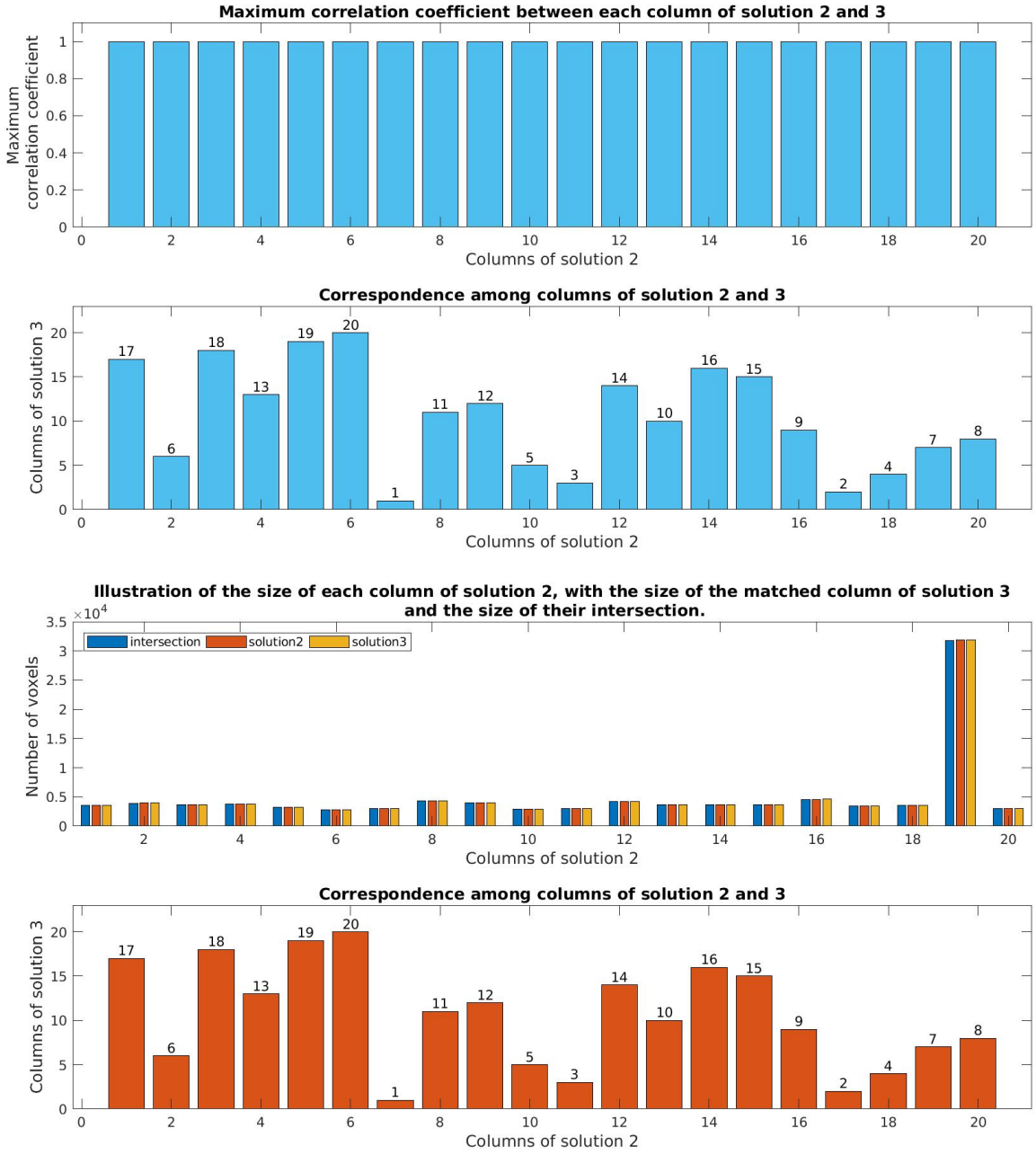


Figure A.116: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on intersection. Solutions 2, 3 correspond to accelerated penalty method and SNR=0 dB.

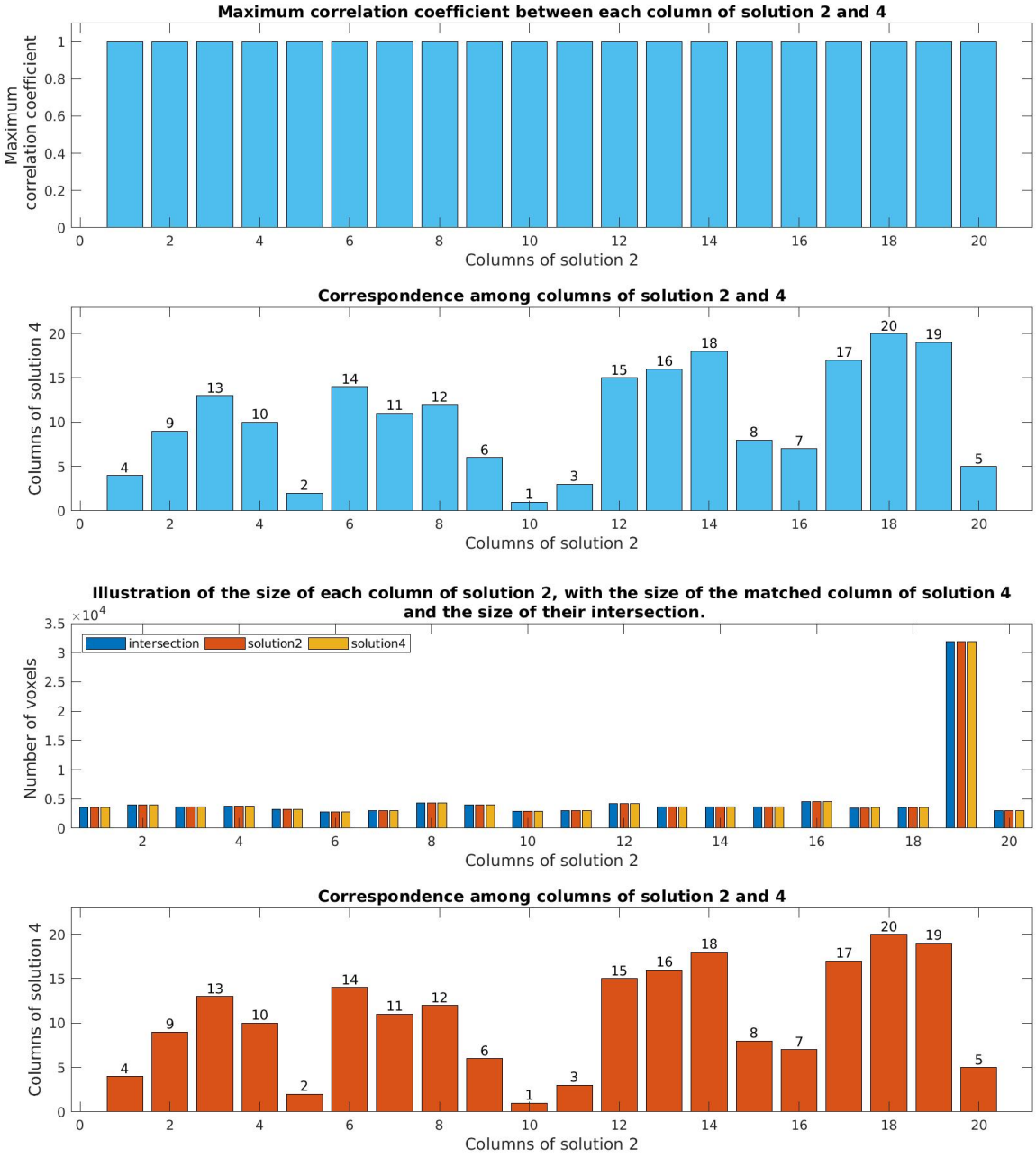


Figure A.117: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on intersection. Solutions 2, 4 correspond to accelerated penalty method and SNR=0 dB.

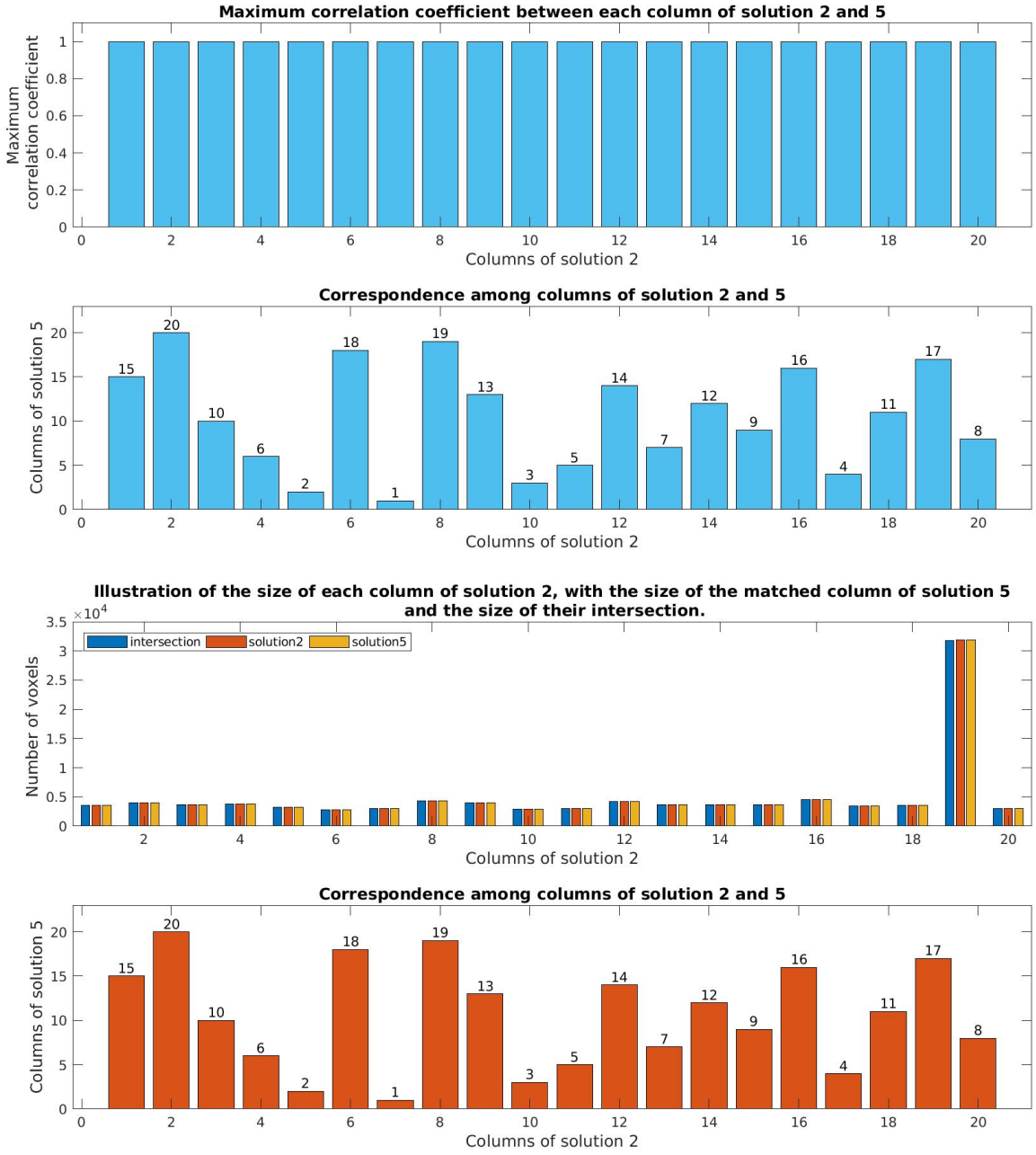


Figure A.118: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 5. We associate the column r of the solution 2 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on intersection. Solutions 2, 5 correspond to accelerated penalty method and SNR=0 dB.

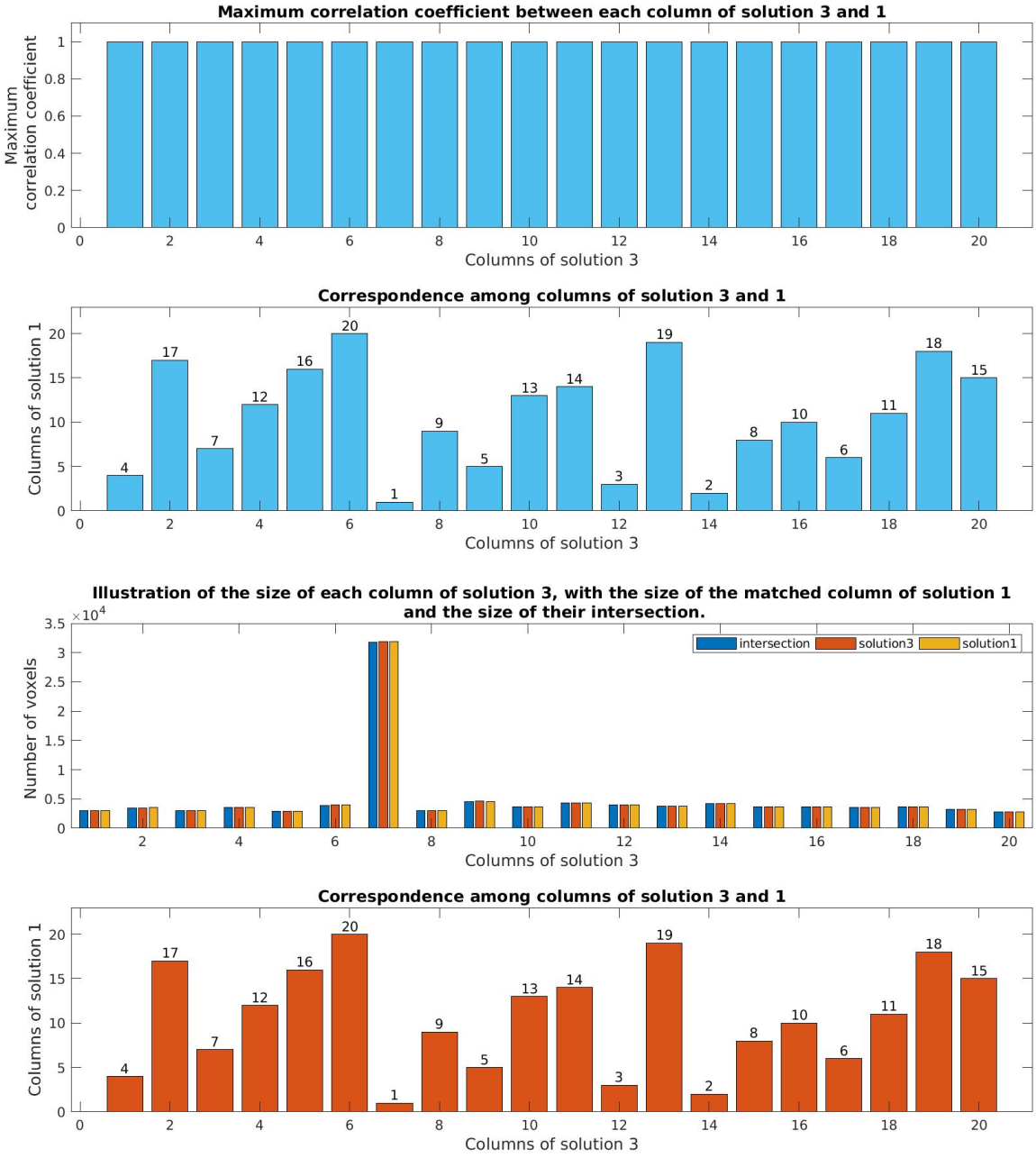


Figure A.119: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on intersection. Solutions 3, 1 correspond to accelerated penalty method and SNR=0 dB.

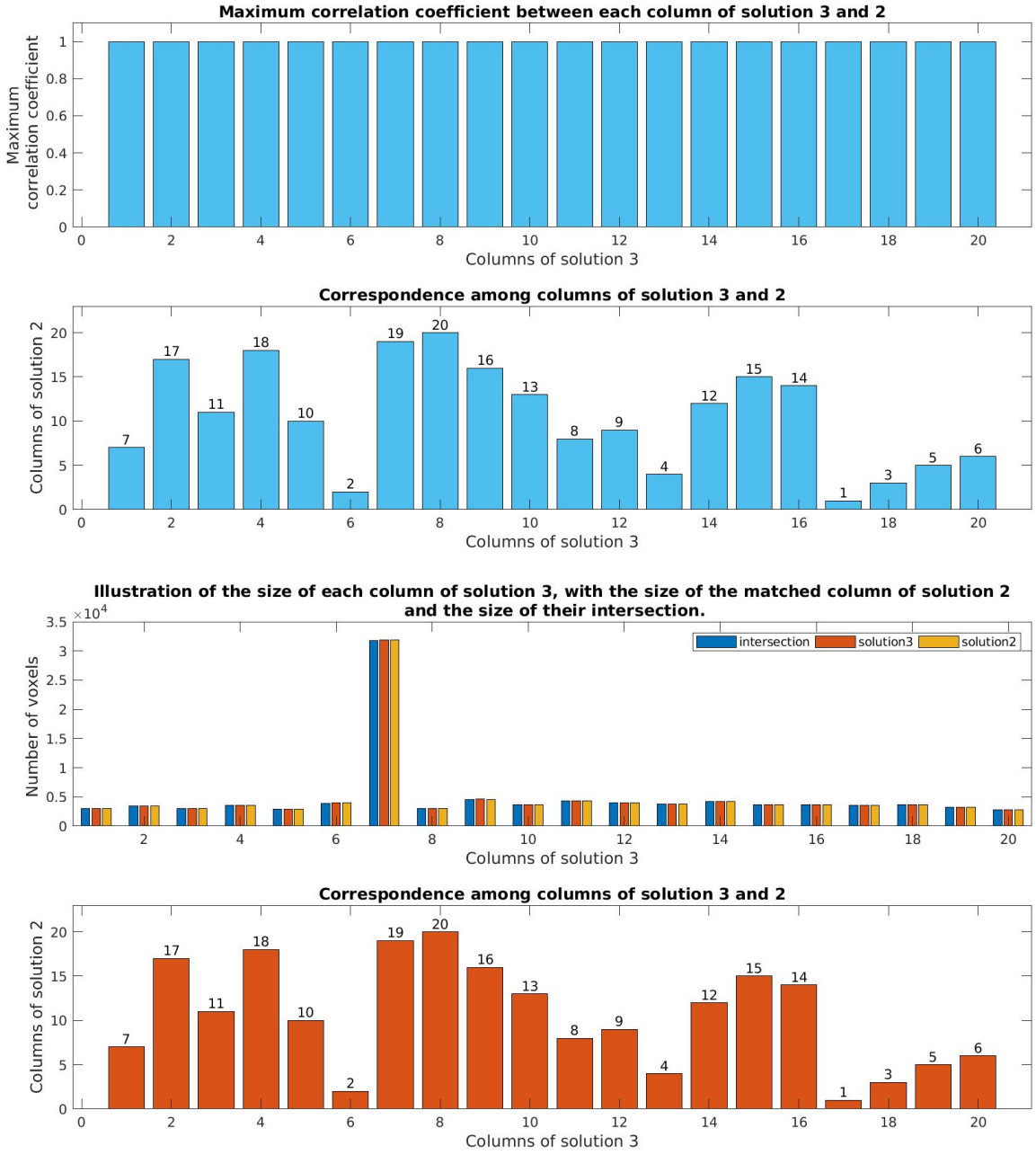


Figure A.120: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on intersection. Solutions 3, 2 correspond to accelerated penalty method and SNR=0 dB.

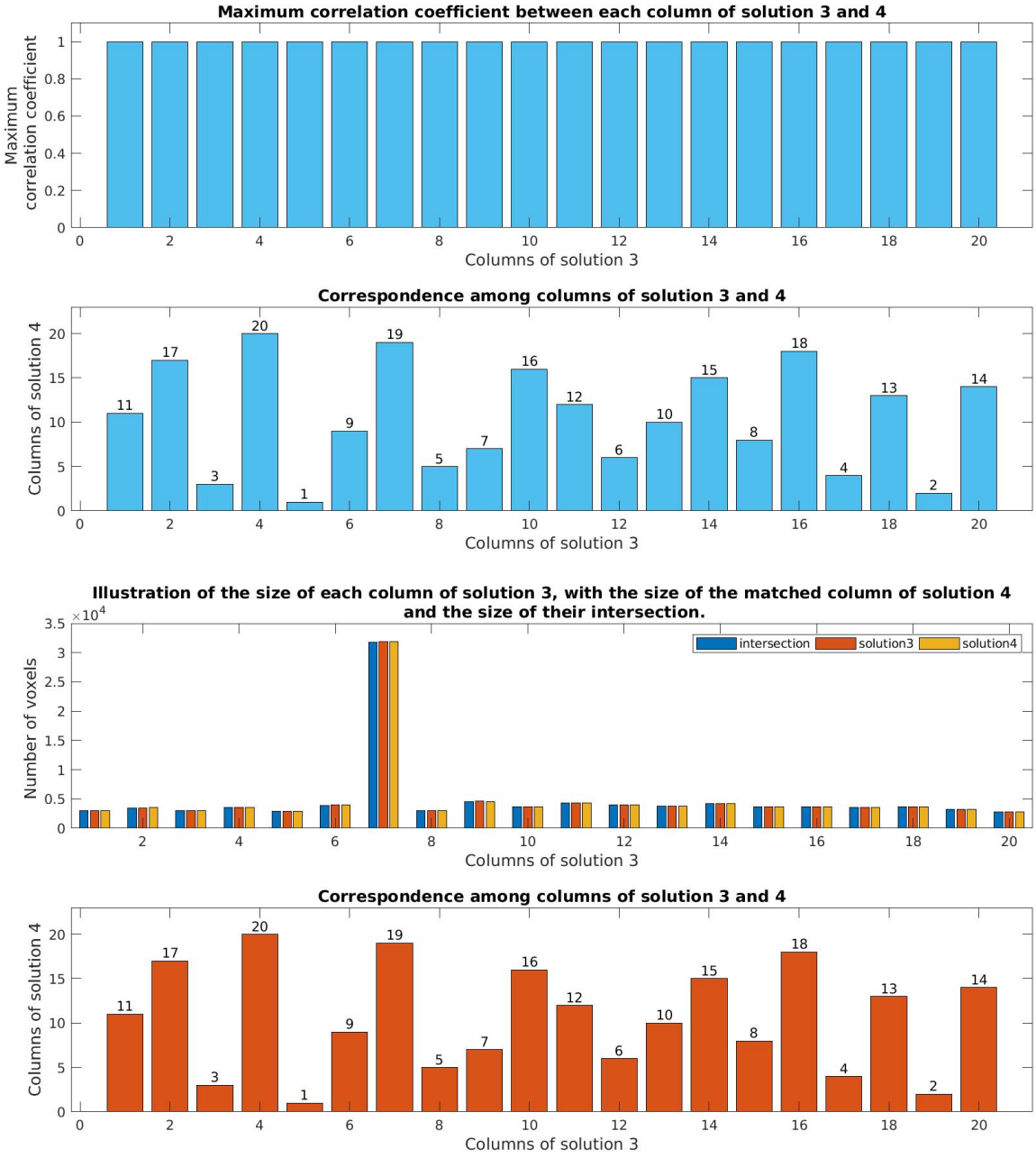


Figure A.121: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on intersection. Solutions 3, 4 correspond to accelerated penalty method and SNR=0 dB.

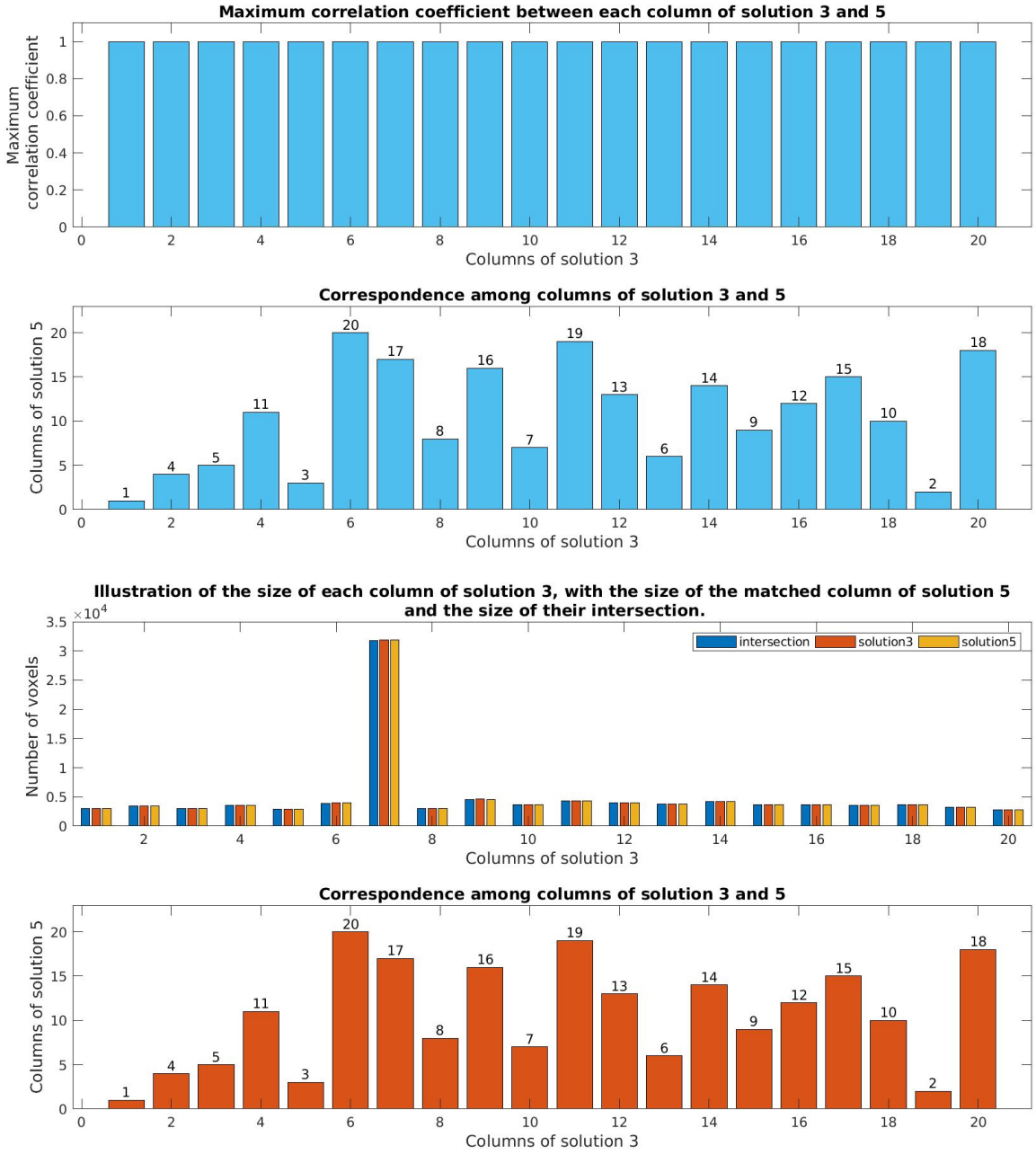


Figure A.122: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 5. We associate the column r of the solution 3 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on intersection. Solutions 3, 5 correspond to accelerated penalty method and SNR=0 dB.

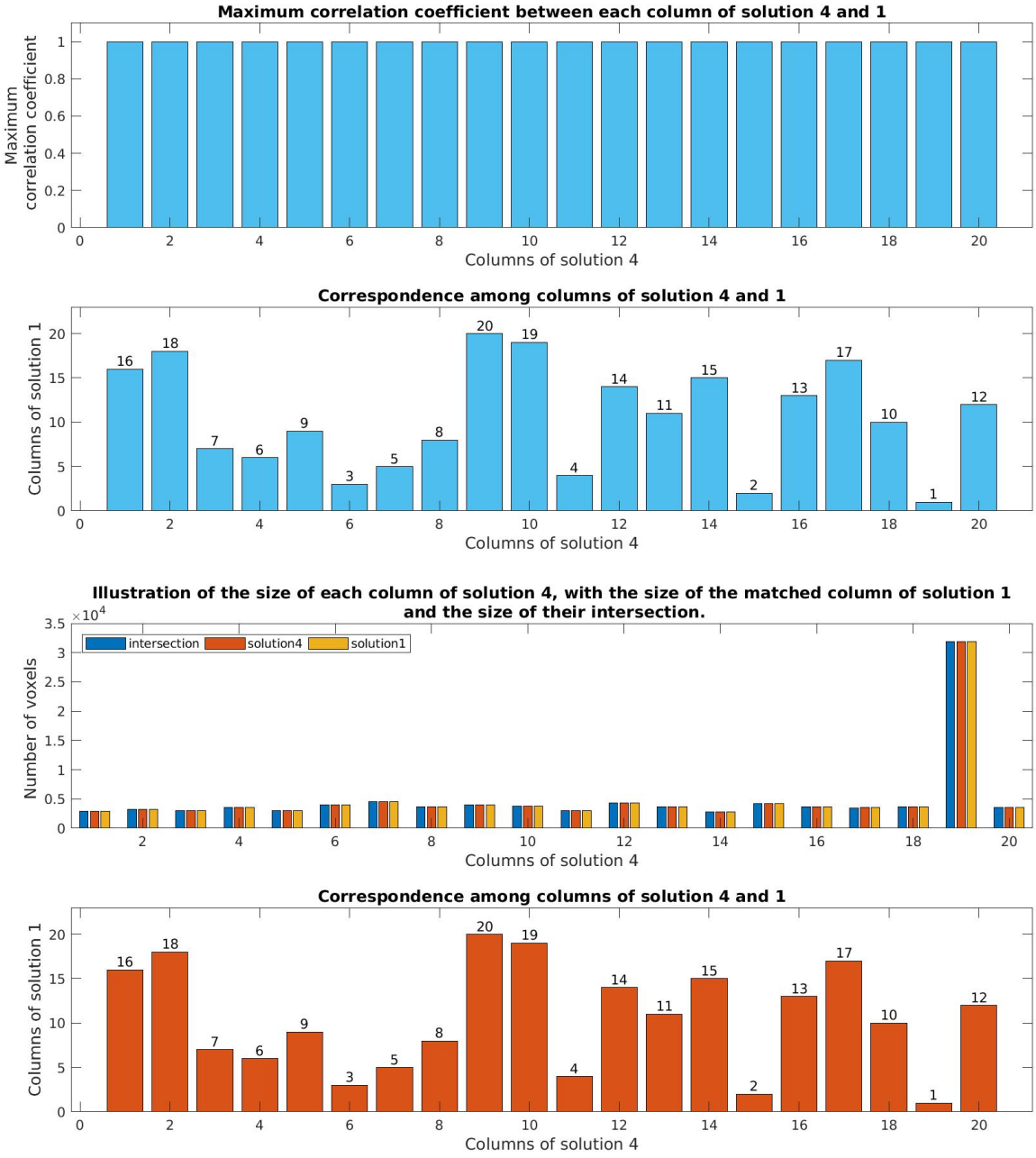


Figure A.123: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on intersection. Solutions 4, 1 correspond to accelerated penalty method and SNR=0 dB.

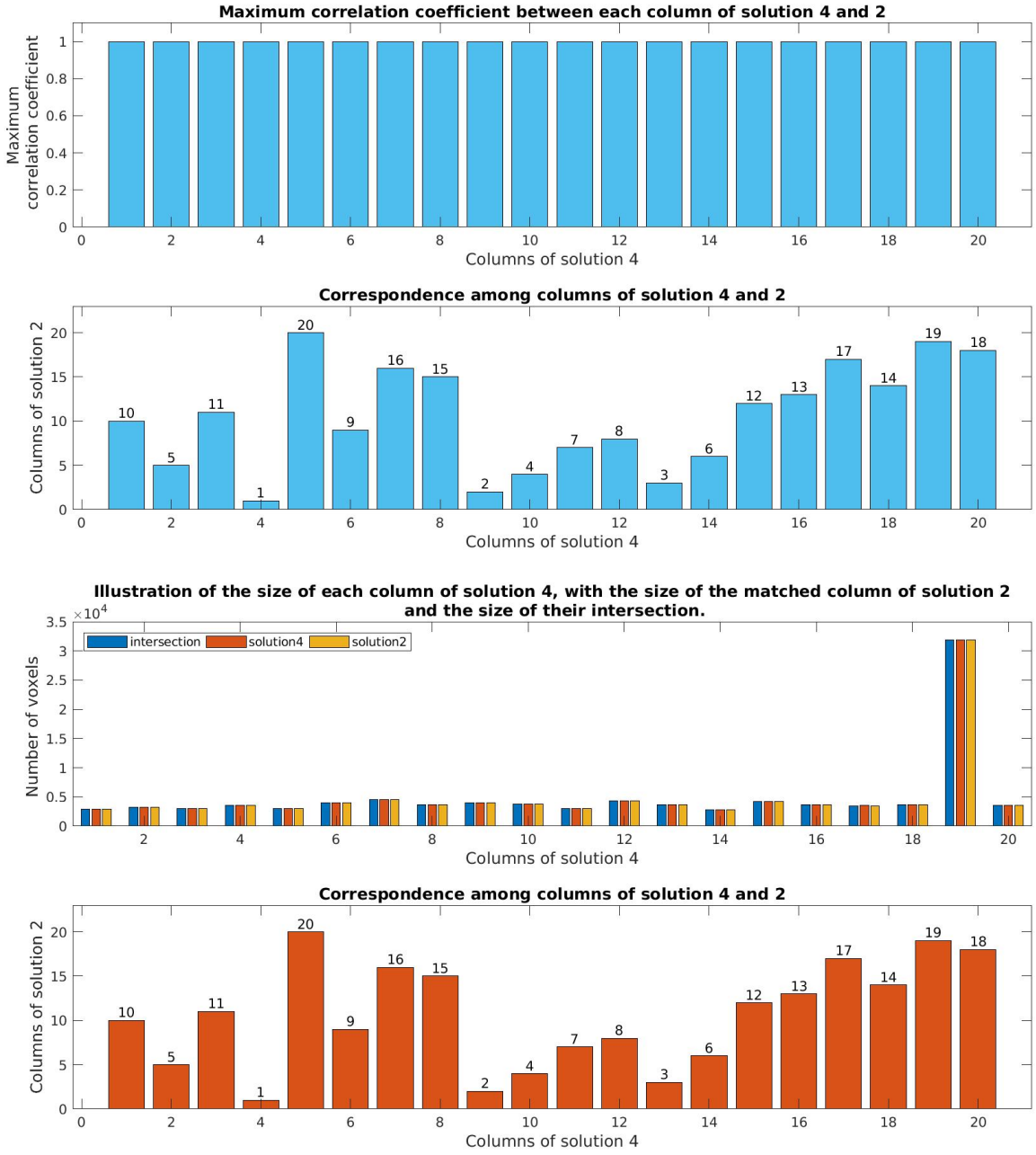


Figure A.124: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on intersection. Solutions 4, 2 correspond to accelerated penalty method and SNR=0 dB.

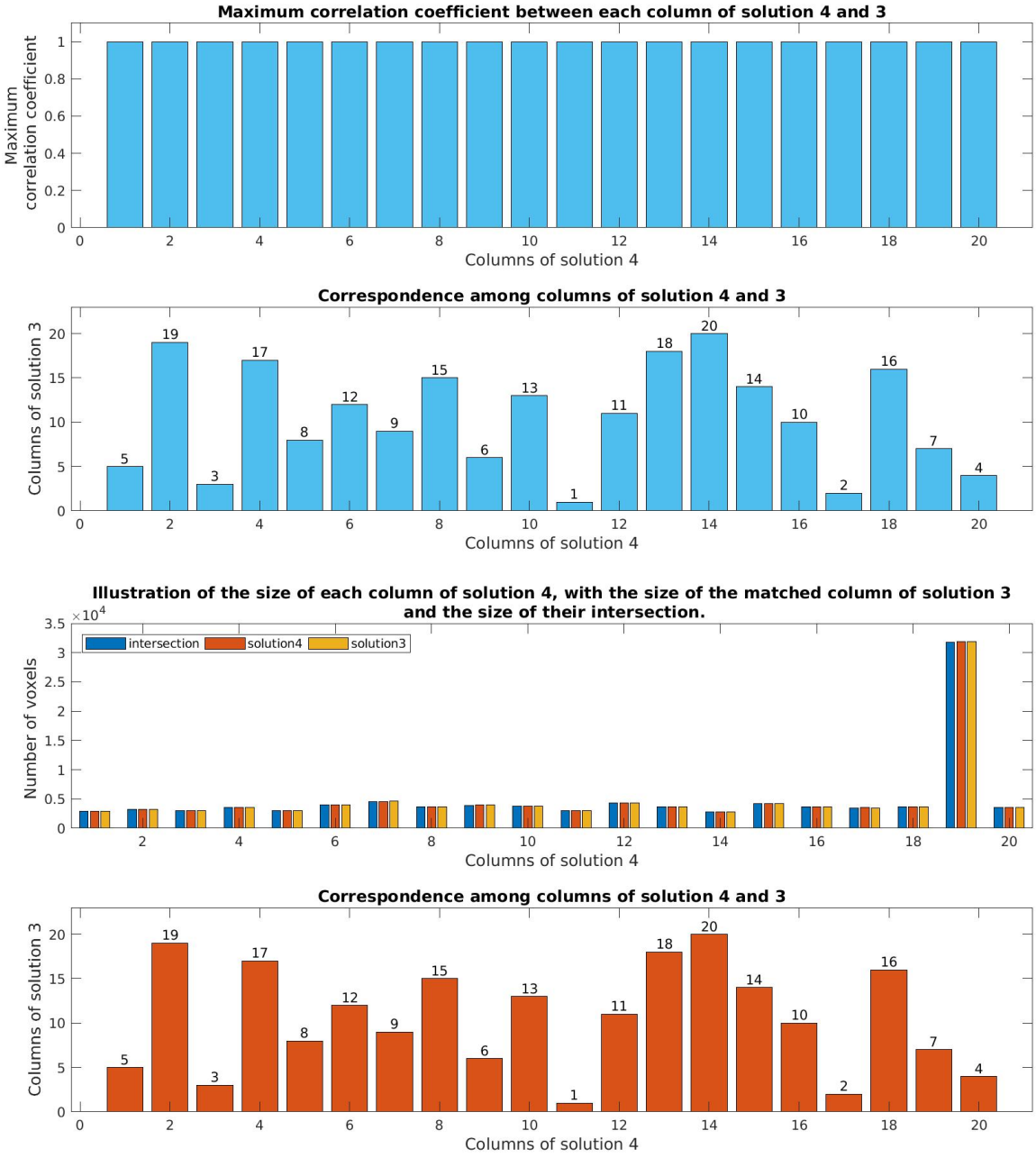


Figure A.125: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on intersection. Solutions 4, 3 correspond to accelerated penalty method and SNR=0 dB.

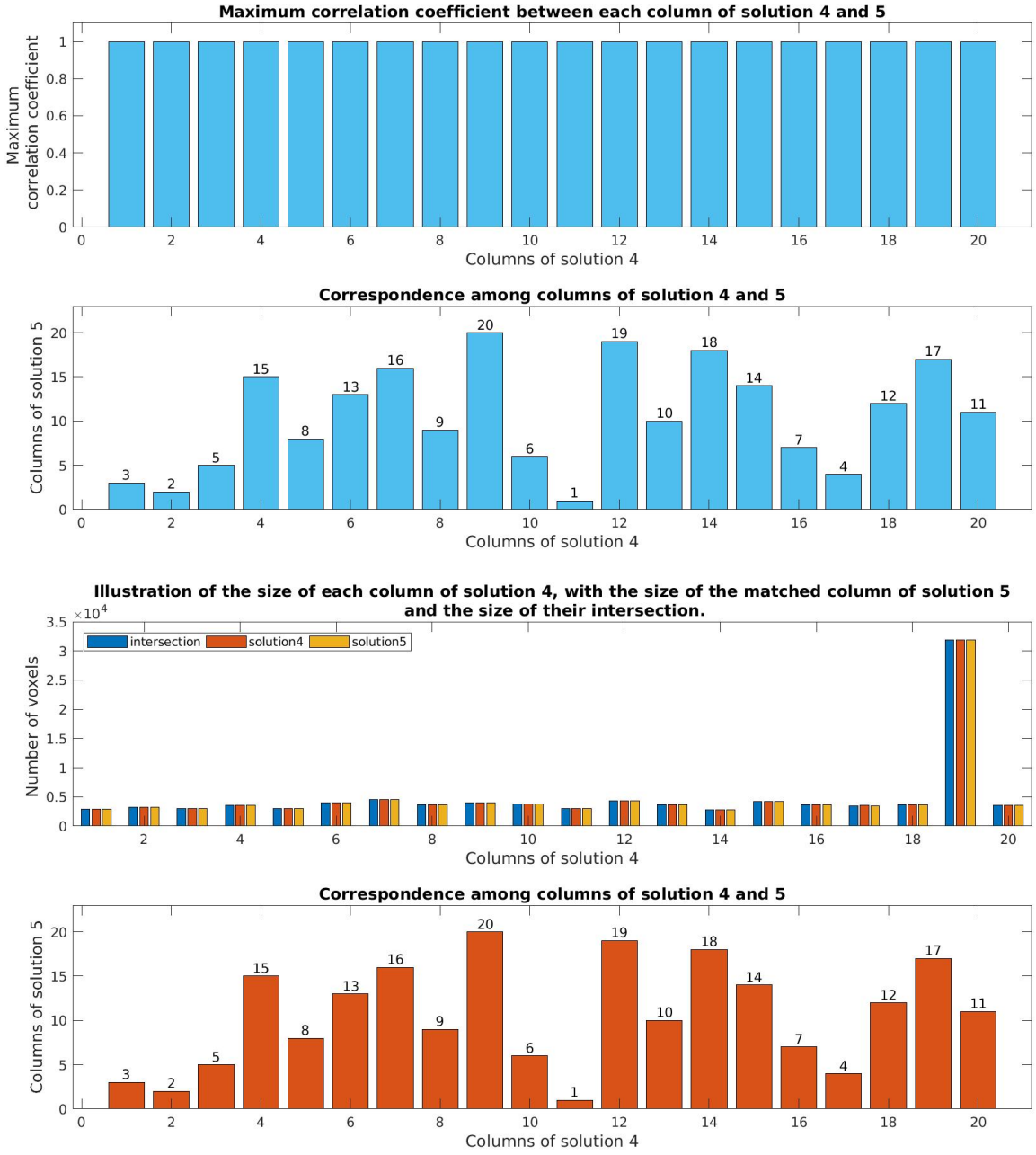


Figure A.126: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 5. We associate the column r of the solution 4 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on intersection. Solutions 4, 5 correspond to accelerated penalty method and SNR=0 dB.

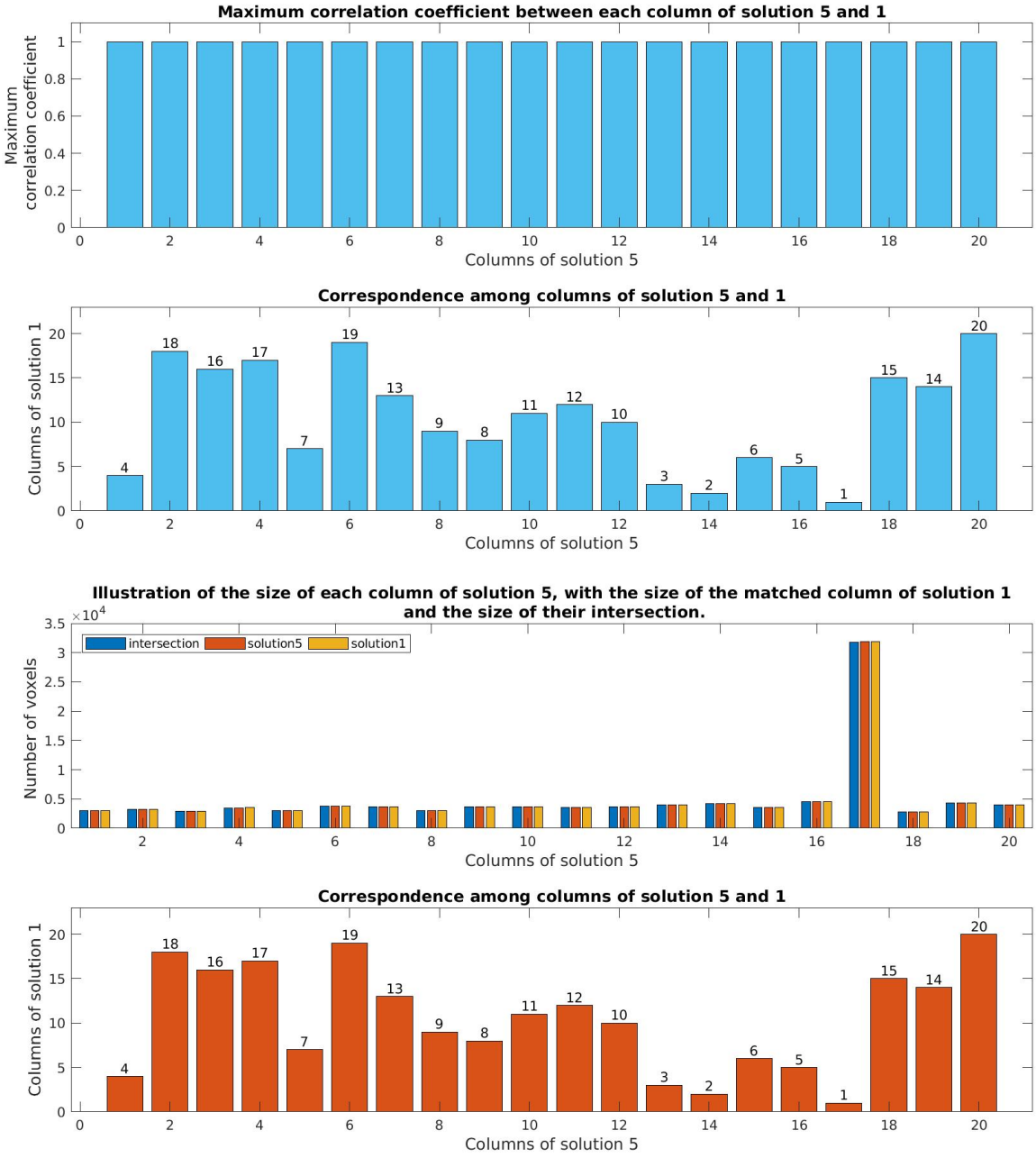


Figure A.127: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 1. We associate the column r of the solution 5 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on intersection. Solutions 5, 1 correspond to accelerated penalty method and SNR=0 dB.

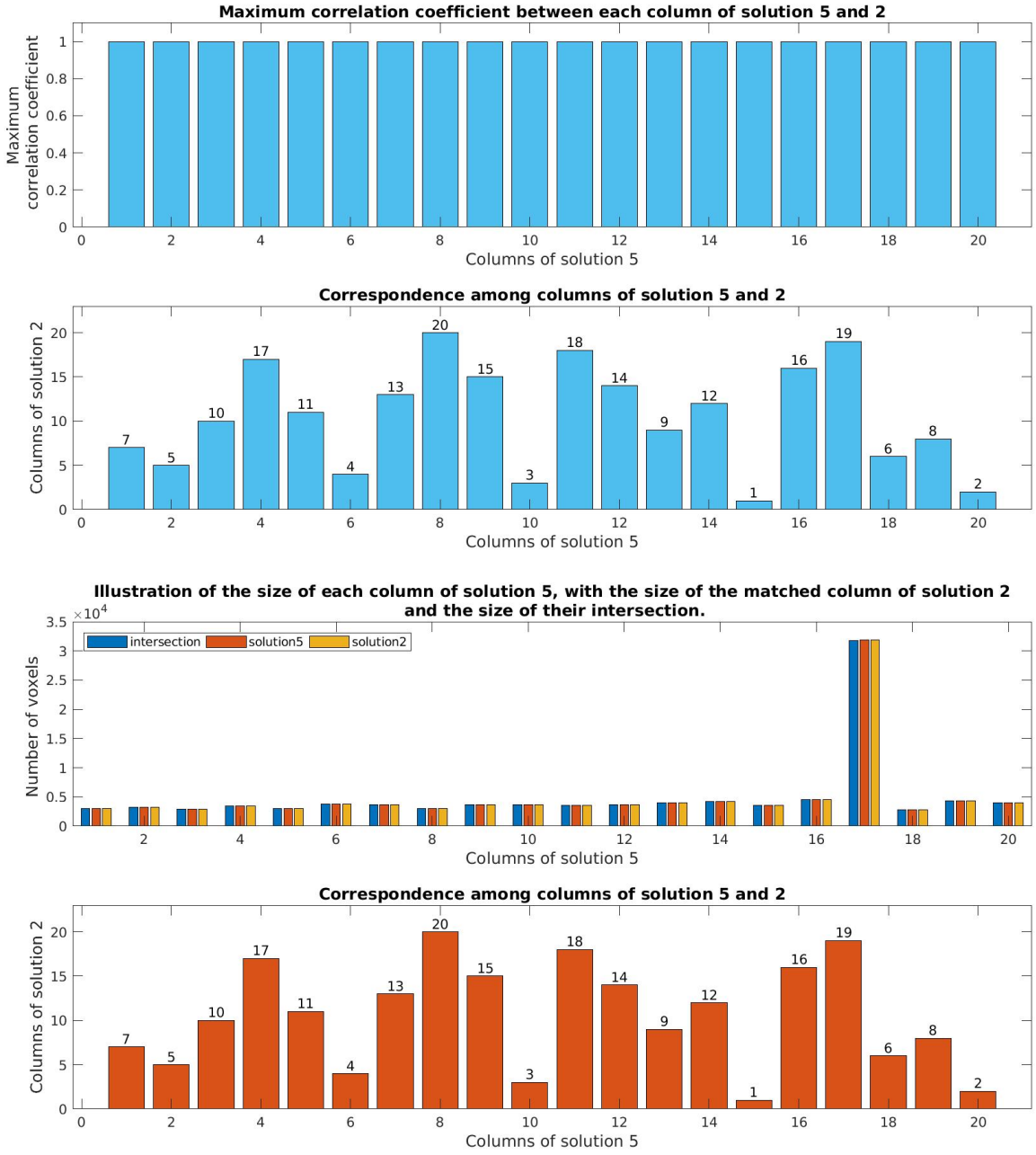


Figure A.128: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 2. We associate the column r of the solution 5 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on intersection. Solutions 5, 2 correspond to accelerated penalty method and SNR=0 dB.

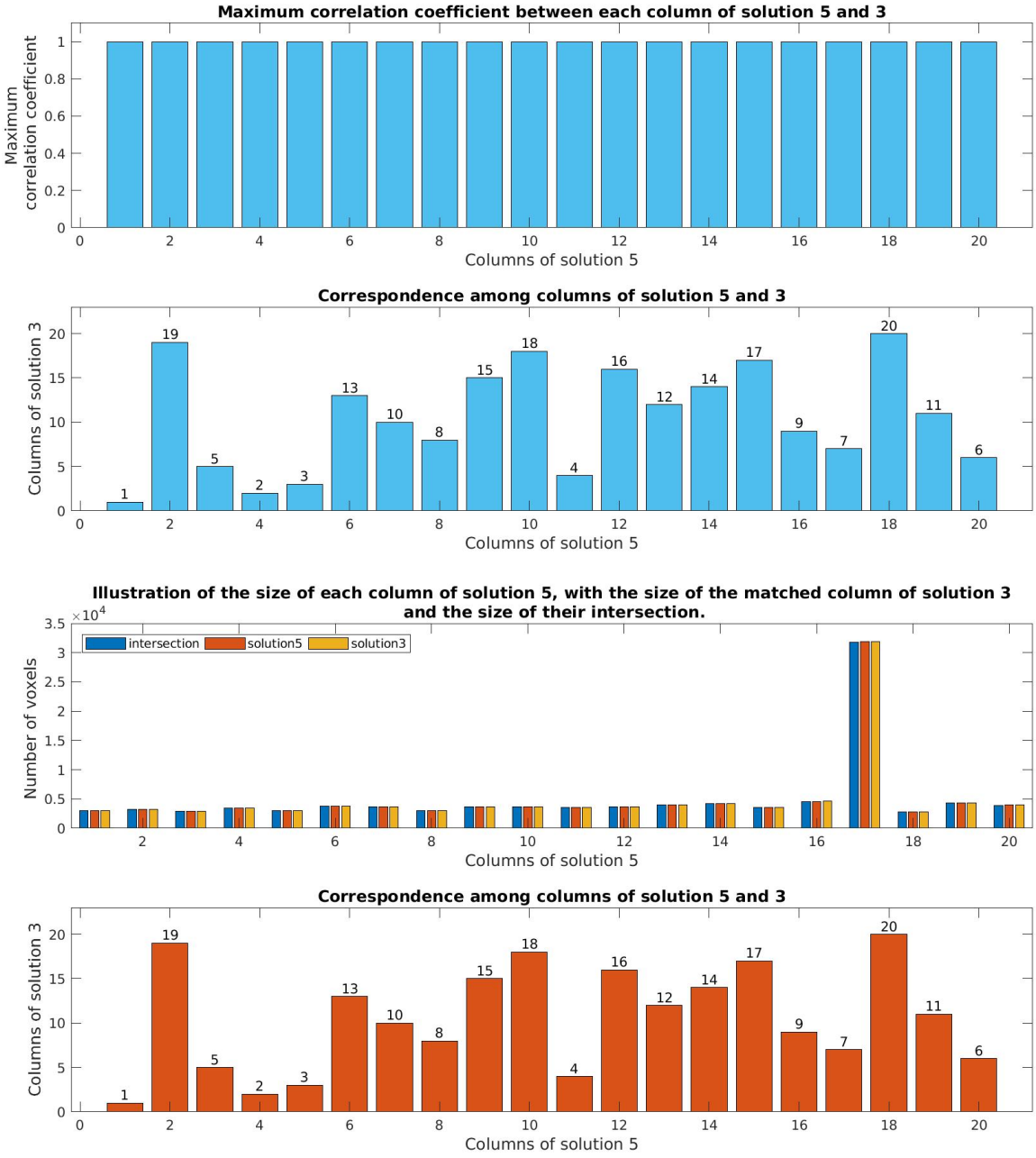


Figure A.129: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 3. We associate the column r of the solution 5 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on intersection. Solutions 5, 3 correspond to accelerated penalty method and SNR=0 dB.

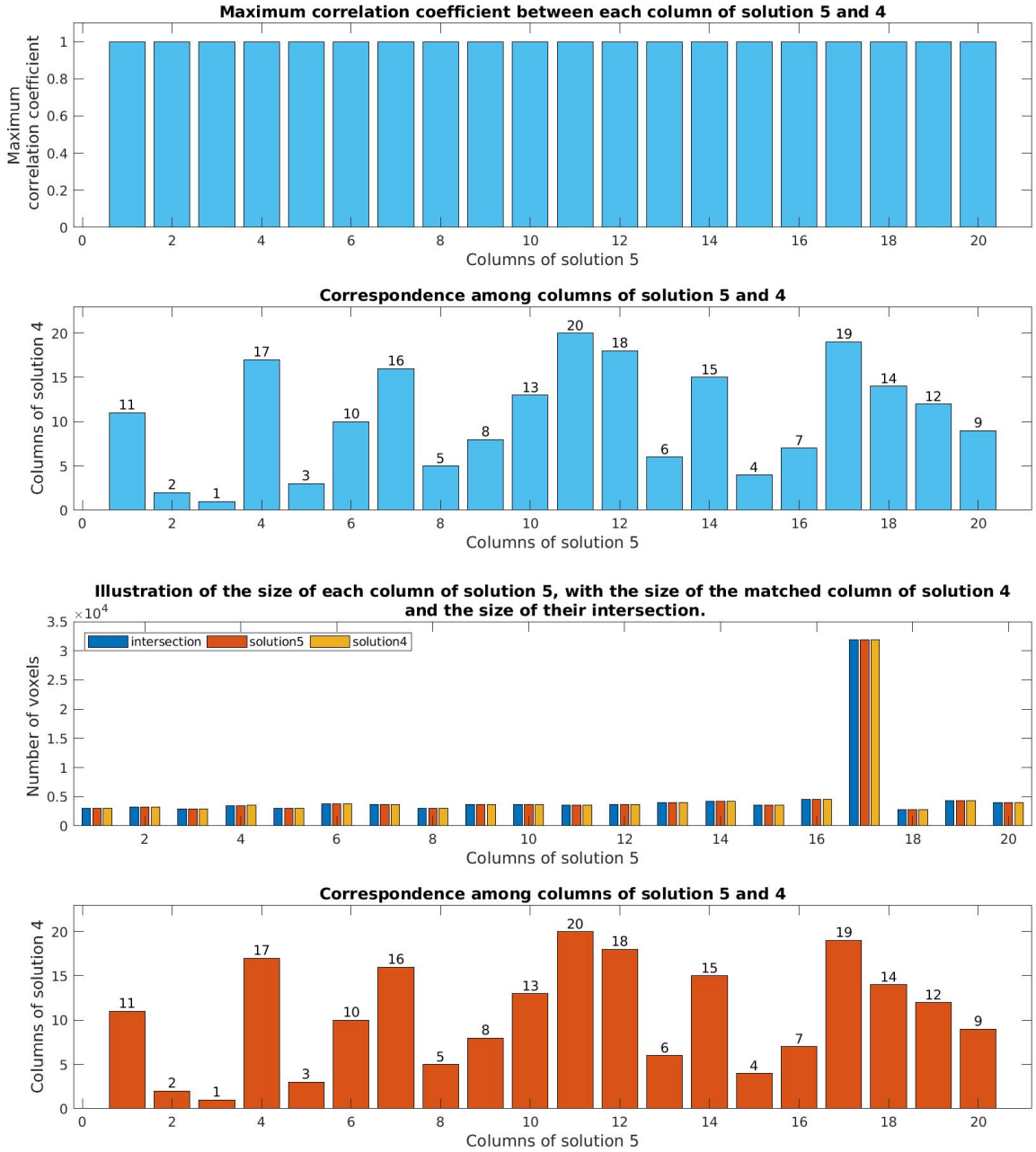


Figure A.130: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 4. We associate the column r of the solution 5 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on intersection. Solutions 5, 4 correspond to accelerated penalty method and SNR=0 dB.

· **Comparison of Accelerated penalty method solutions to real factor A.**

· **Noiseless case**

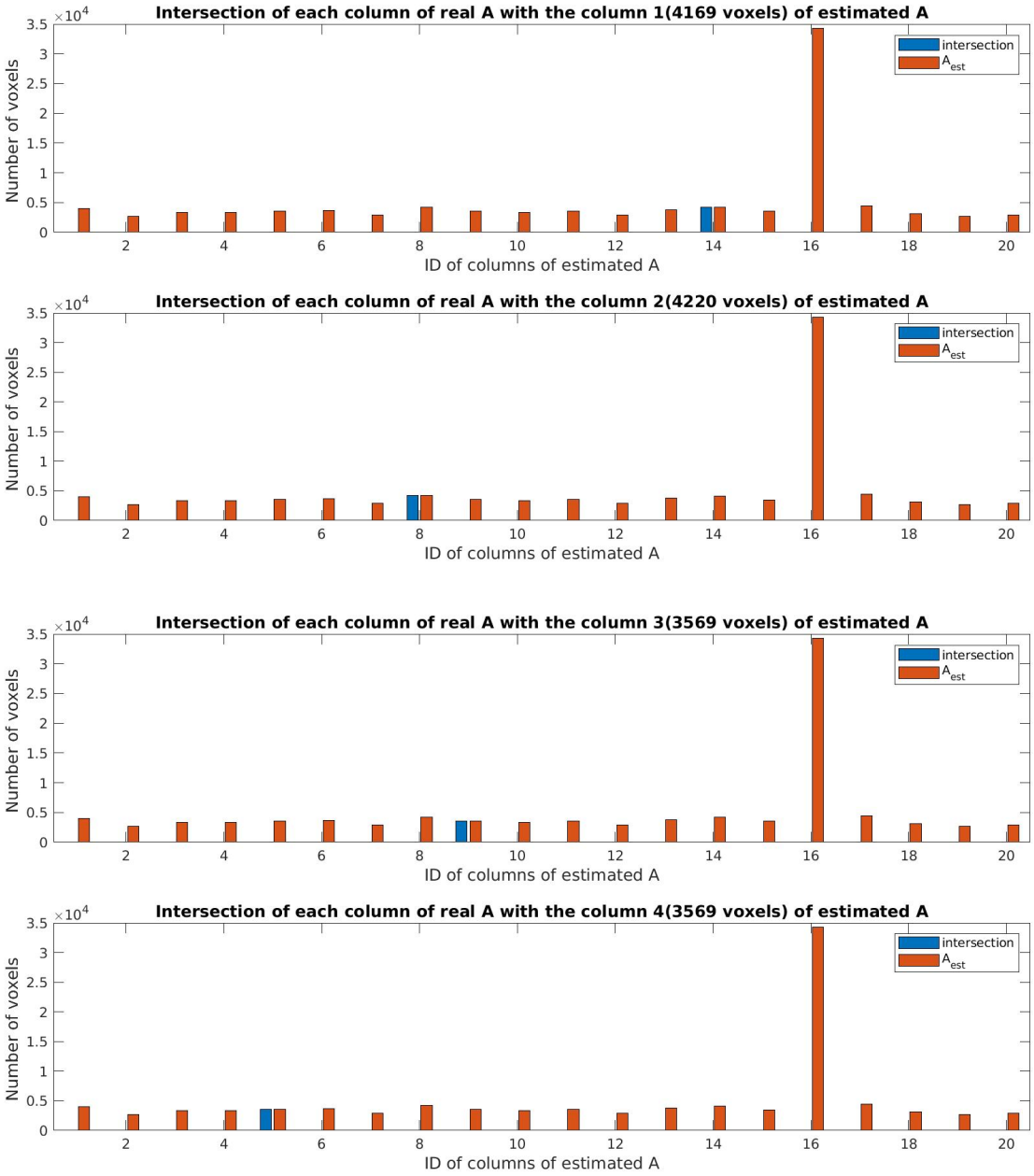


Figure A.131: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

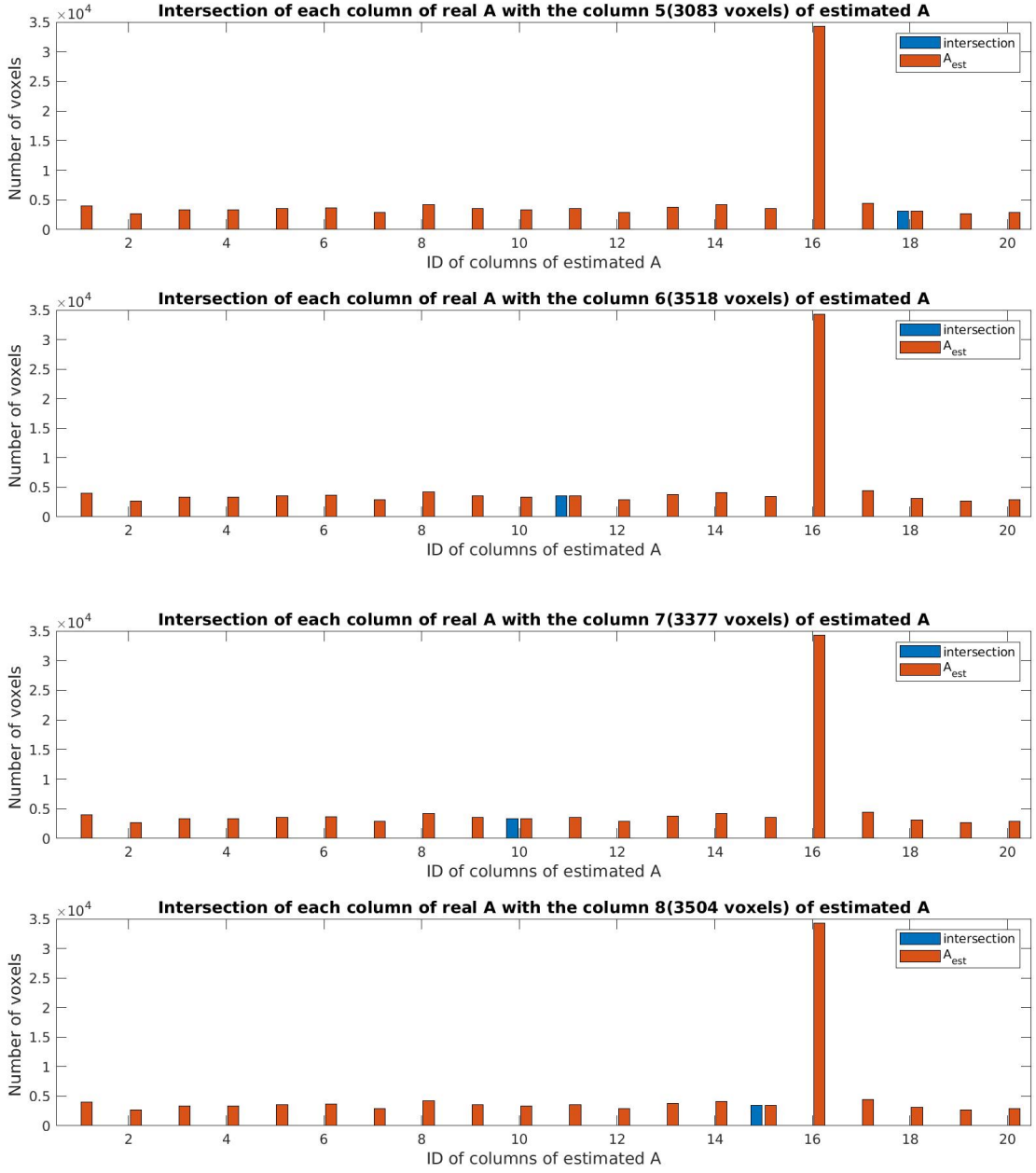


Figure A.132: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

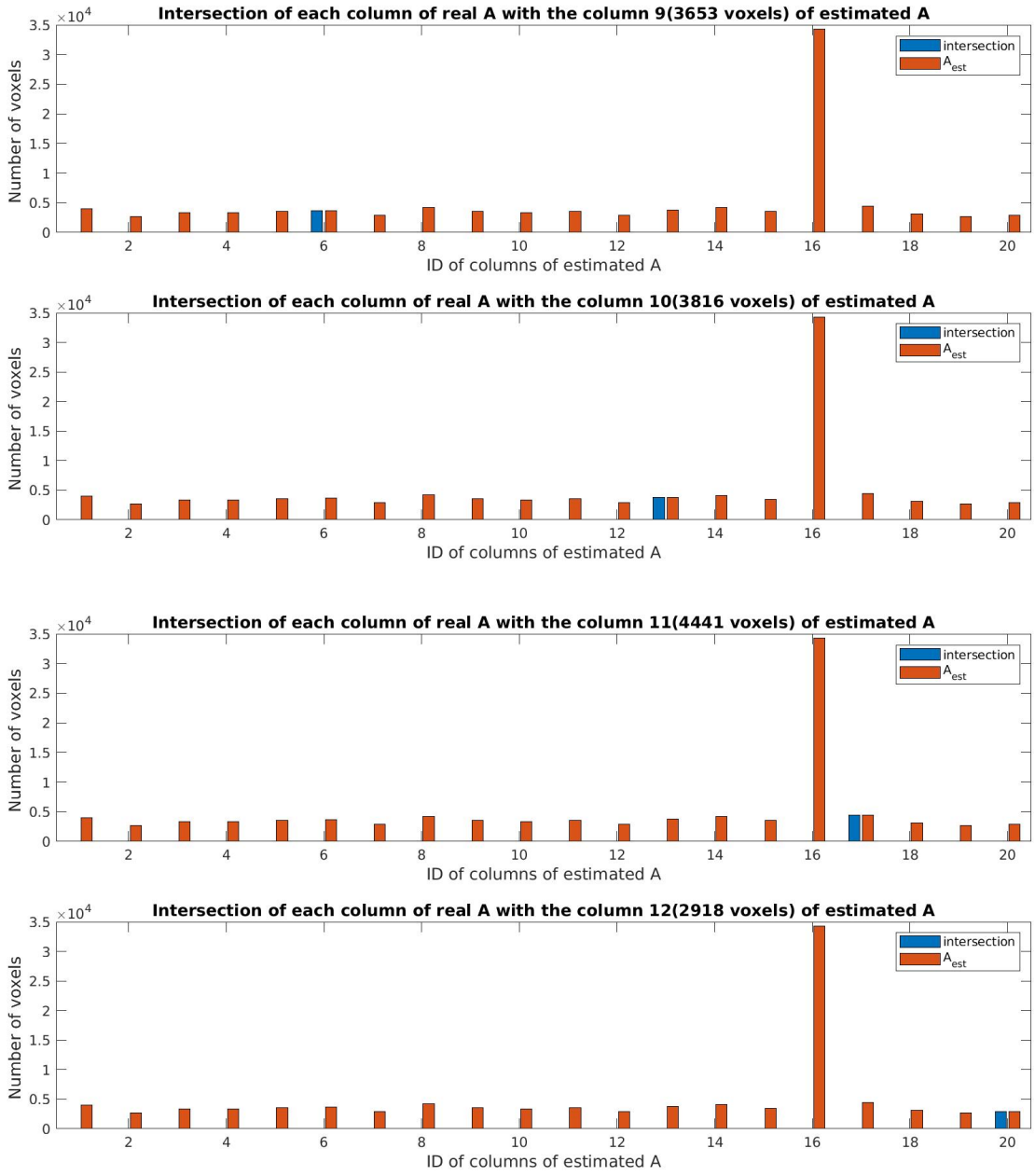


Figure A.133: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

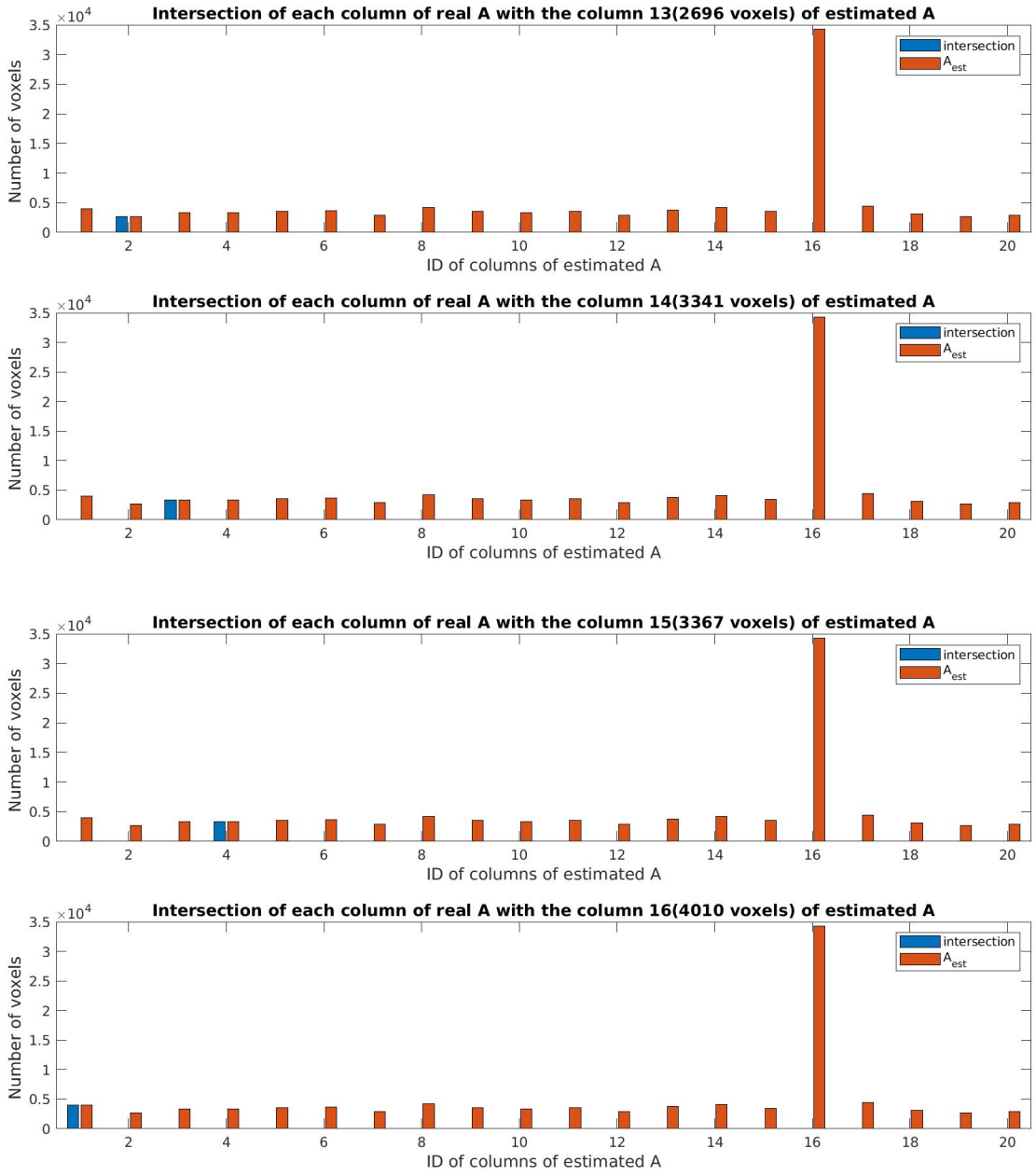


Figure A.134: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

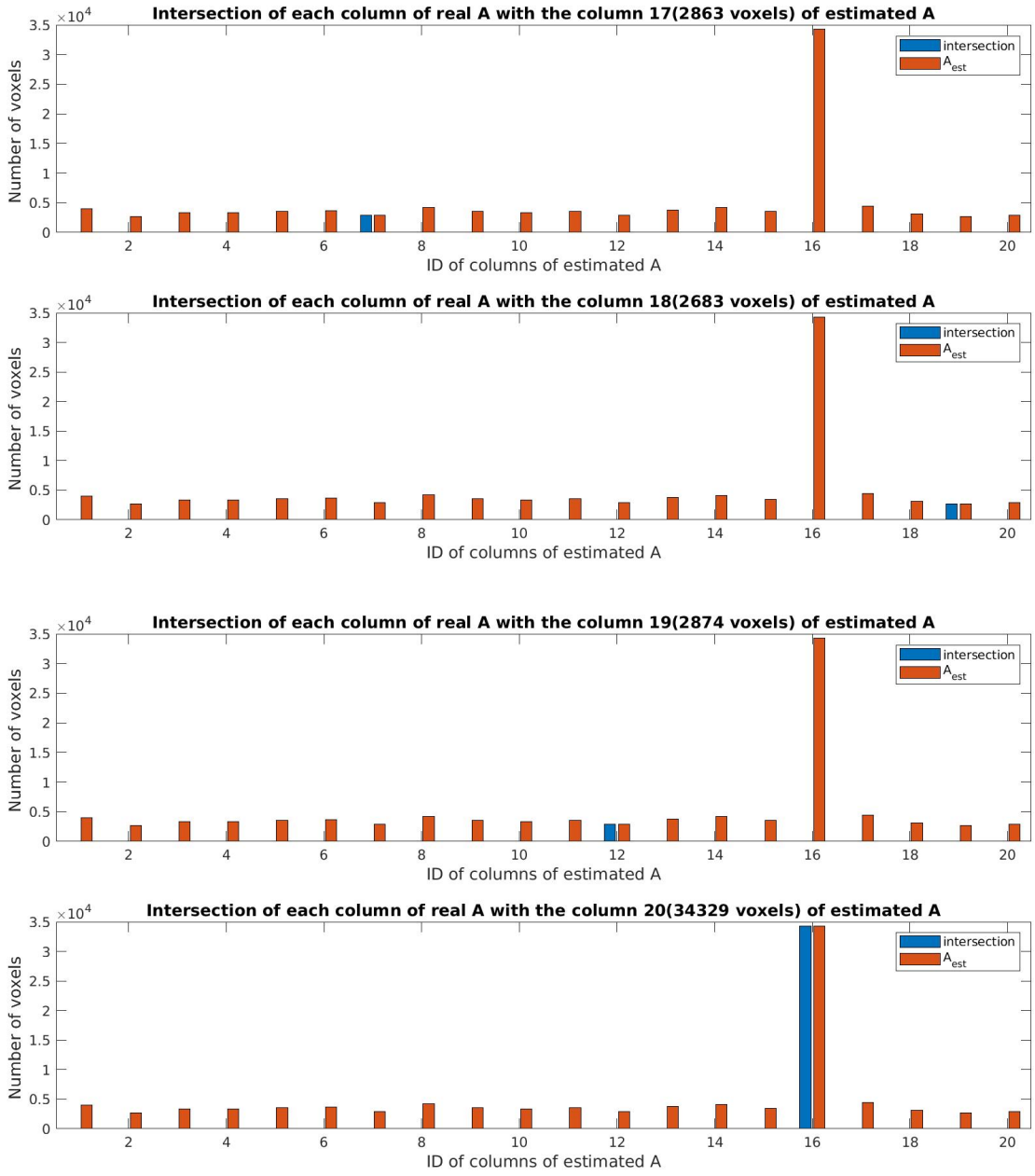


Figure A.135: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

· **SNR=0 dB**

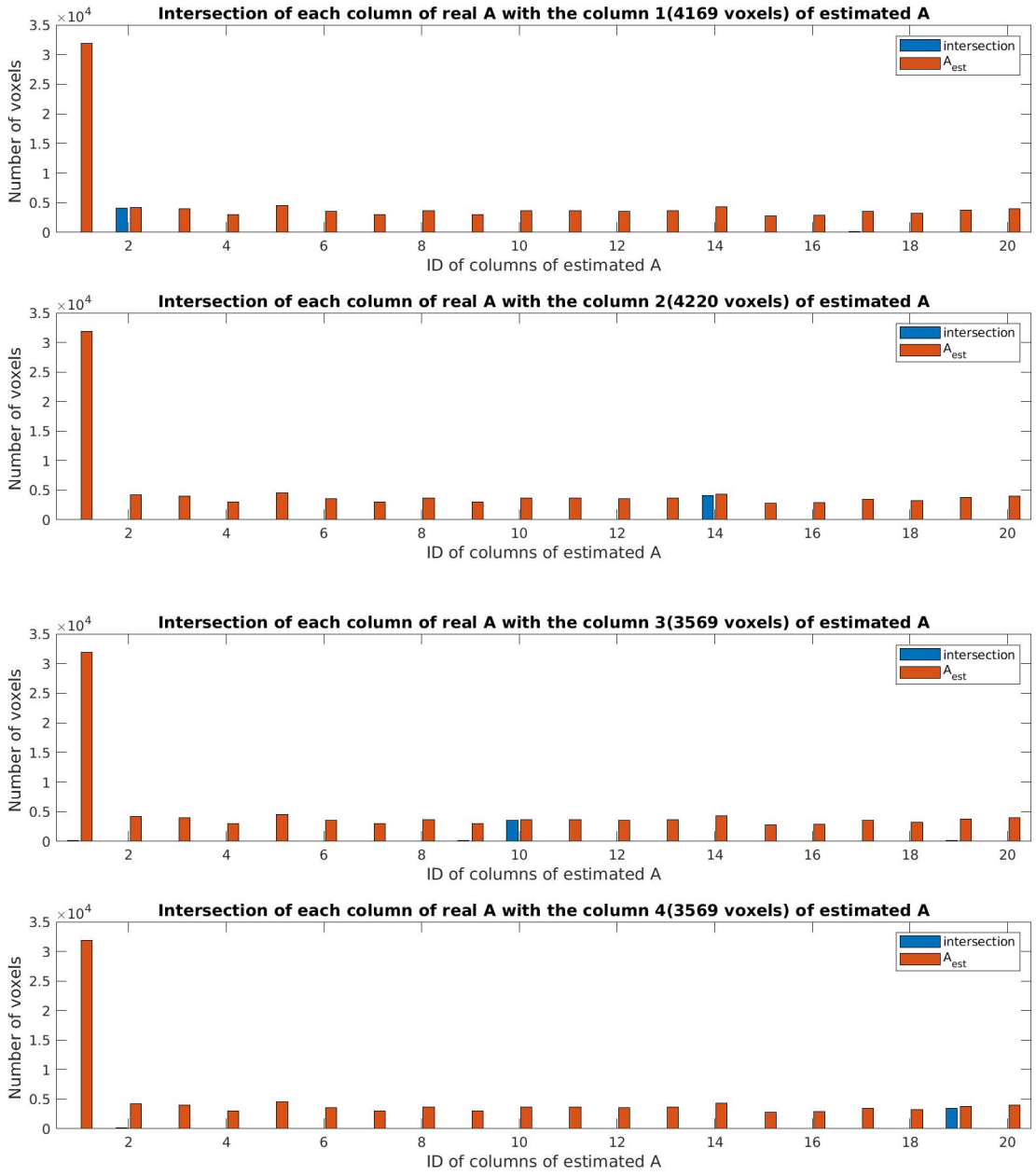


Figure A.136: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

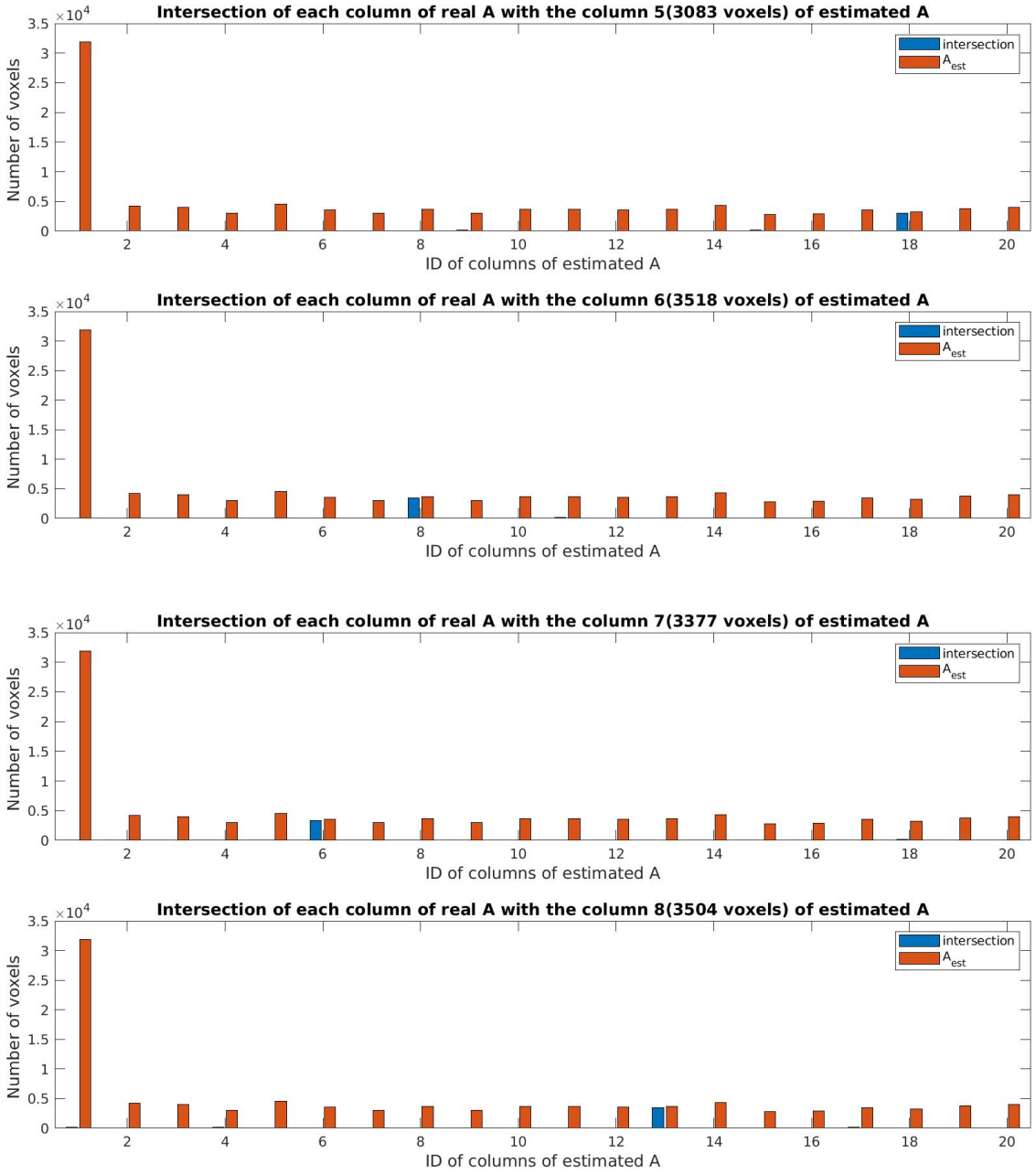


Figure A.137: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

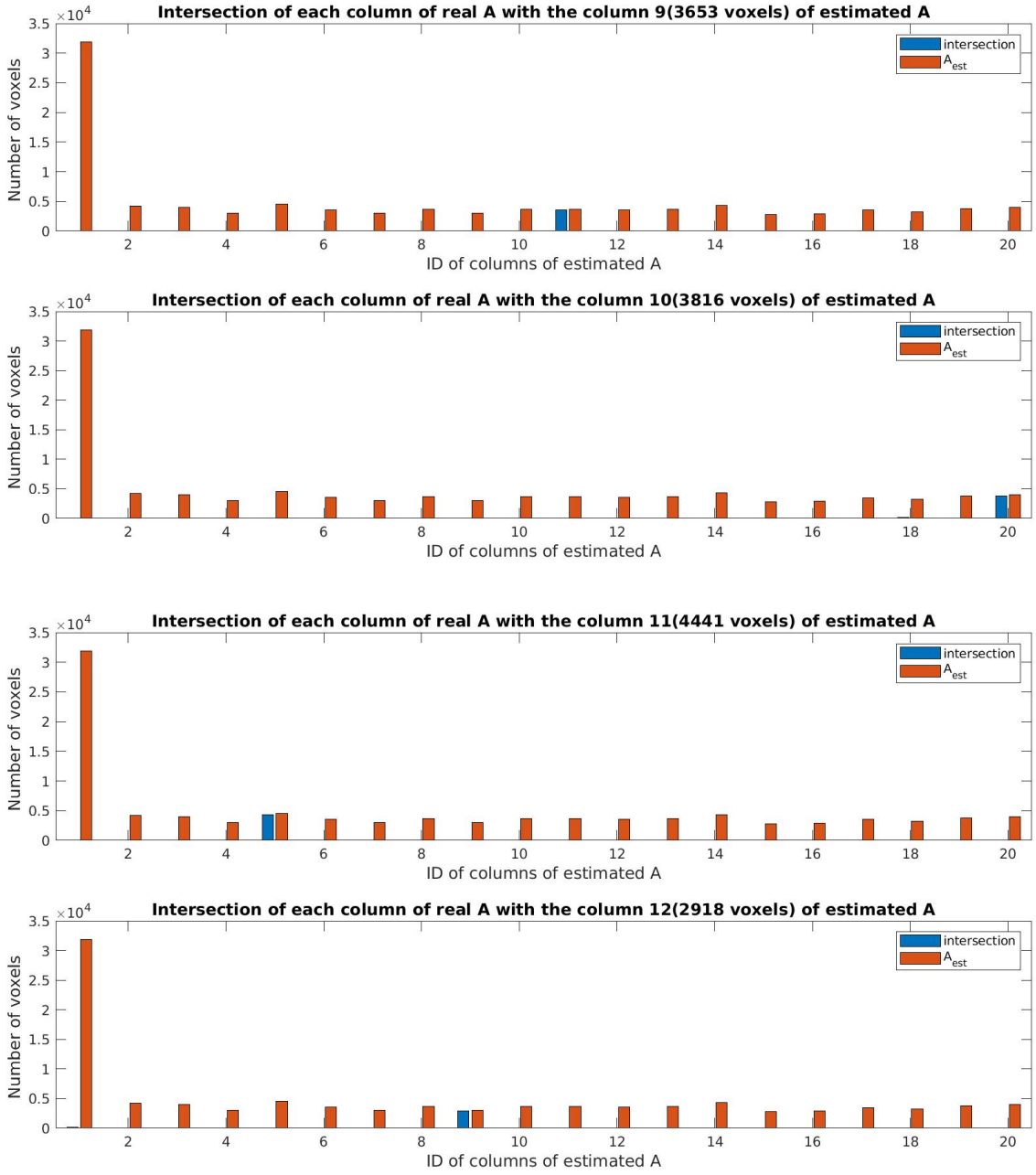


Figure A.138: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

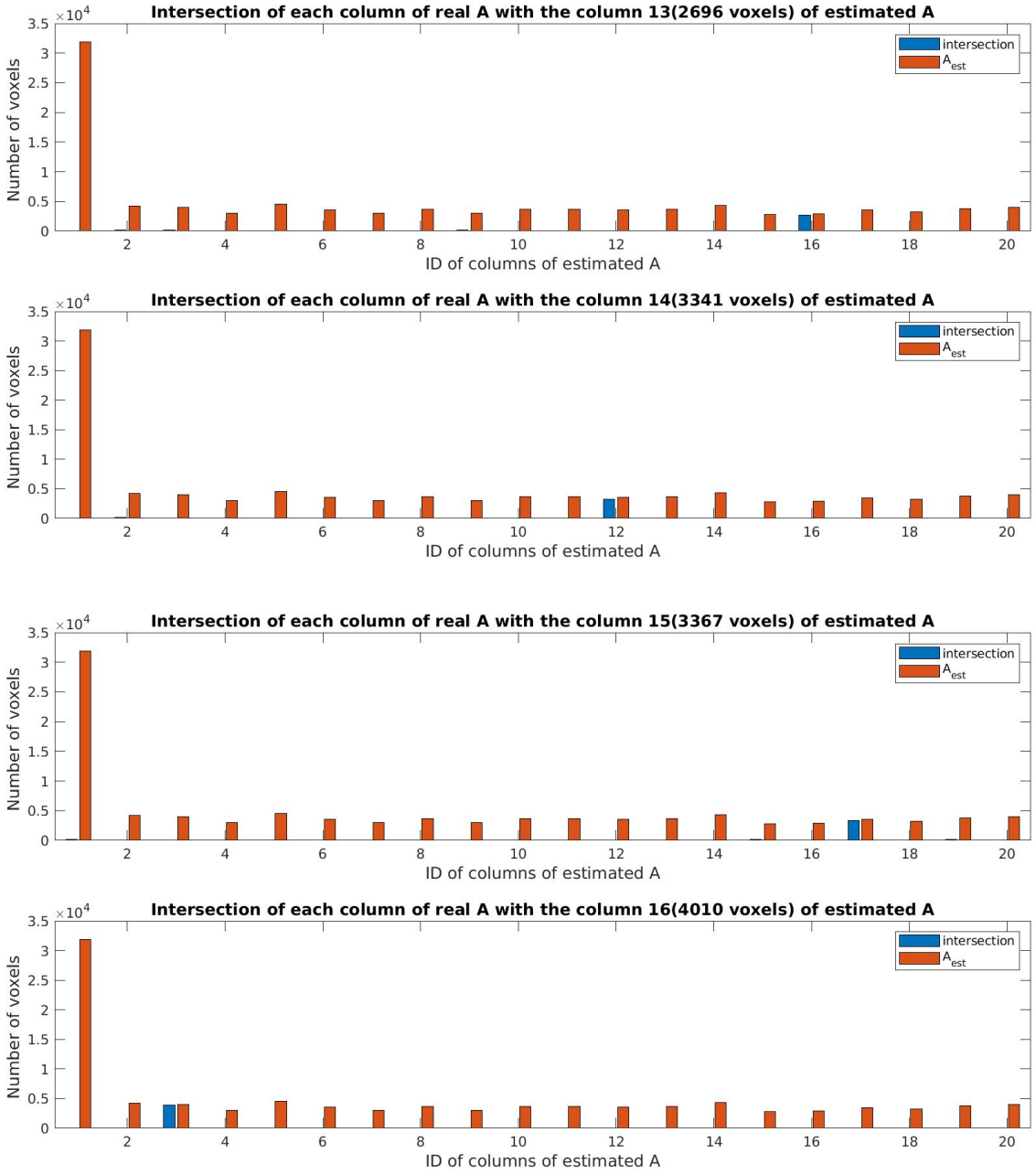


Figure A.139: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

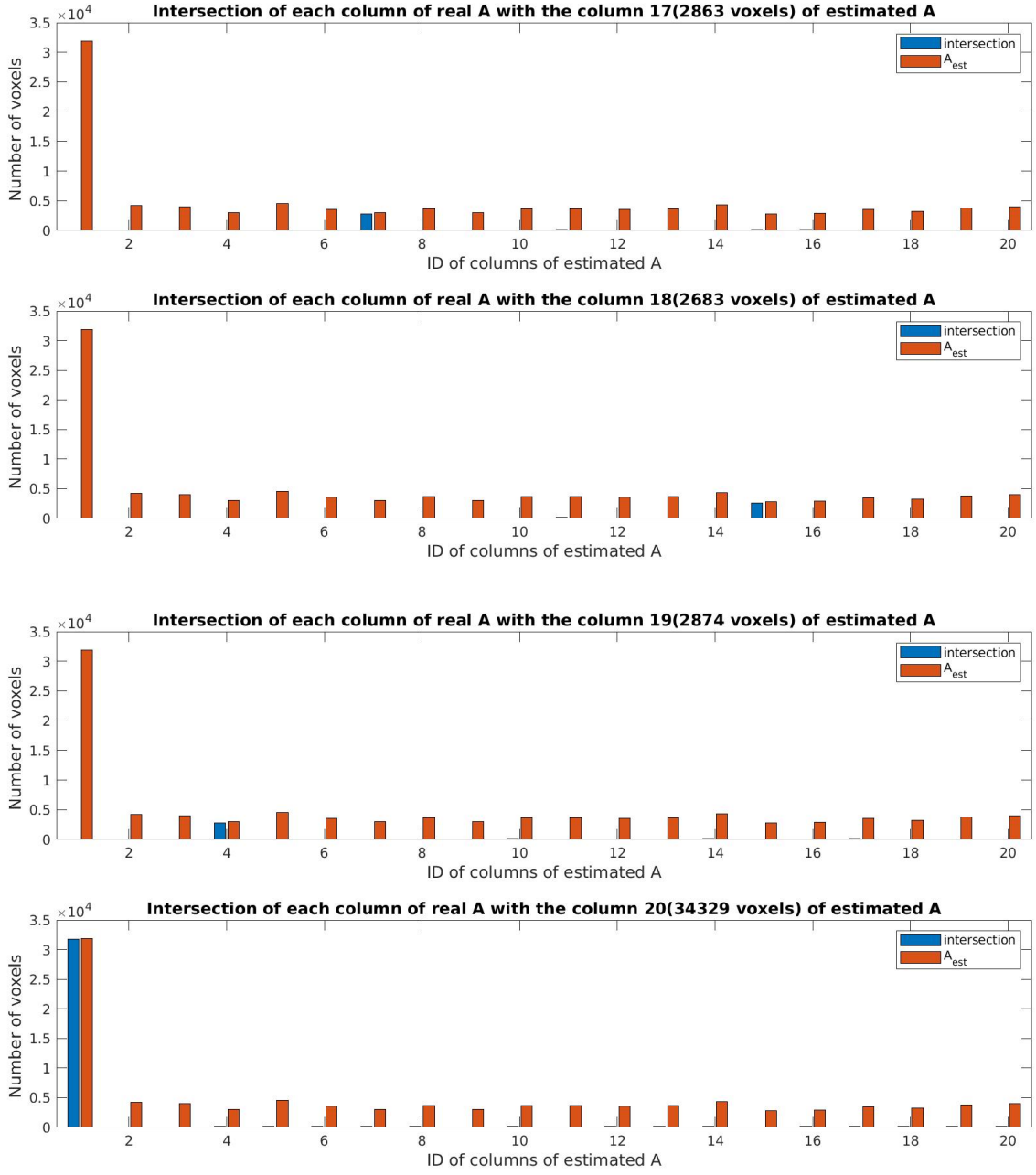


Figure A.140: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

· **Histograms for SNR=0 dB**

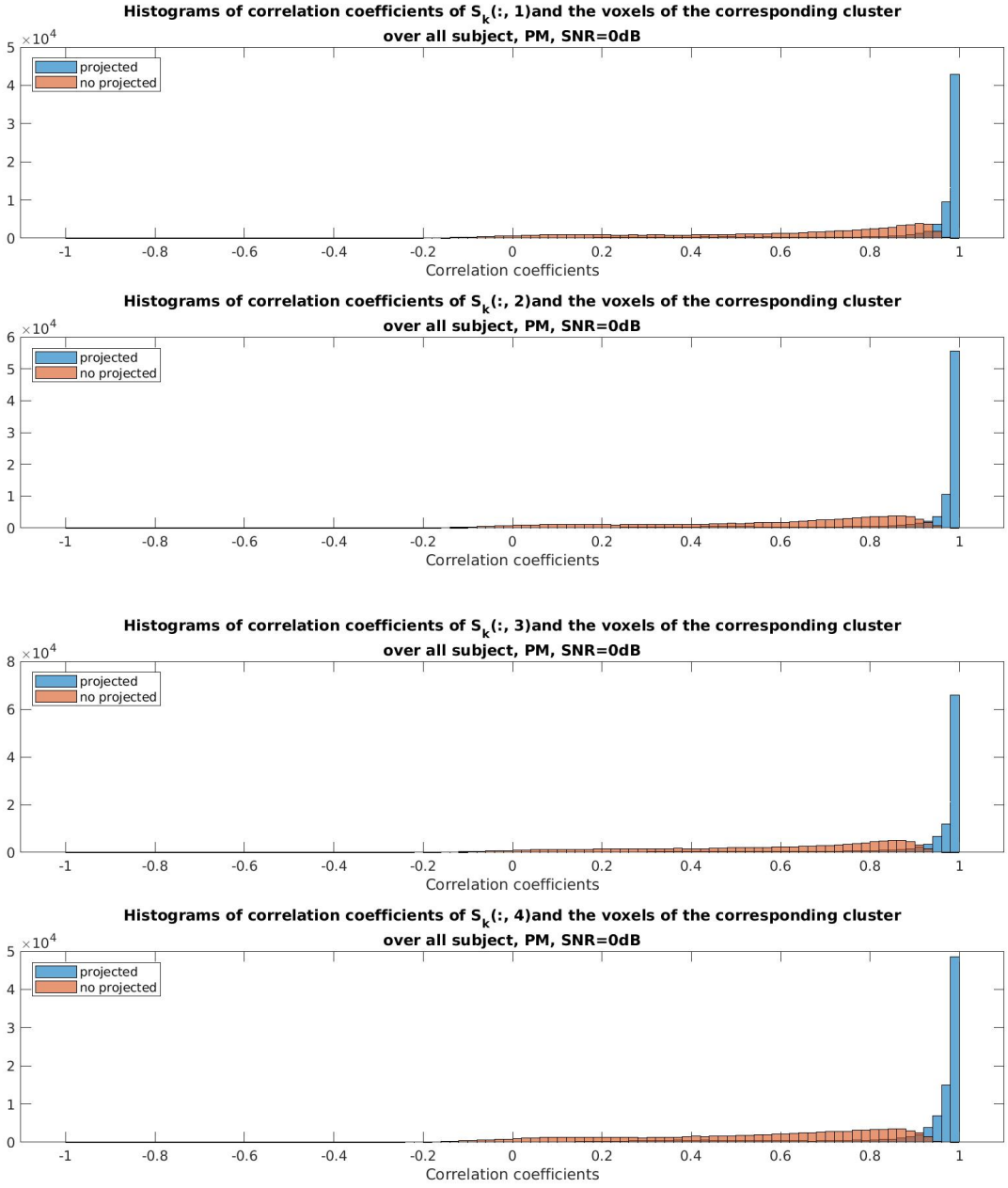


Figure A.141: For the column r and subject k, for $r = 1, \ldots, 4$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 1 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 2 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 3 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 4 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error.

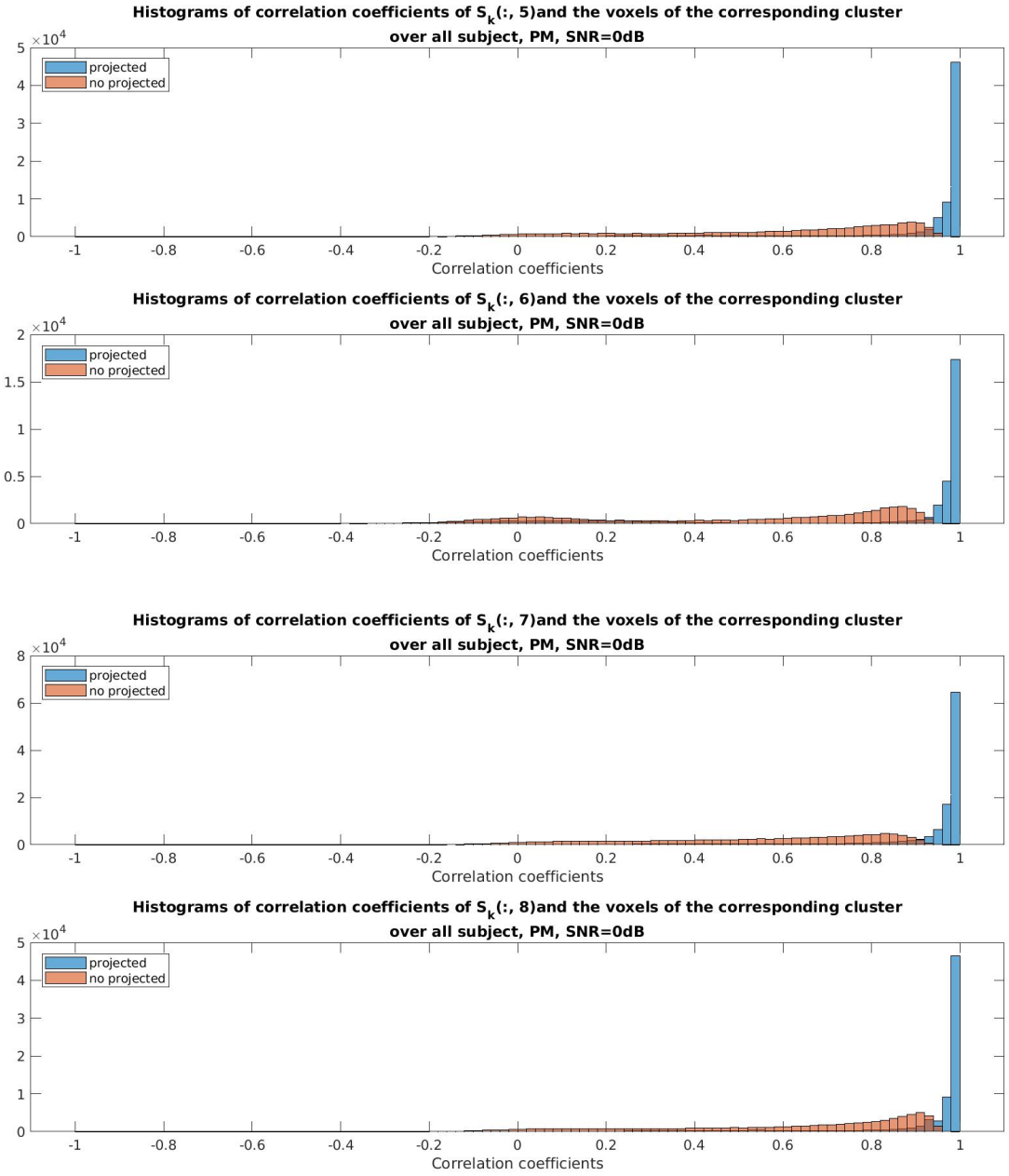


Figure A.142: For the column r and subject k, for $r = 5, \ldots, 8$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 5 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 6 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 7 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 8 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error.

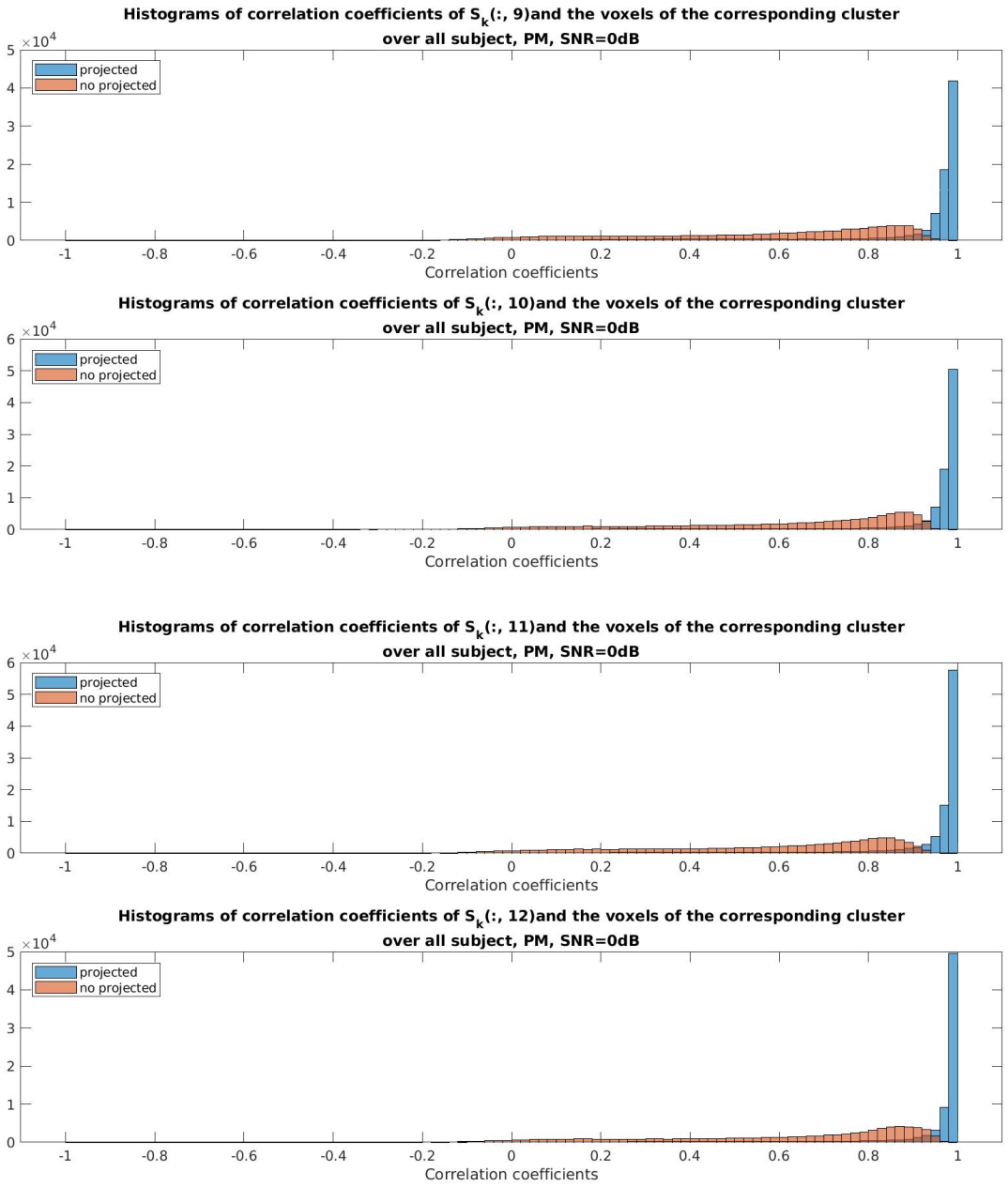


Figure A.143: For the column r and subject k, for $r = 9, \ldots, 12$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 9 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 10 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 11 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 12 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error.

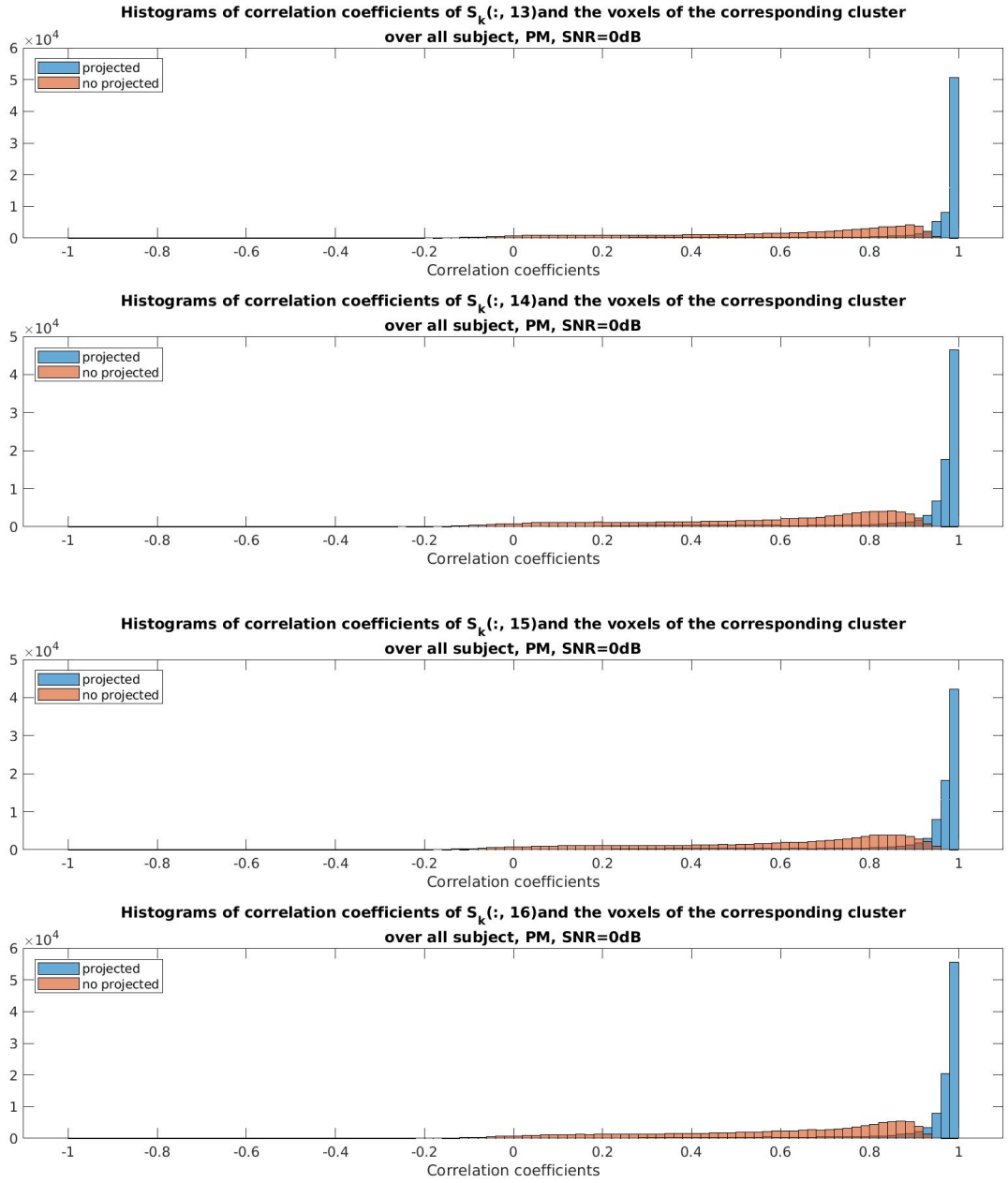


Figure A.144: For the column r and subject k, for $r = 13, \ldots, 16$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 13 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 14 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 15 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 16 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error.

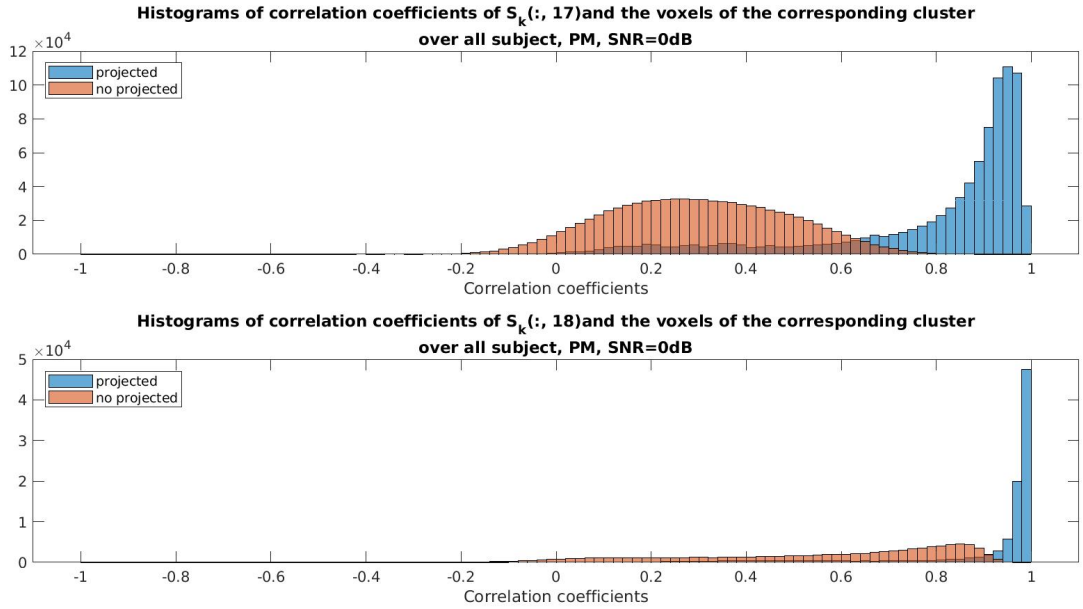


Figure A.145: For the column r and subject k, for $r = 17$, 18 and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 17 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 18 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by penalty method and corresponds to the best solution based on relative error.

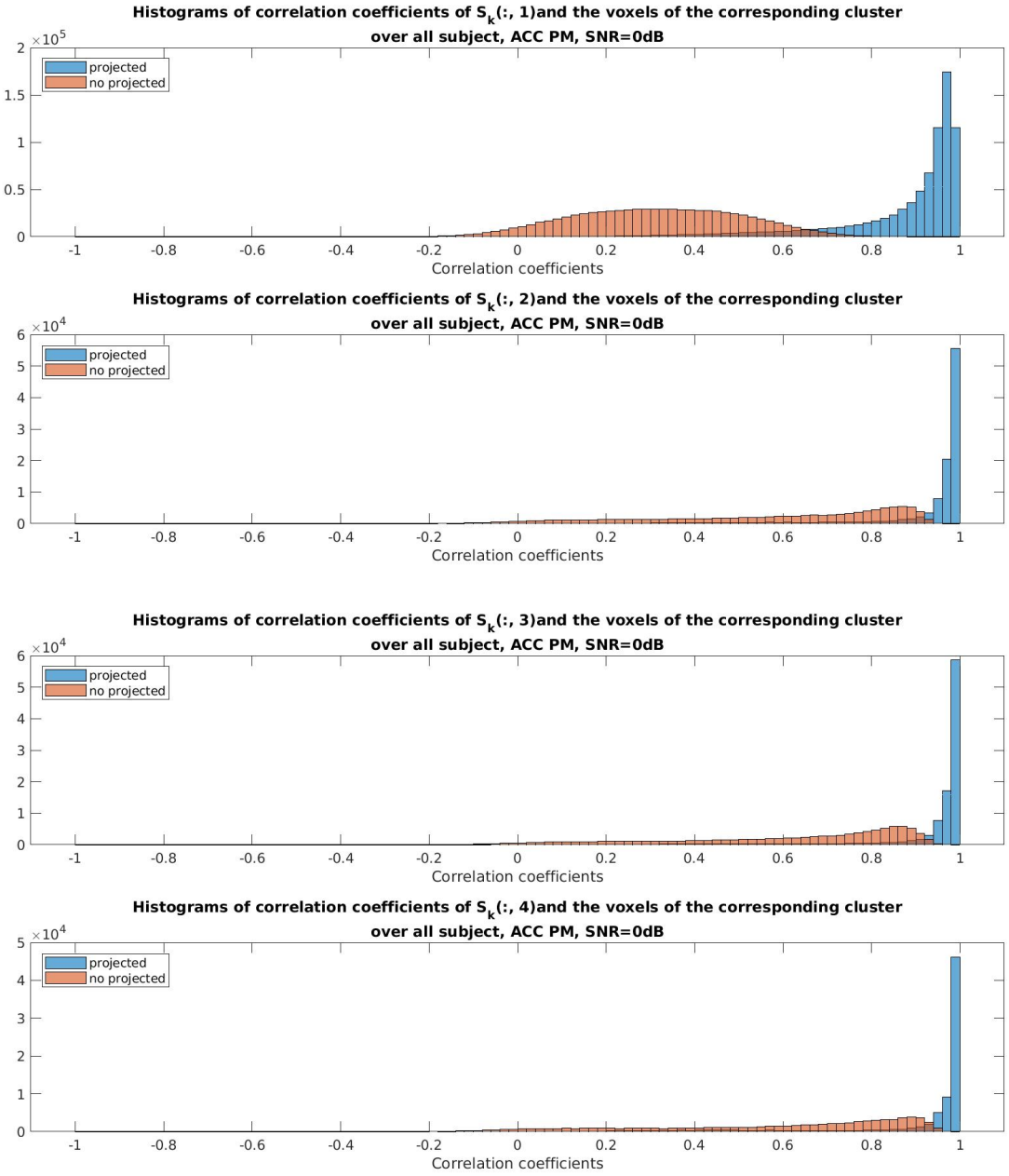


Figure A.146: For the column r and subject k, for $r = 1, \ldots, 4$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 1 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 2 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 3 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 4 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

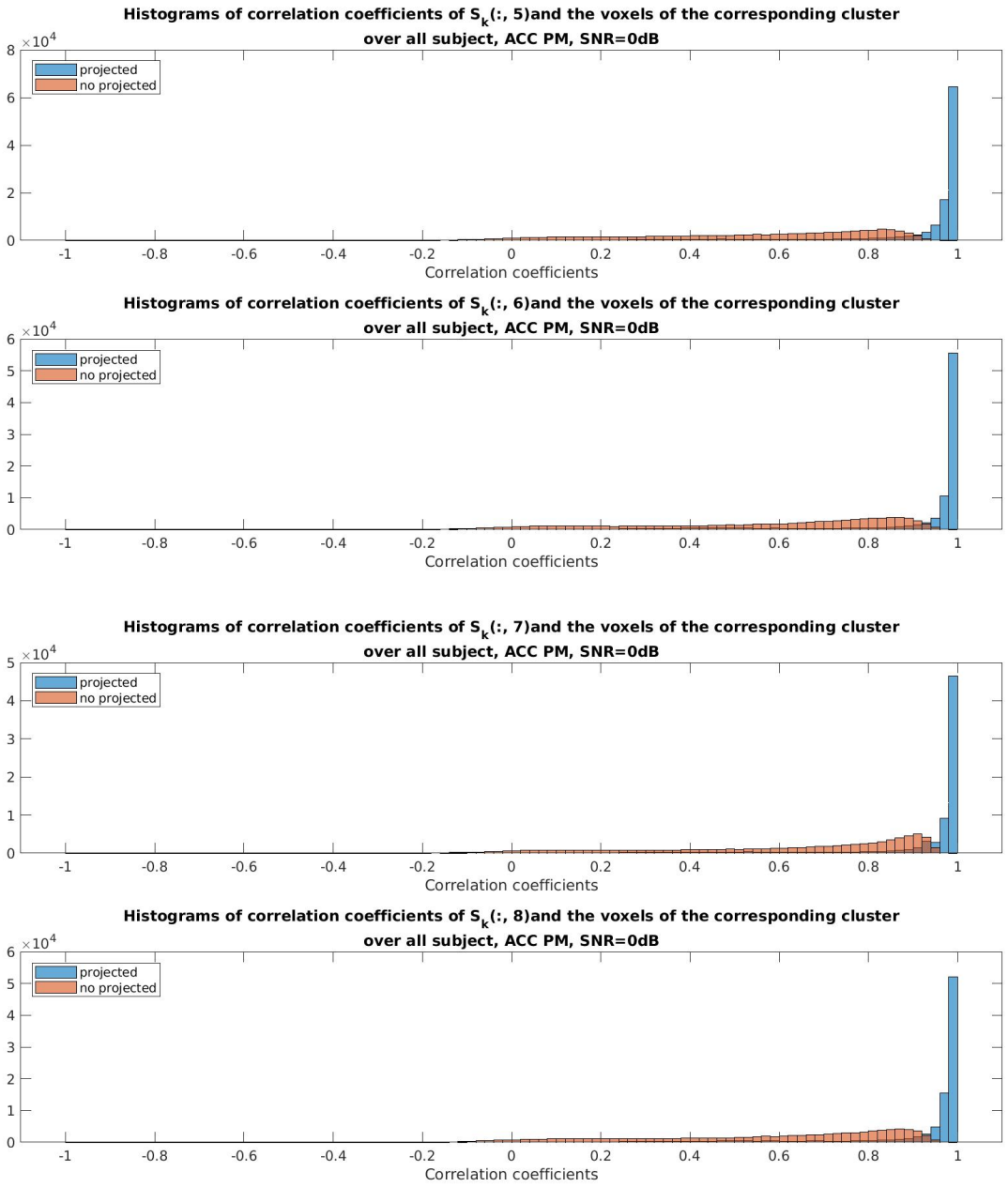


Figure A.147: For the column r and subject k, for $r = 5, \ldots, 8$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 5 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 6 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 7 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 8 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

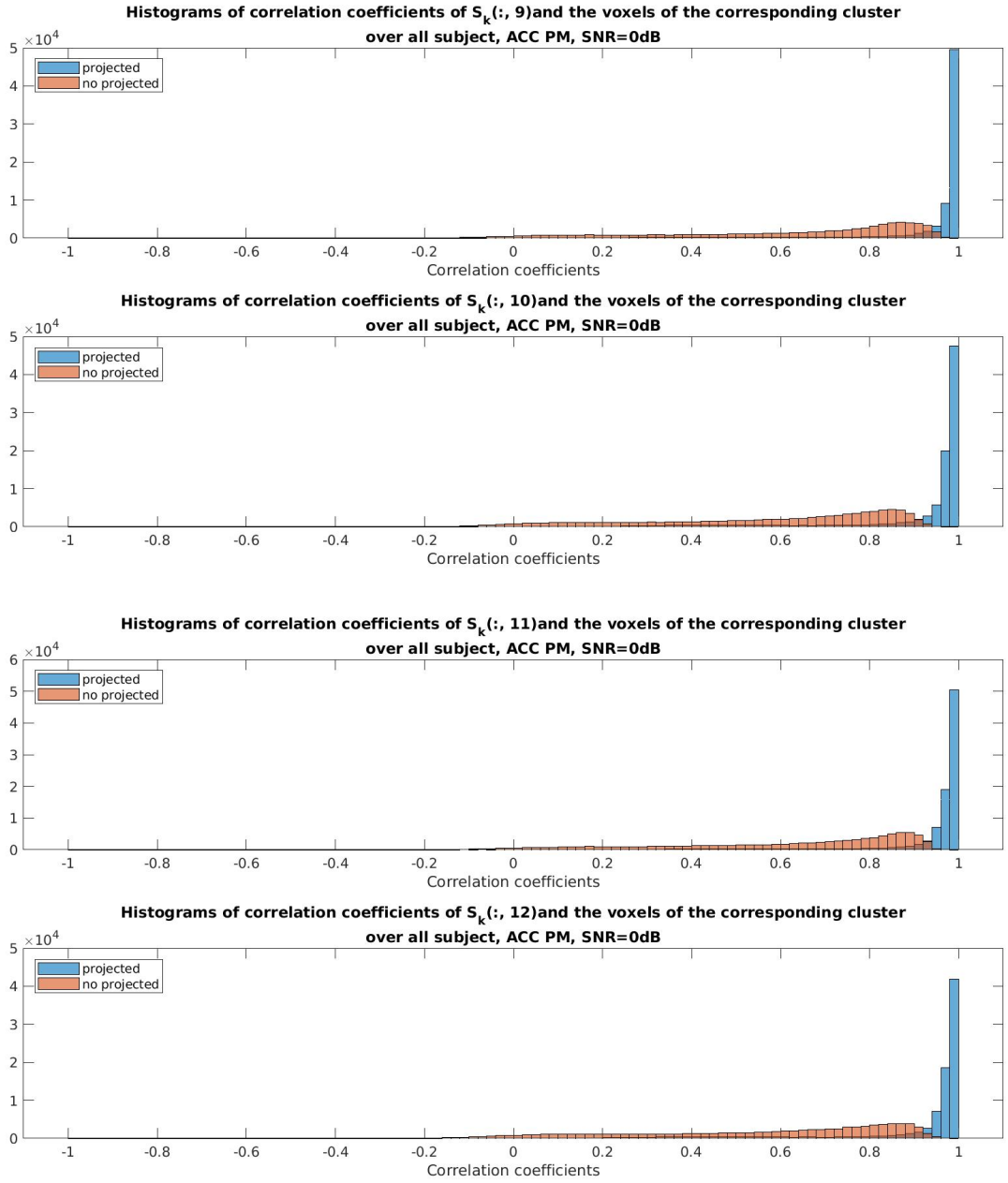


Figure A.148: For the column r and subject k, for $r = 9, \ldots, 12$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 9 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 10 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 11 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 12 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

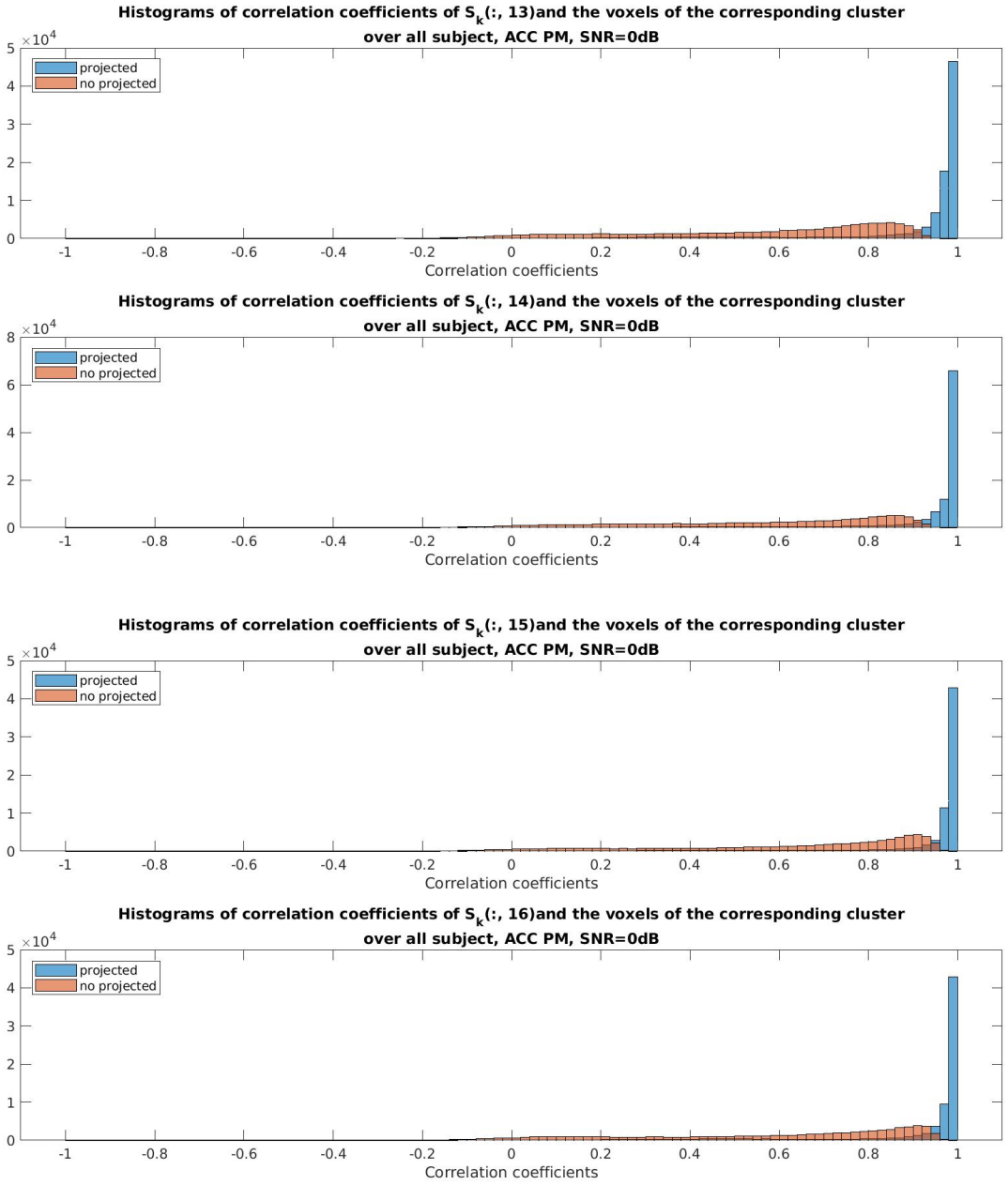


Figure A.149: For the column r and subject k, for $r = 13, \ldots, 16$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 13 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 14 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 15 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 16 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

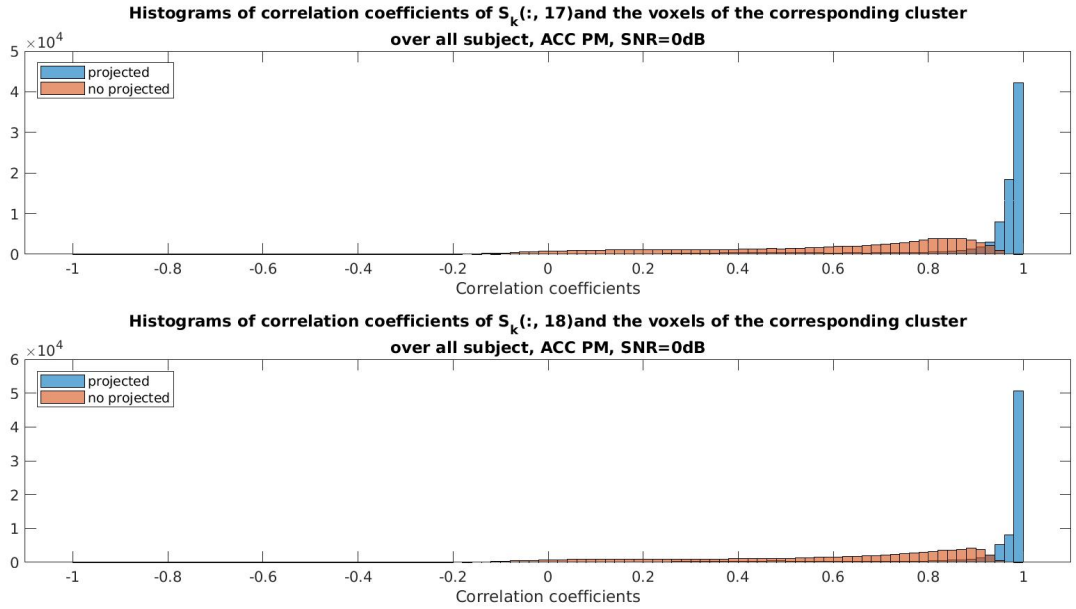


Figure A.150: For the column r and subject k, for $r = 17$, 18 and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 17 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 18 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

A.3.2 Case 2: Common orthonormal matrix A with each subject having correlated temporal components S_k.

Accelerated Penalty method

· **Comparison of Accelerated Penalty method solutions**

· **Noiseless data**

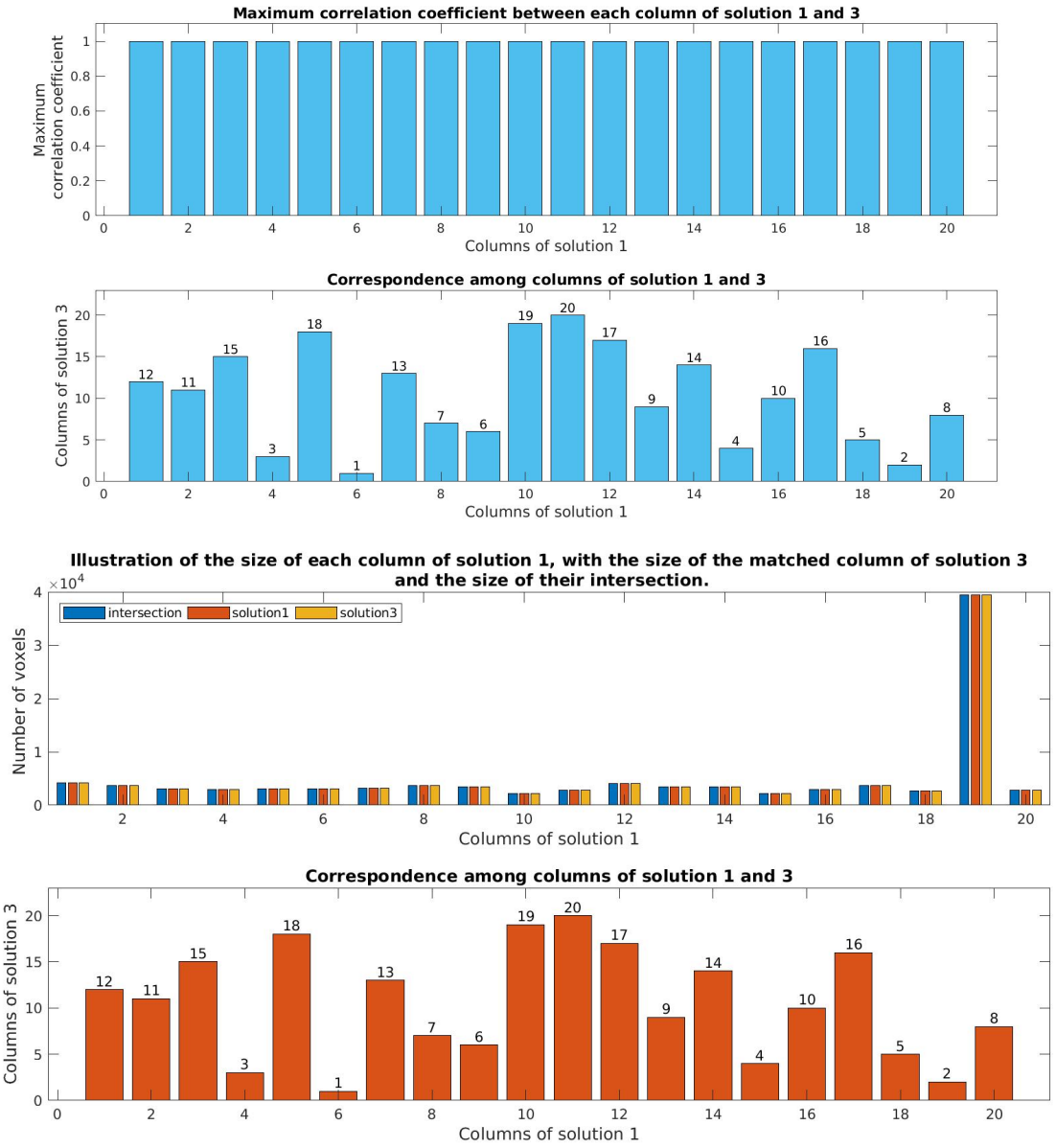


Figure A.151: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on intersection. Solutions 1, 3 correspond to accelerated penalty method and noiseless data.

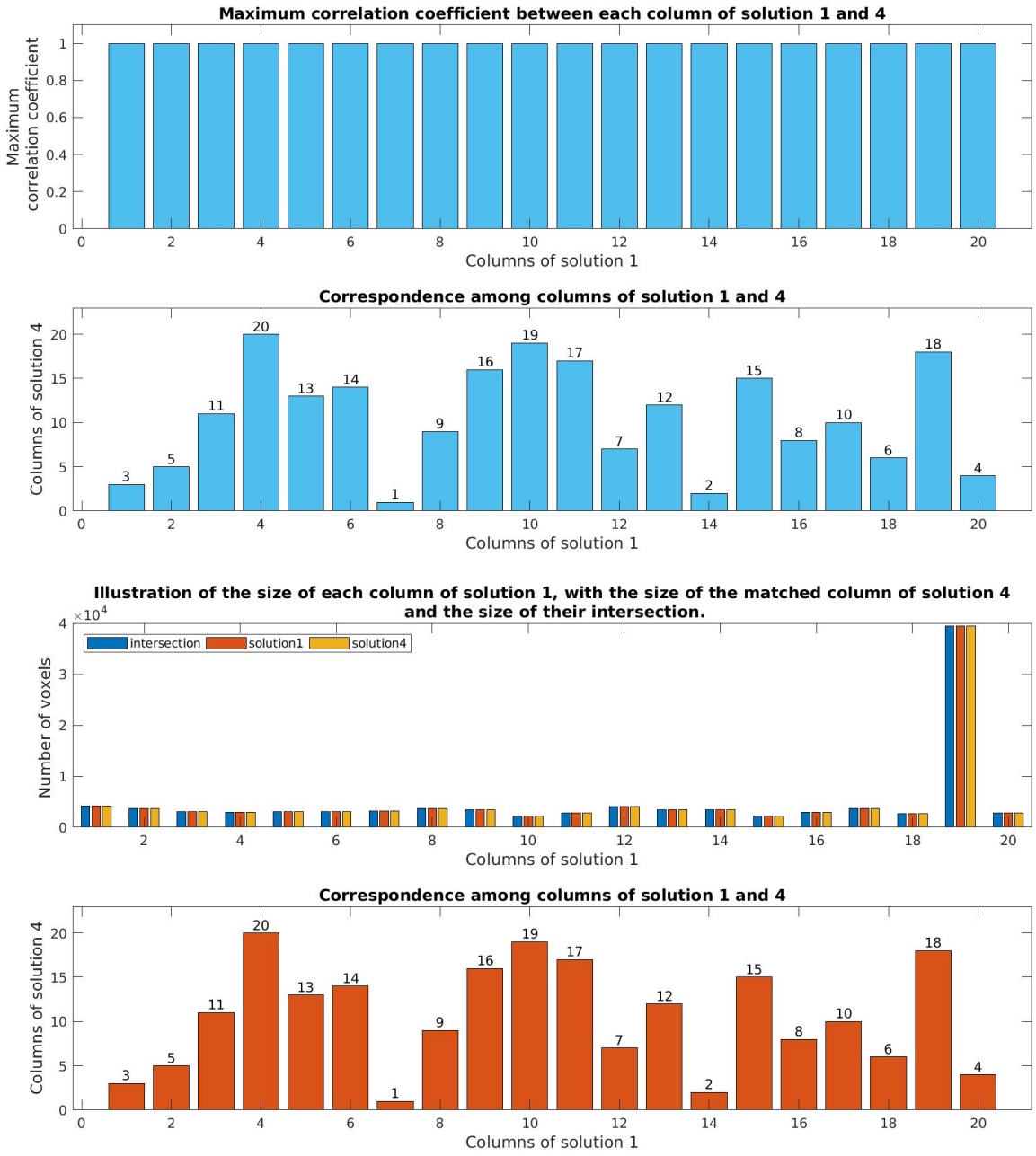


Figure A.152: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on intersection. Solutions 1, 4 correspond to accelerated penalty method and noiseless data.

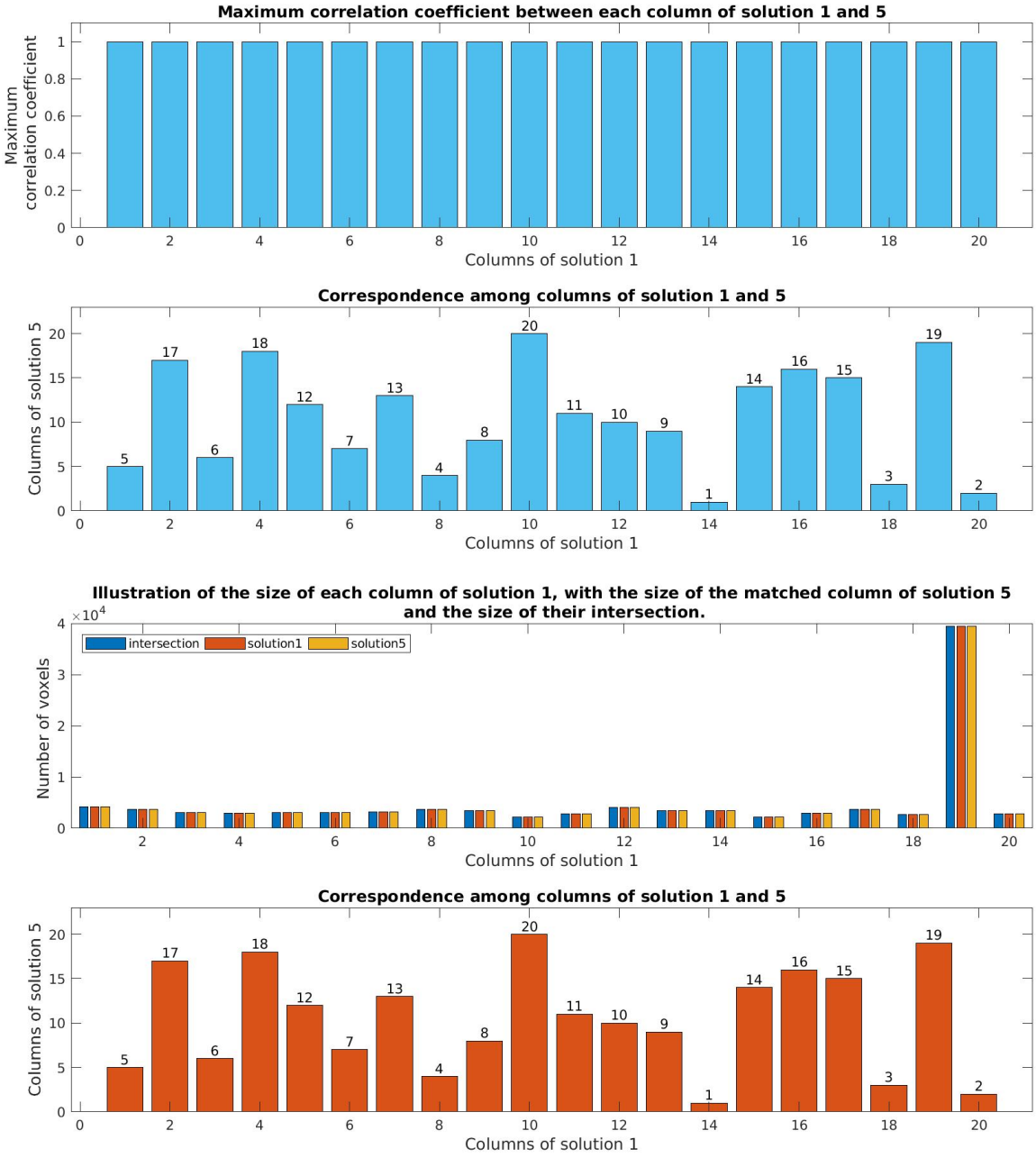


Figure A.153: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 5. We associate the column r of the solution 1 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on intersection. Solutions 1, 5 correspond to accelerated penalty method and noiseless data.

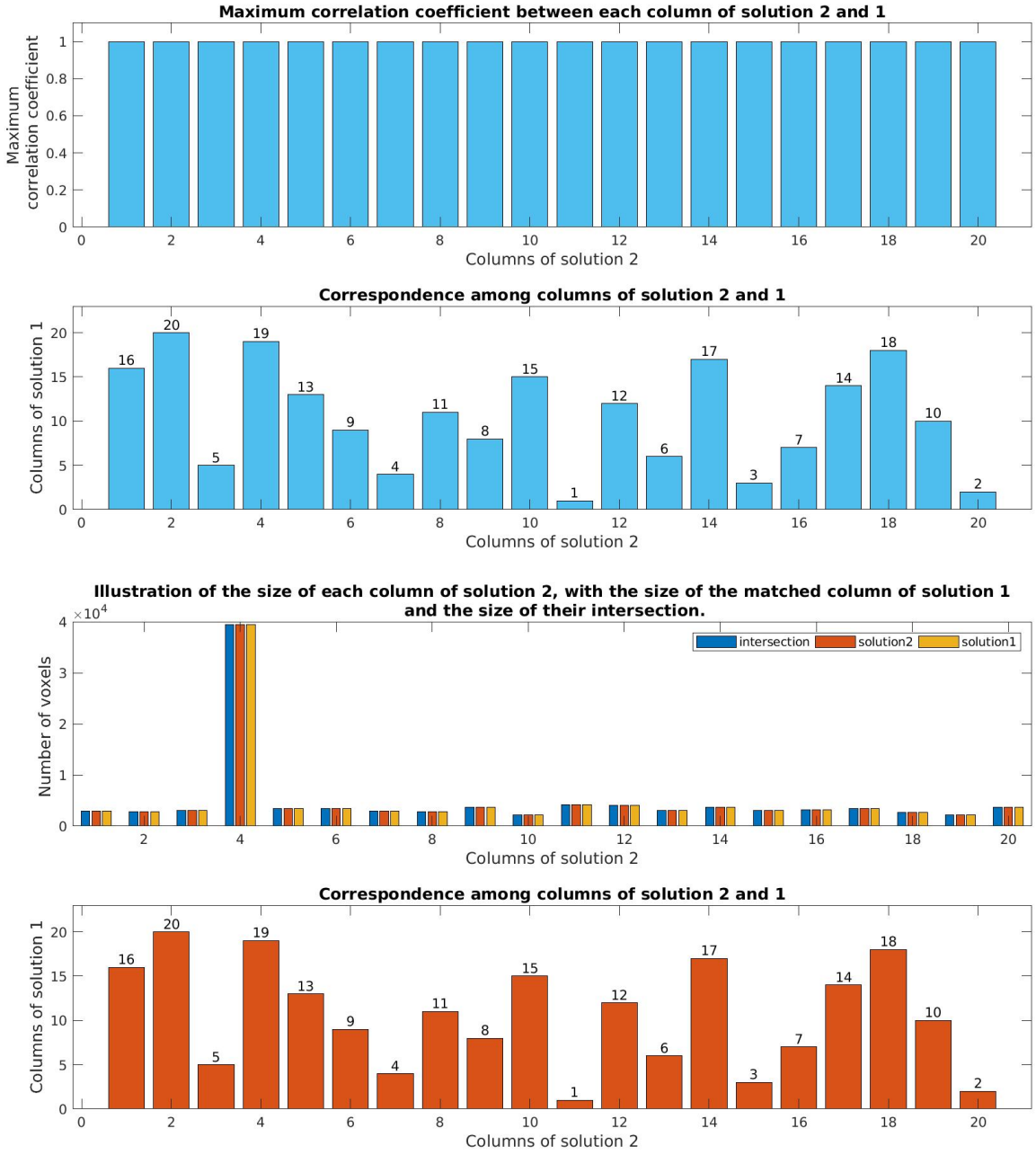


Figure A.154: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on intersection. Solutions 2, 1 correspond to accelerated penalty method and noiseless data.

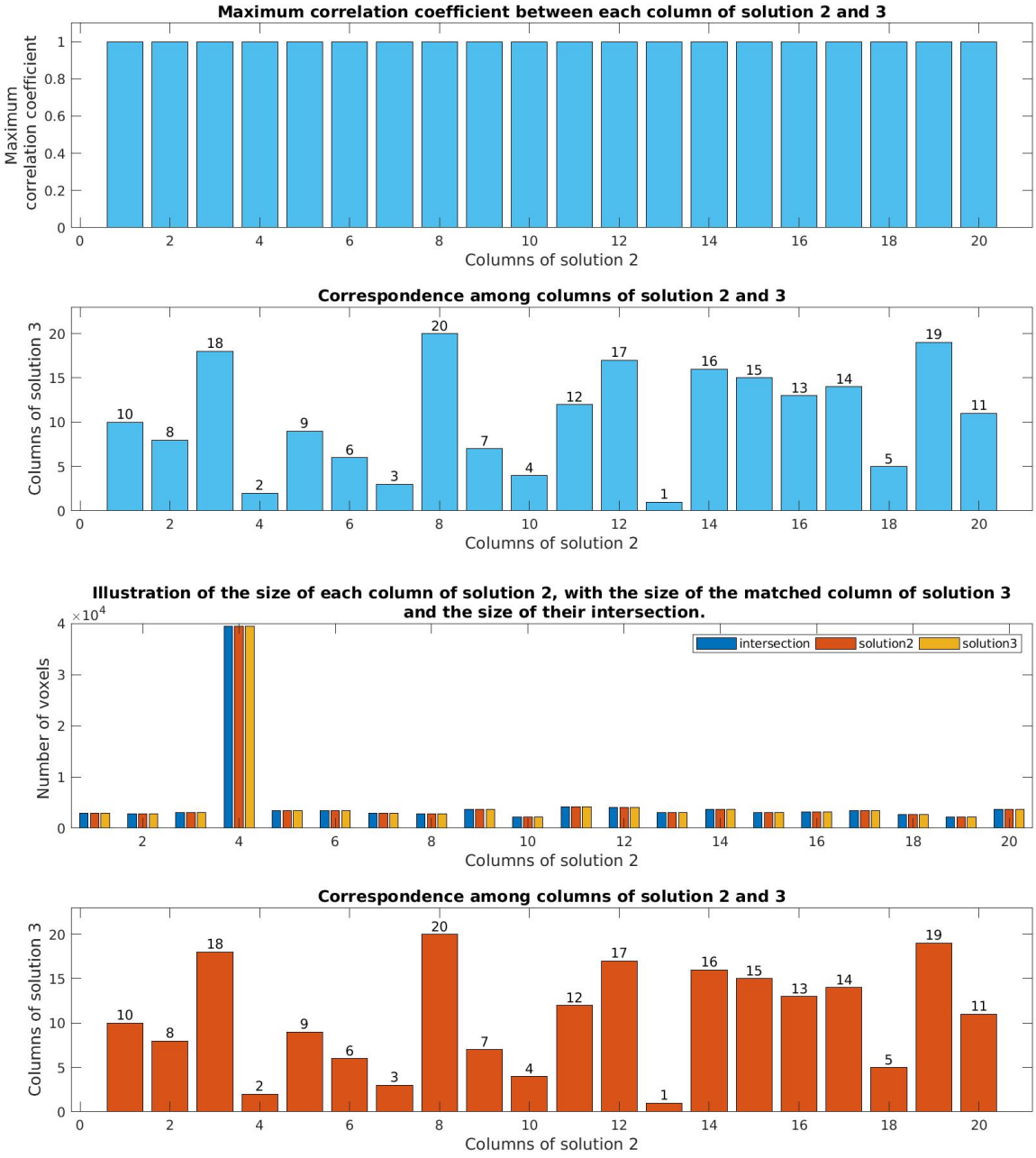


Figure A.155: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on intersection. Solutions 2, 3 correspond to accelerated penalty method and noiseless data.

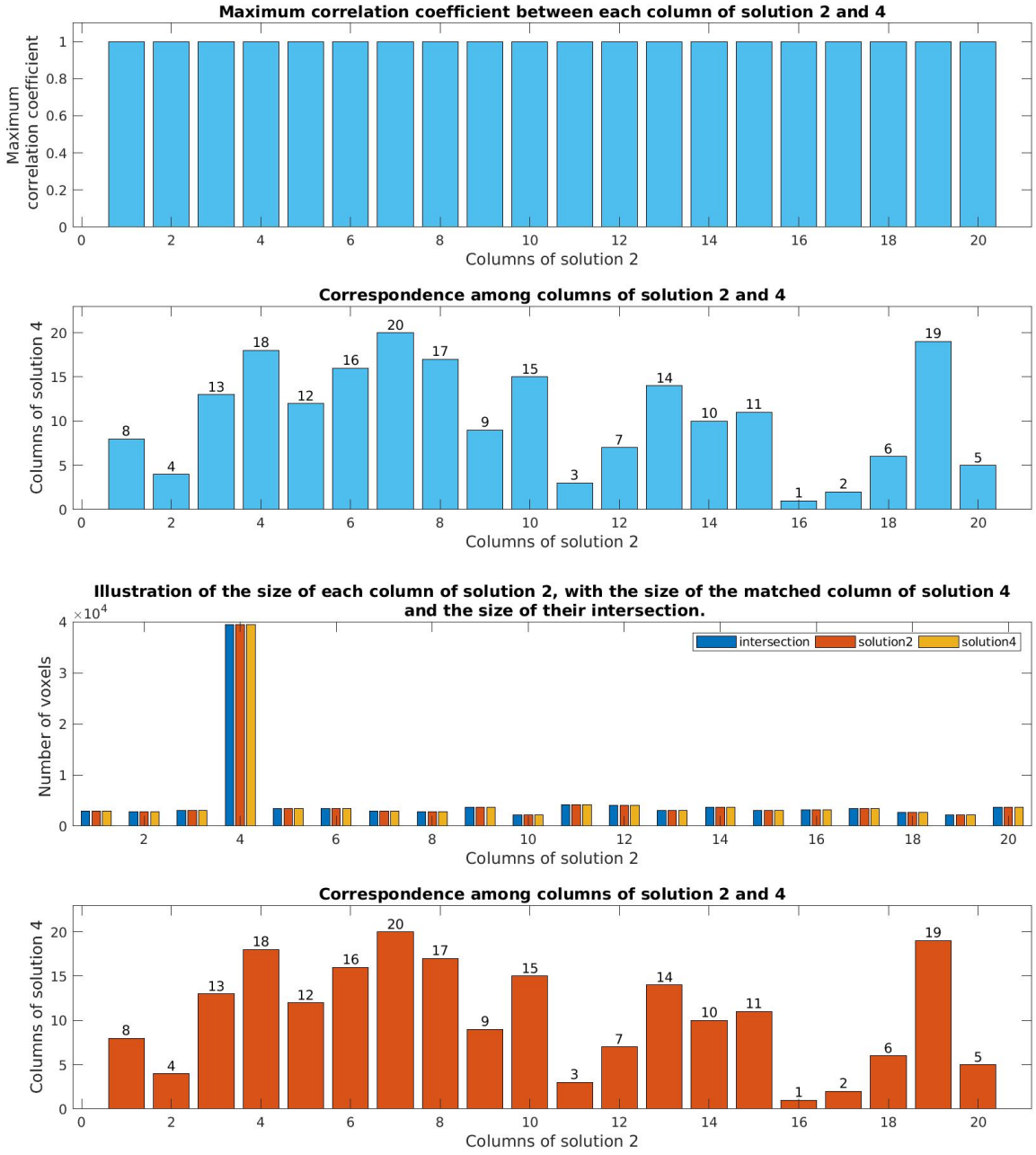


Figure A.156: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on intersection. Solutions 2, 4 correspond to accelerated penalty method and noiseless data.

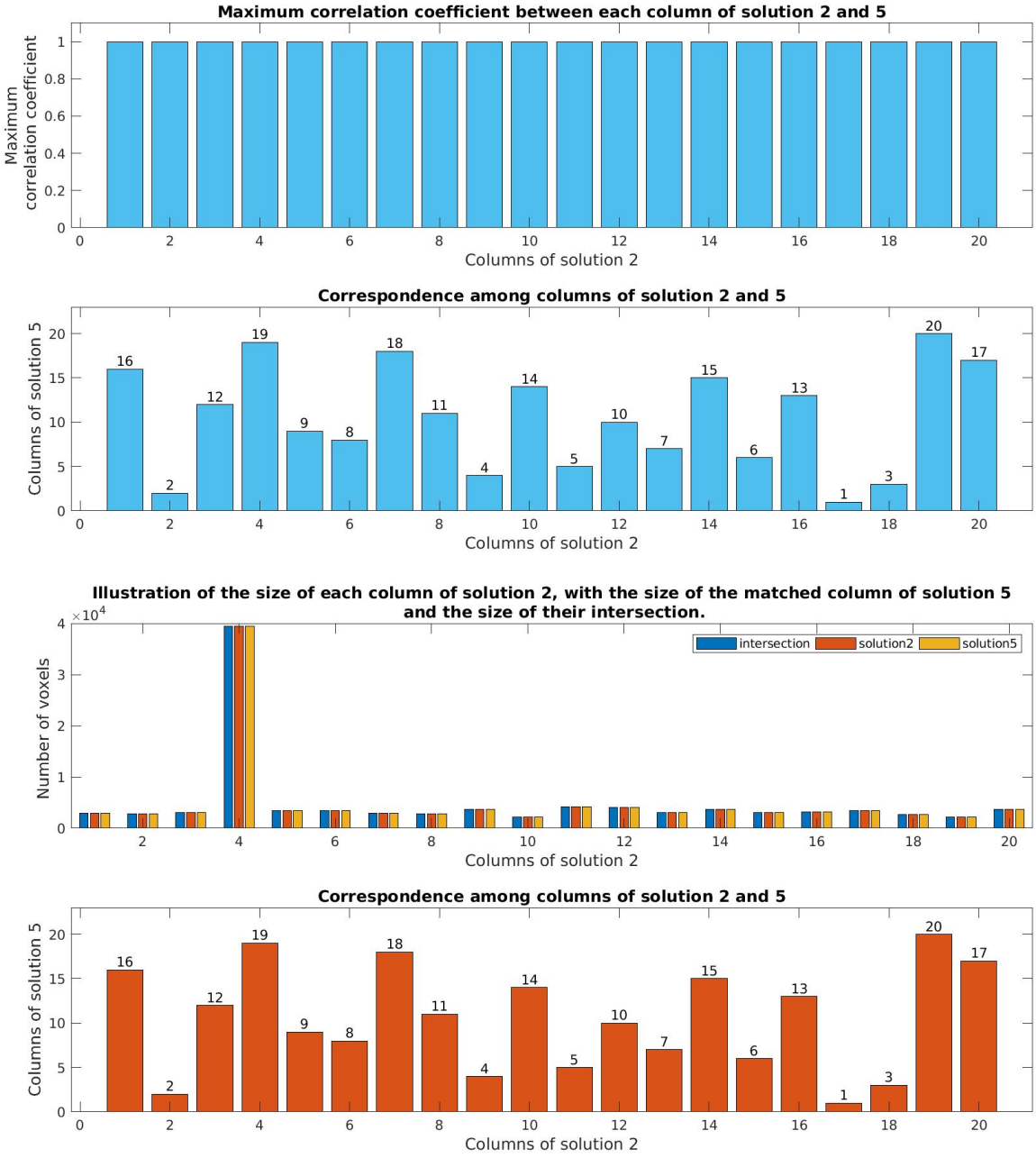


Figure A.157: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 5. We associate the column r of the solution 2 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on intersection. Solutions 2, 5 correspond to accelerated penalty method and noiseless data.

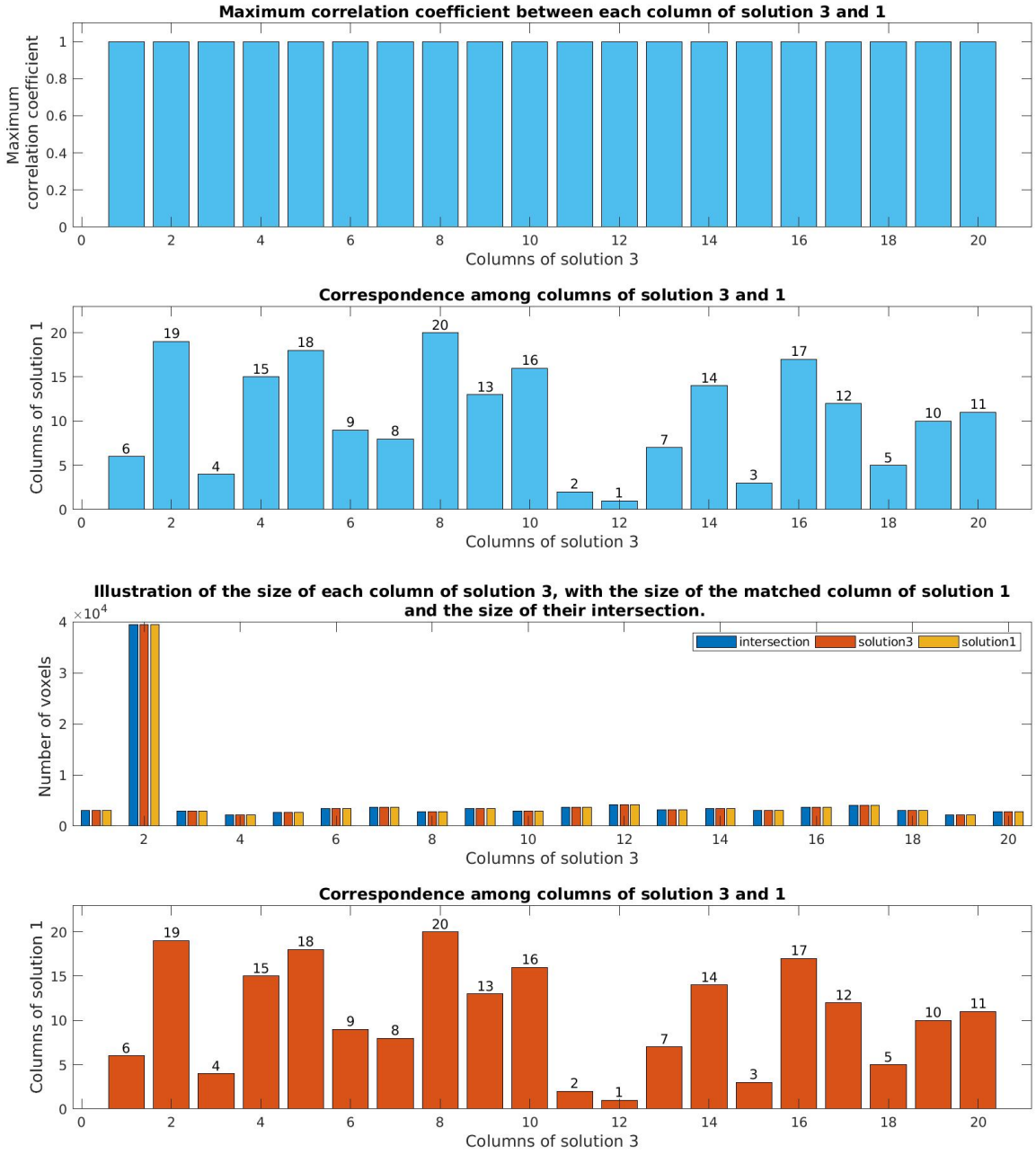


Figure A.158: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on intersection. Solutions 3, 1 correspond to accelerated penalty method and noiseless data.

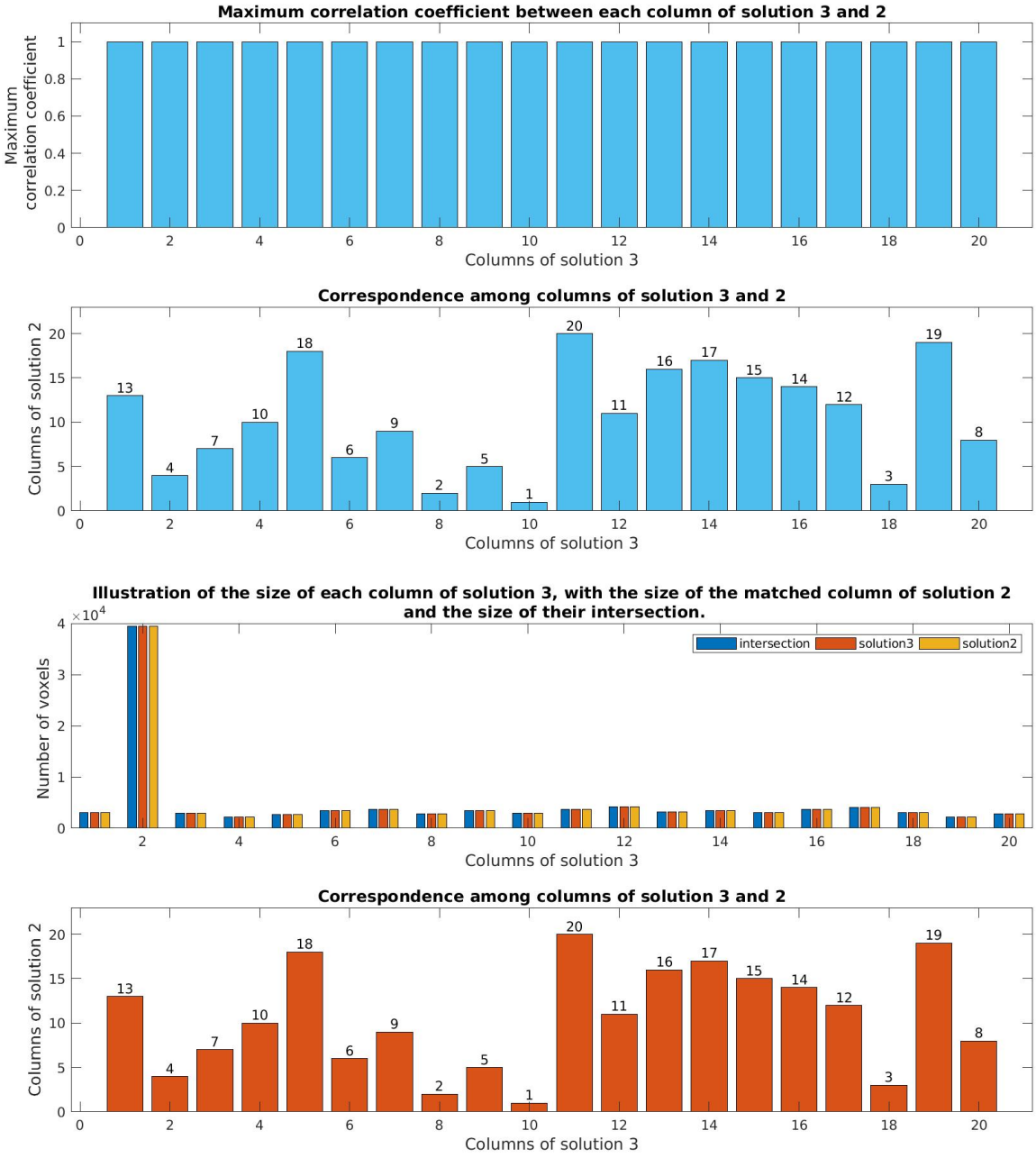


Figure A.159: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on intersection. Solutions 3, 2 correspond to accelerated penalty method and noiseless data.

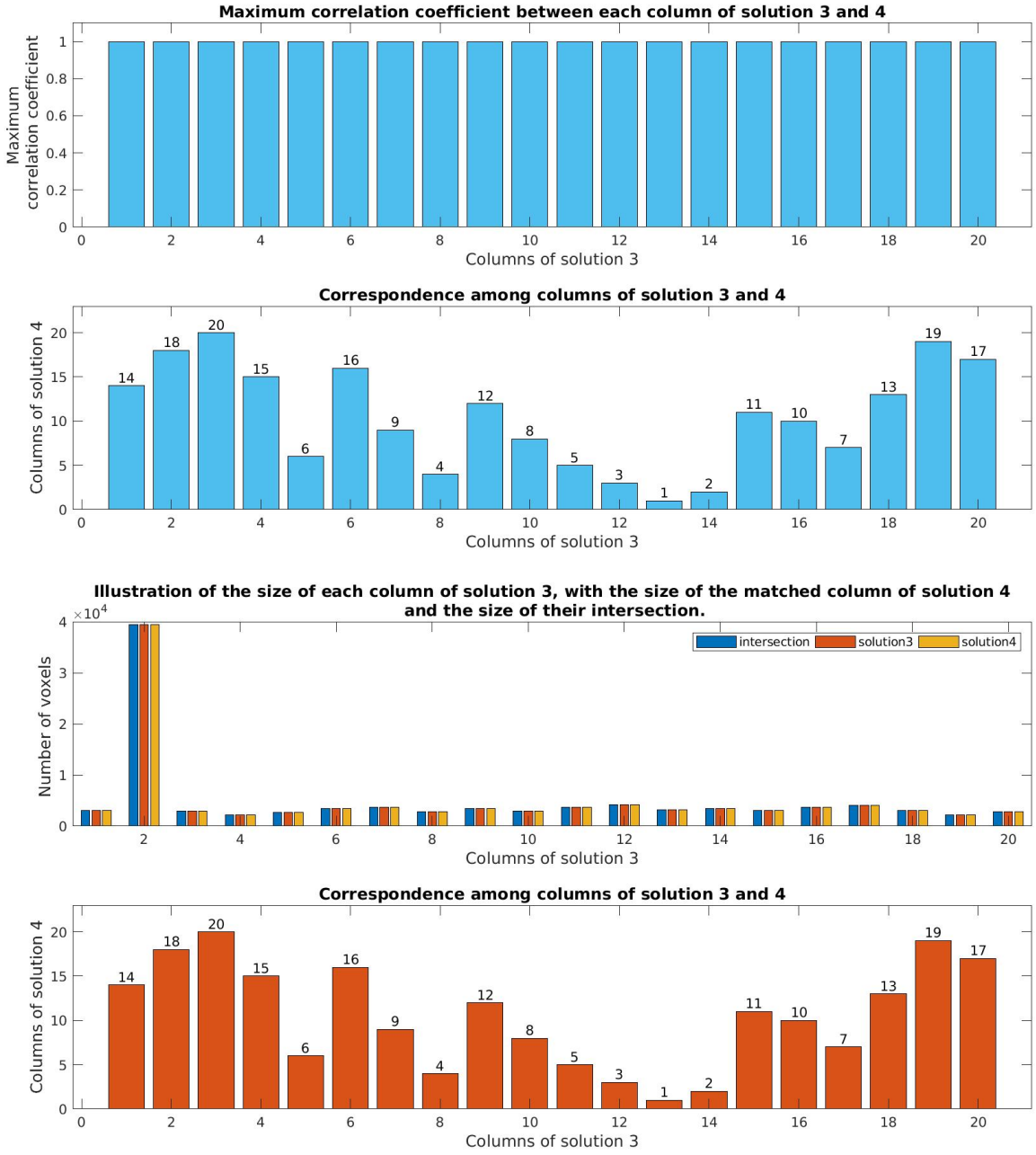


Figure A.160: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on intersection. Solutions 3, 4 correspond to accelerated penalty method and noiseless data.

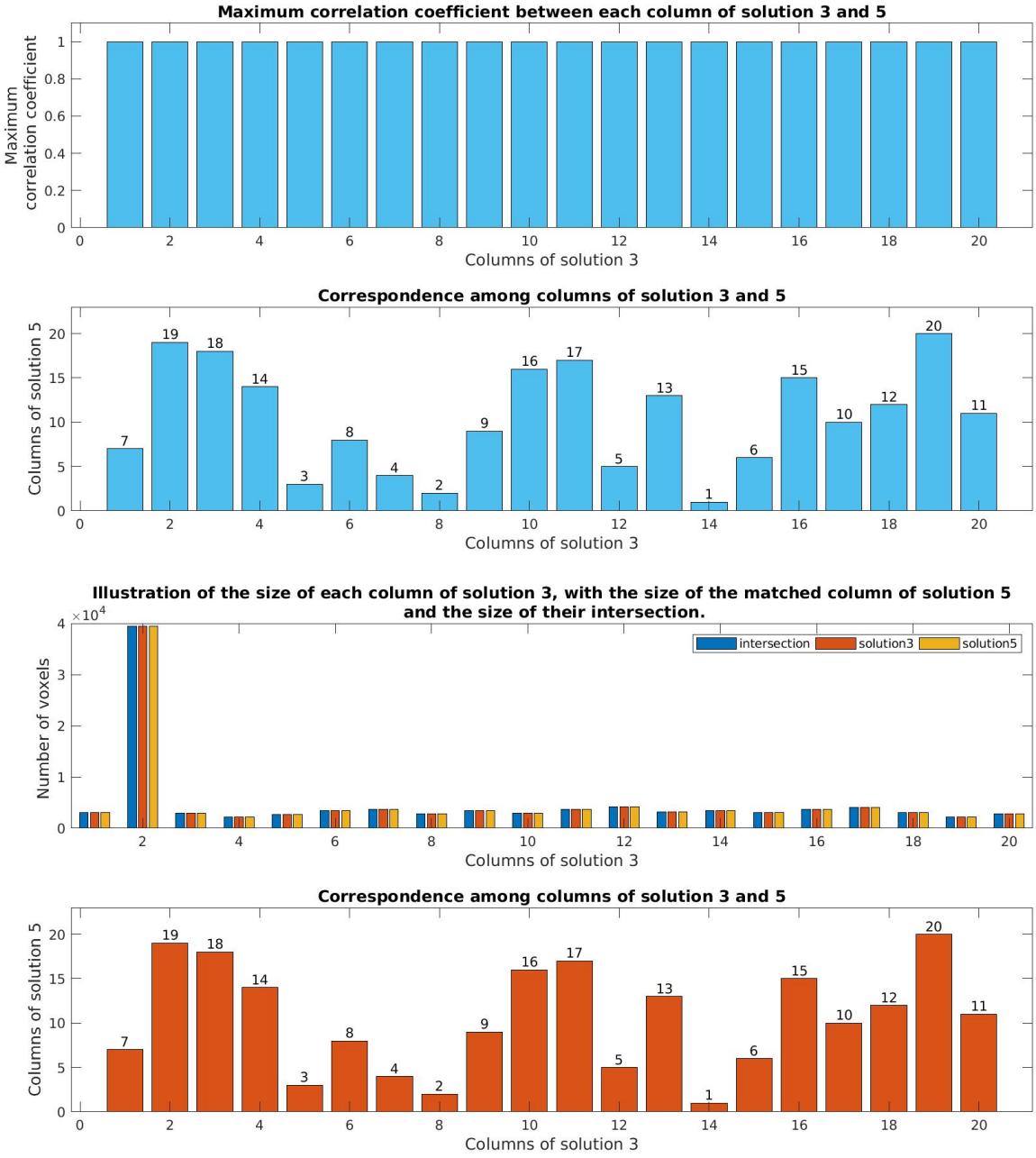


Figure A.161: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 5. We associate the column r of the solution 3 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on intersection. Solutions 3, 5 correspond to accelerated penalty method and noiseless data.

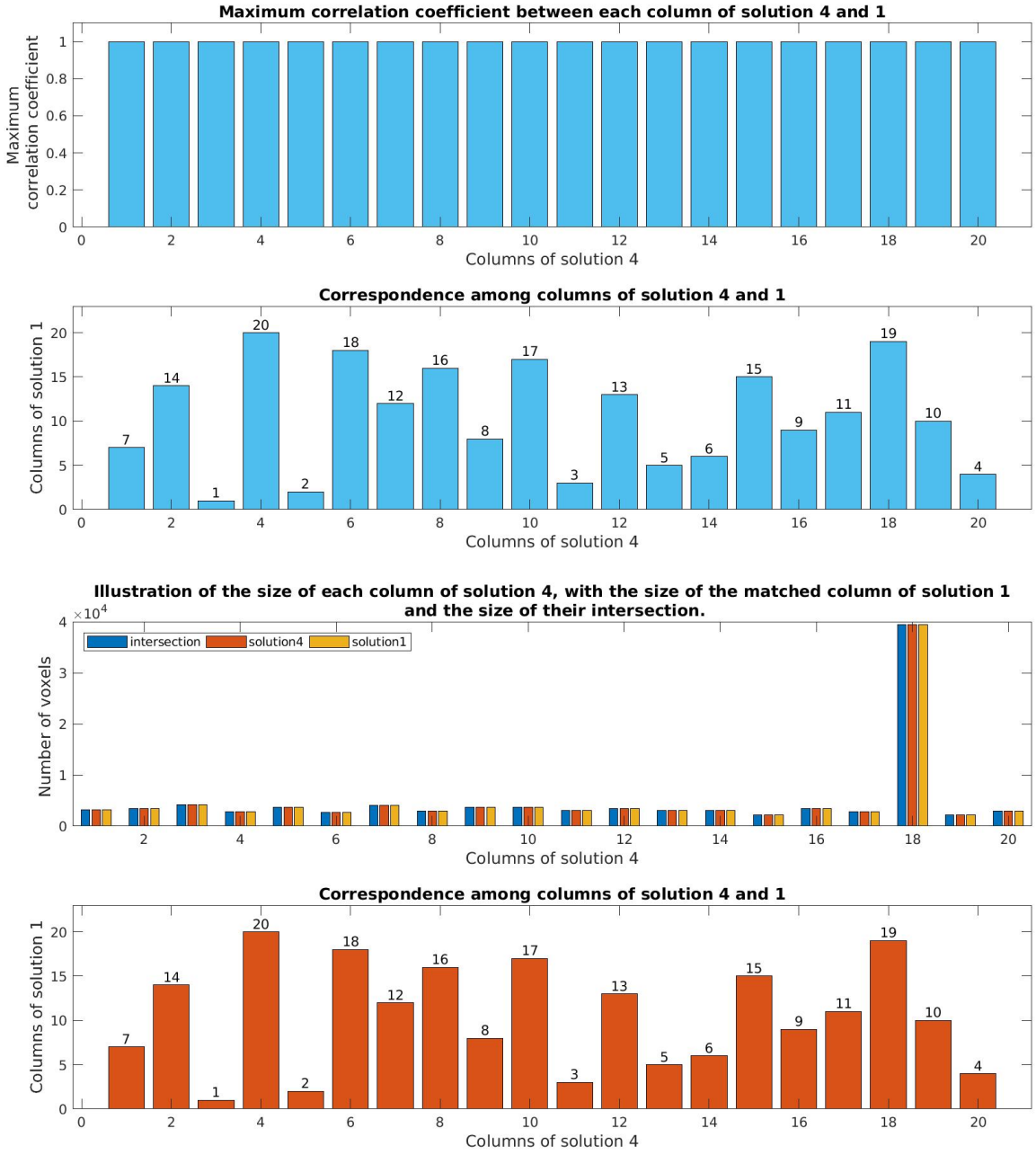


Figure A.162: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on intersection. Solutions 4, 1 correspond to accelerated penalty method and noiseless data.

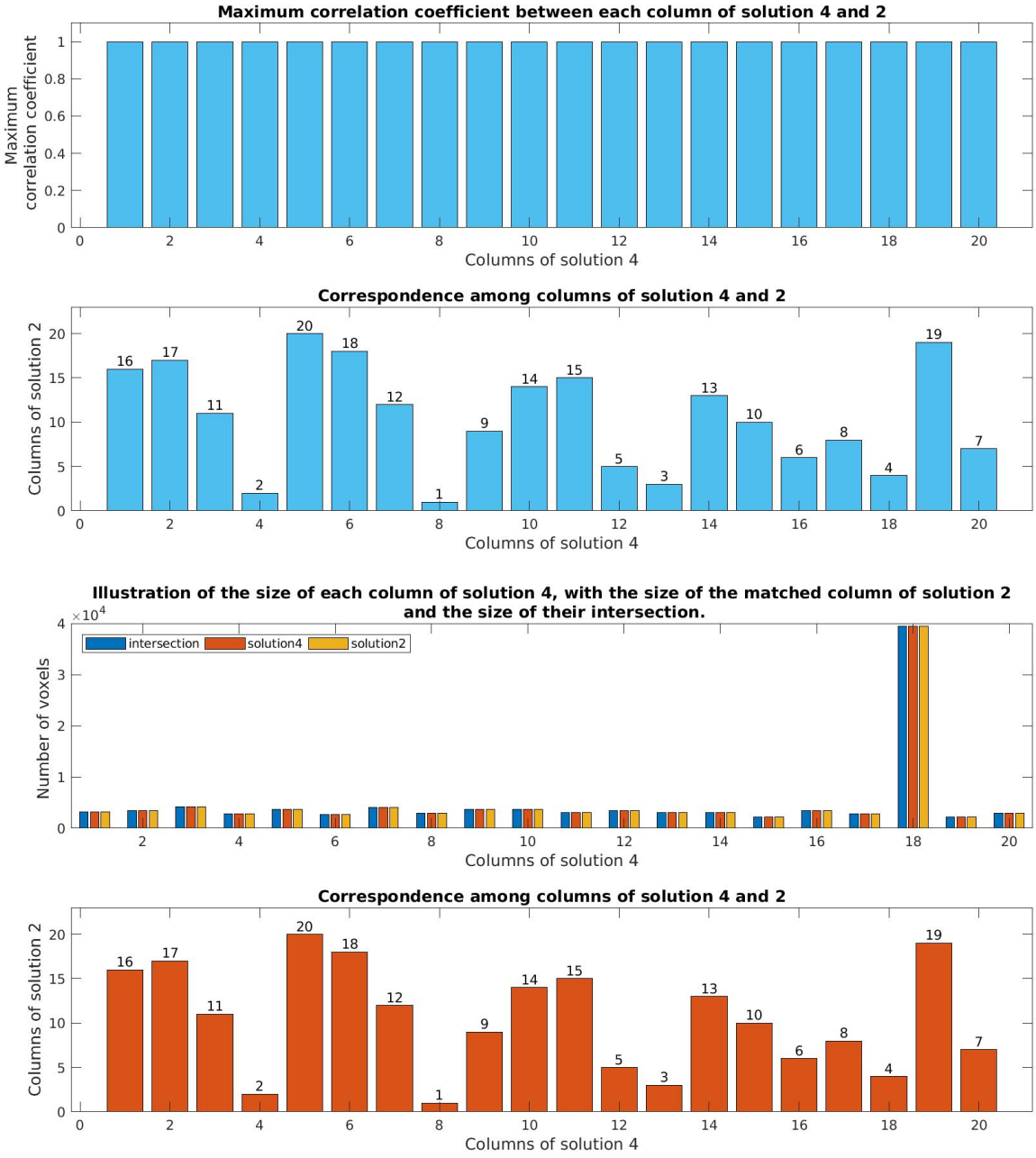


Figure A.163: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on intersection. Solutions 4, 2 correspond to accelerated penalty method and noiseless data.

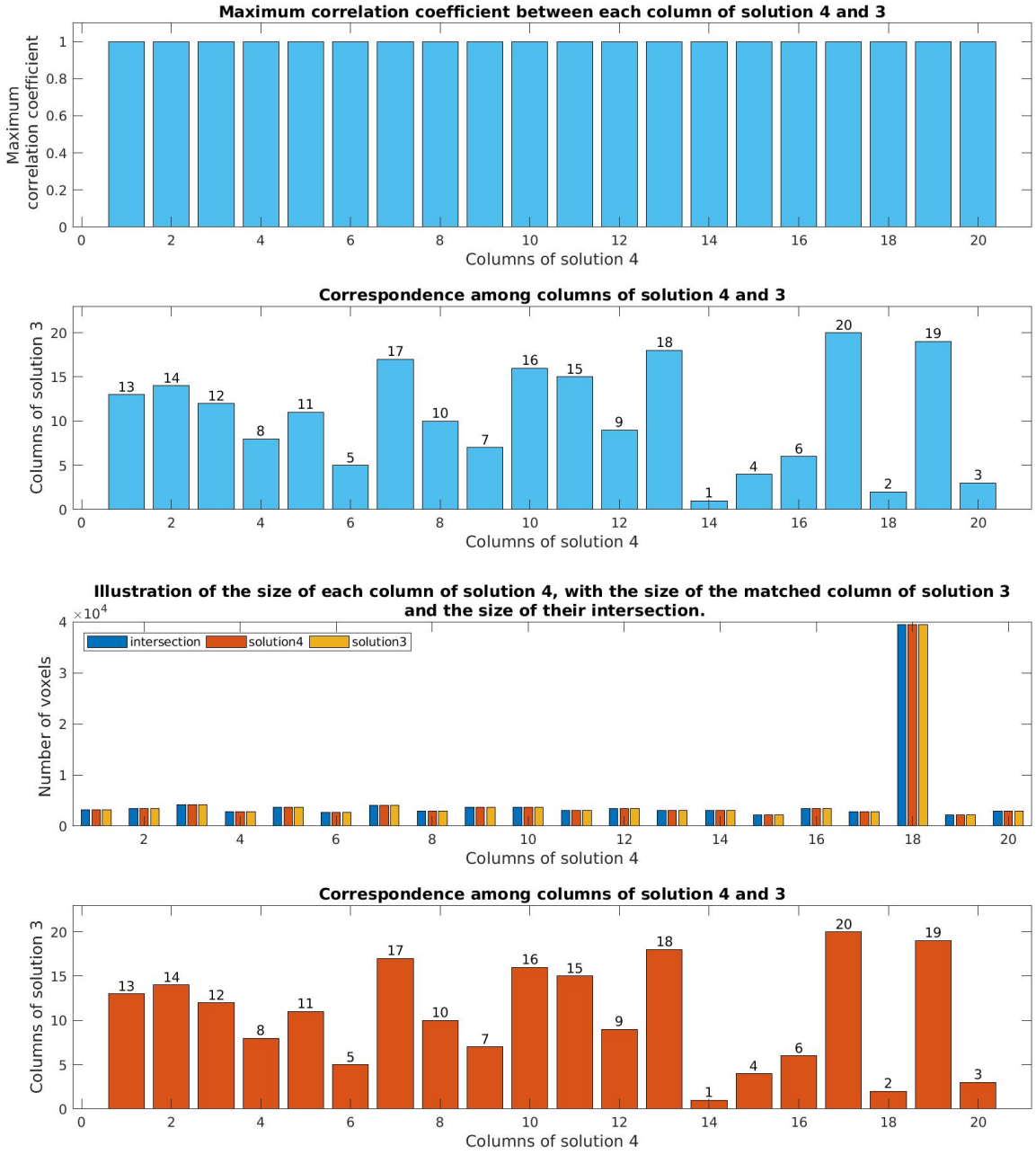


Figure A.164: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on intersection. Solutions 4, 3 correspond to accelerated penalty method and noiseless data.

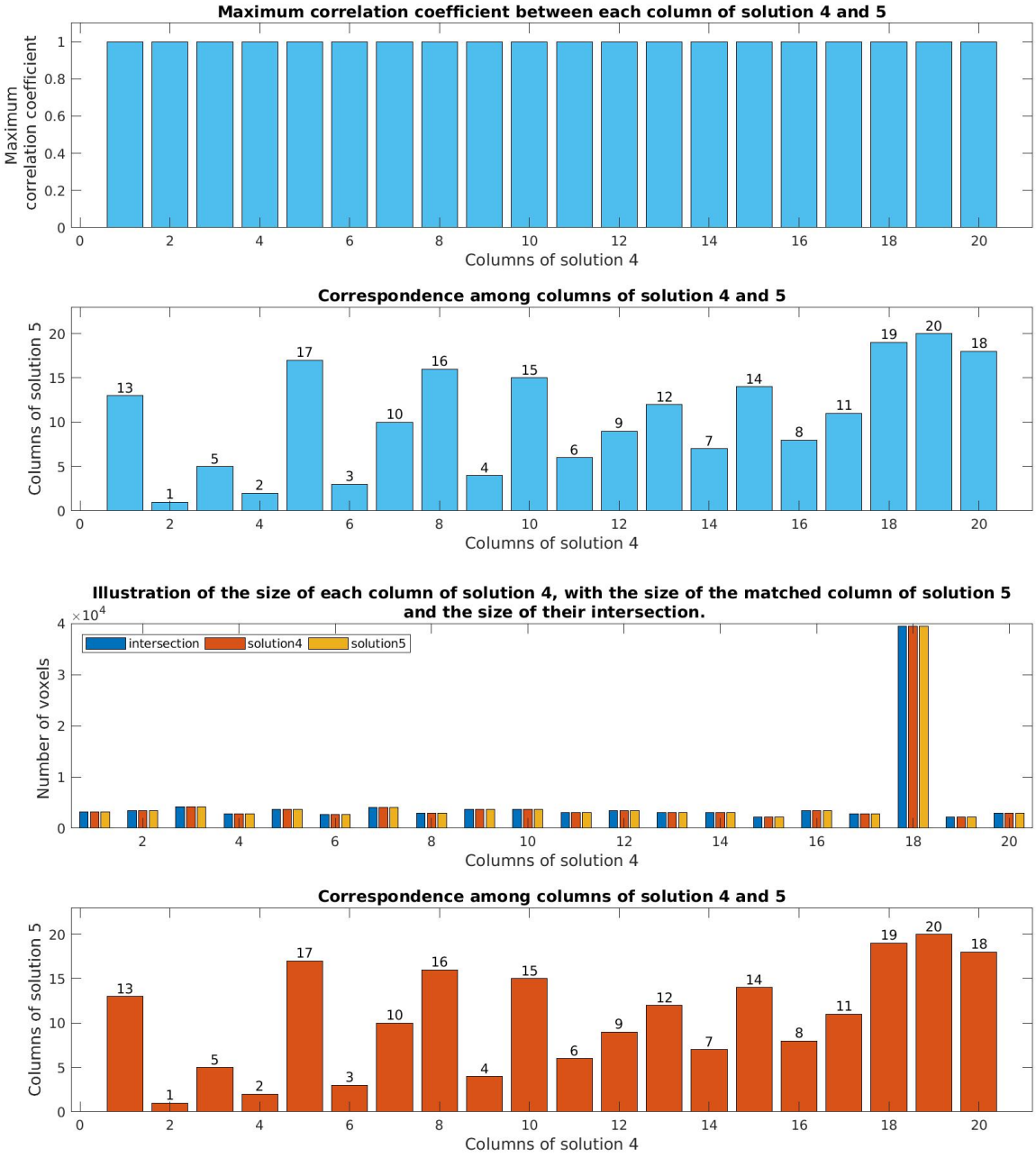


Figure A.165: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 5. We associate the column r of the solution 4 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on intersection. Solutions 4, 5 correspond to accelerated penalty method and noiseless data.

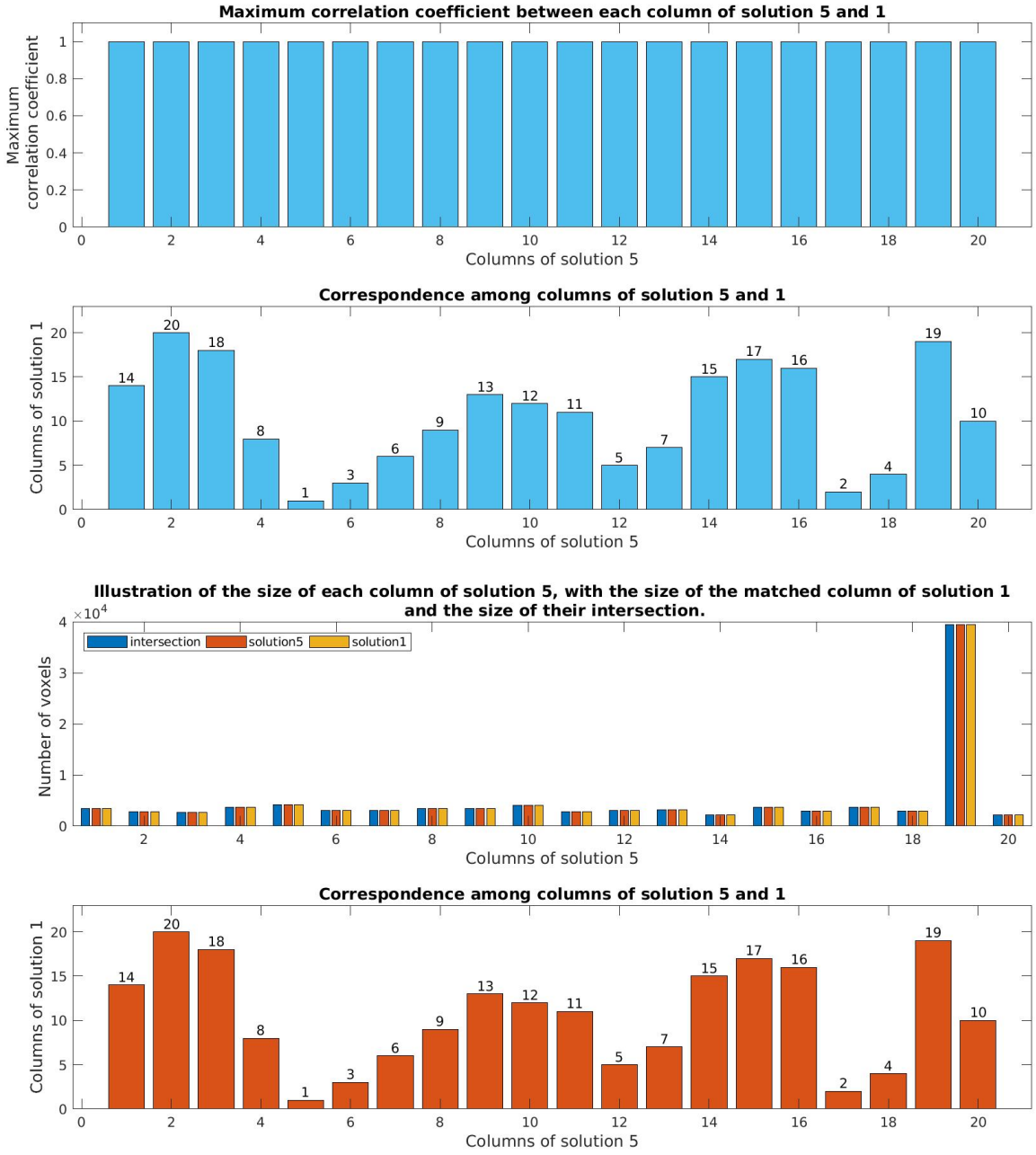


Figure A.166: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 1. We associate the column r of the solution 5 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on intersection. Solutions 5, 1 correspond to accelerated penalty method and noiseless data.

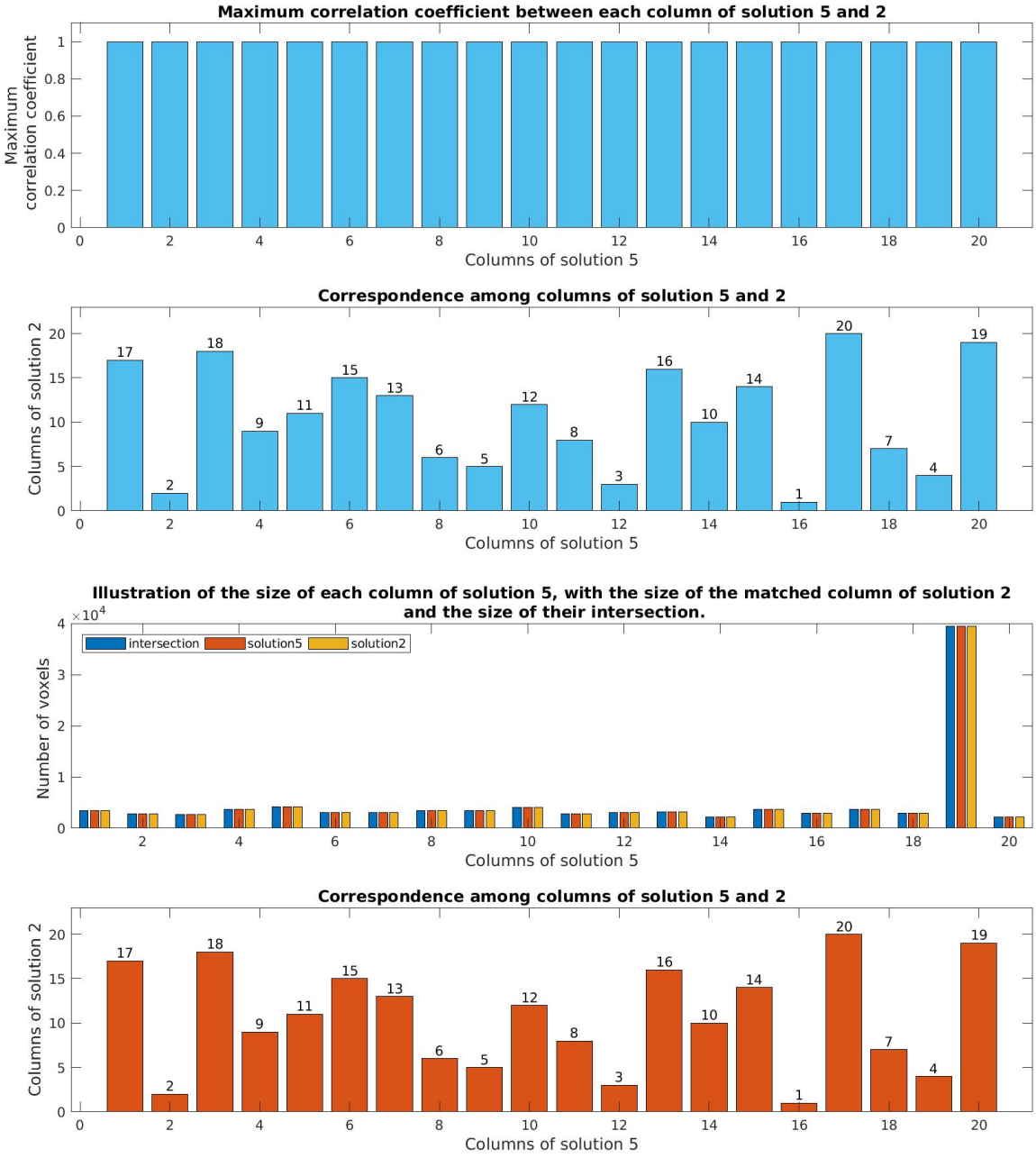


Figure A.167: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 2. We associate the column r of the solution 5 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on intersection. Solutions 5, 2 correspond to accelerated penalty method and noiseless data.

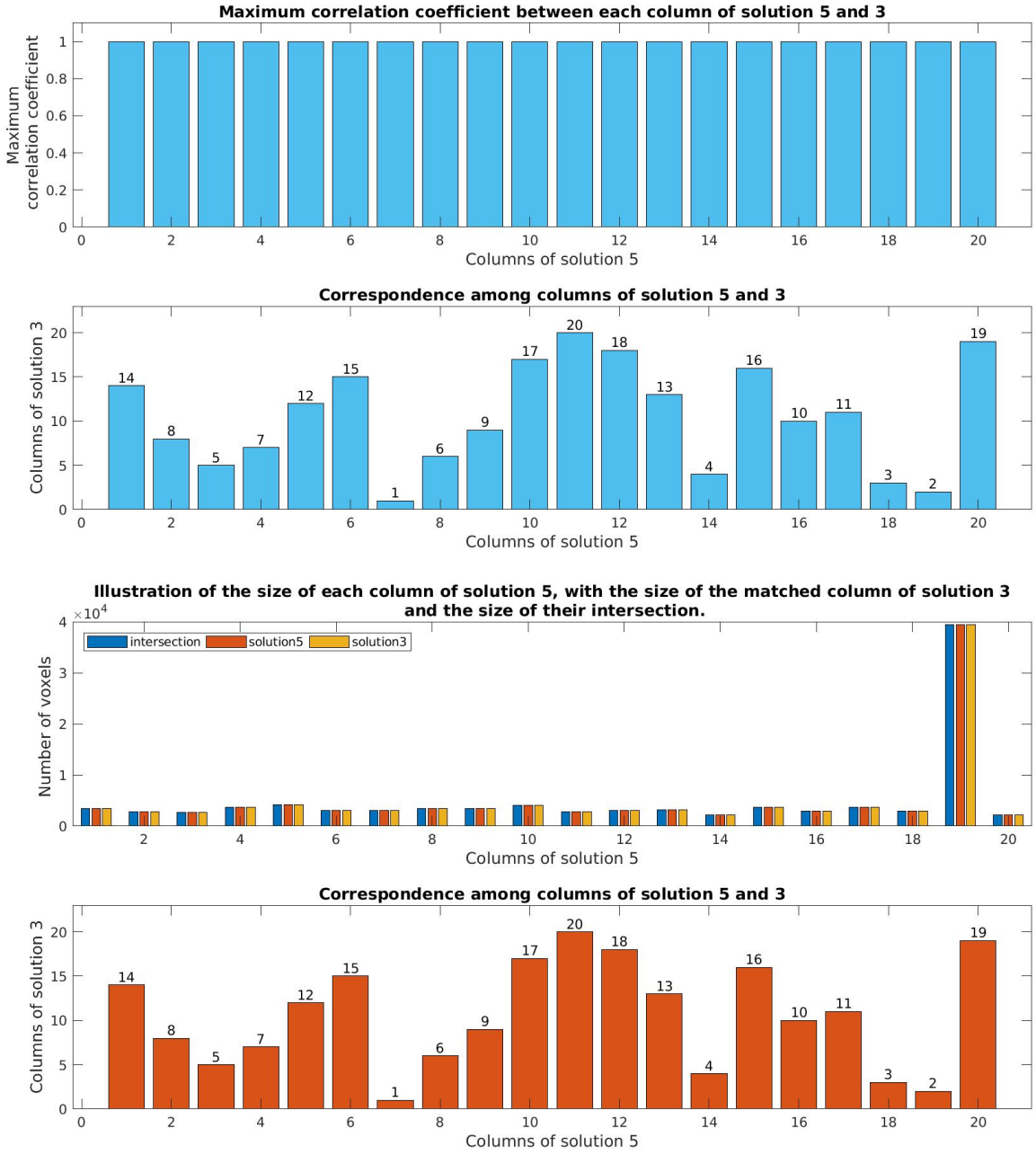


Figure A.168: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 3. We associate the column r of the solution 5 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on intersection. Solutions 5, 3 correspond to accelerated penalty method and noiseless data.

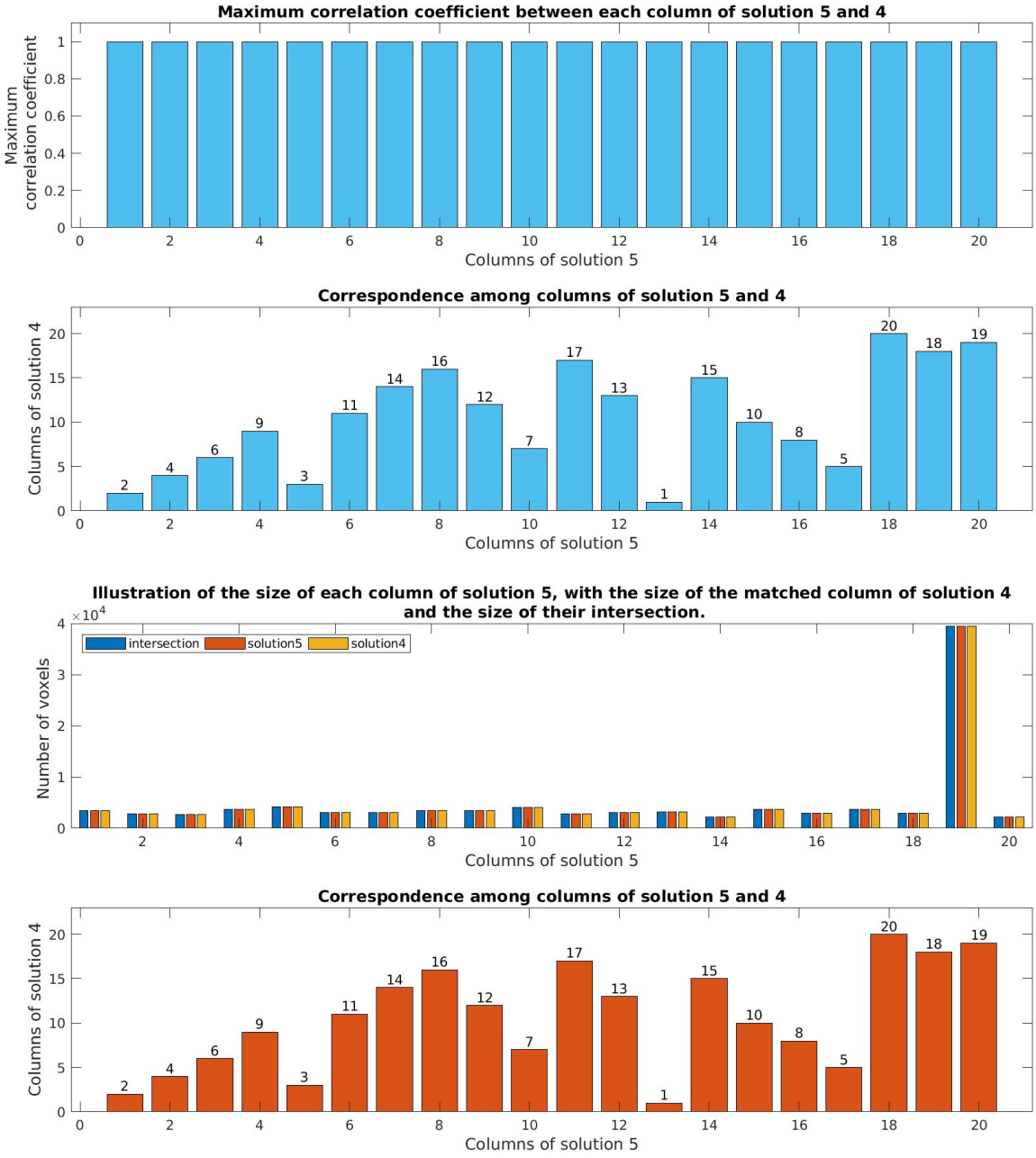


Figure A.169: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 4. We associate the column r of the solution 5 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on intersection. Solutions 5, 4 correspond to accelerated penalty method and noiseless data.

· **SNR= 0 dB**

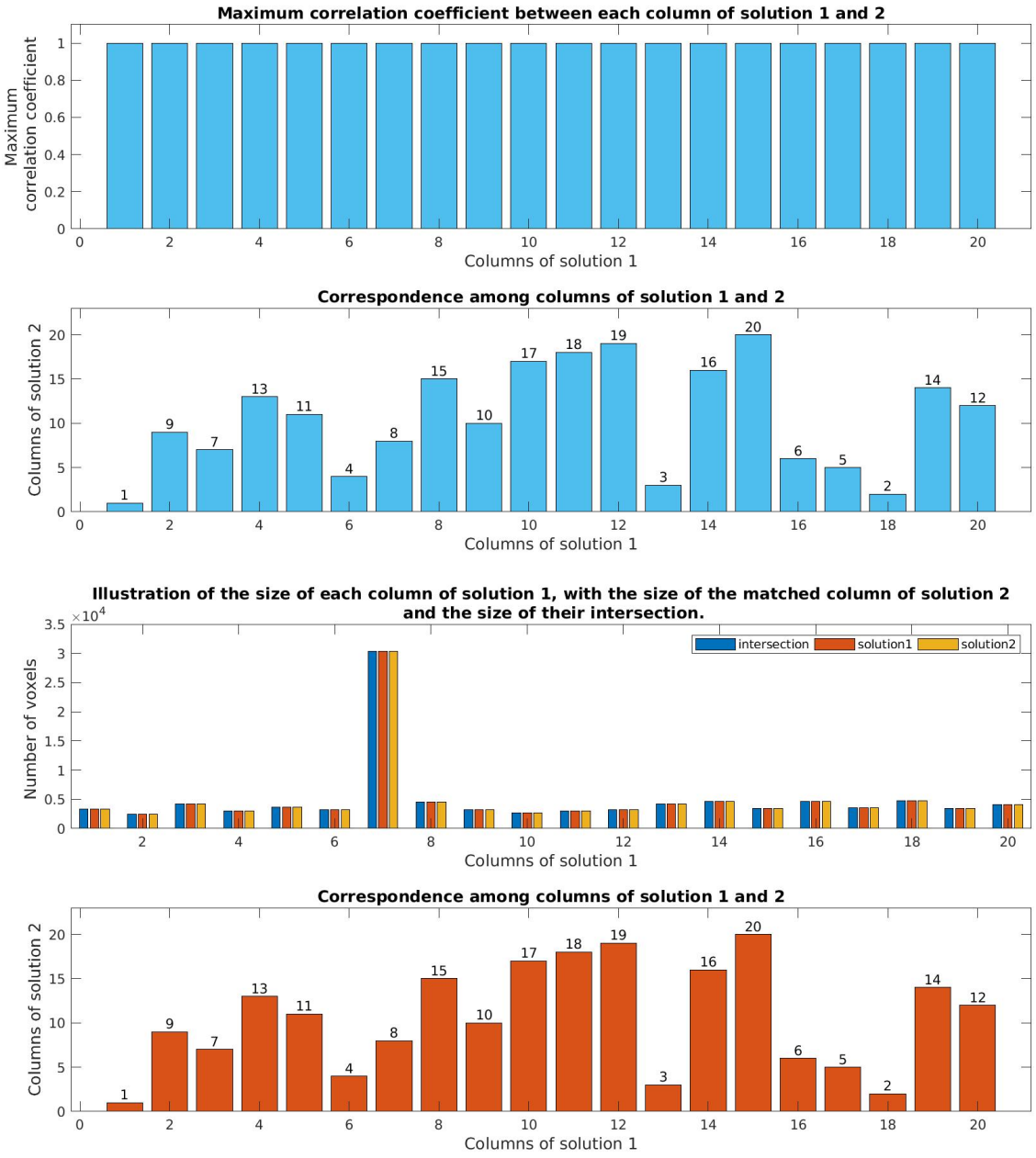


Figure A.170: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated, based on intersection. Solutions 1, 2 correspond to accelerated penalty method and SNR=0 dB.

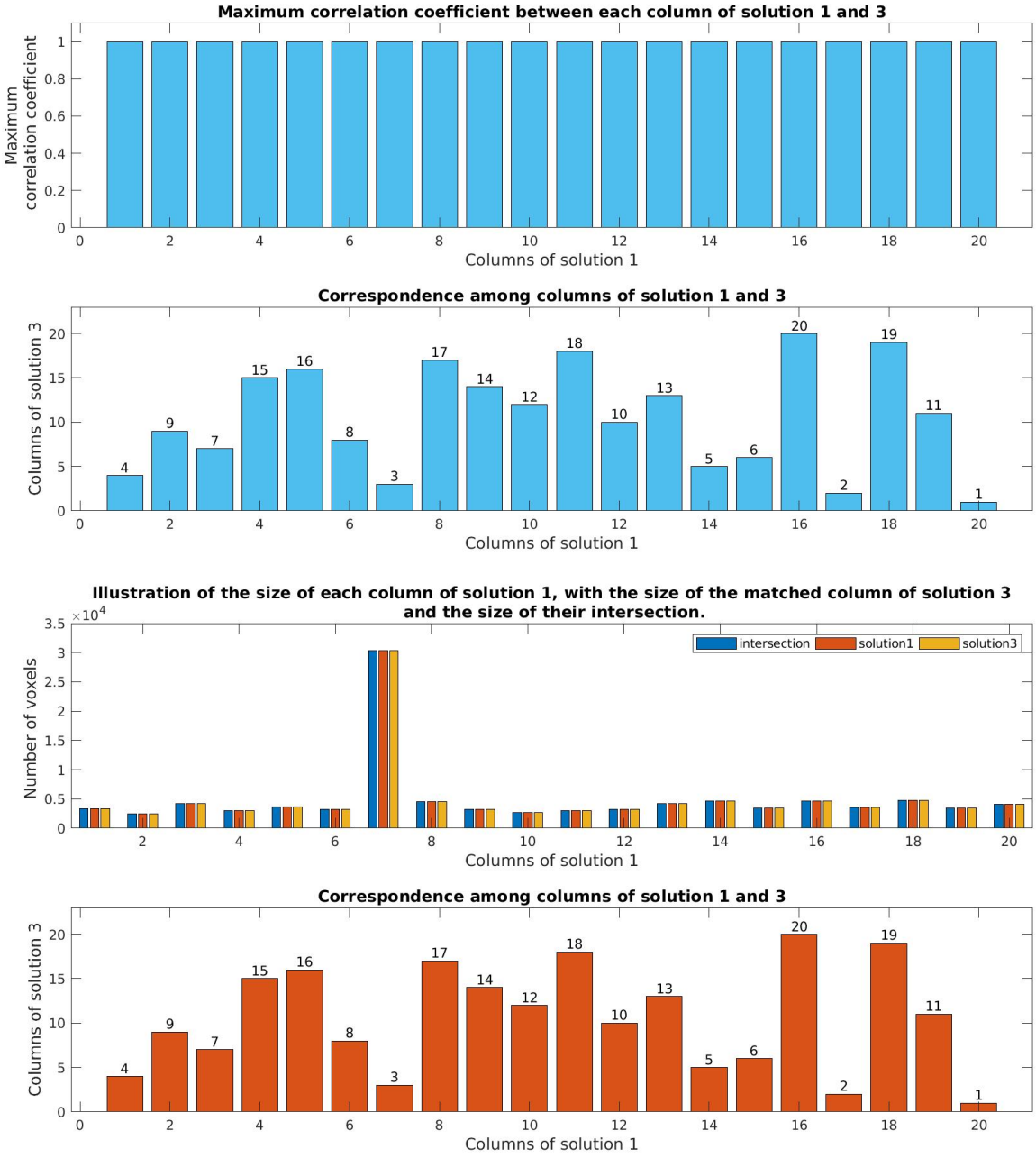


Figure A.171: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on intersection. Solutions 1, 3 correspond to accelerated penalty method and SNR=0 dB.

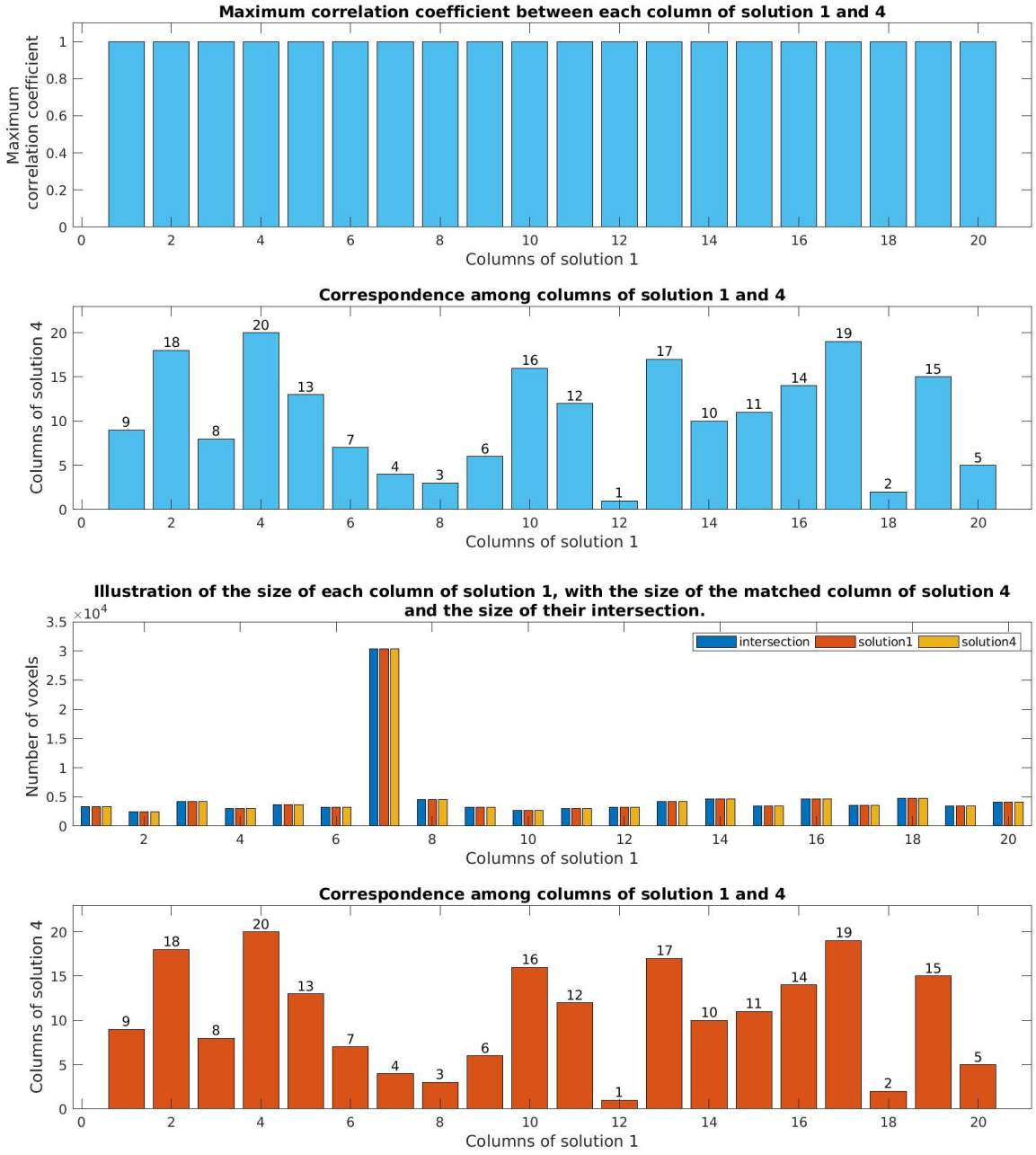


Figure A.172: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on intersection. Solutions 1, 4 correspond to accelerated penalty method and SNR=0 dB.

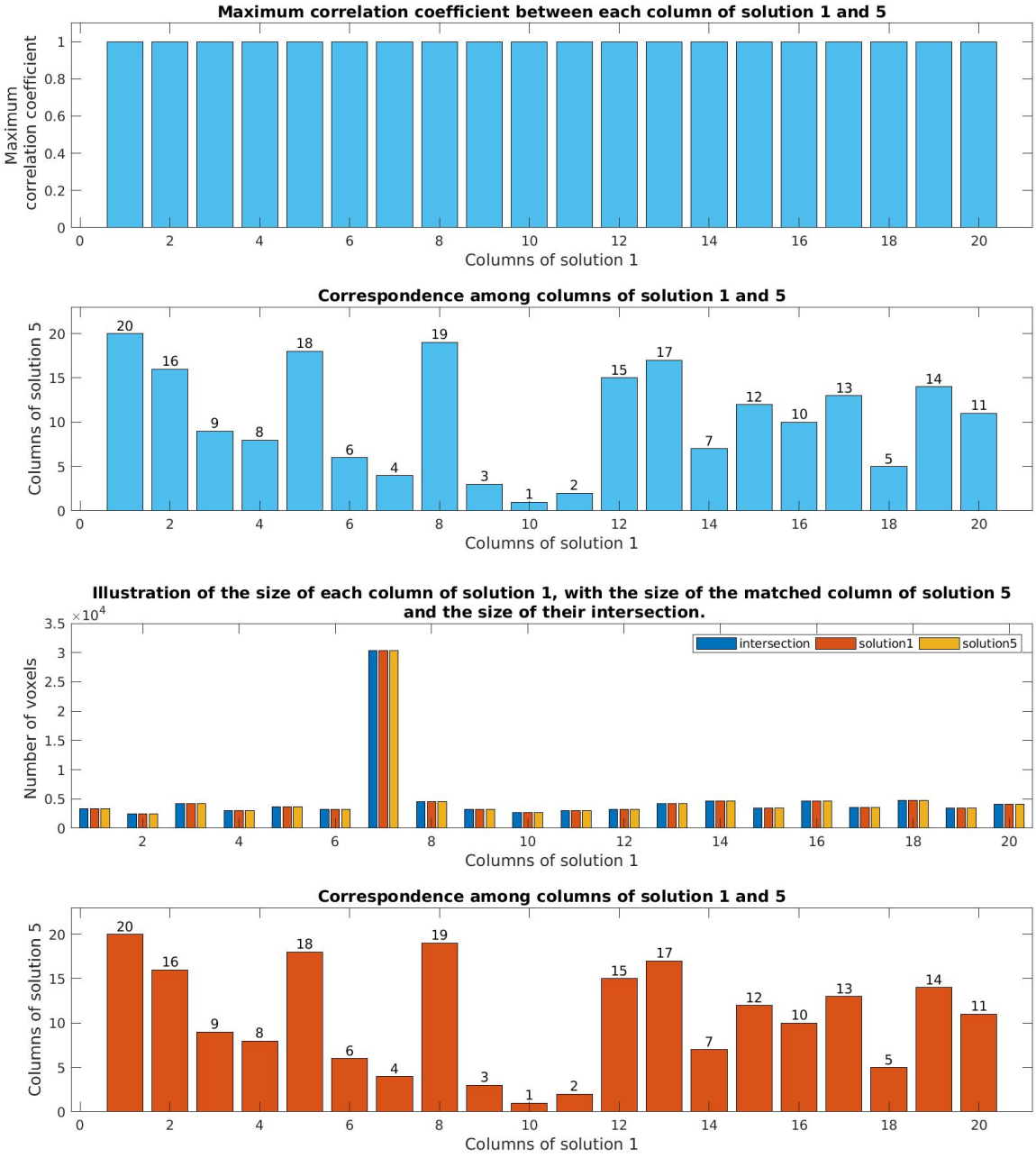


Figure A.173: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 5. We associate the column r of the solution 1 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on intersection. Solutions 1, 5 correspond to accelerated penalty method and SNR=0 dB.

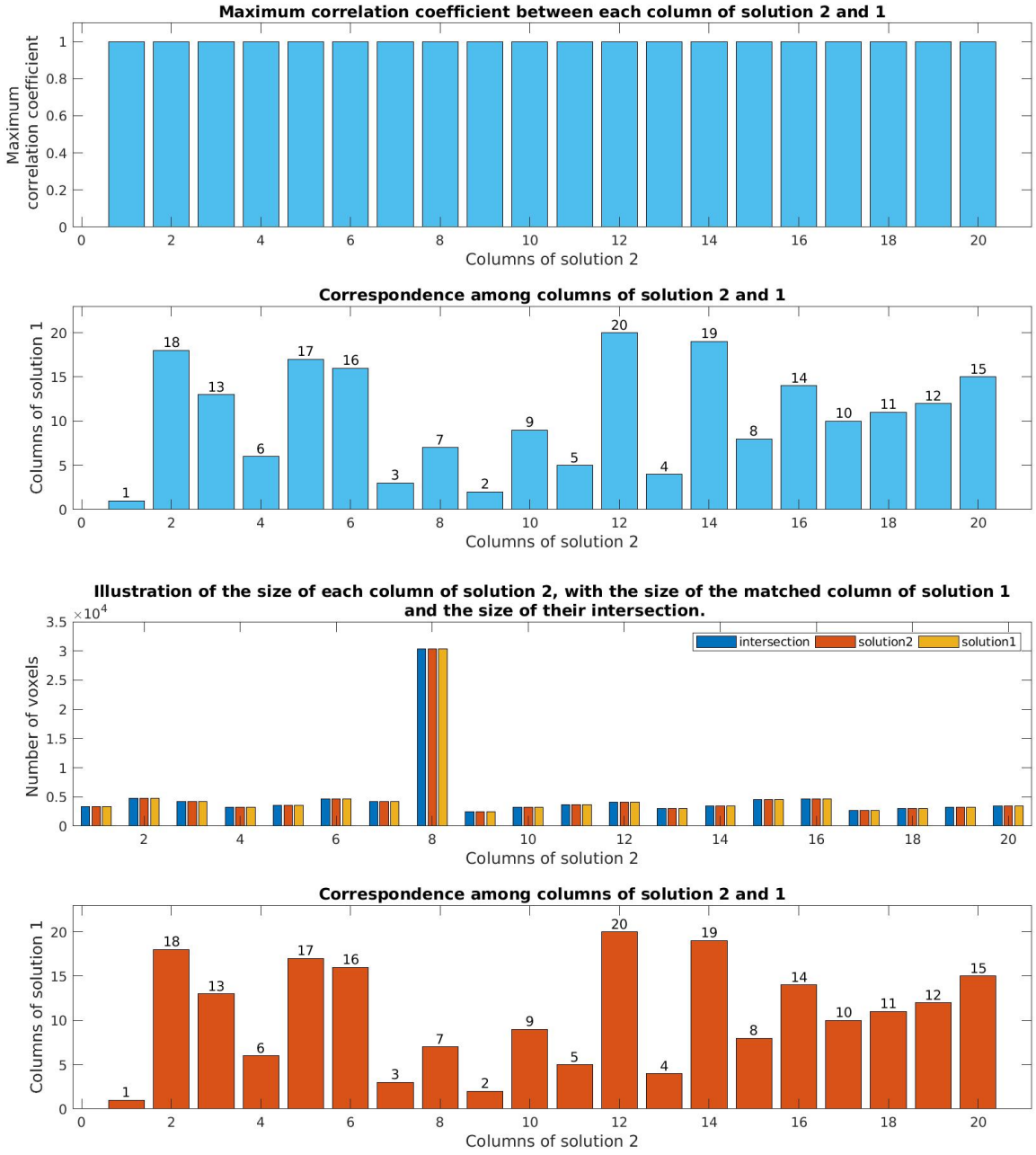


Figure A.174: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on intersection. Solutions 2, 1 correspond to accelerated penalty method and SNR=0 dB.

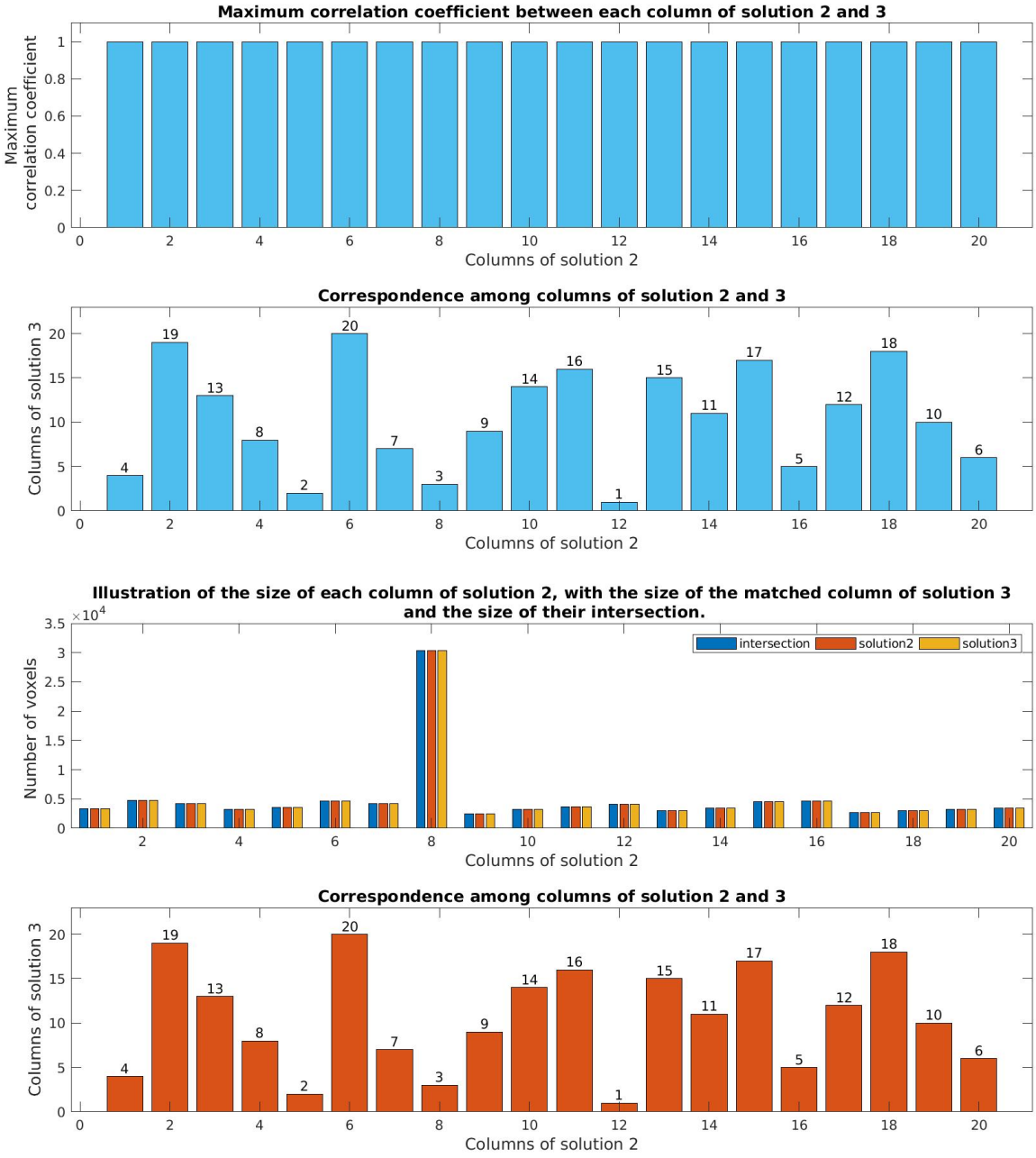


Figure A.175: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on intersection. Solutions 2, 3 correspond to accelerated penalty method and SNR=0 dB.

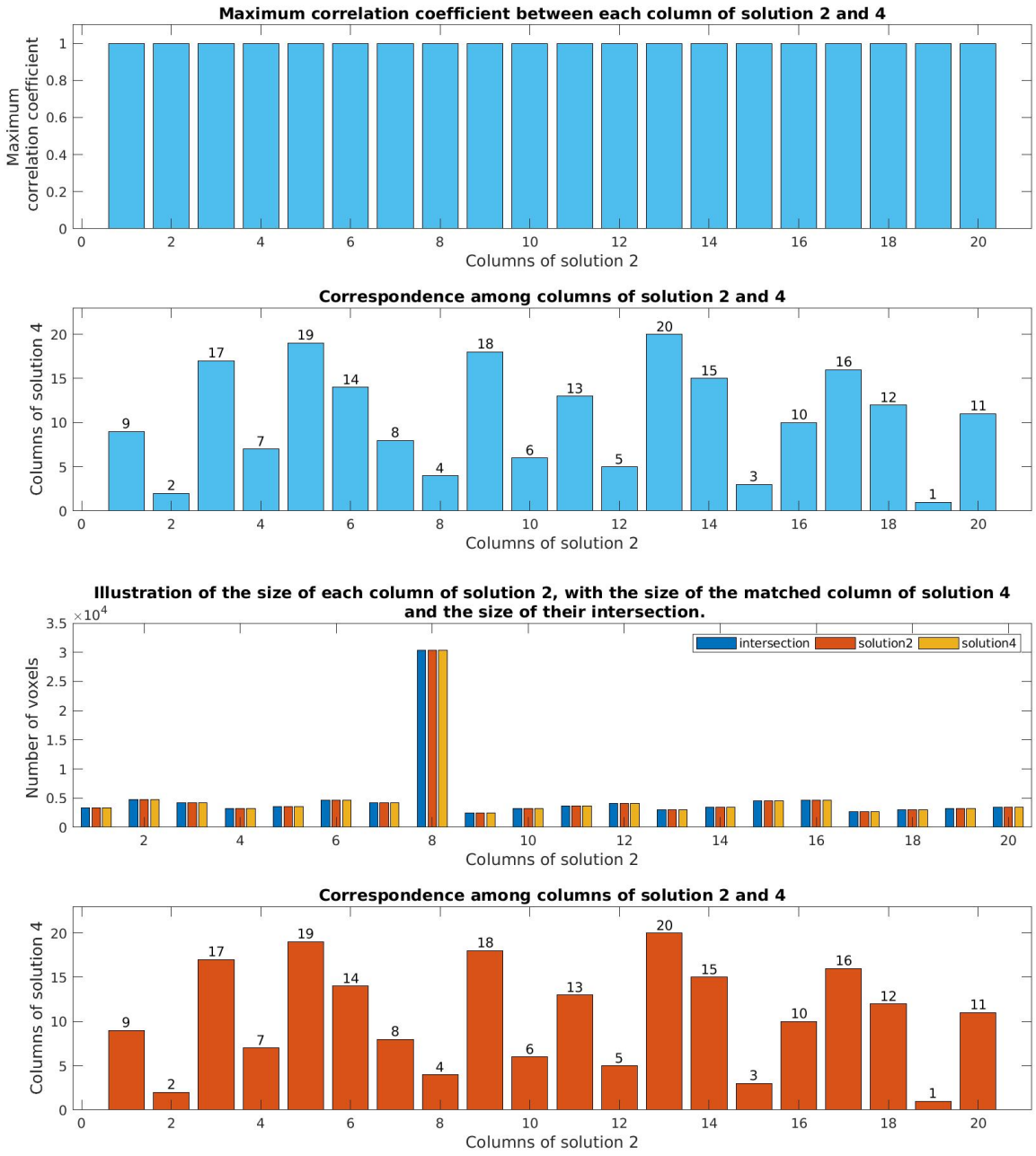


Figure A.176: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on intersection. Solutions 2, 4 correspond to accelerated penalty method and SNR=0 dB.

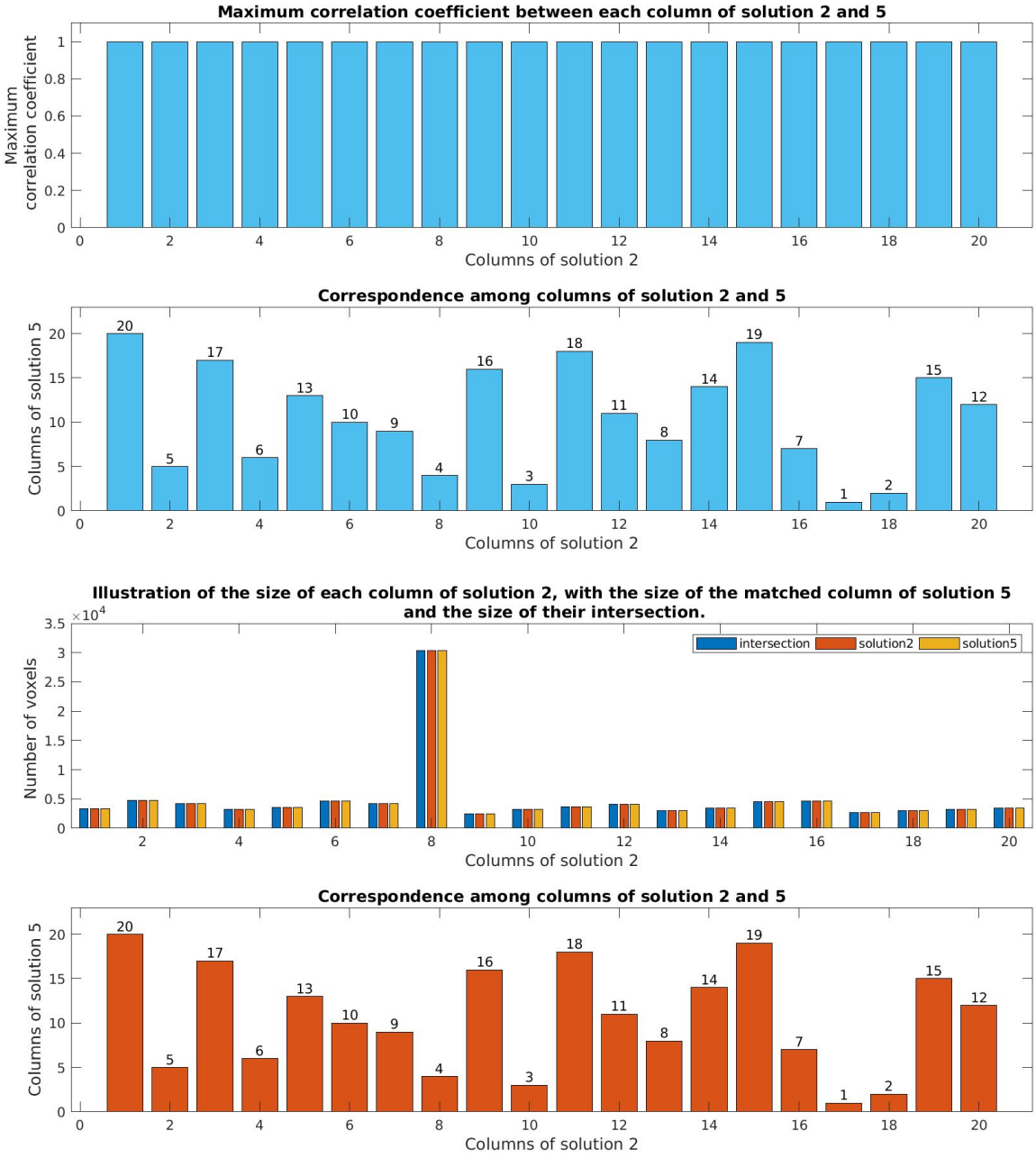


Figure A.177: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 5. We associate the column r of the solution 2 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on intersection. Solutions 2, 5 correspond to accelerated penalty method and SNR=0 dB.

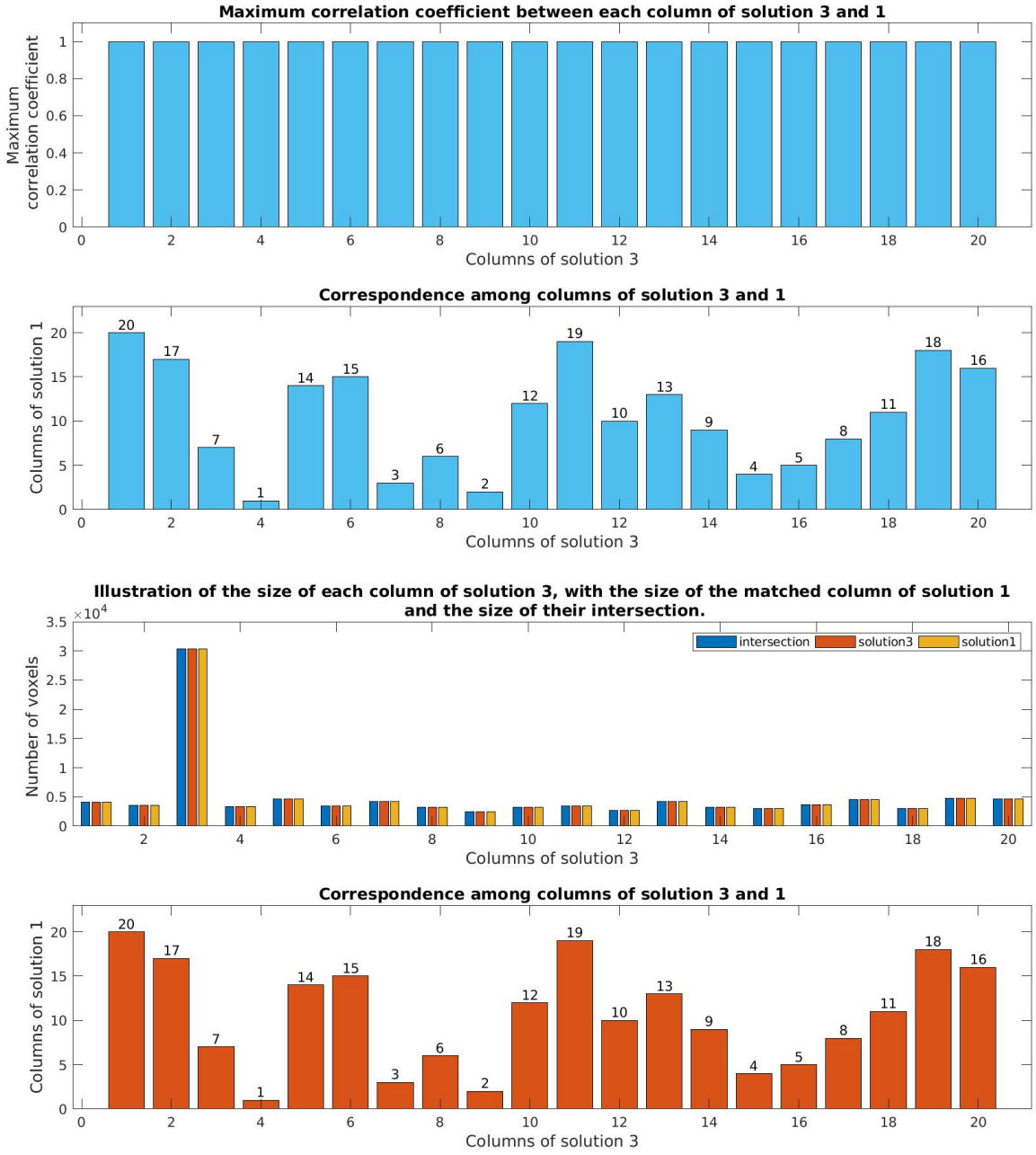


Figure A.178: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on intersection. Solutions 3, 1 correspond to accelerated penalty method and SNR=0 dB.

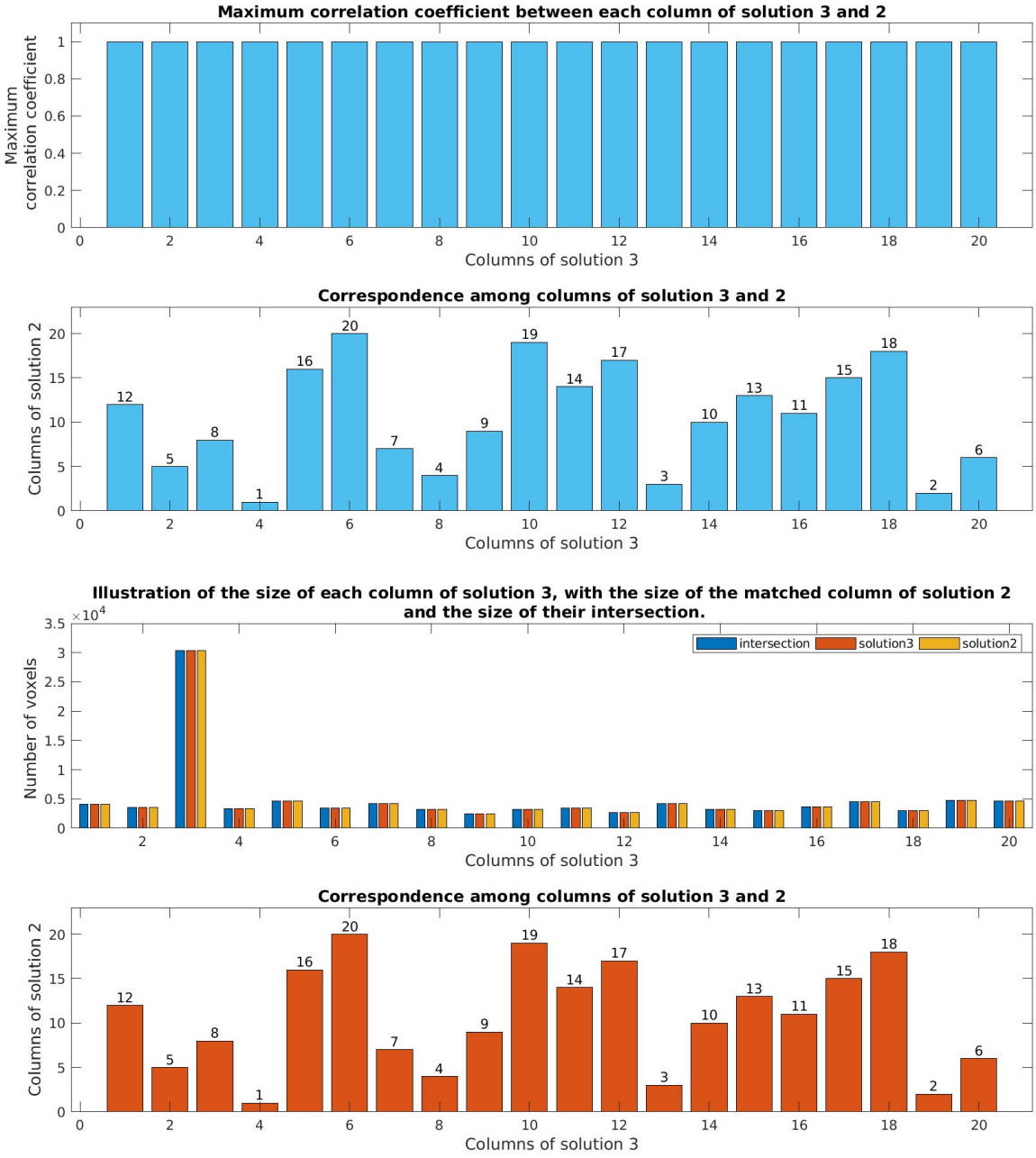


Figure A.179: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on intersection. Solutions 3, 2 correspond to accelerated penalty method and SNR=0 dB.

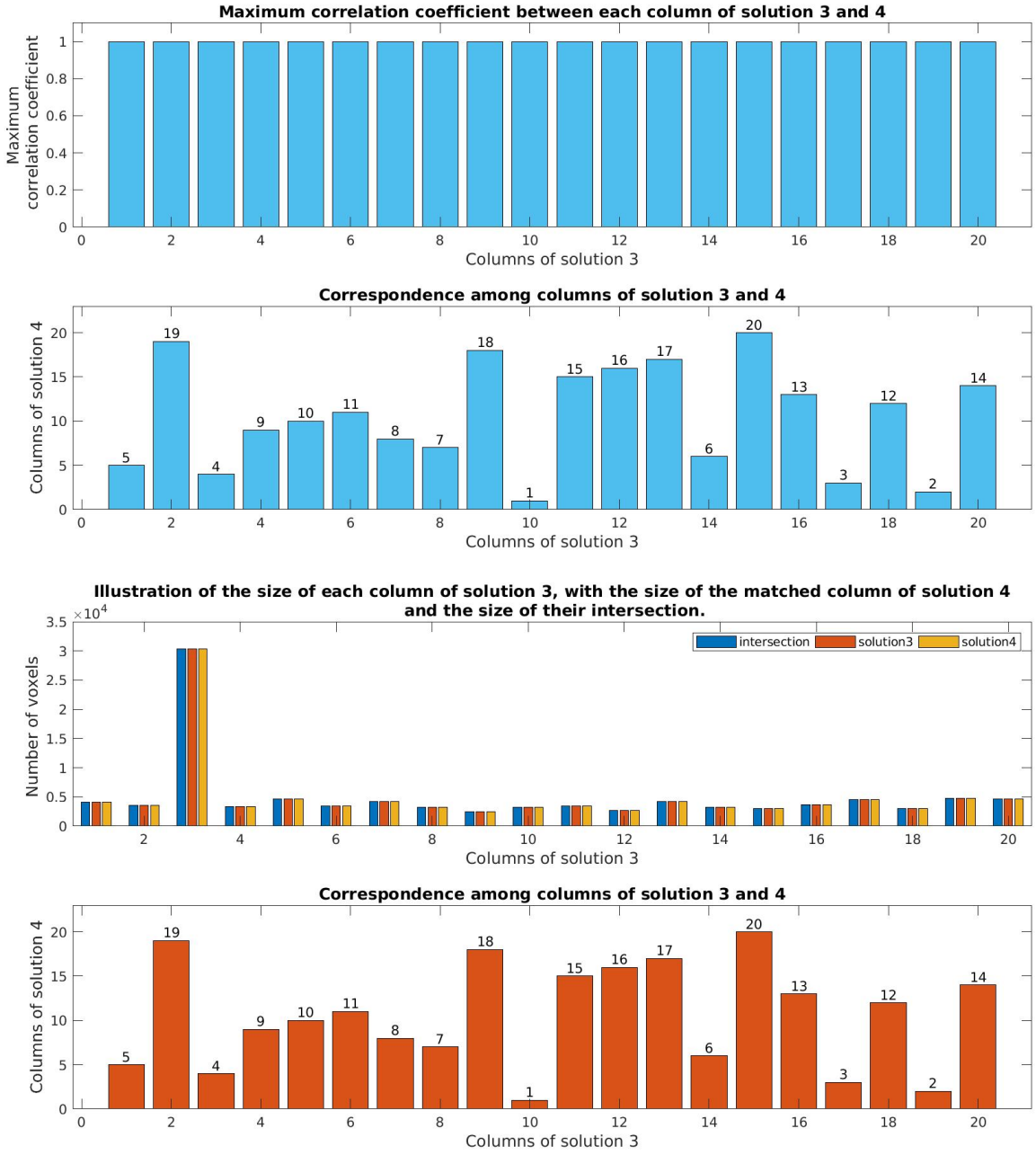


Figure A.180: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on intersection. Solutions 3, 4 correspond to accelerated penalty method and SNR=0 dB.

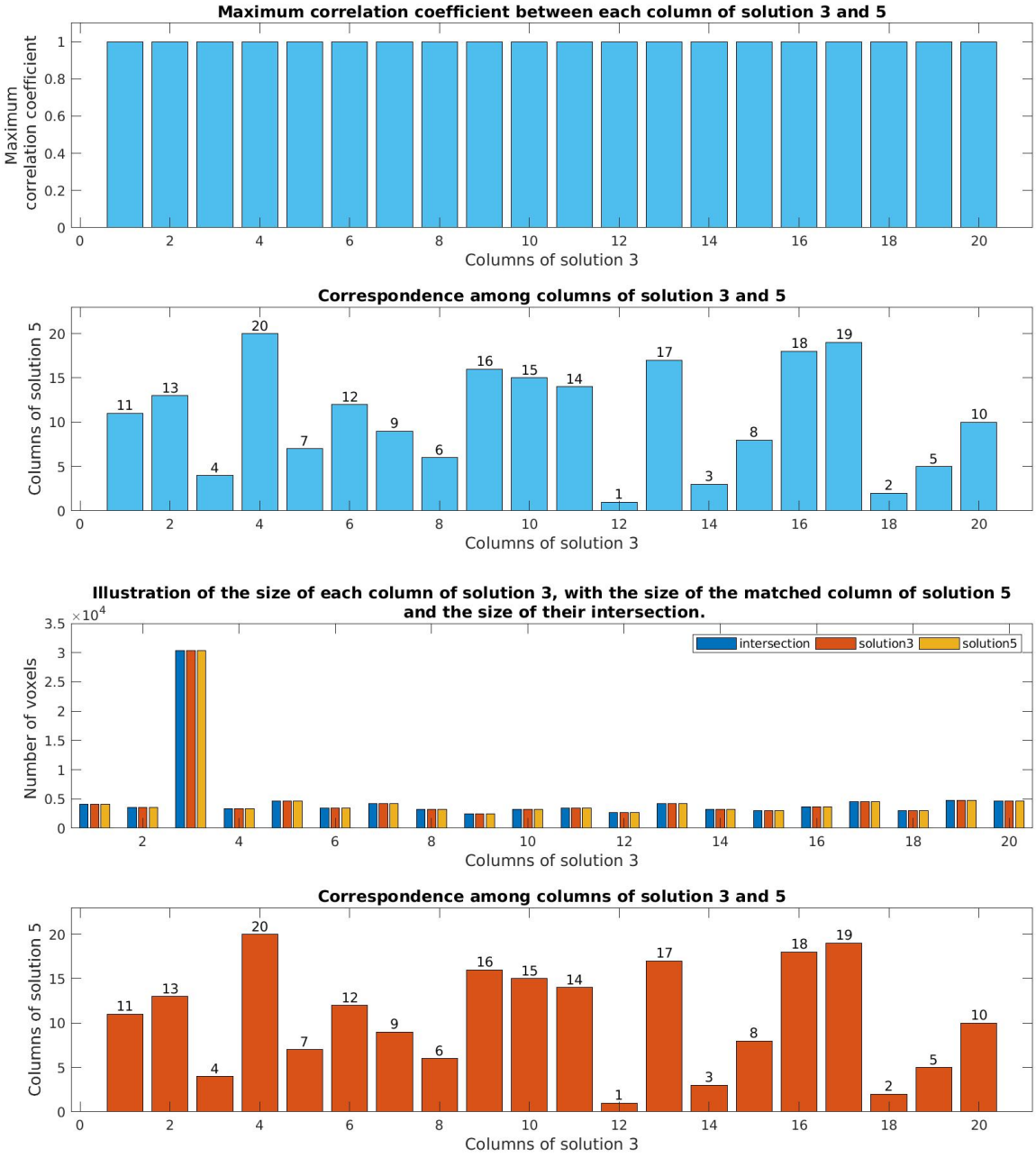


Figure A.181: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 5. We associate the column r of the solution 3 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on intersection. Solutions 3, 5 correspond to accelerated penalty method and SNR=0 dB.

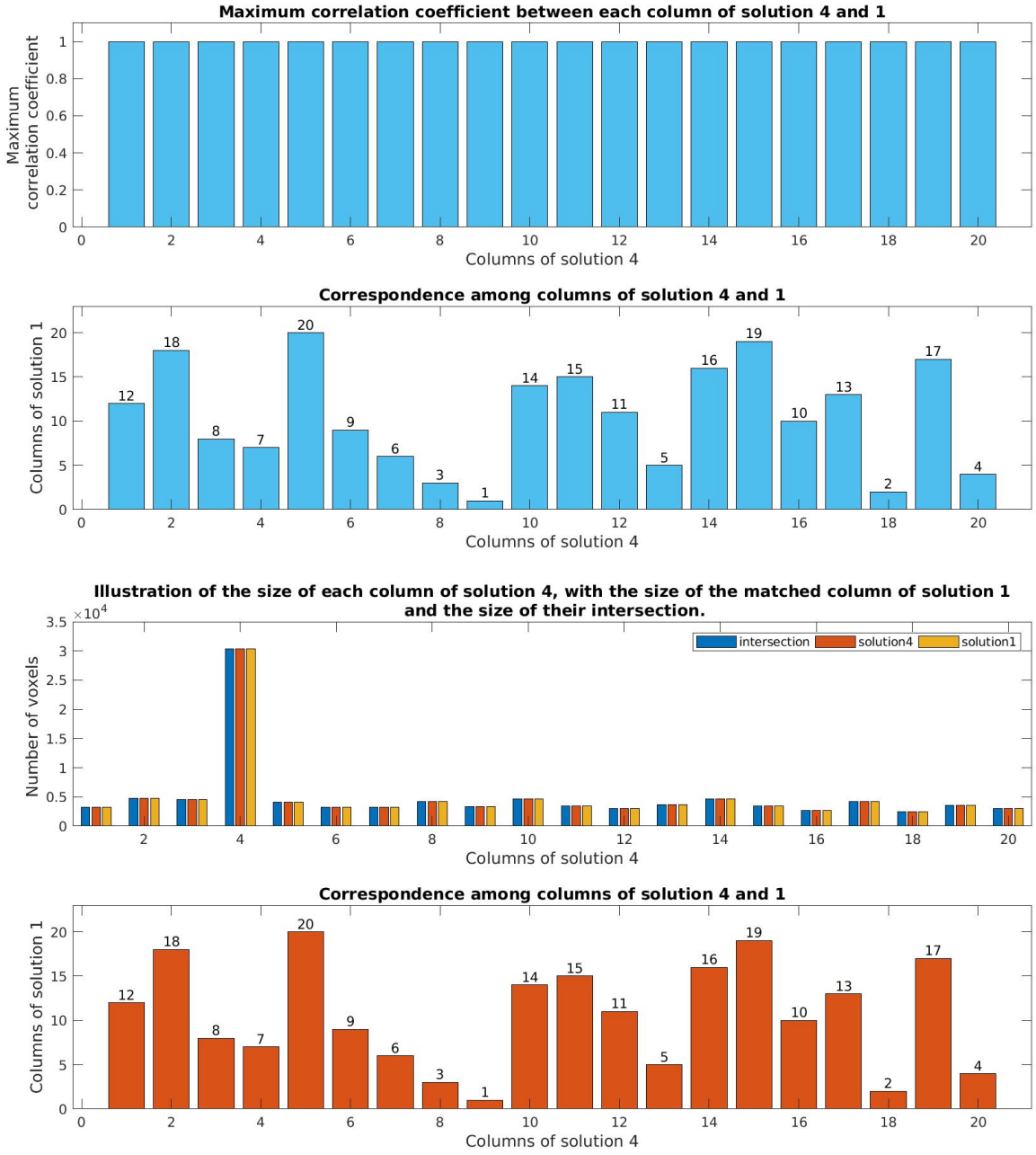


Figure A.182: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on intersection. Solutions 4, 1 correspond to accelerated penalty method and SNR=0 dB.

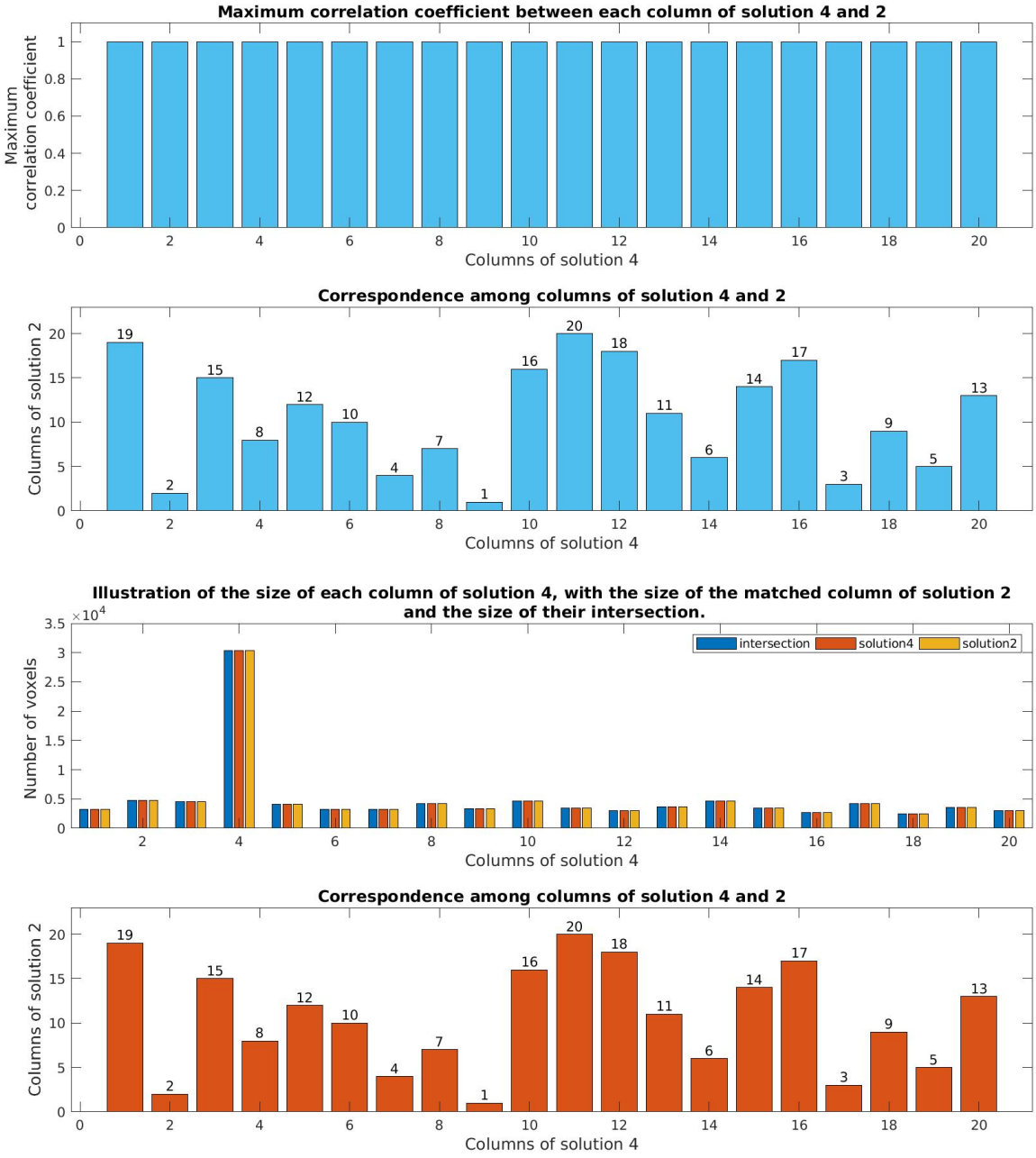


Figure A.183: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on intersection. Solutions 4, 2 correspond to accelerated penalty method and SNR=0 dB.

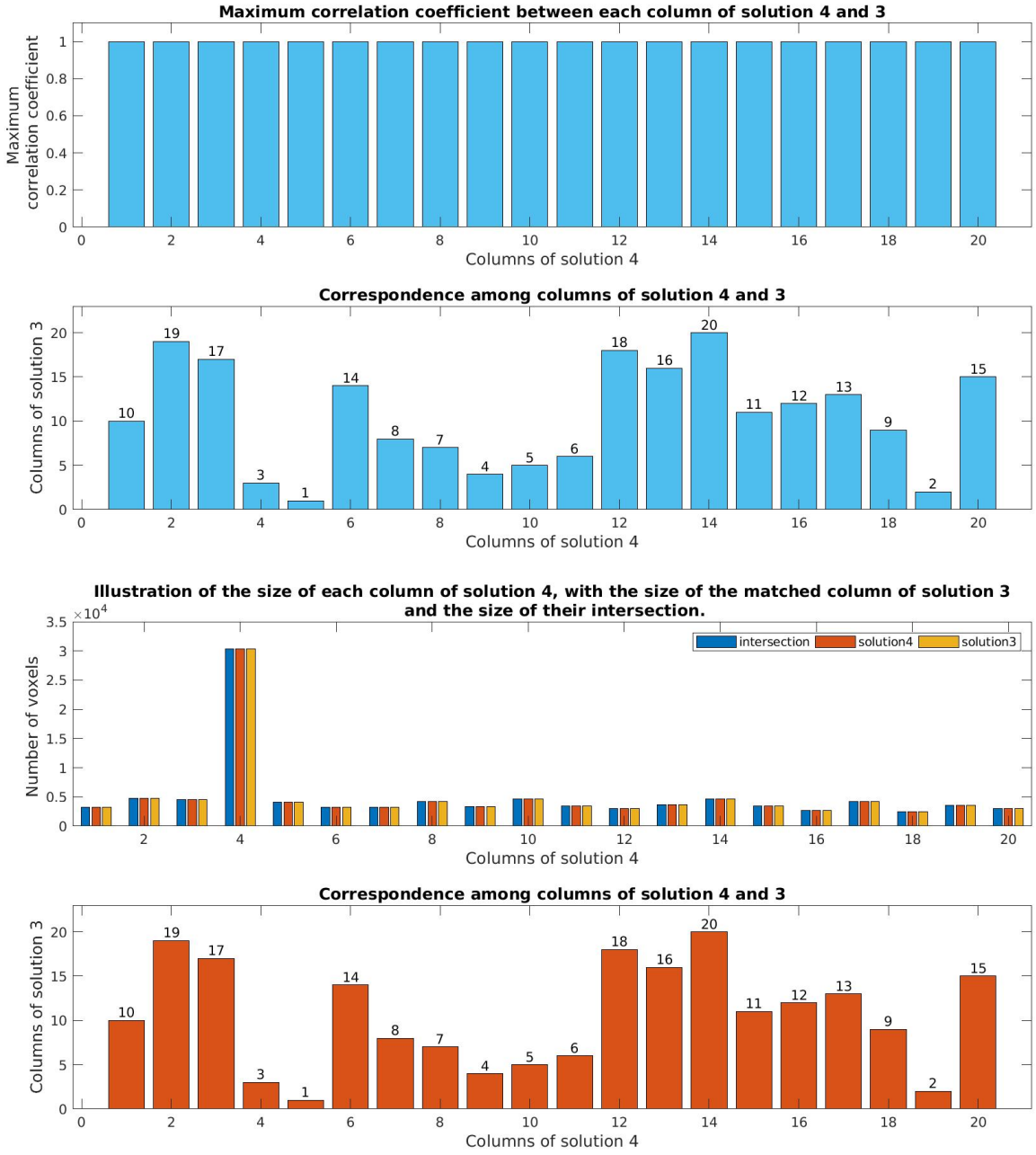


Figure A.184: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on intersection. Solutions 4, 3 correspond to accelerated penalty method and SNR=0 dB.

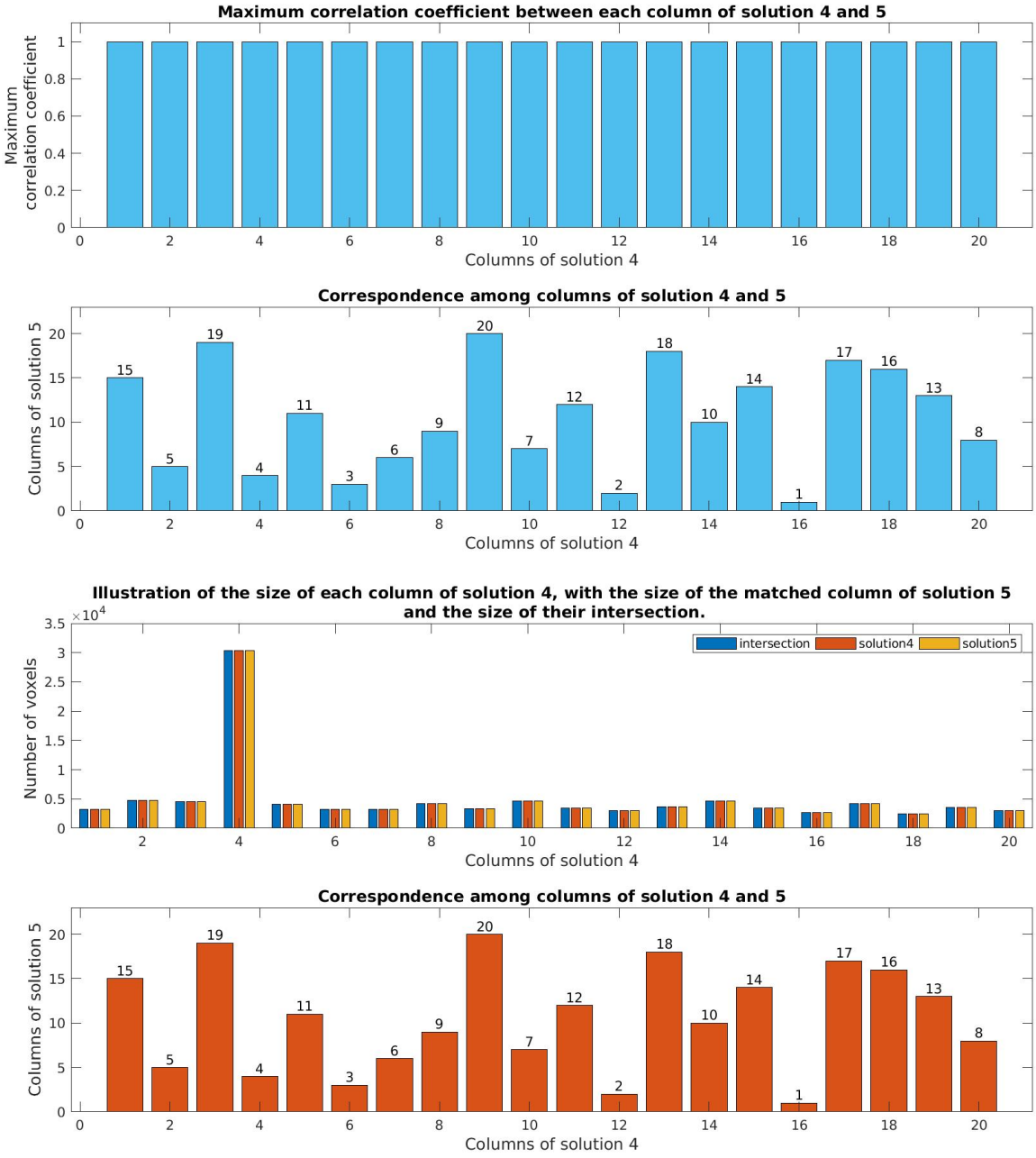


Figure A.185: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 5. We associate the column r of the solution 4 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on intersection. Solutions 4, 5 correspond to accelerated penalty method and SNR=0 dB.

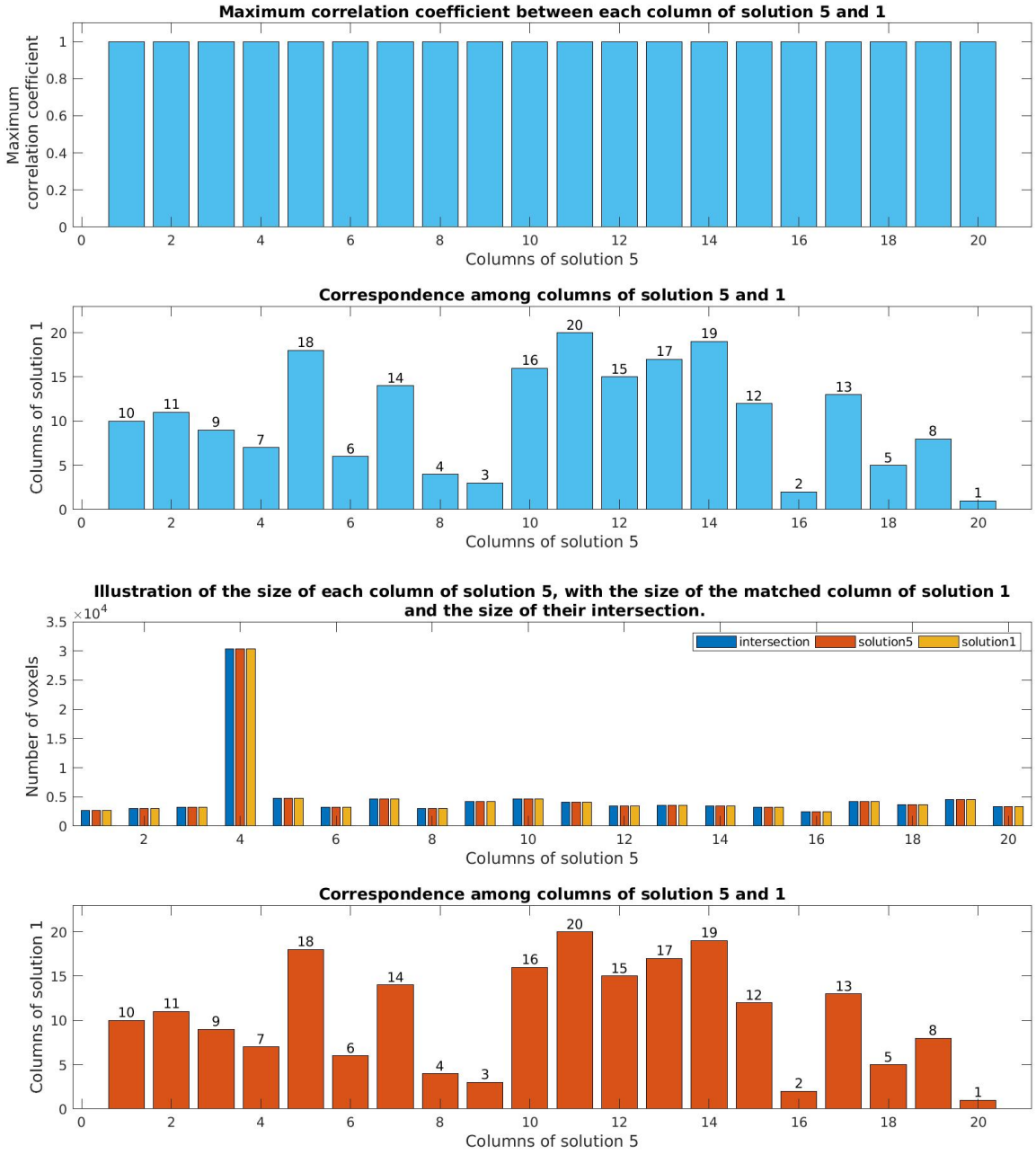


Figure A.186: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 1. We associate the column r of the solution 5 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on intersection. Solutions 5, 1 correspond to accelerated penalty method and SNR=0 dB.

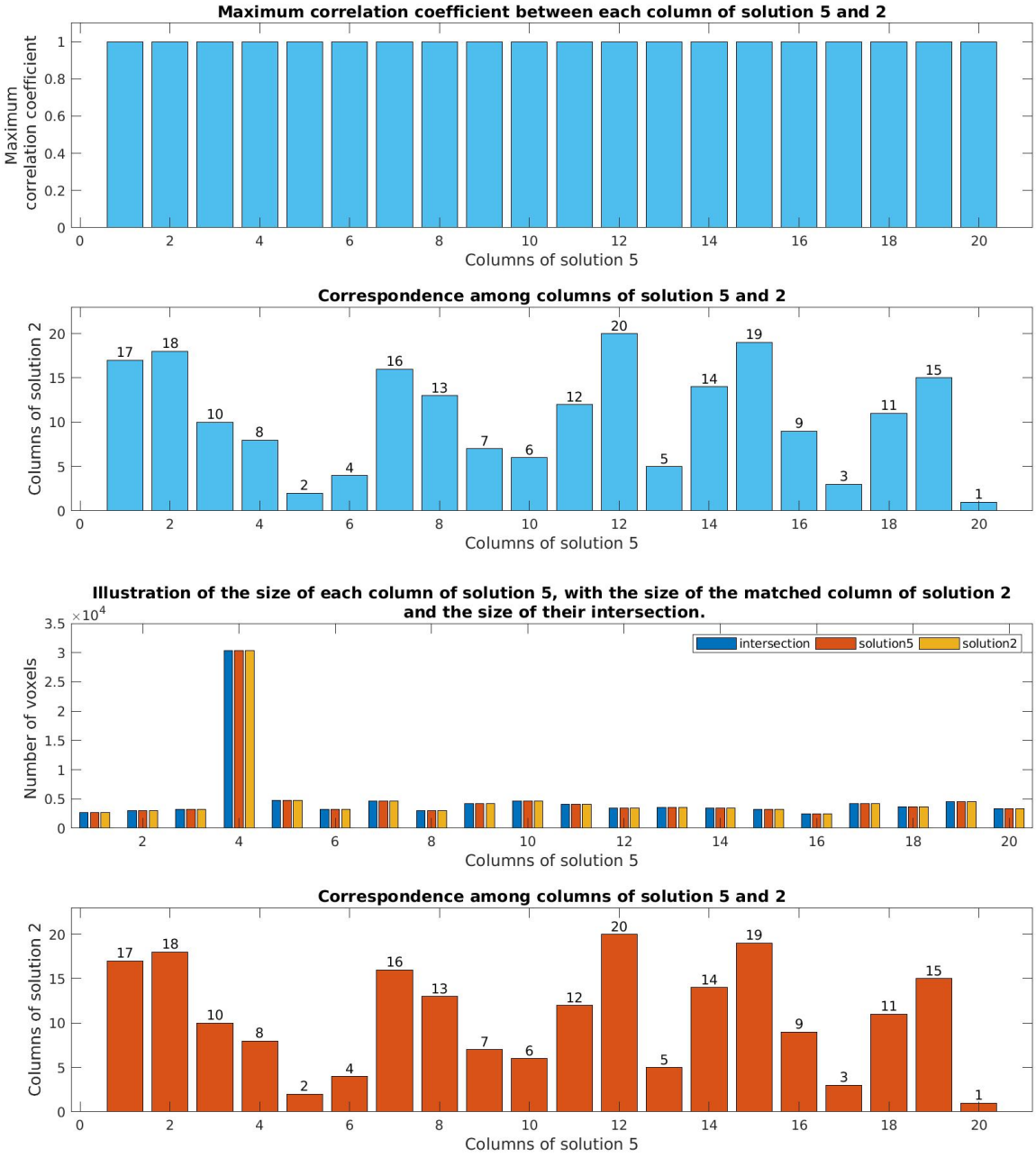


Figure A.187: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 2. We associate the column r of the solution 5 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on intersection. Solutions 5, 2 correspond to accelerated penalty method and SNR=0 dB.

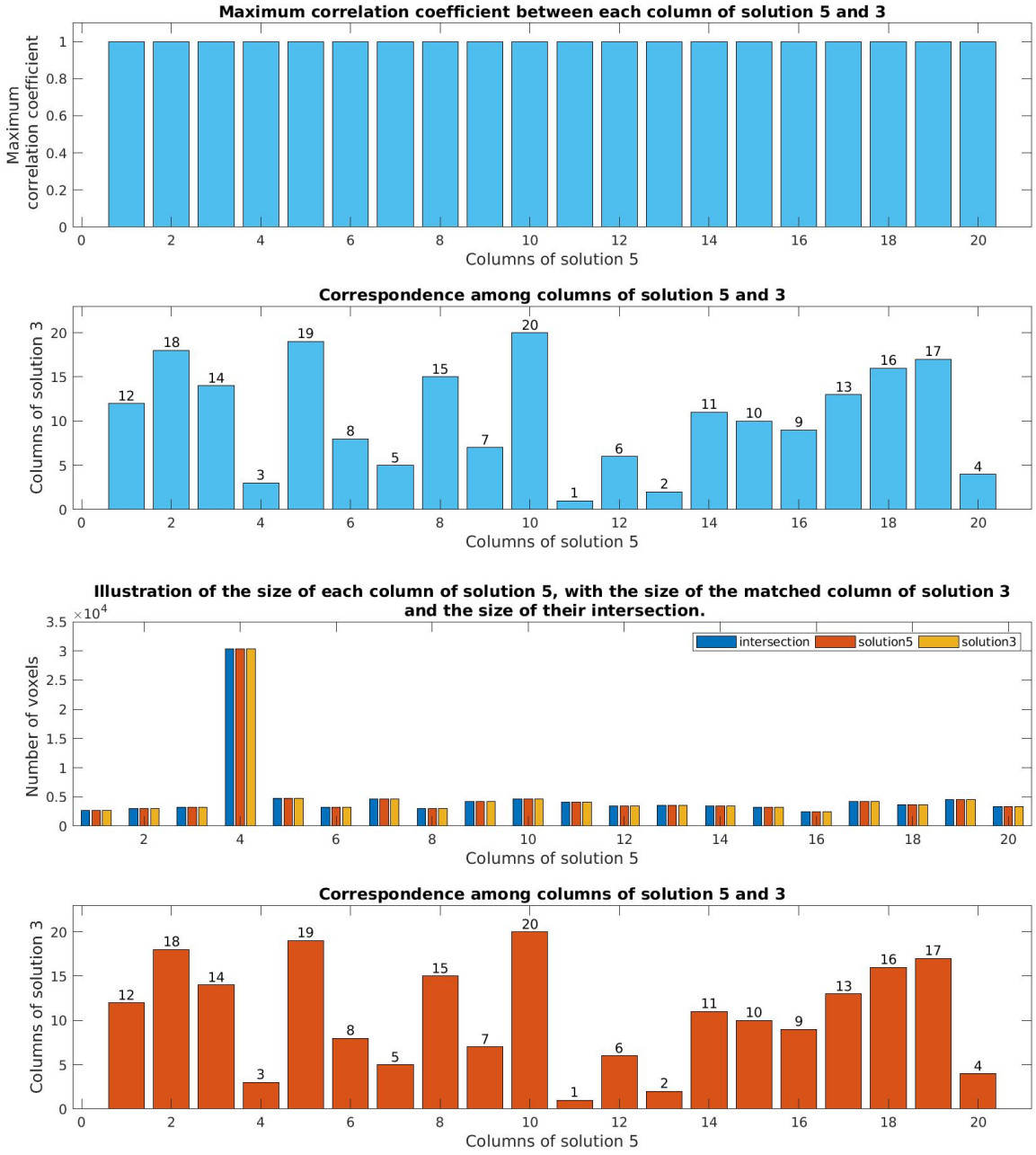


Figure A.188: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 3. We associate the column r of the solution 5 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on intersection. Solutions 5, 3 correspond to accelerated penalty method and SNR=0 dB.

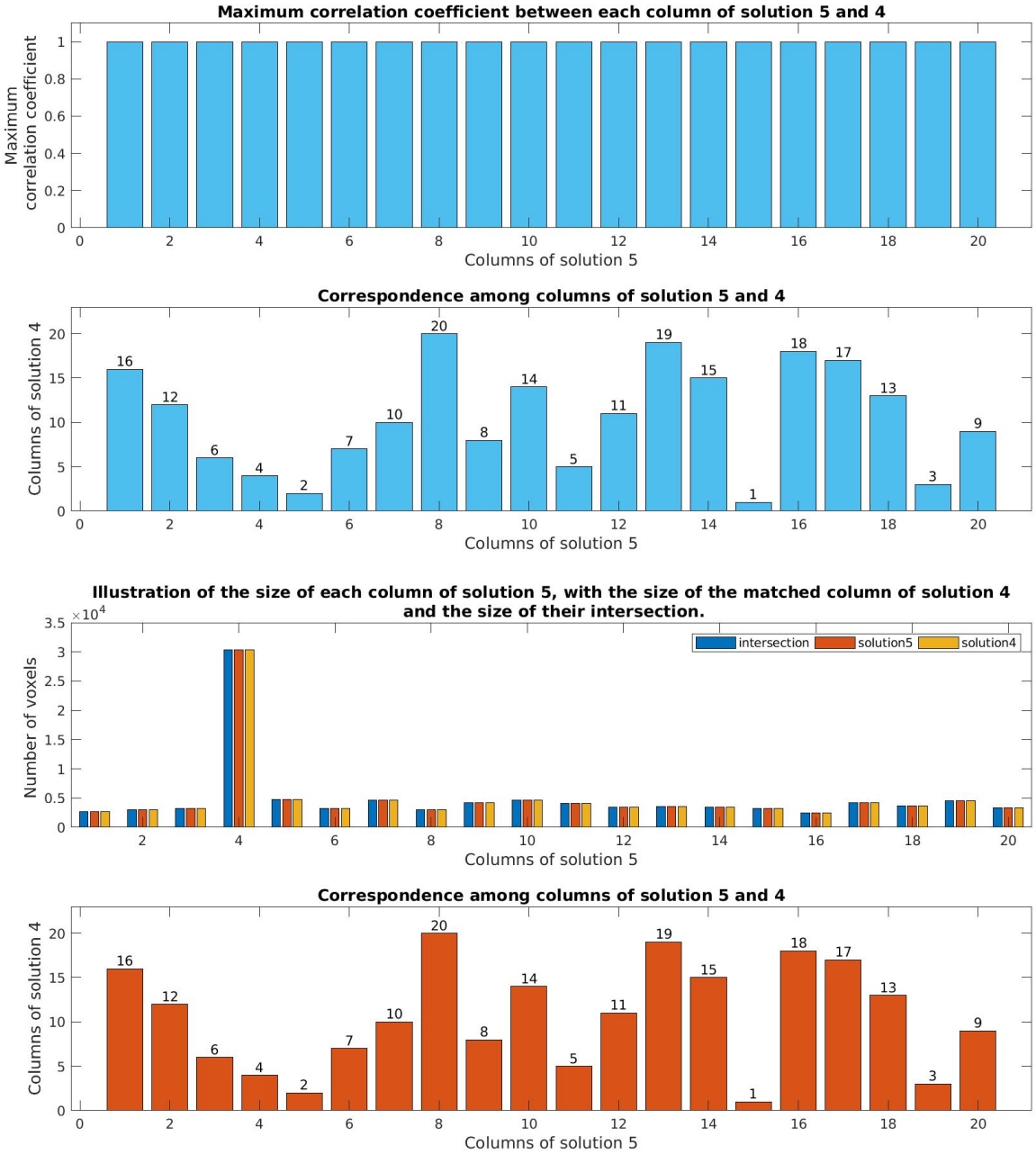


Figure A.189: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 4. We associate the column r of the solution 5 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on intersection. Solutions 5, 4 correspond to accelerated penalty method and SNR=0 dB.

· **Comparison of Accelerated penalty method solutions to real factor A.**

· **Noiseless case**

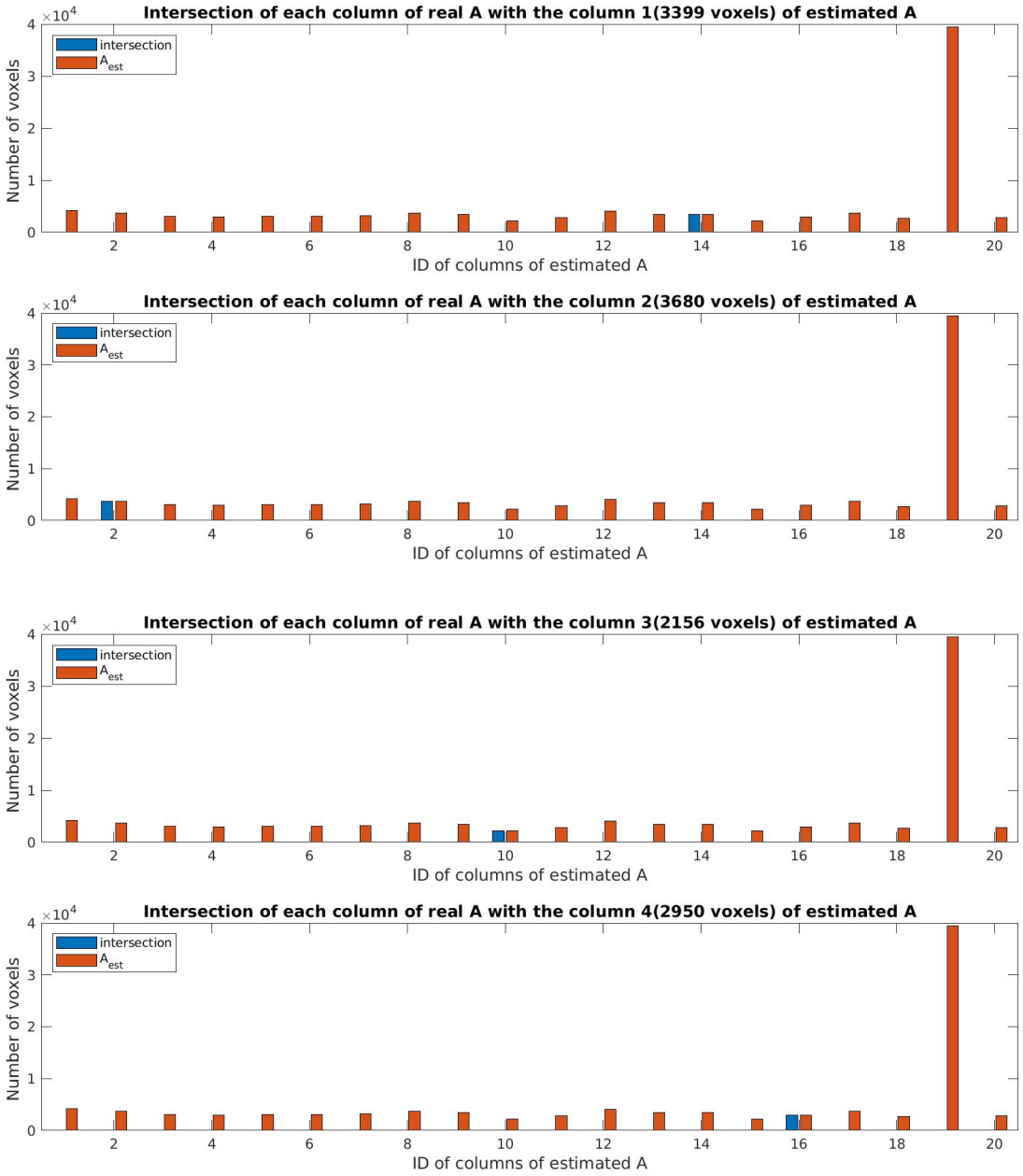


Figure A.190: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

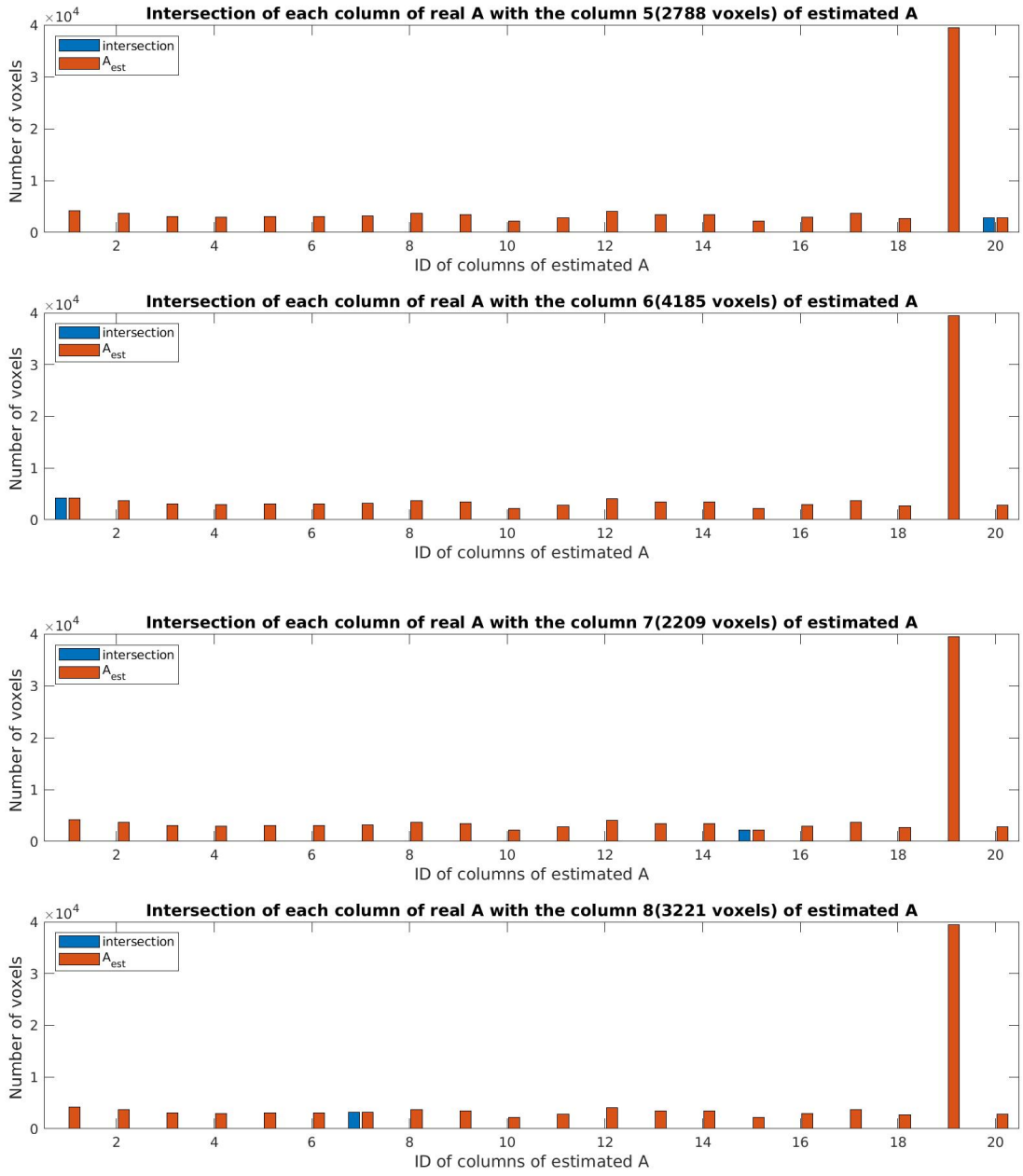


Figure A.191: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

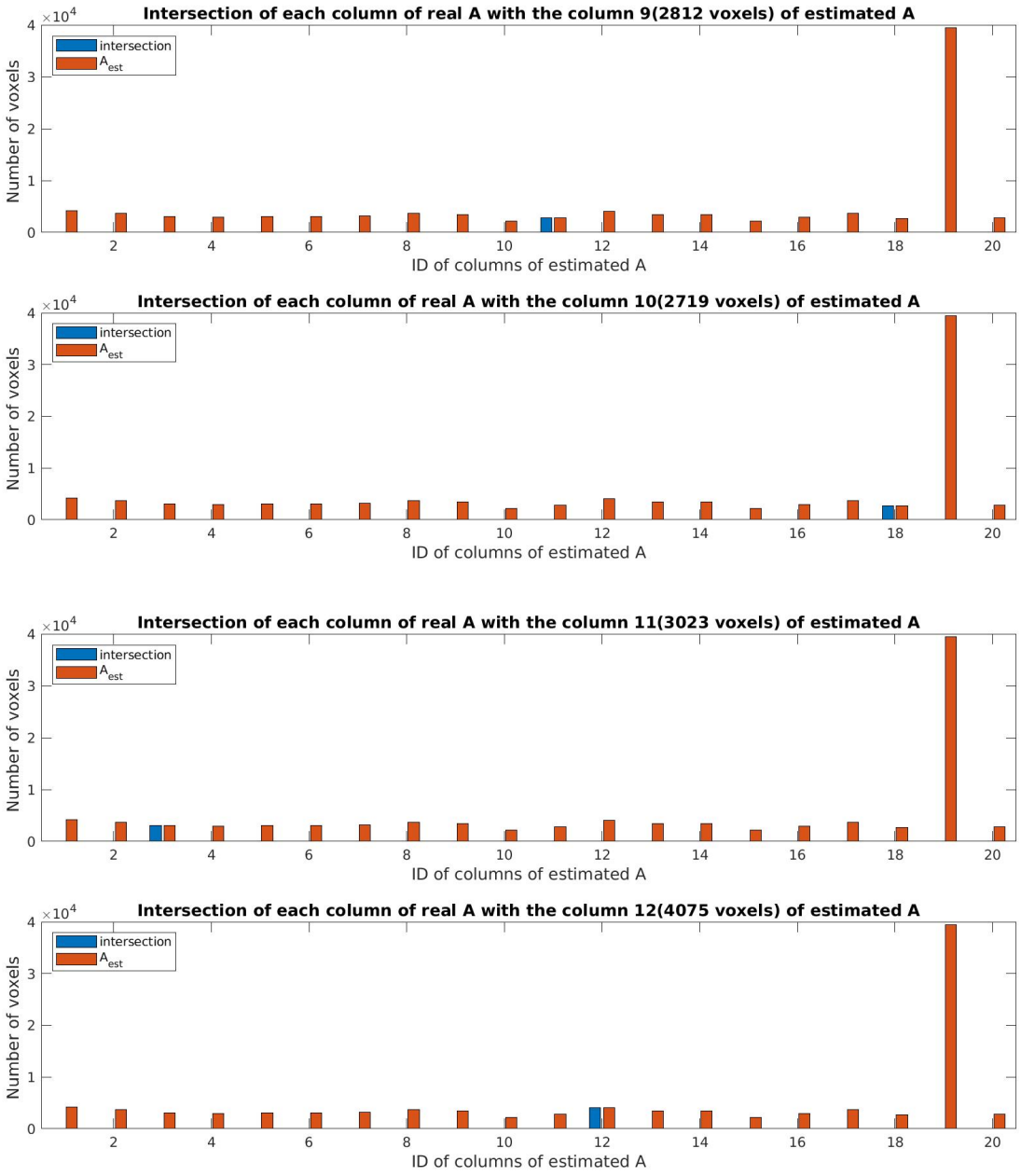


Figure A.192: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

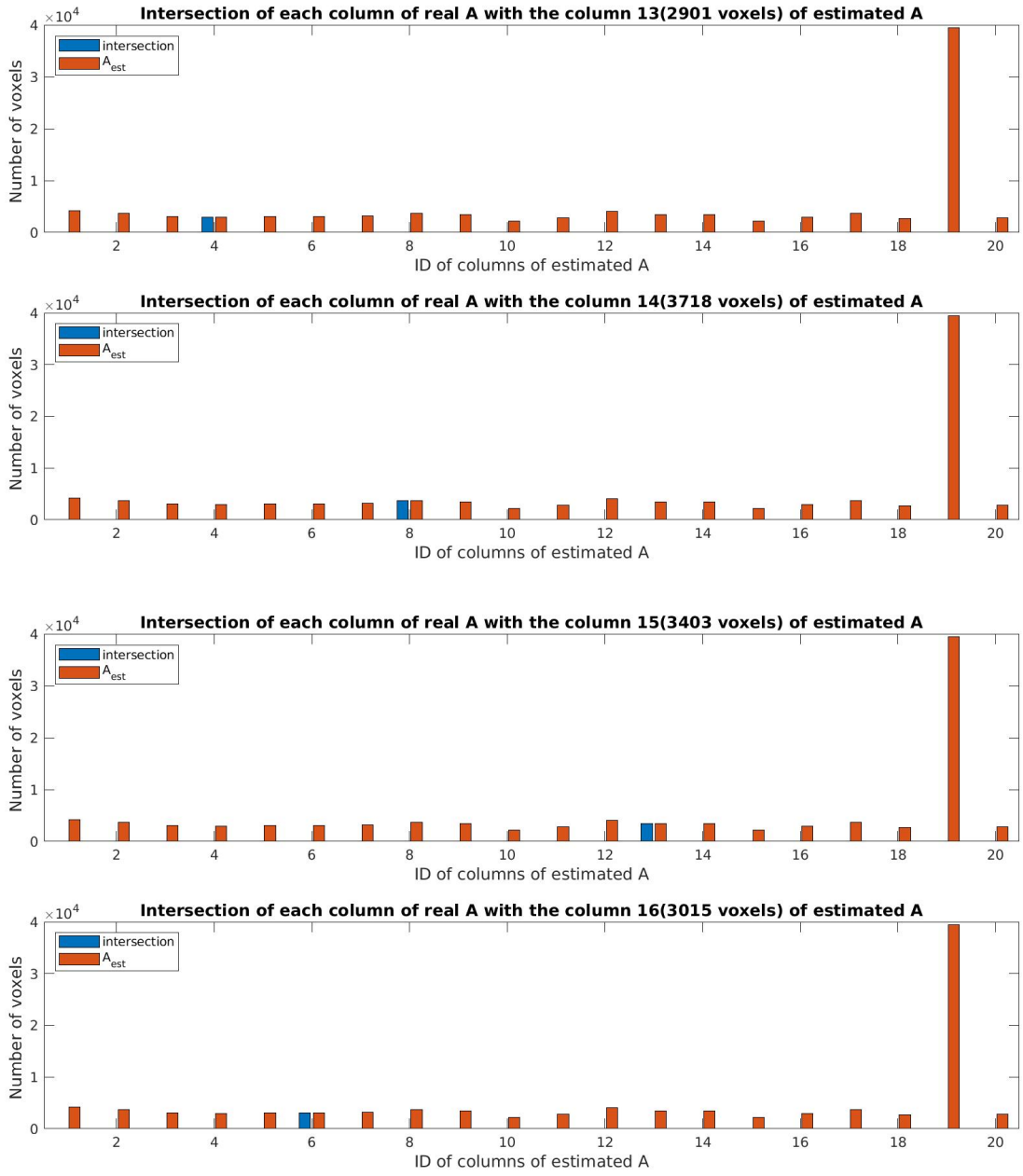


Figure A.193: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

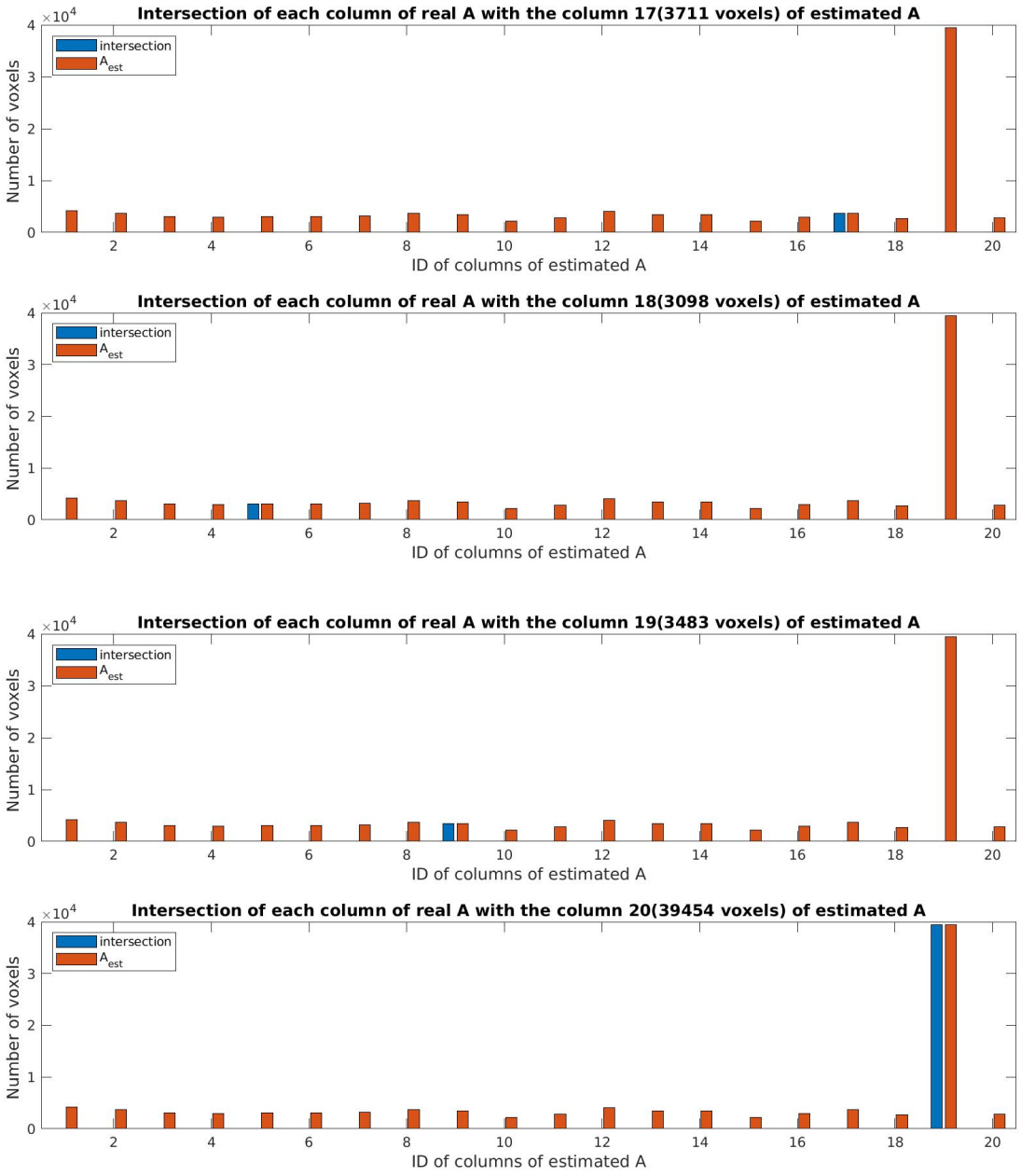


Figure A.194: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

· **SNR= 0 dB**

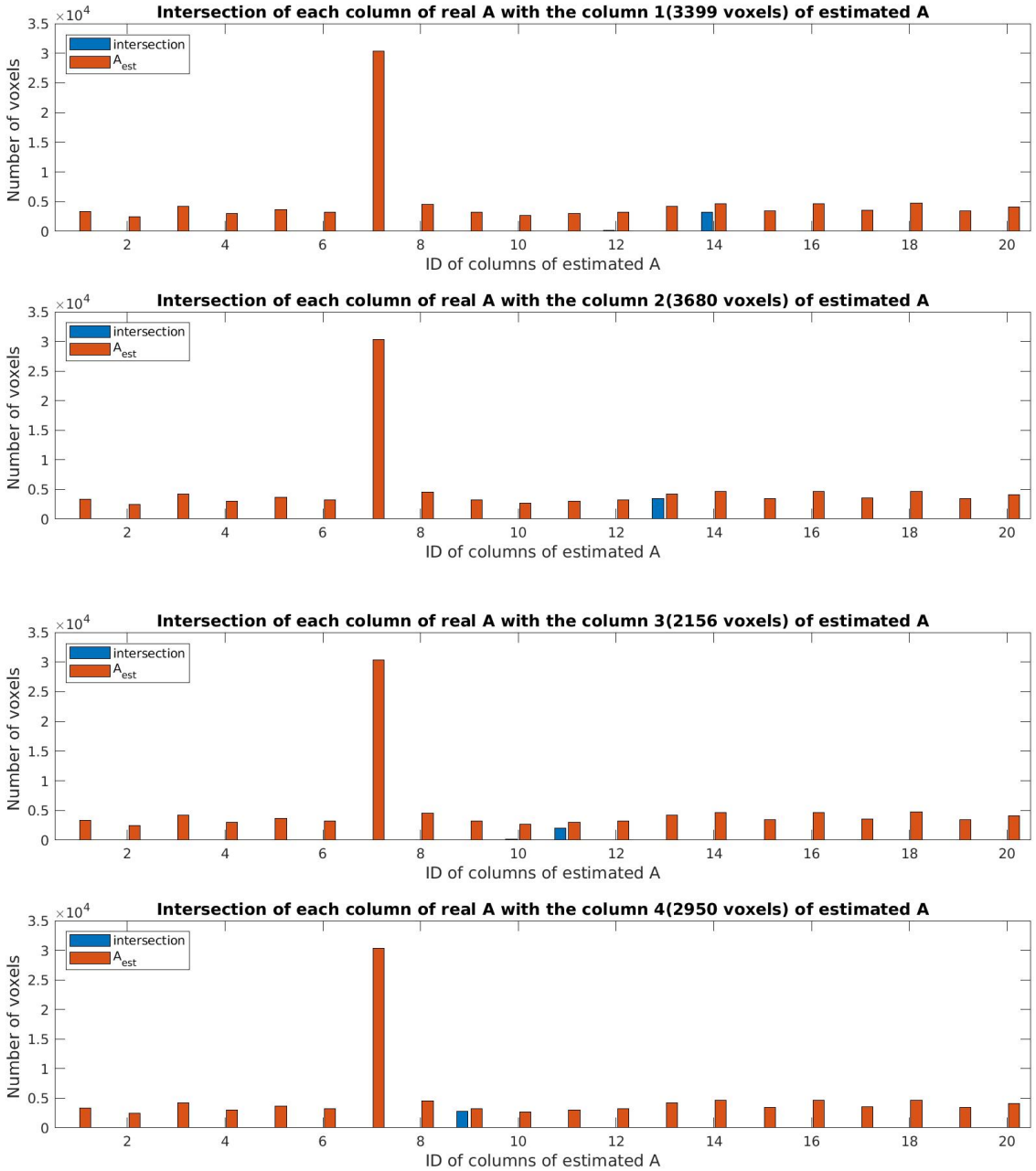


Figure A.195: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

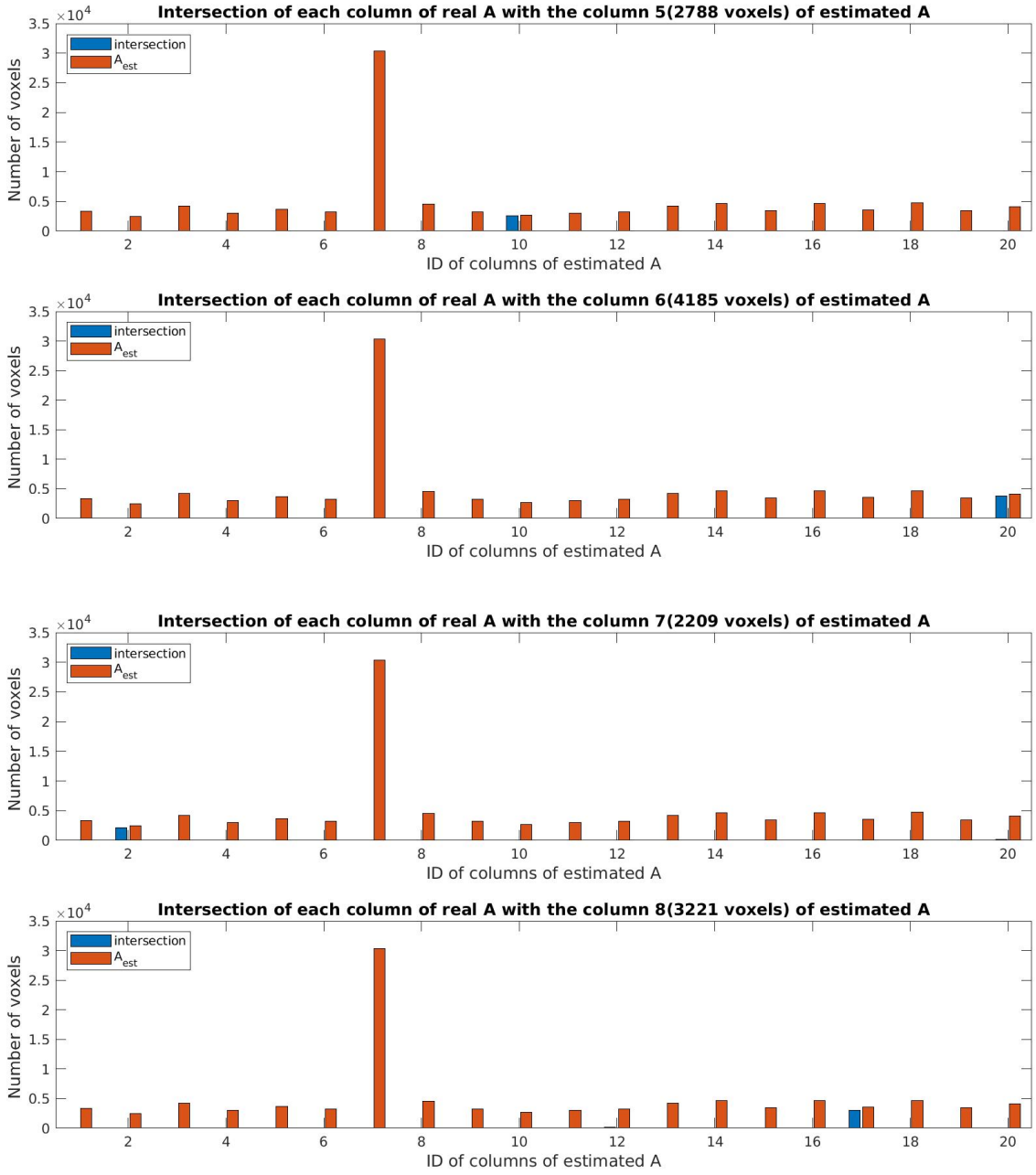


Figure A.196: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

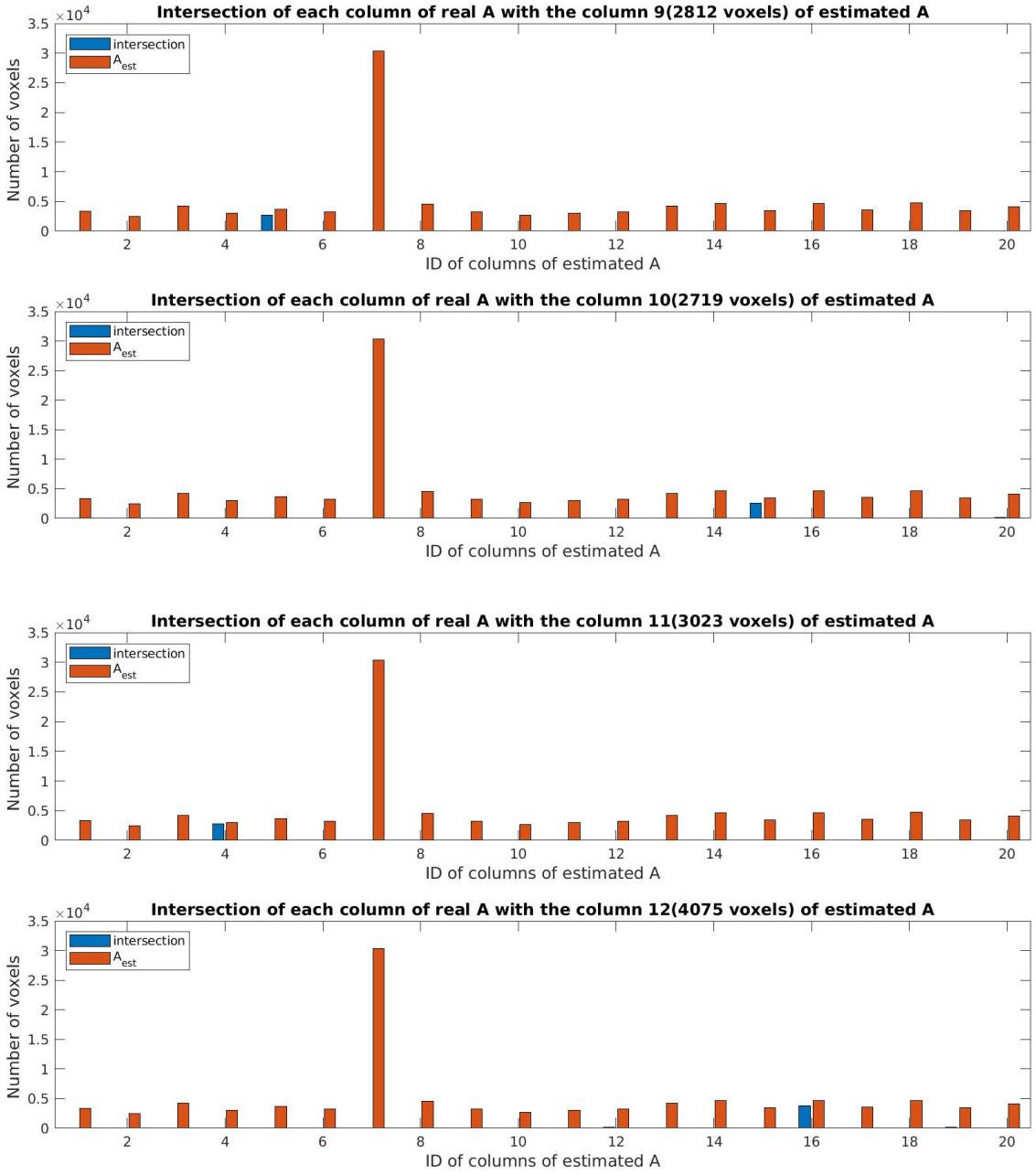


Figure A.197: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

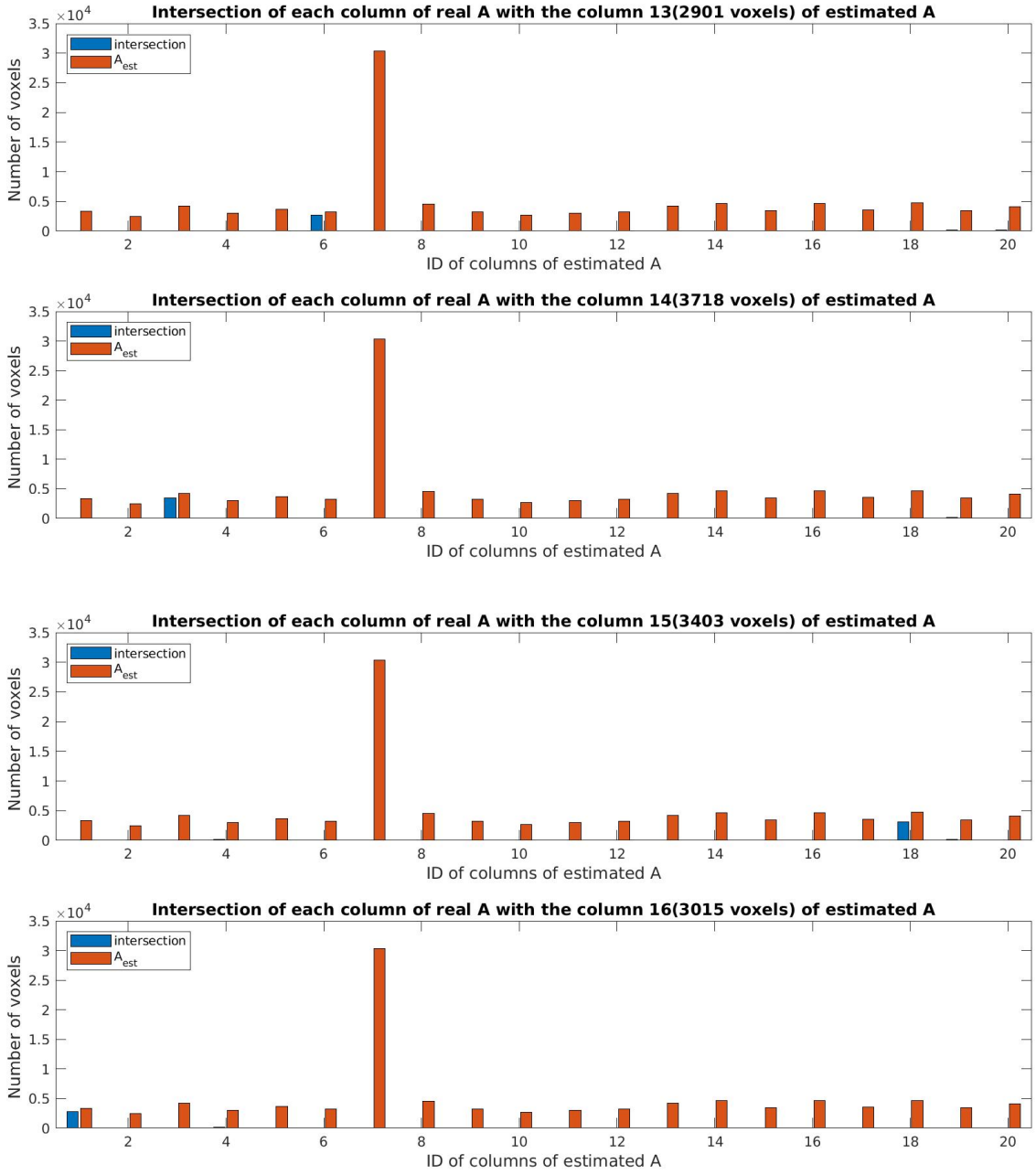


Figure A.198: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

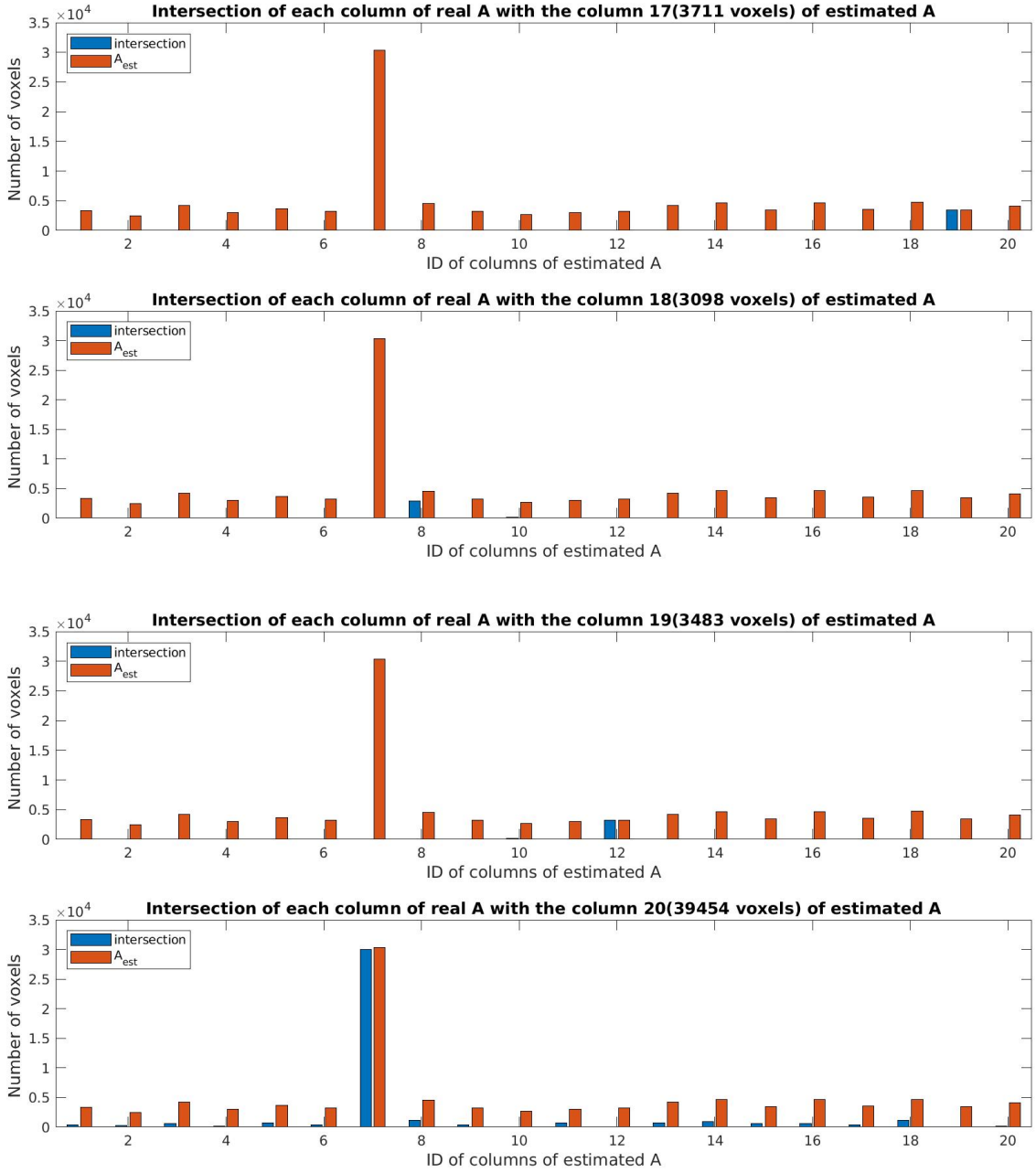


Figure A.199: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

· **Histograms for SNR=0 dB**

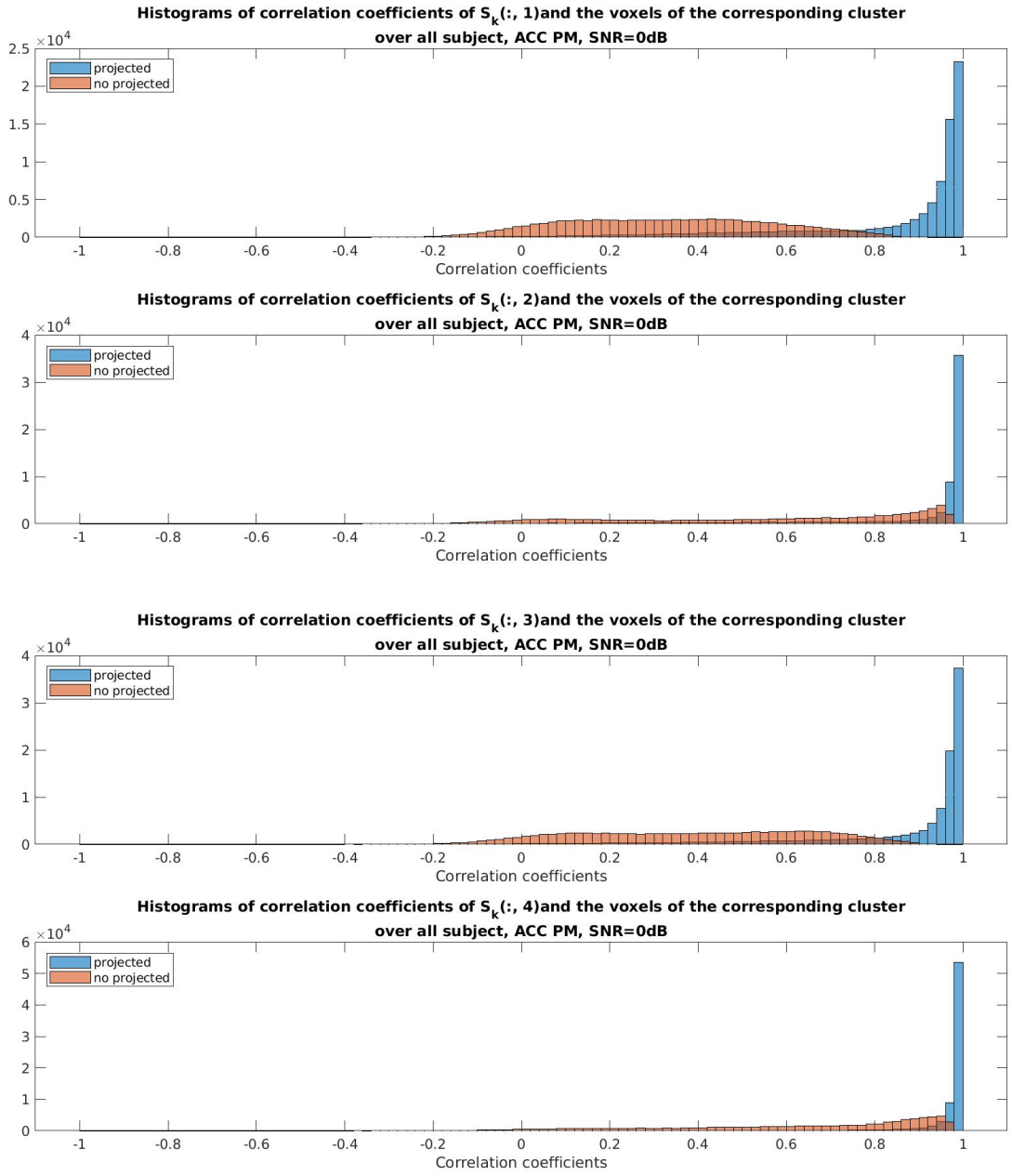


Figure A.200: For the column r and subject k, for $r = 1, \ldots, 4$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 1 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 2 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 3 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 4 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

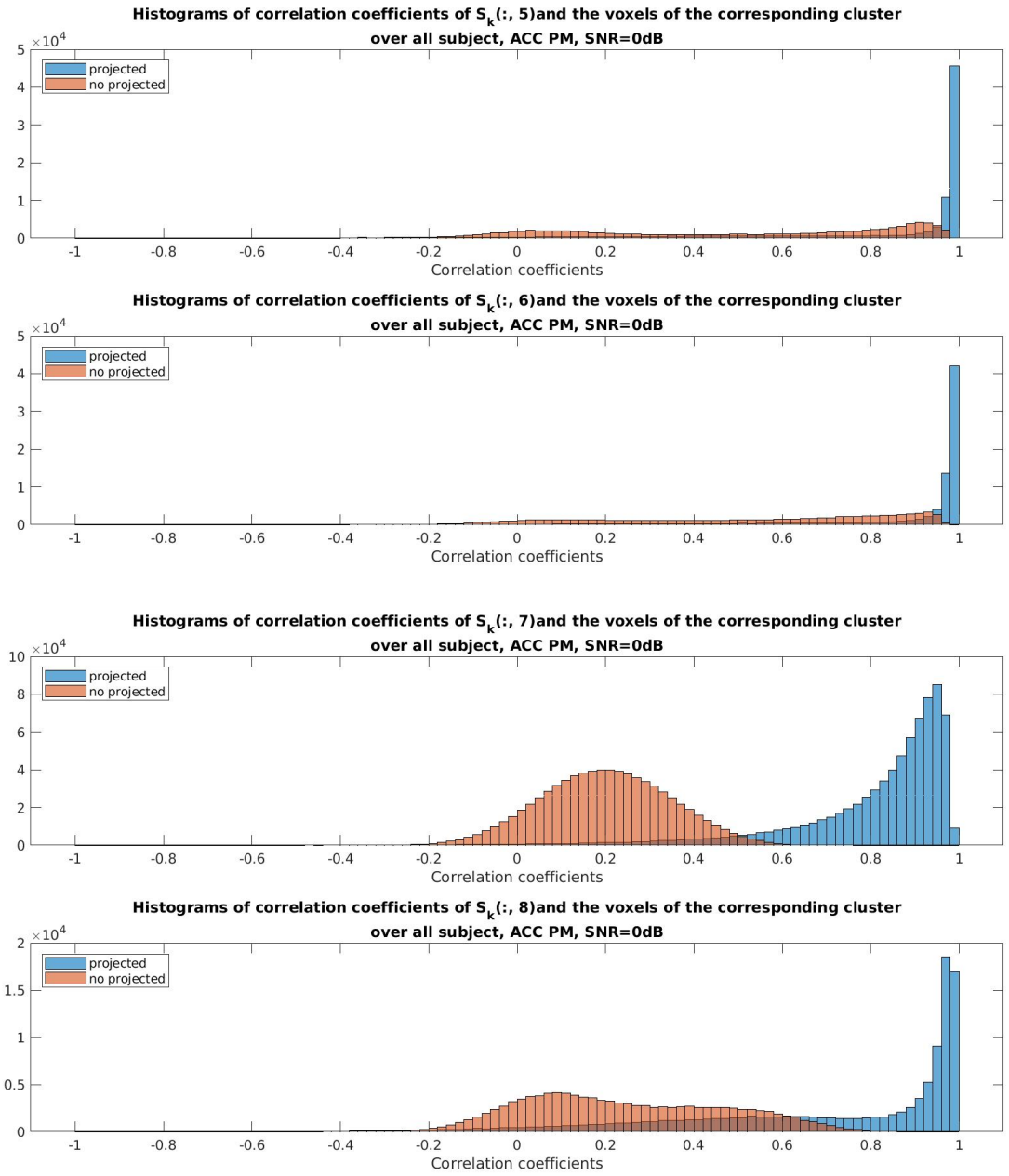


Figure A.201: For the column r and subject k, for $r = 5, \ldots, 8$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 5 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 6 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 7 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 8 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

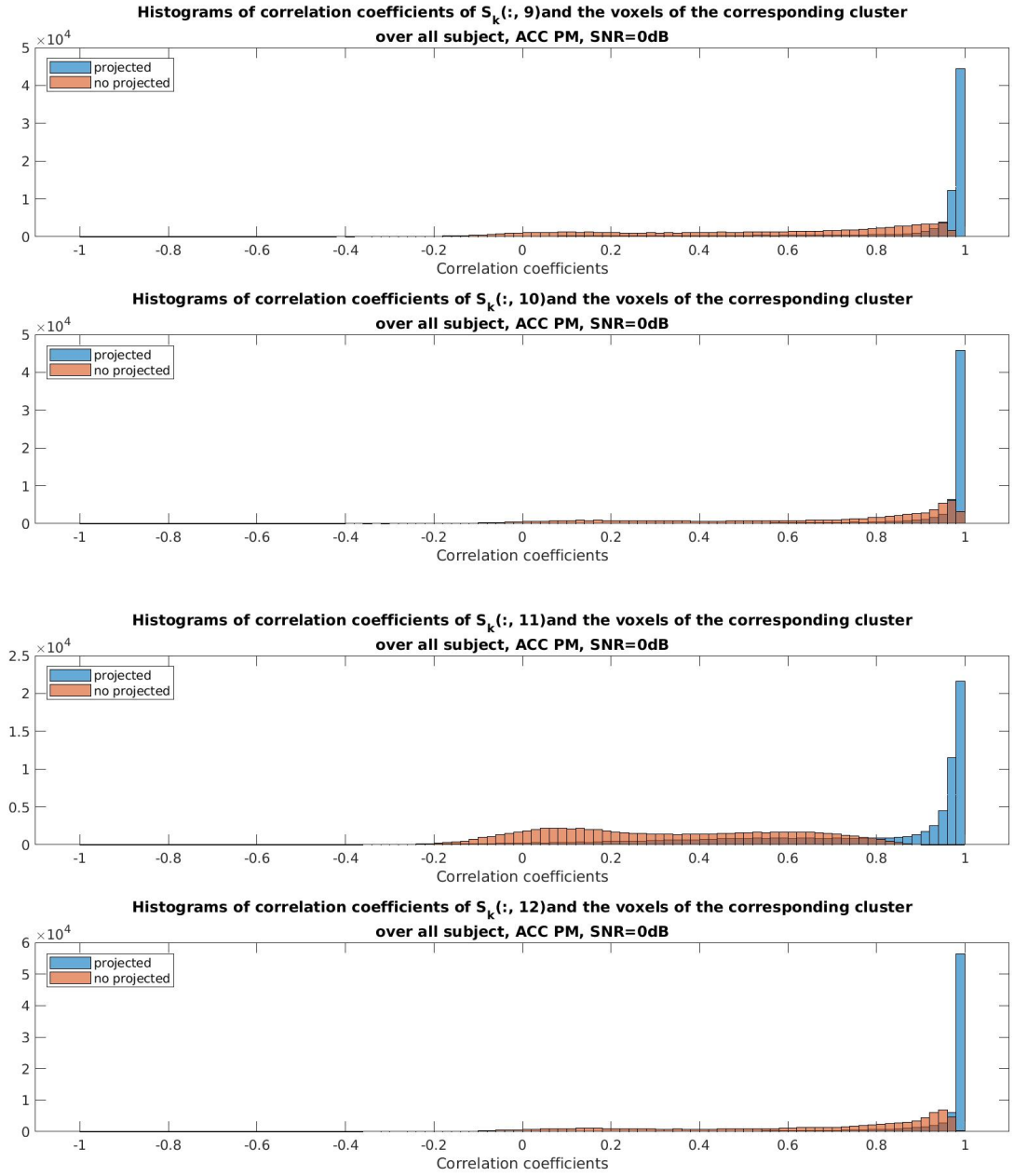


Figure A.202: For the column r and subject k, for $r = 9, \ldots, 12$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 9 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 10 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 11 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 12 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

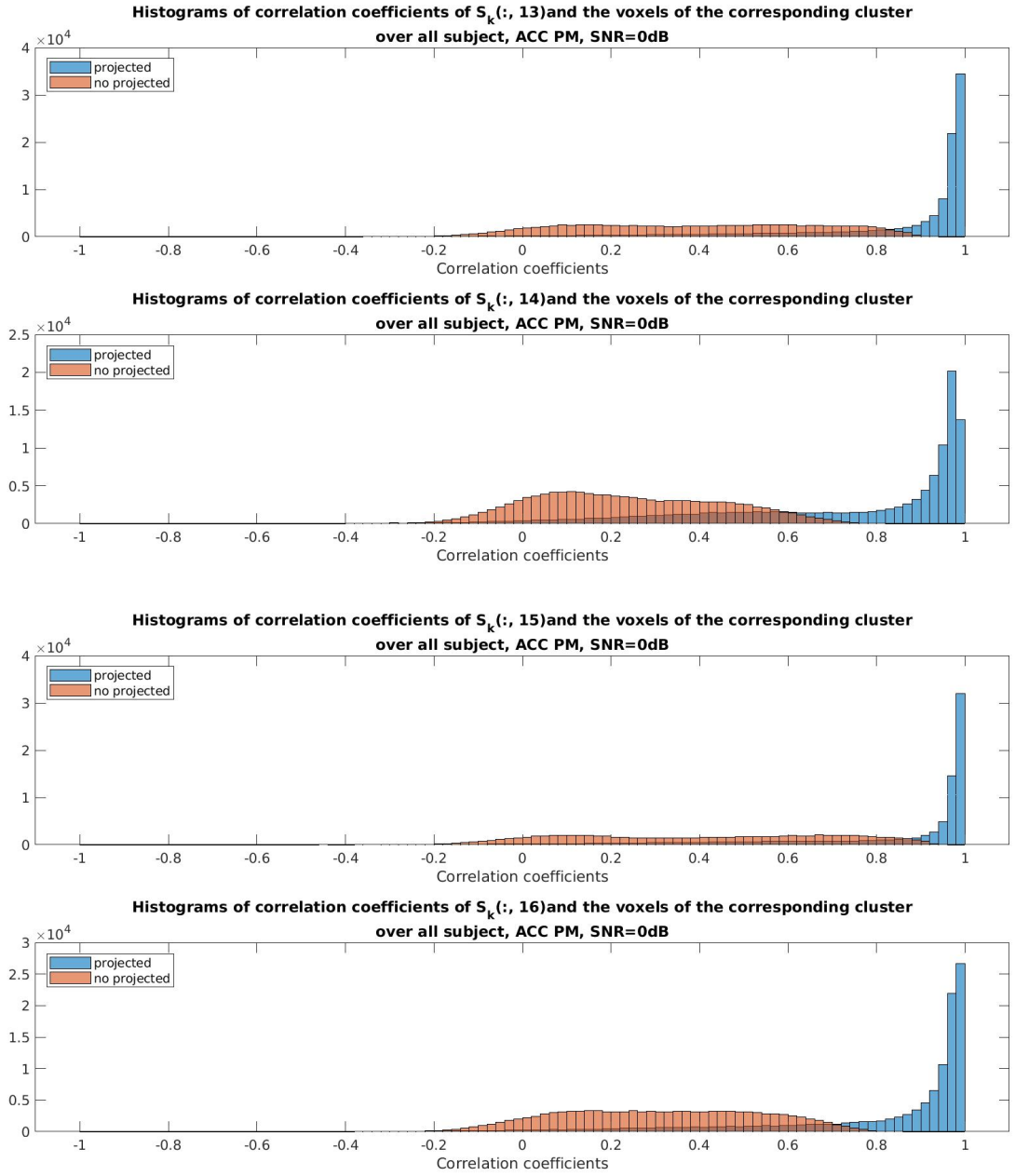


Figure A.203: For the column r and subject k, for $r = 13, \ldots, 16$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 13 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 14 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 15 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 16 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

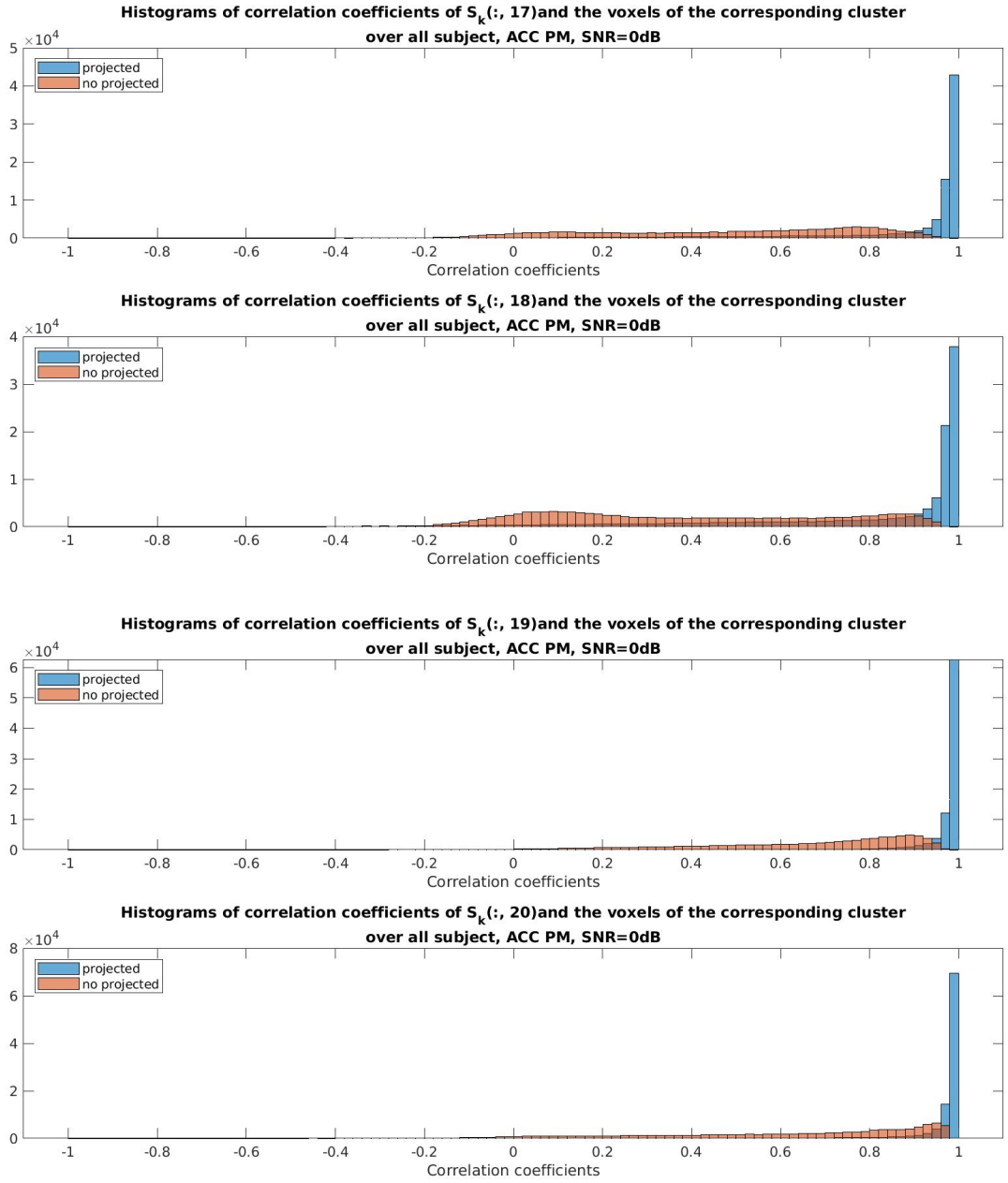


Figure A.204: For the column r and subject k, for $r = 17, \ldots, 20$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 17 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 18 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 19 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 20 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

A.3.3 Case 3: Corrupted common orthonormal matrix A with each subject having correlated temporal components S_k.

Accelerated Penalty method

· **Comparison of Accelerated Penalty method solutions**

· **Noiseless data**

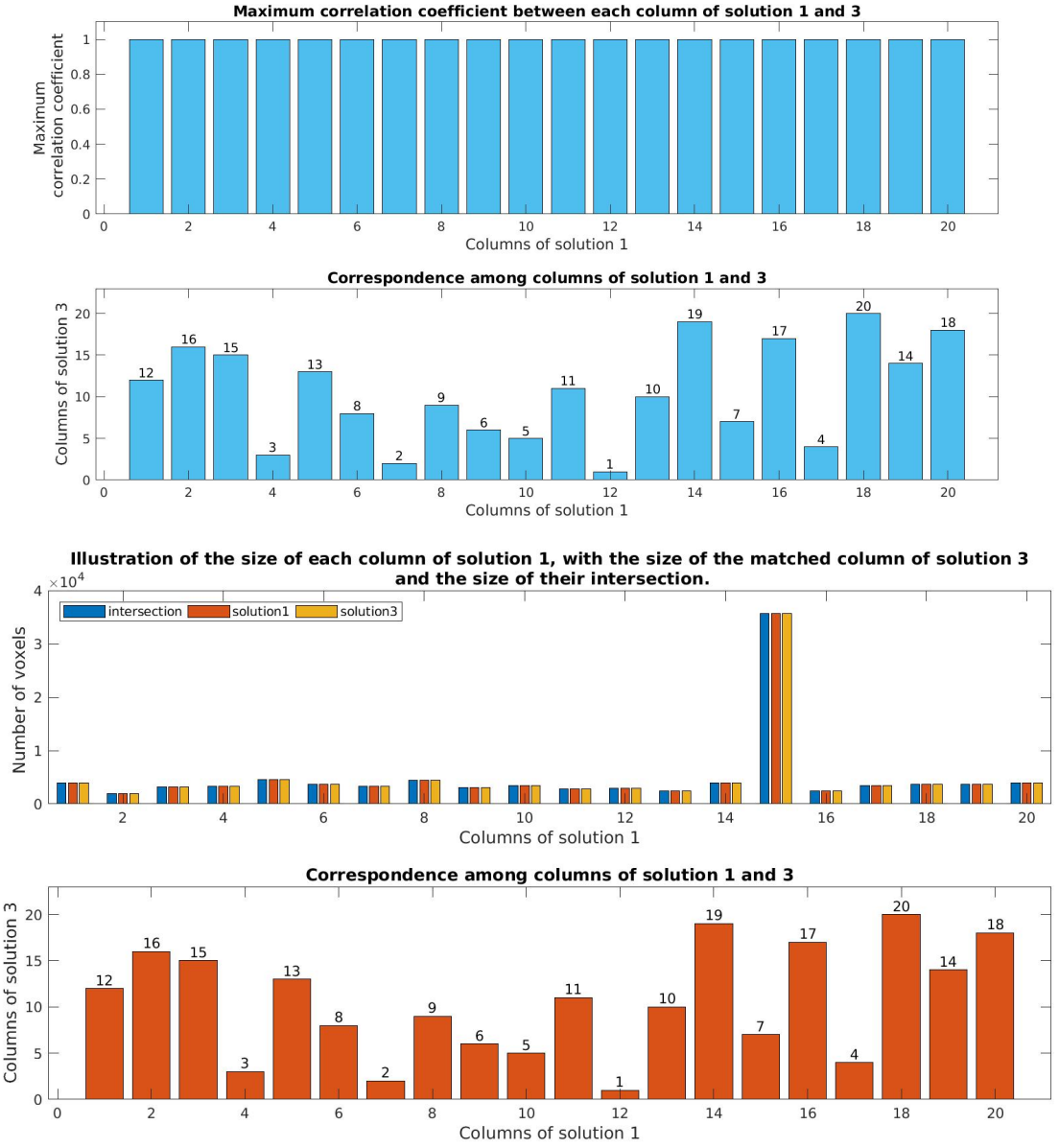


Figure A.205: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on intersection. Solutions 1, 3 correspond to accelerated penalty method and noiseless data.

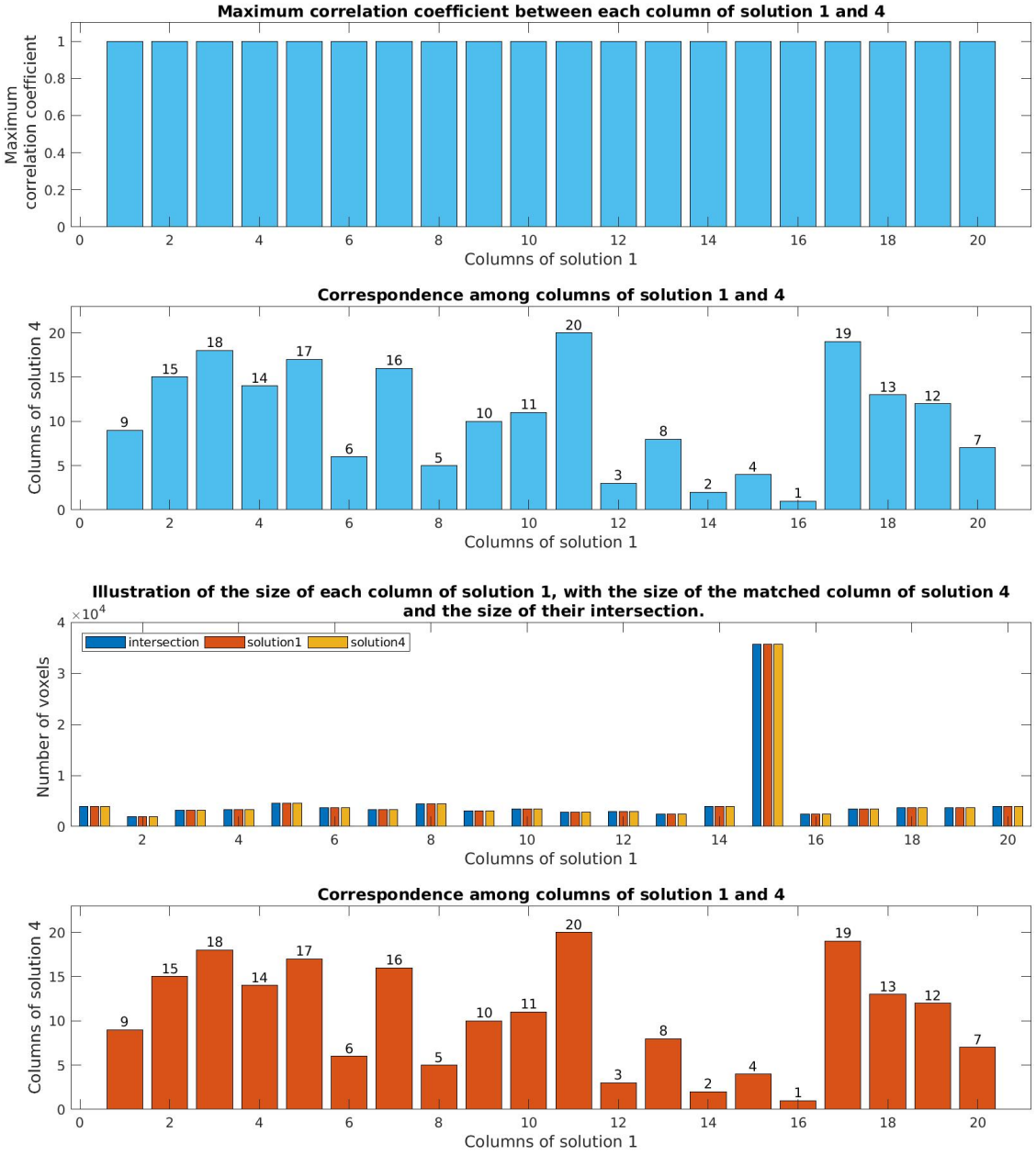


Figure A.206: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on intersection. Solutions 1, 4 correspond to accelerated penalty method and noiseless data.

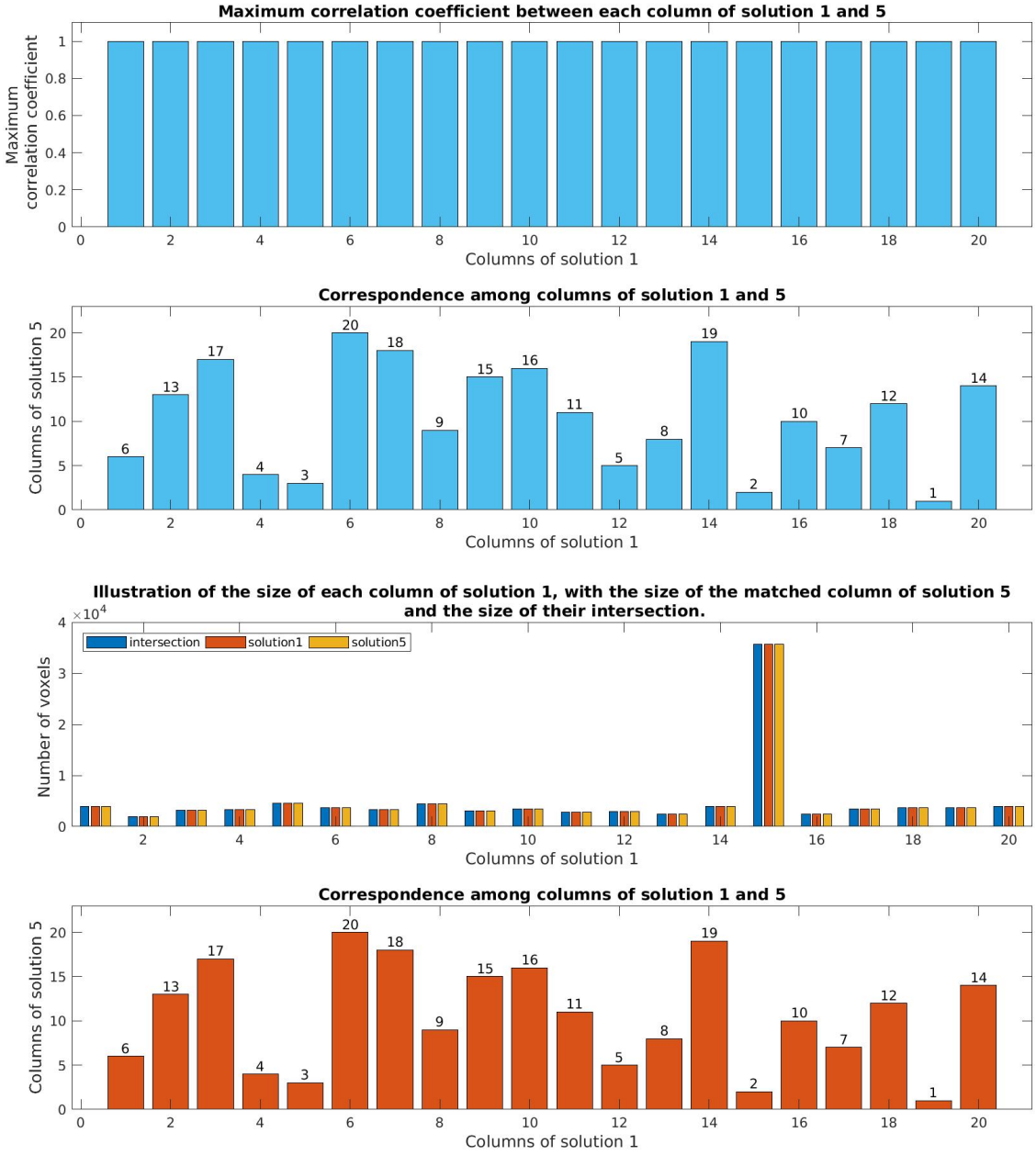


Figure A.207: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 5. We associate the column r of the solution 1 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on intersection. Solutions 1, 5 correspond to accelerated penalty method and noiseless data.

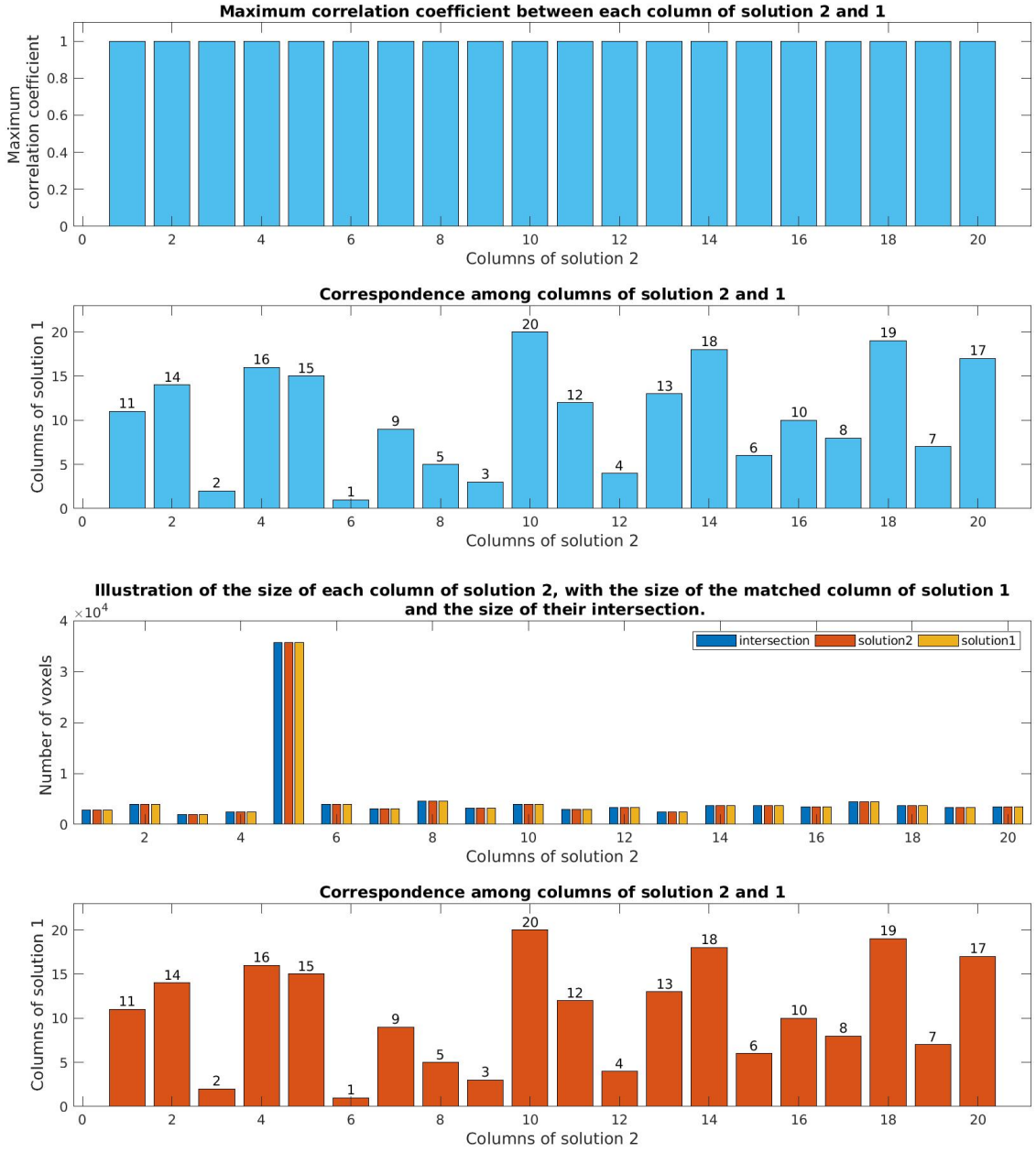


Figure A.208: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on intersection. Solutions 2, 1 correspond to accelerated penalty method and noiseless data.

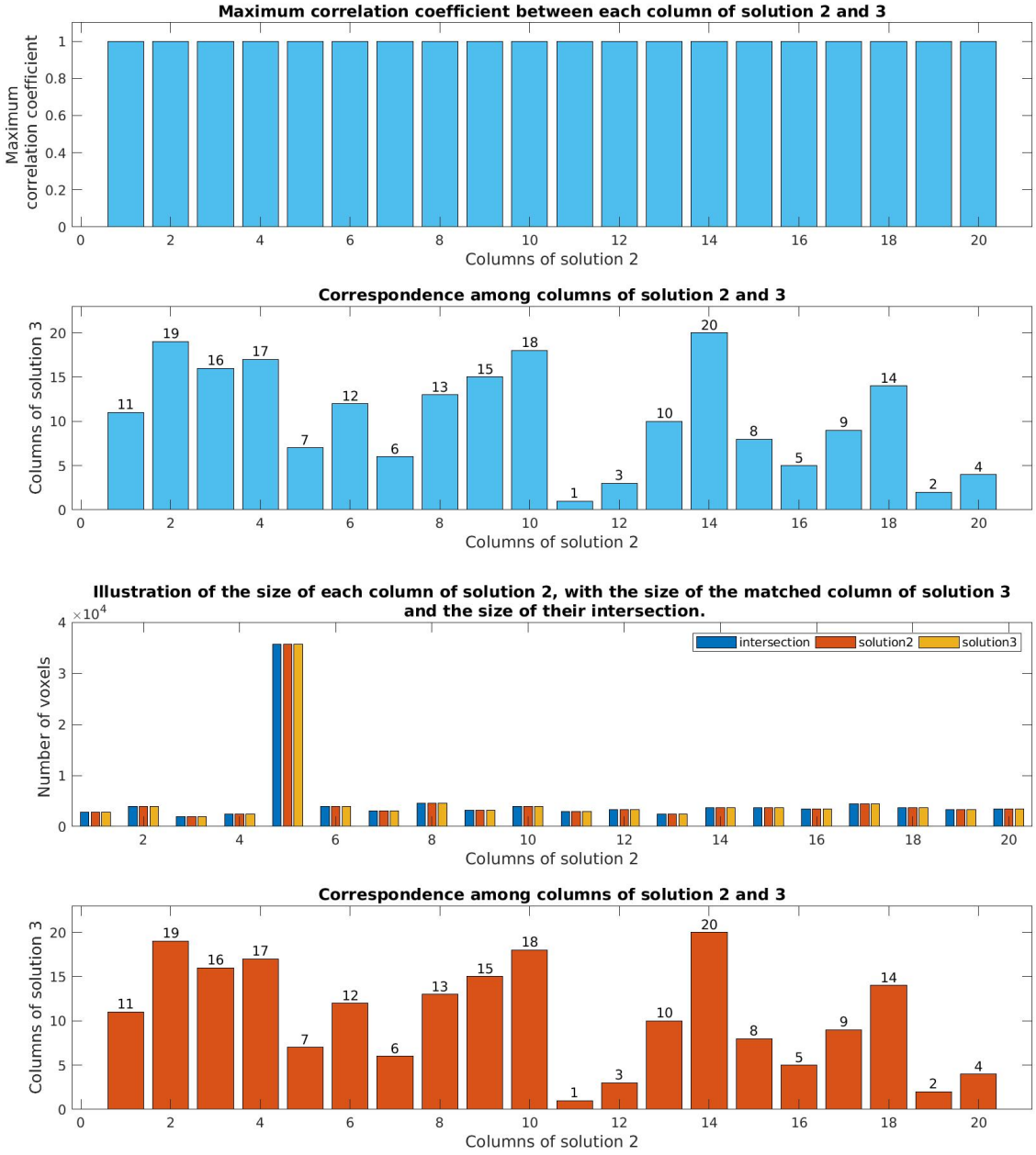


Figure A.209: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on intersection. Solutions 2, 3 correspond to accelerated penalty method and noiseless data.

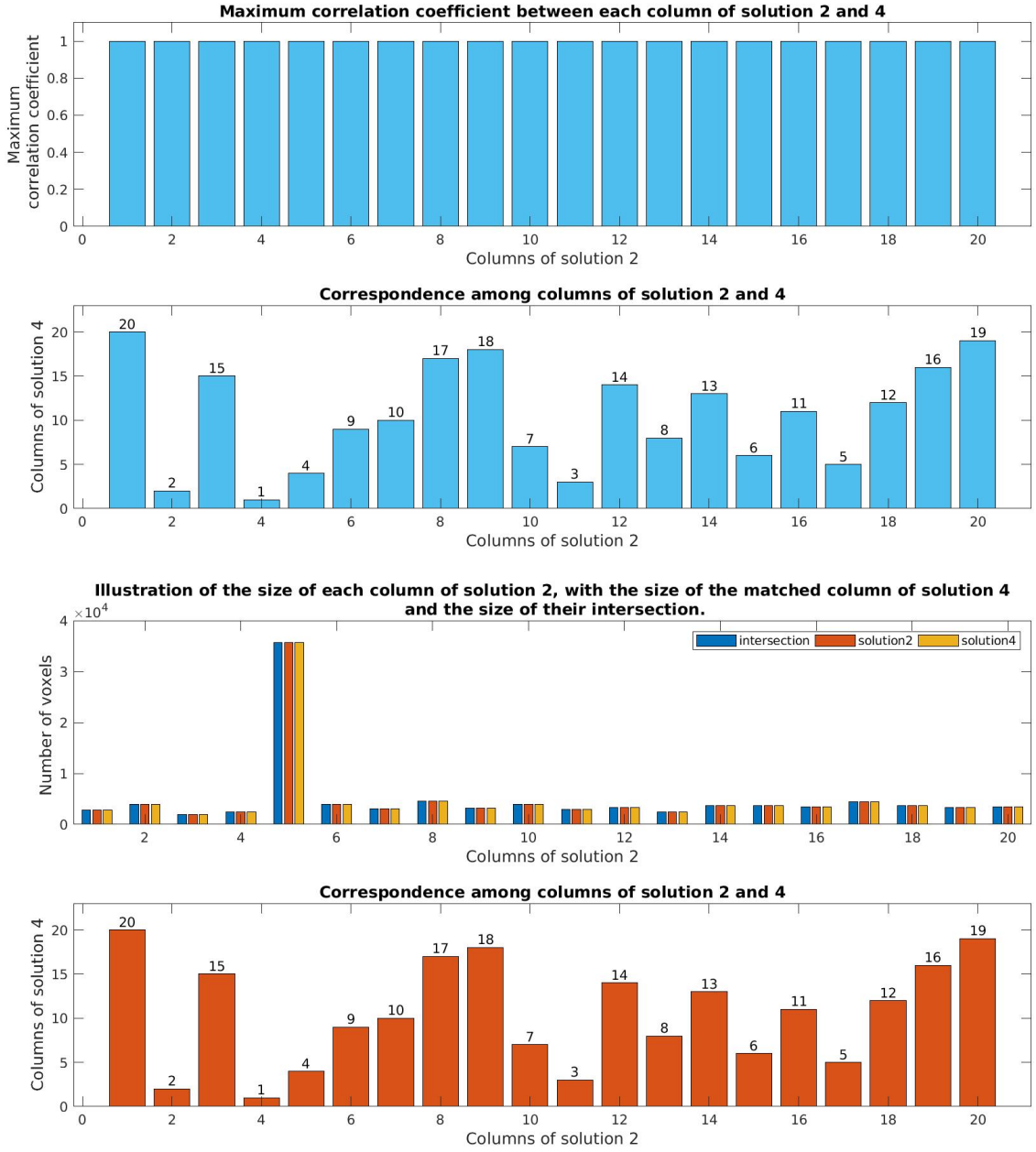


Figure A.210: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on intersection. Solutions 2, 4 correspond to accelerated penalty method and noiseless data.

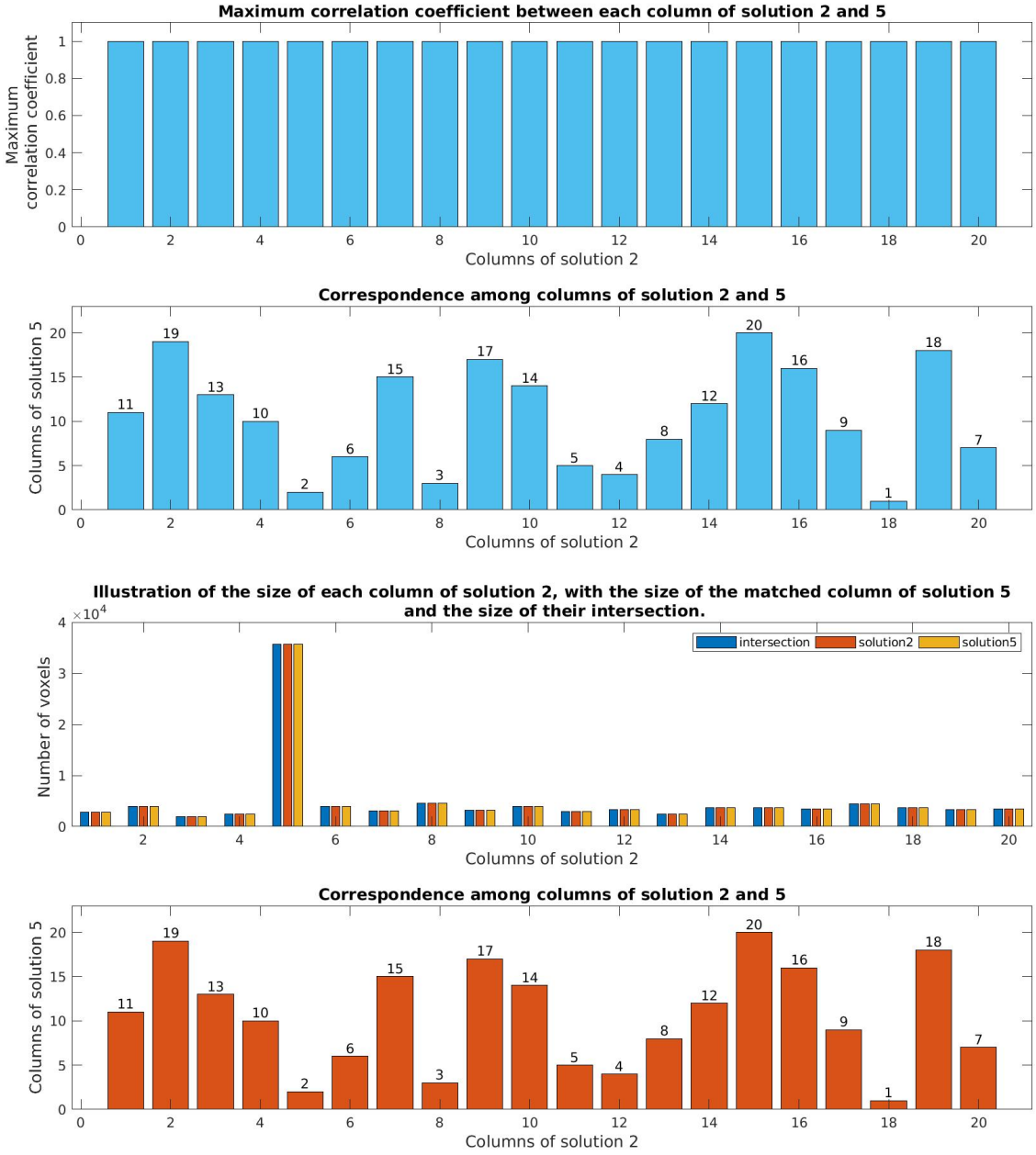


Figure A.211: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 5. We associate the column r of the solution 2 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on intersection. Solutions 2, 5 correspond to accelerated penalty method and noiseless data.

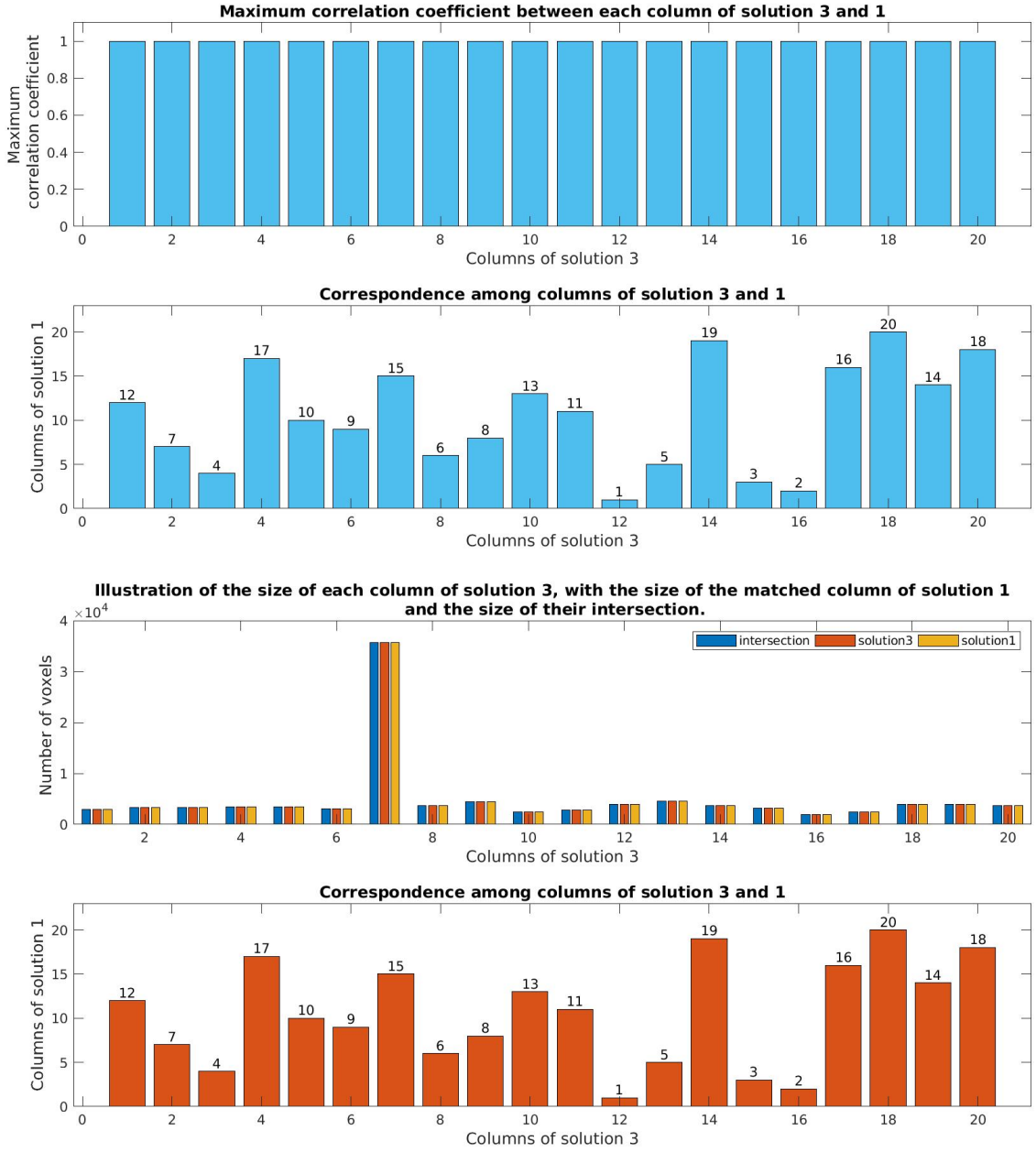


Figure A.212: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on intersection. Solutions 3, 1 correspond to accelerated penalty method and noiseless data.

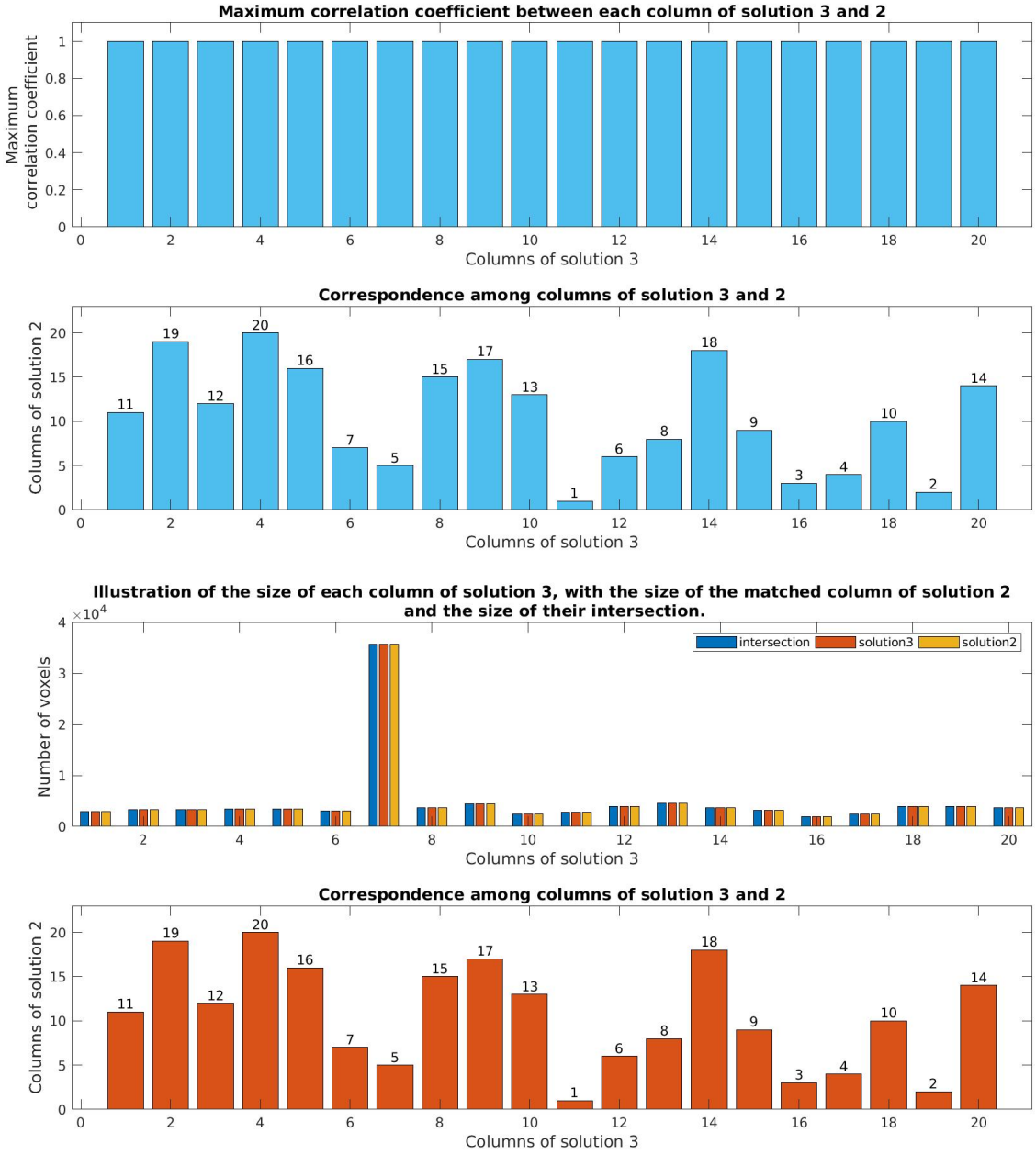


Figure A.213: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on intersection. Solutions 3, 2 correspond to accelerated penalty method and noiseless data.

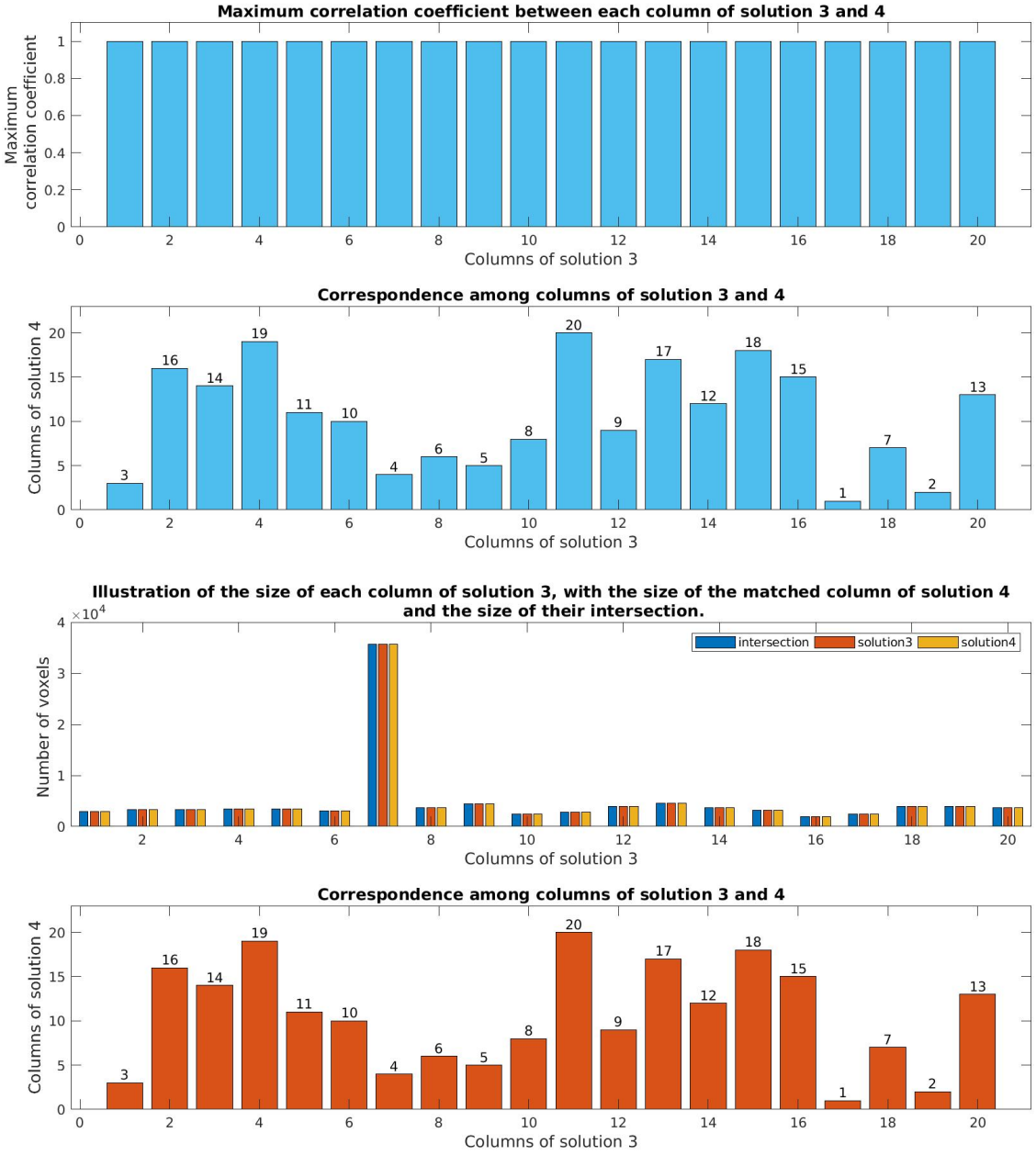


Figure A.214: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on intersection. Solutions 3, 4 correspond to accelerated penalty method and noiseless data.

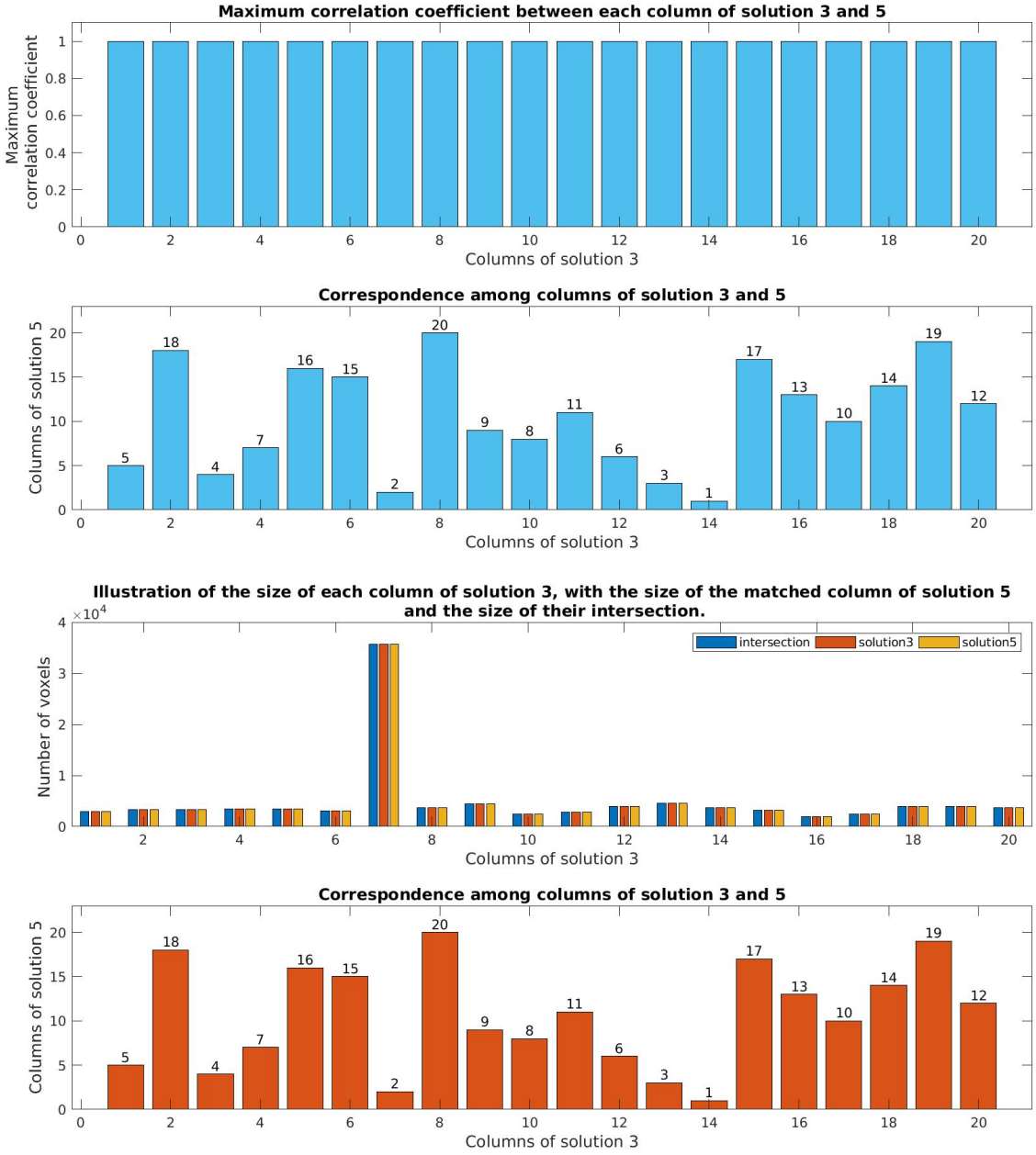


Figure A.215: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 5. We associate the column r of the solution 3 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on intersection. Solutions 3, 5 correspond to accelerated penalty method and noiseless data.

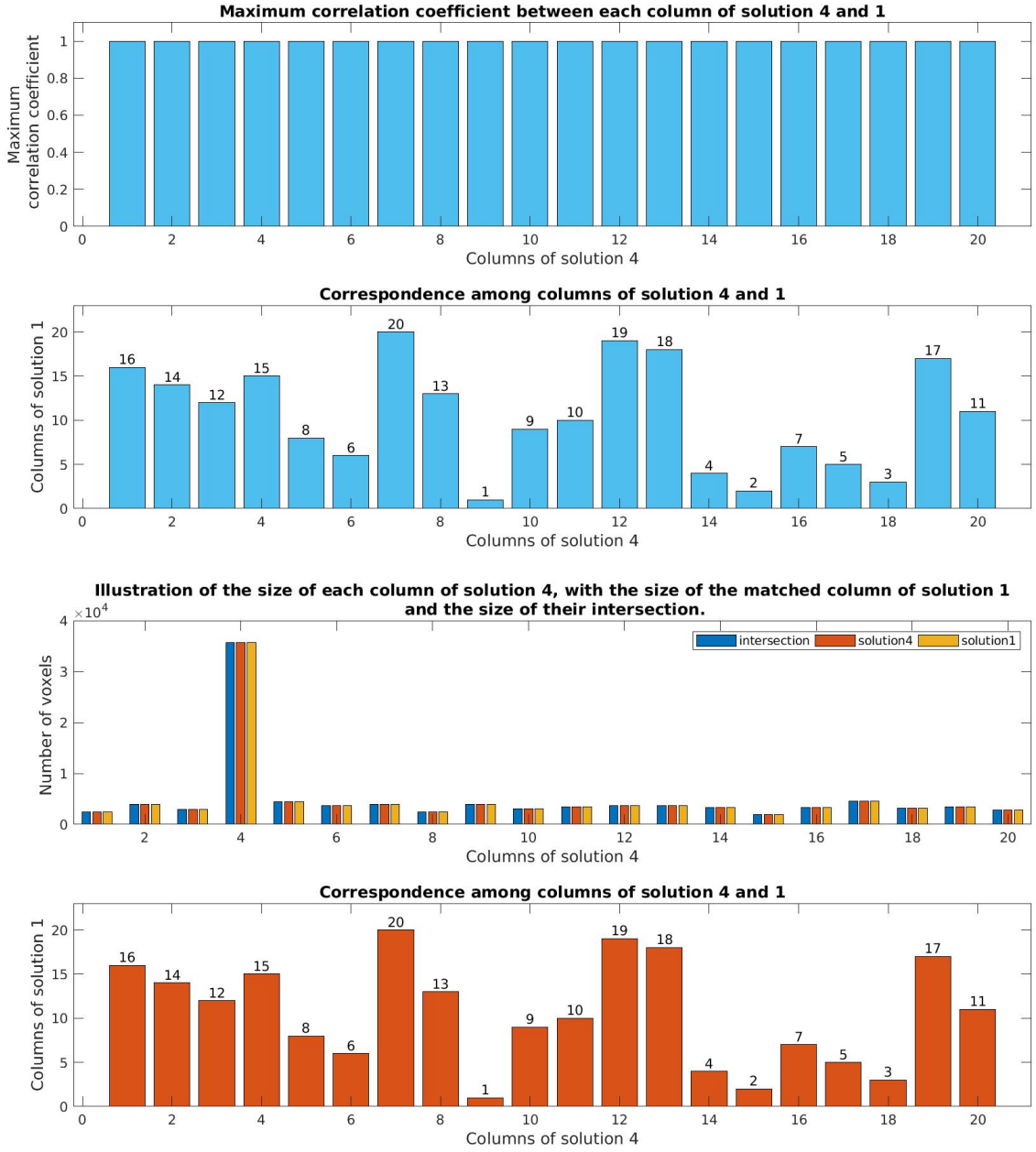


Figure A.216: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on intersection. Solutions 4, 1 correspond to accelerated penalty method and noiseless data.

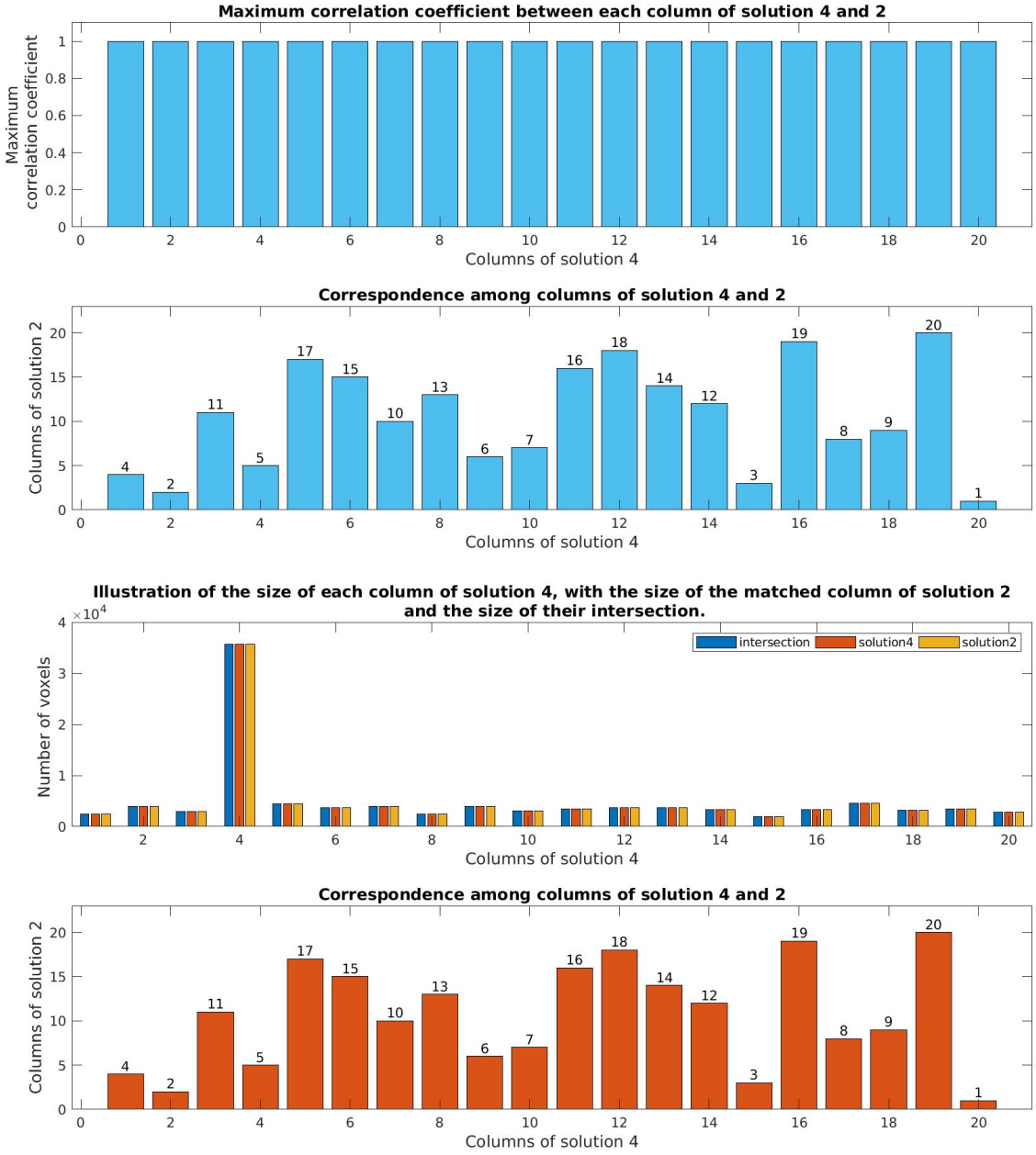


Figure A.217: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on intersection. Solutions 4, 2 correspond to accelerated penalty method and noiseless data.

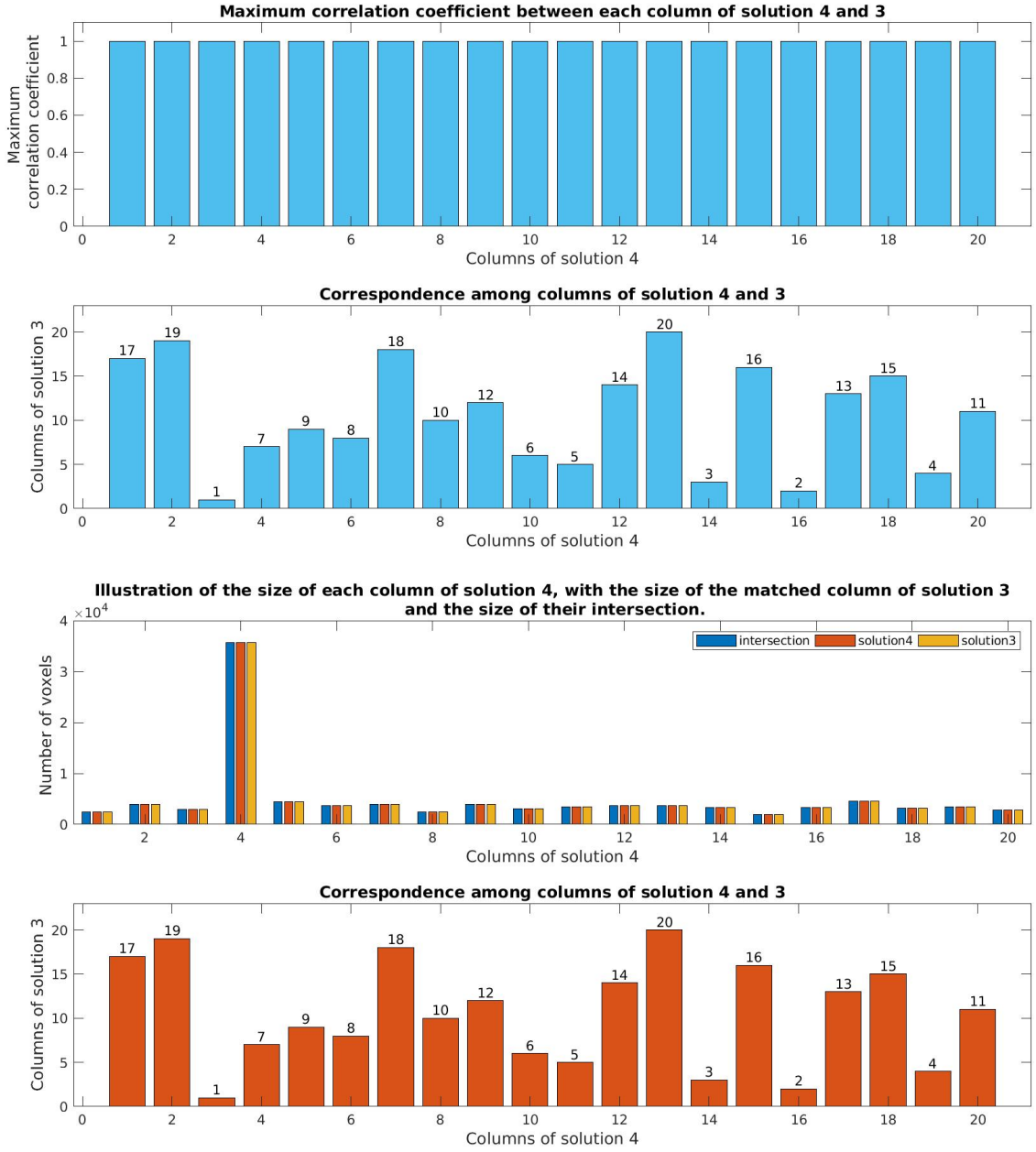


Figure A.218: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on intersection. Solutions 4, 3 correspond to accelerated penalty method and noiseless data.

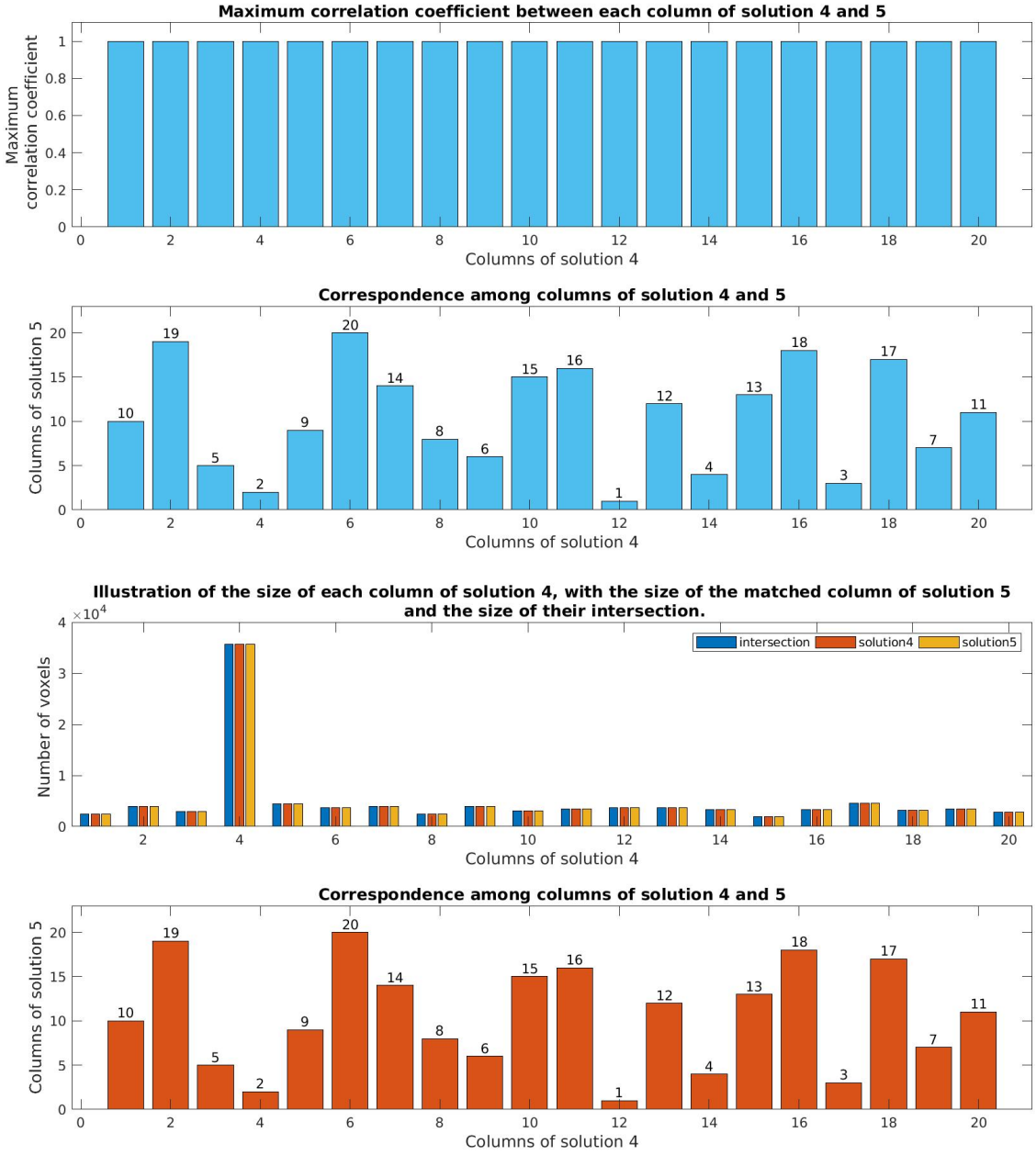


Figure A.219: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 5. We associate the column r of the solution 4 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on intersection. Solutions 4, 5 correspond to accelerated penalty method and noiseless data.

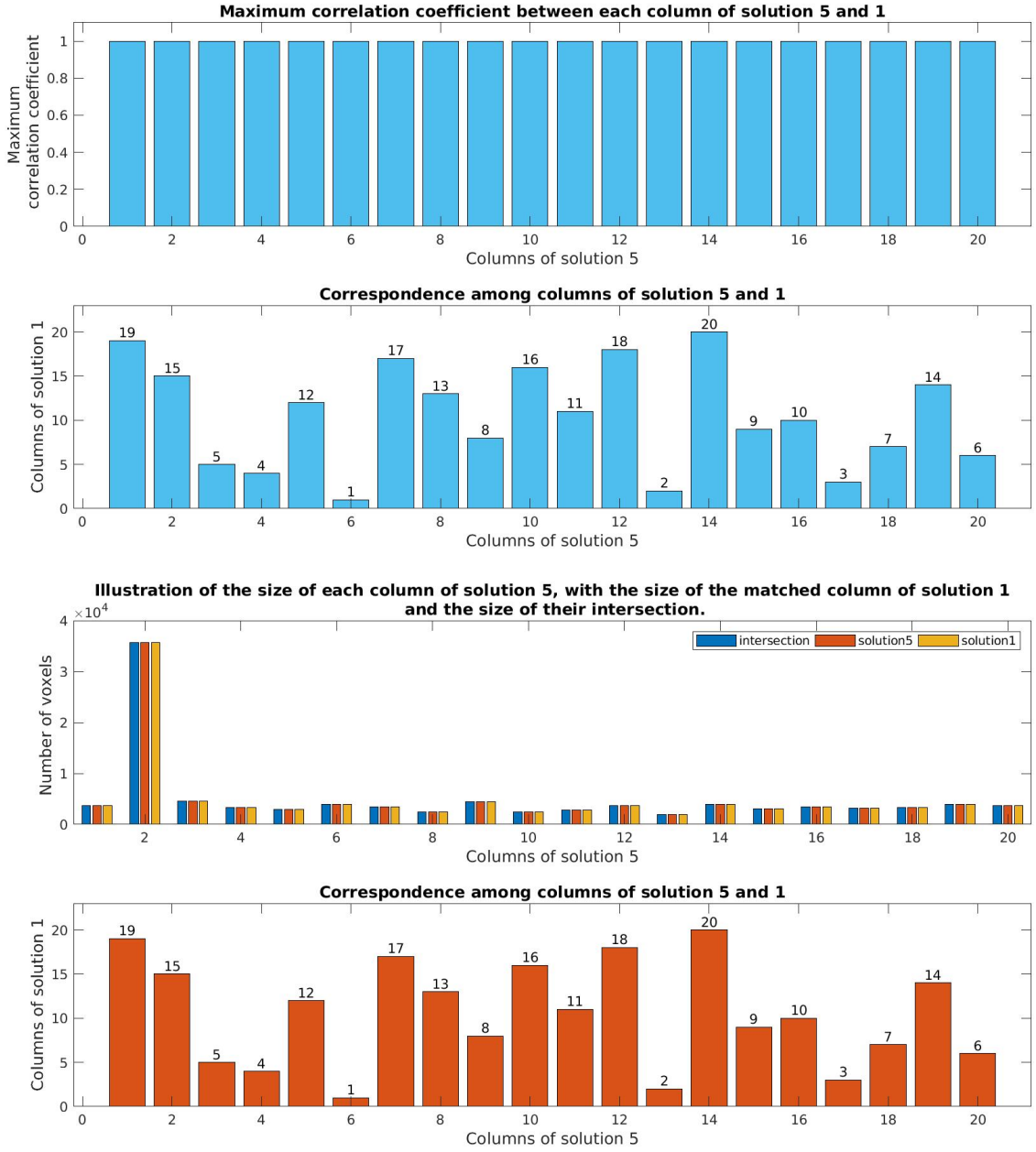


Figure A.220: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 1. We associate the column r of the solution 5 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on intersection. Solutions 5, 1 correspond to accelerated penalty method and noiseless data.

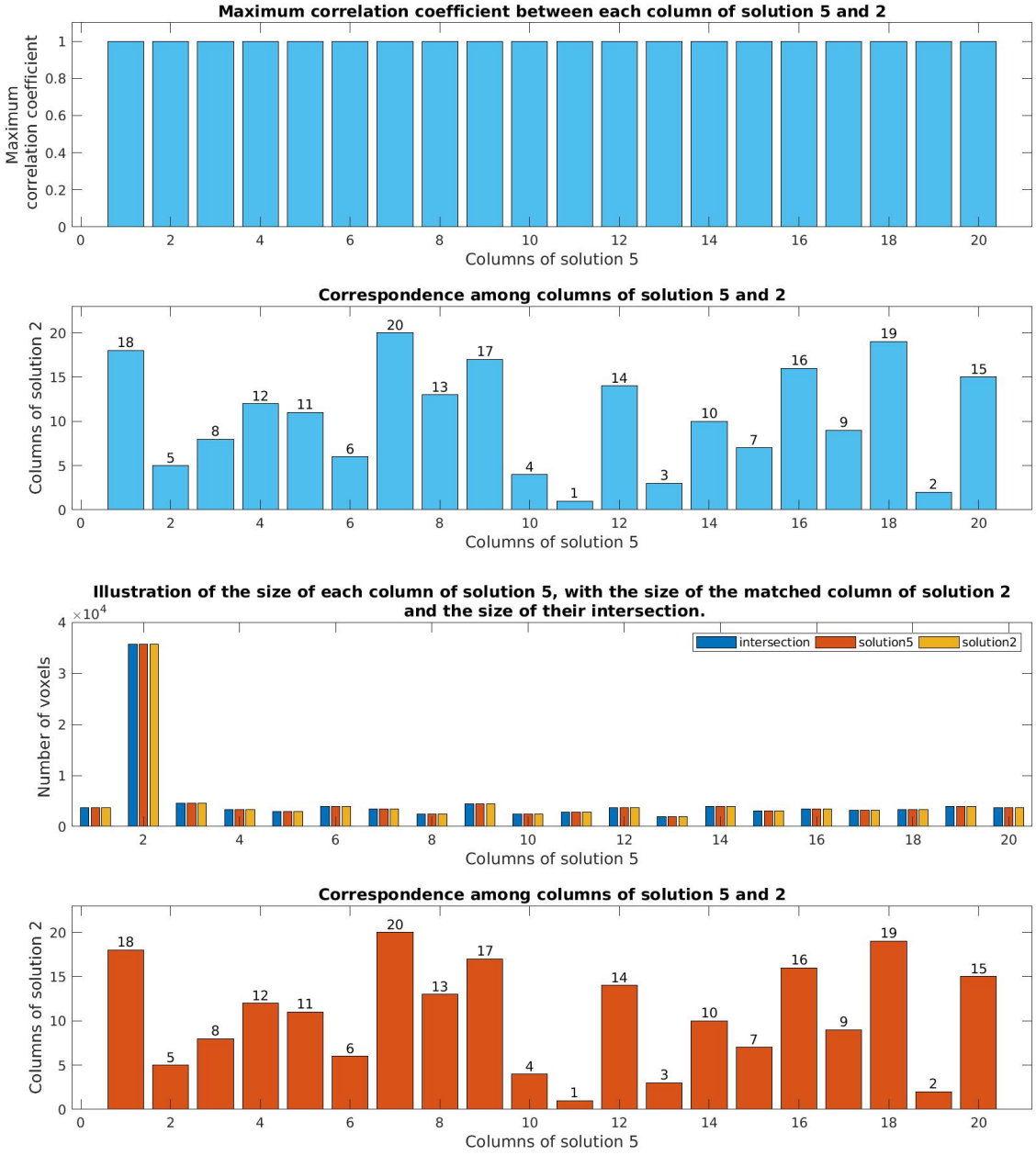


Figure A.221: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 2. We associate the column r of the solution 5 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on intersection. Solutions 5, 2 correspond to accelerated penalty method and noiseless data.

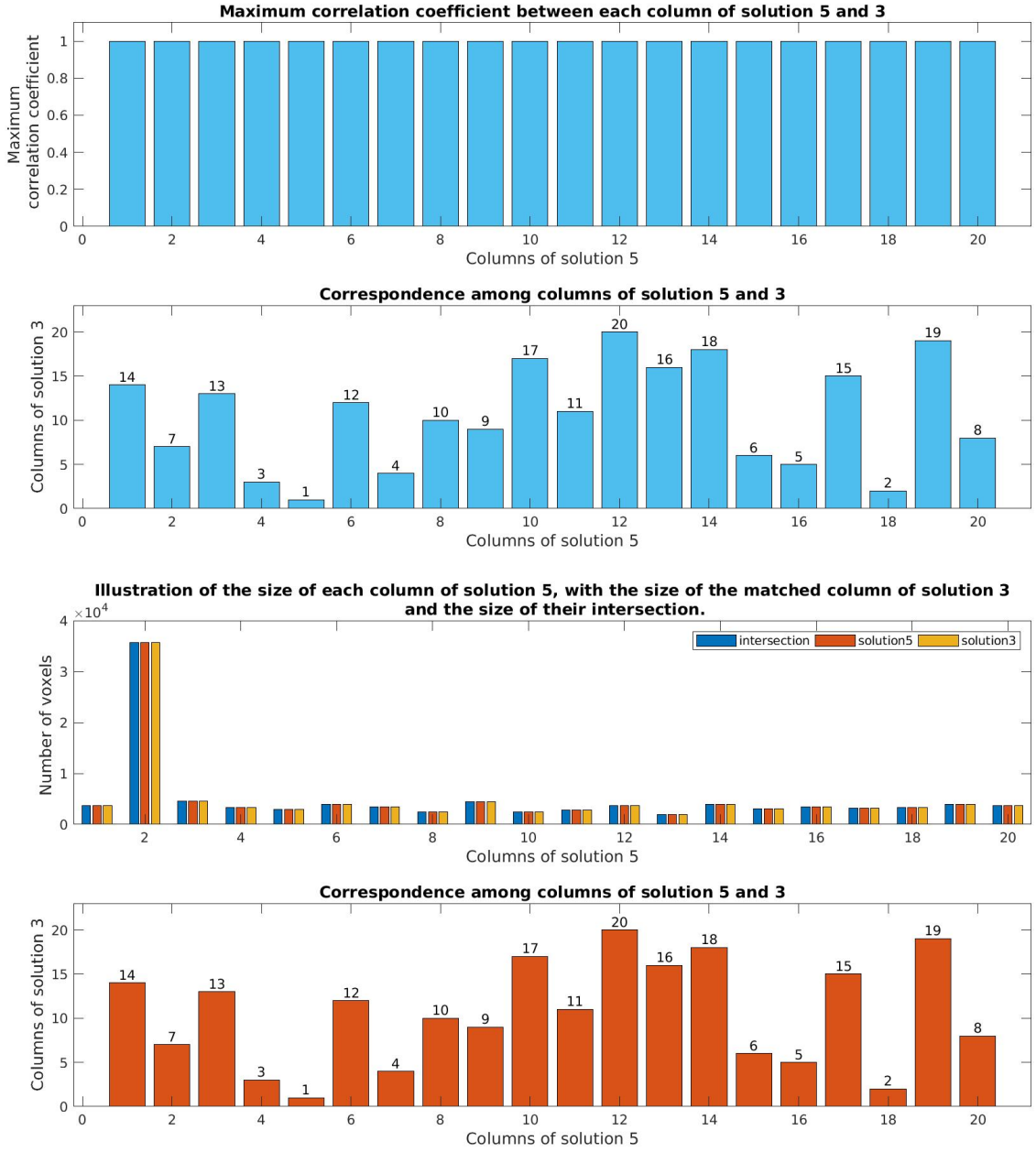


Figure A.222: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 3. We associate the column r of the solution 5 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on intersection. Solutions 5, 3 correspond to accelerated penalty method and noiseless data.

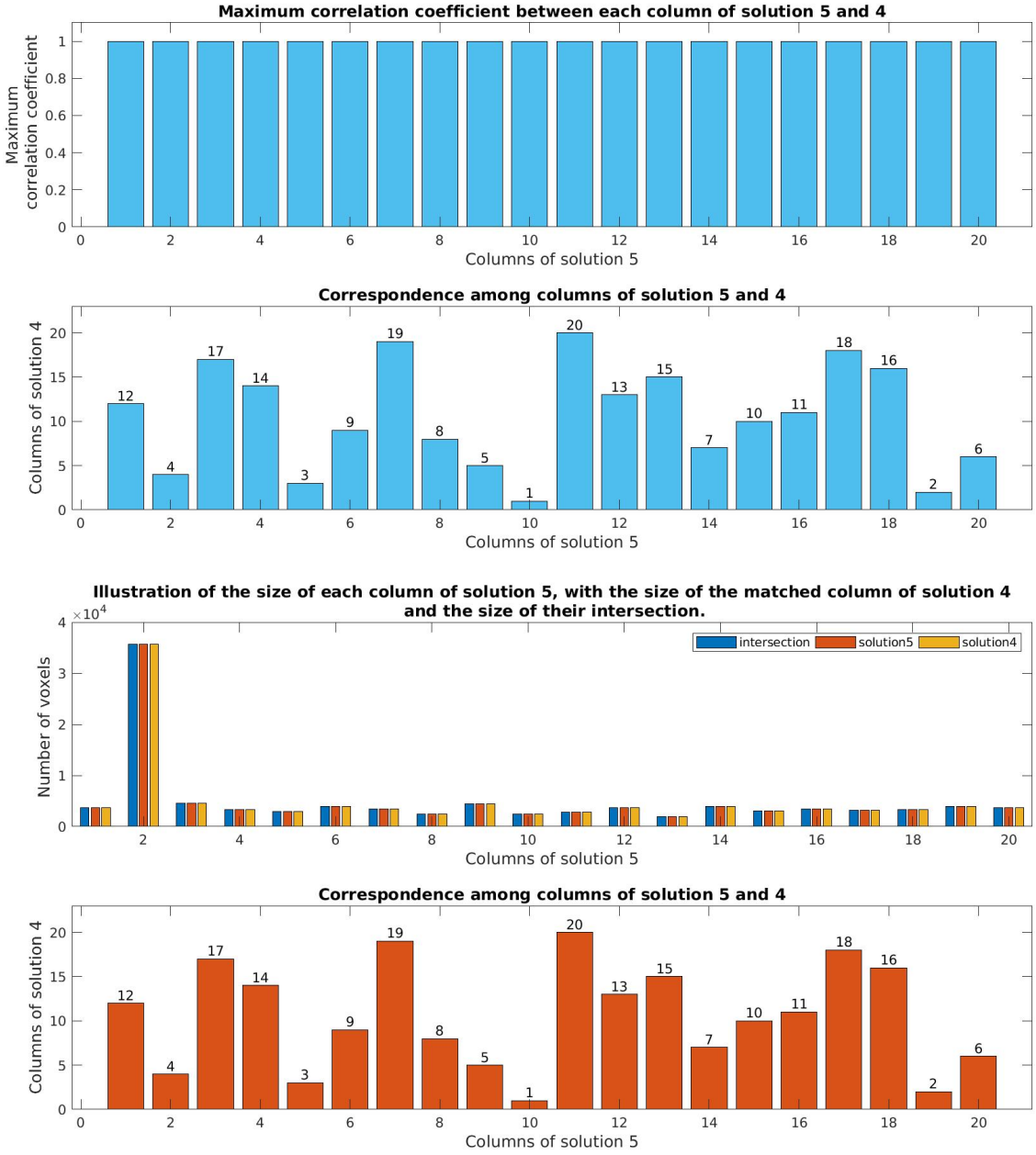


Figure A.223: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 4. We associate the column r of the solution 5 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on intersection. Solutions 5, 4 correspond to accelerated penalty method and noiseless data.

· **SNR= 0 dB**

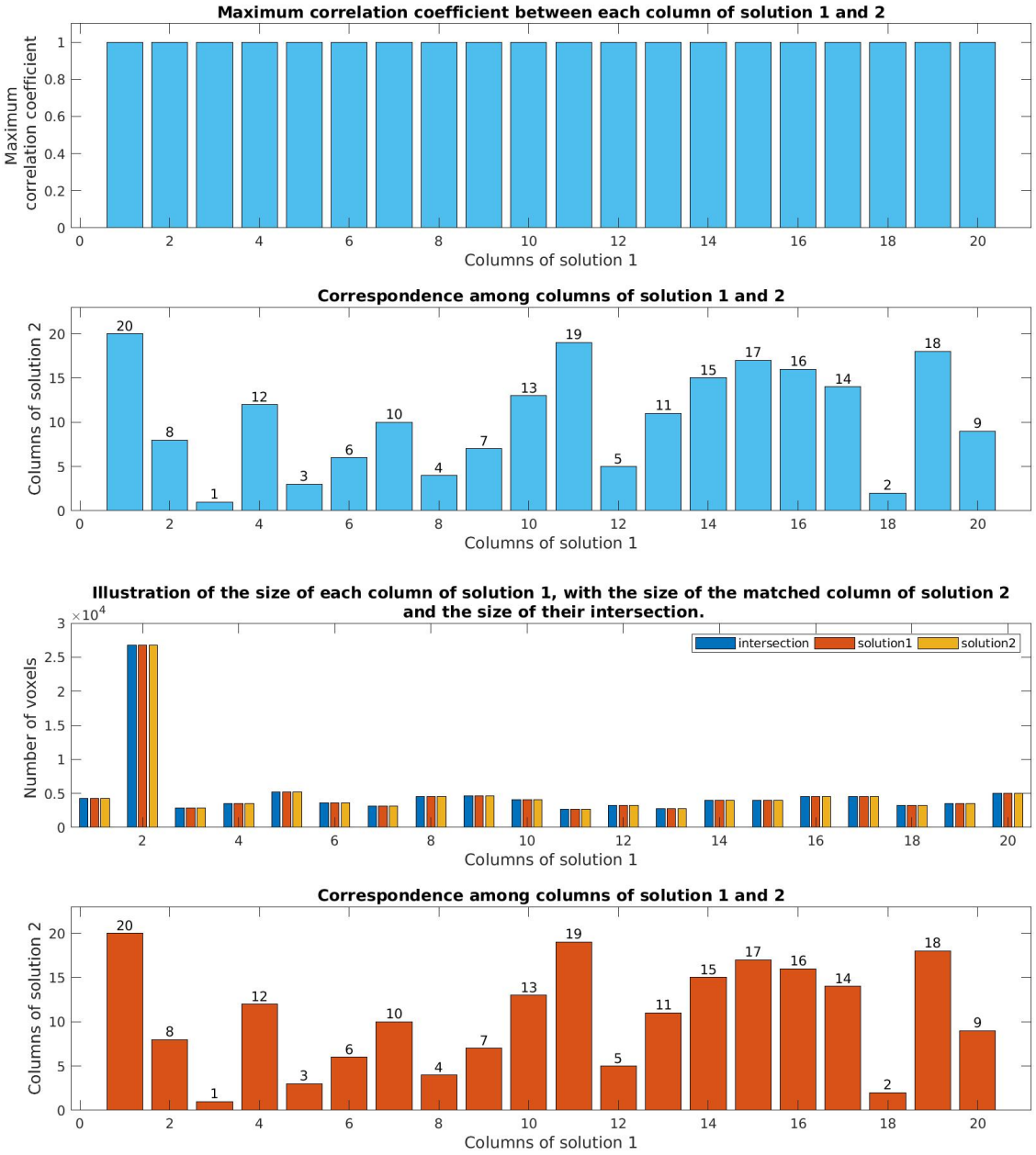


Figure A.224: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated, based on intersection. Solutions 1, 2 correspond to accelerated penalty method and SNR=0 dB.

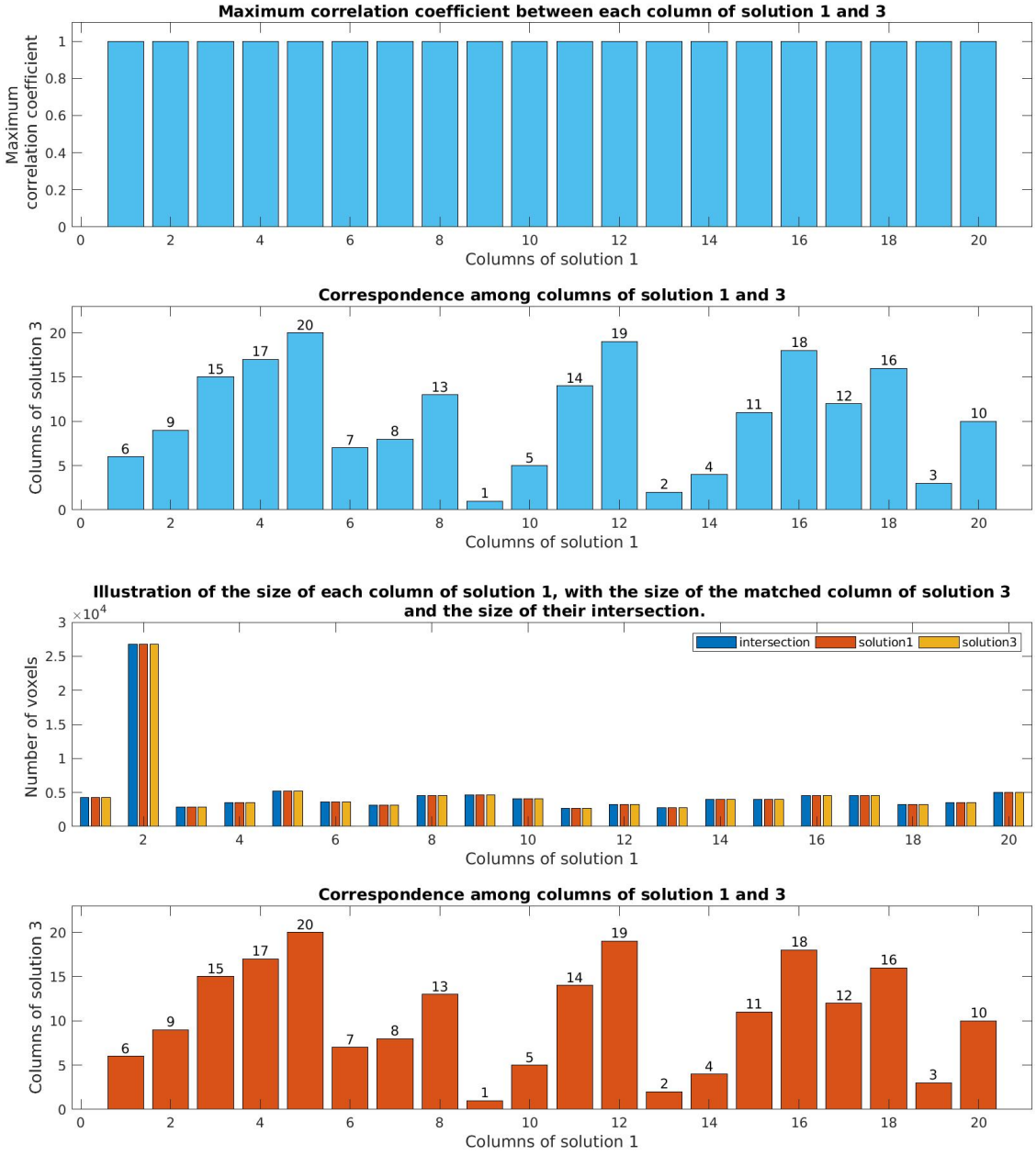


Figure A.225: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated, based on intersection. Solutions 1, 3 correspond to accelerated penalty method and SNR=0 dB.

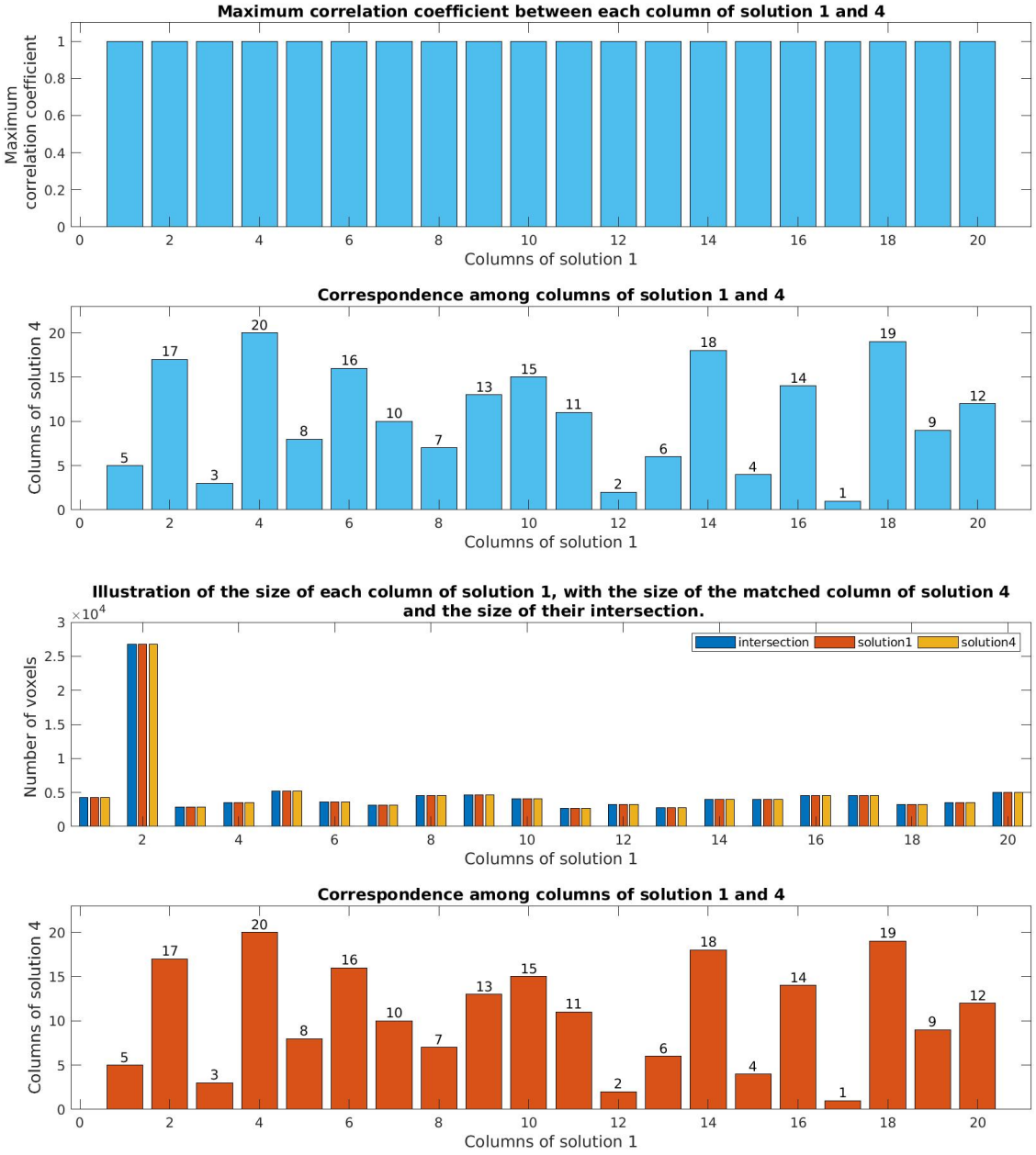


Figure A.226: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated, based on intersection. Solutions 1, 4 correspond to accelerated penalty method and SNR=0 dB.

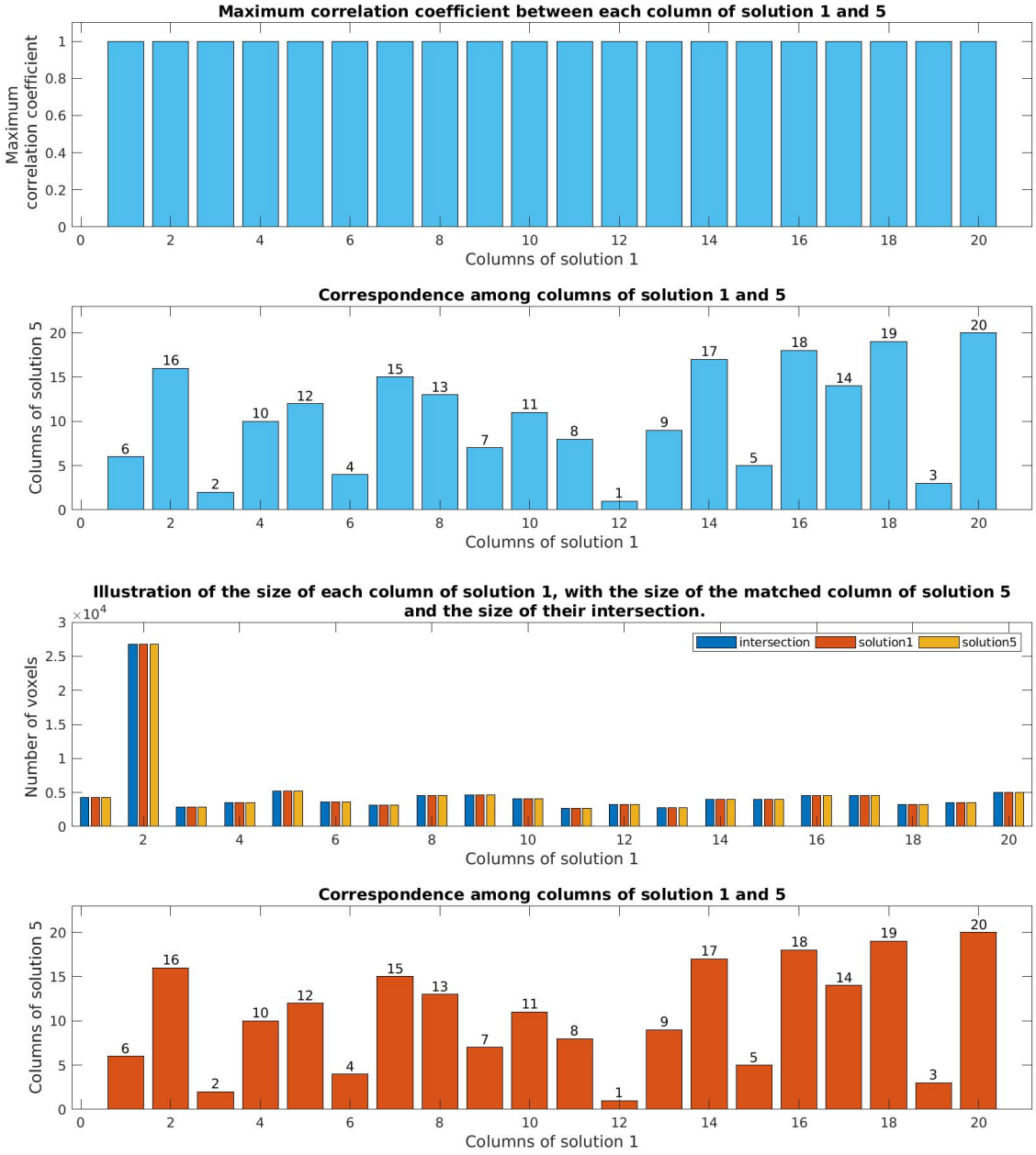


Figure A.227: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 5. We associate the column r of the solution 1 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 1 is associated, based on intersection. Solutions 1, 5 correspond to accelerated penalty method and SNR=0 dB.



Figure A.228: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated, based on intersection. Solutions 2, 1 correspond to accelerated penalty method and SNR=0 dB.

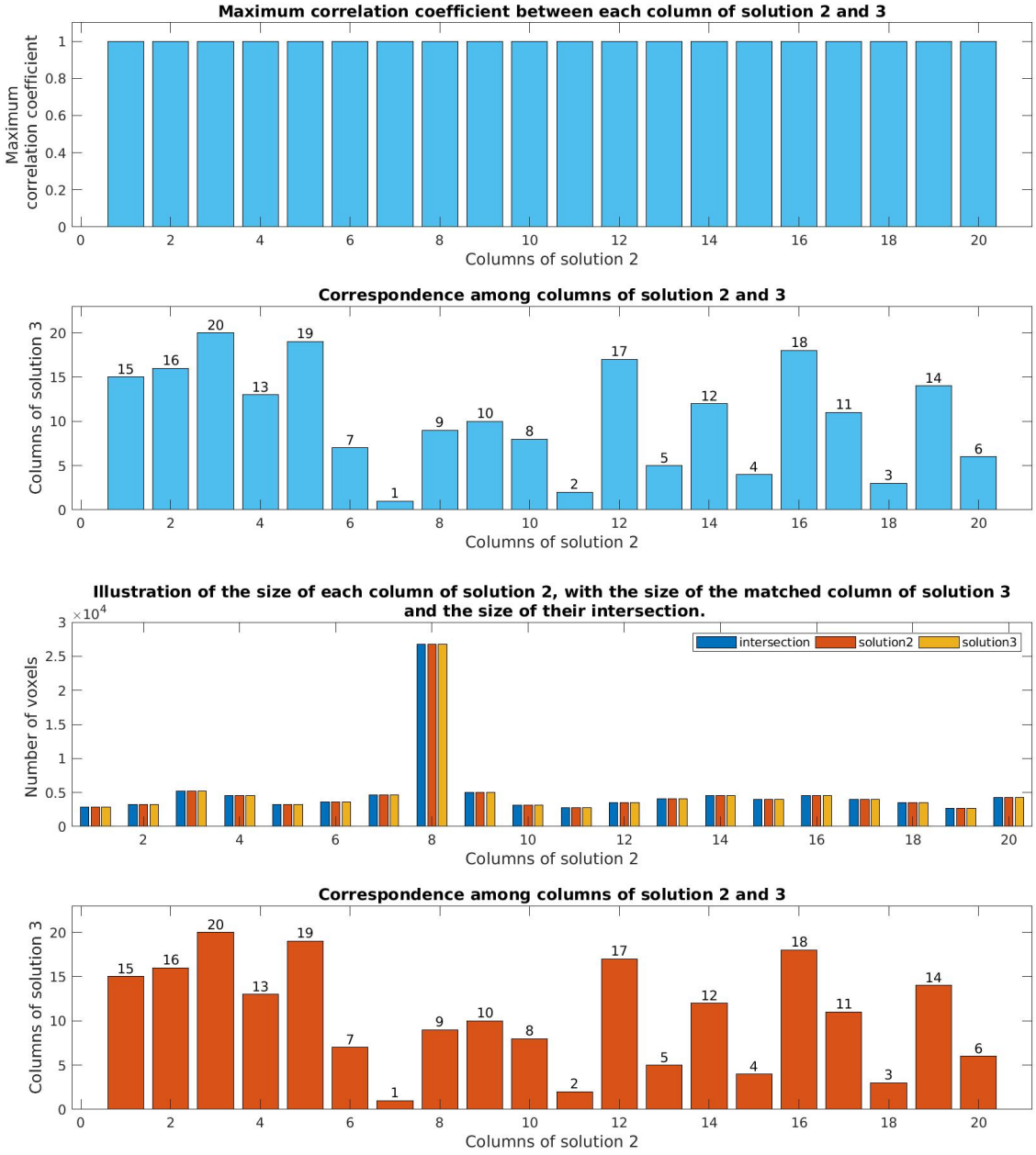


Figure A.229: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated, based on intersection. Solutions 2, 3 correspond to accelerated penalty method and SNR=0 dB.

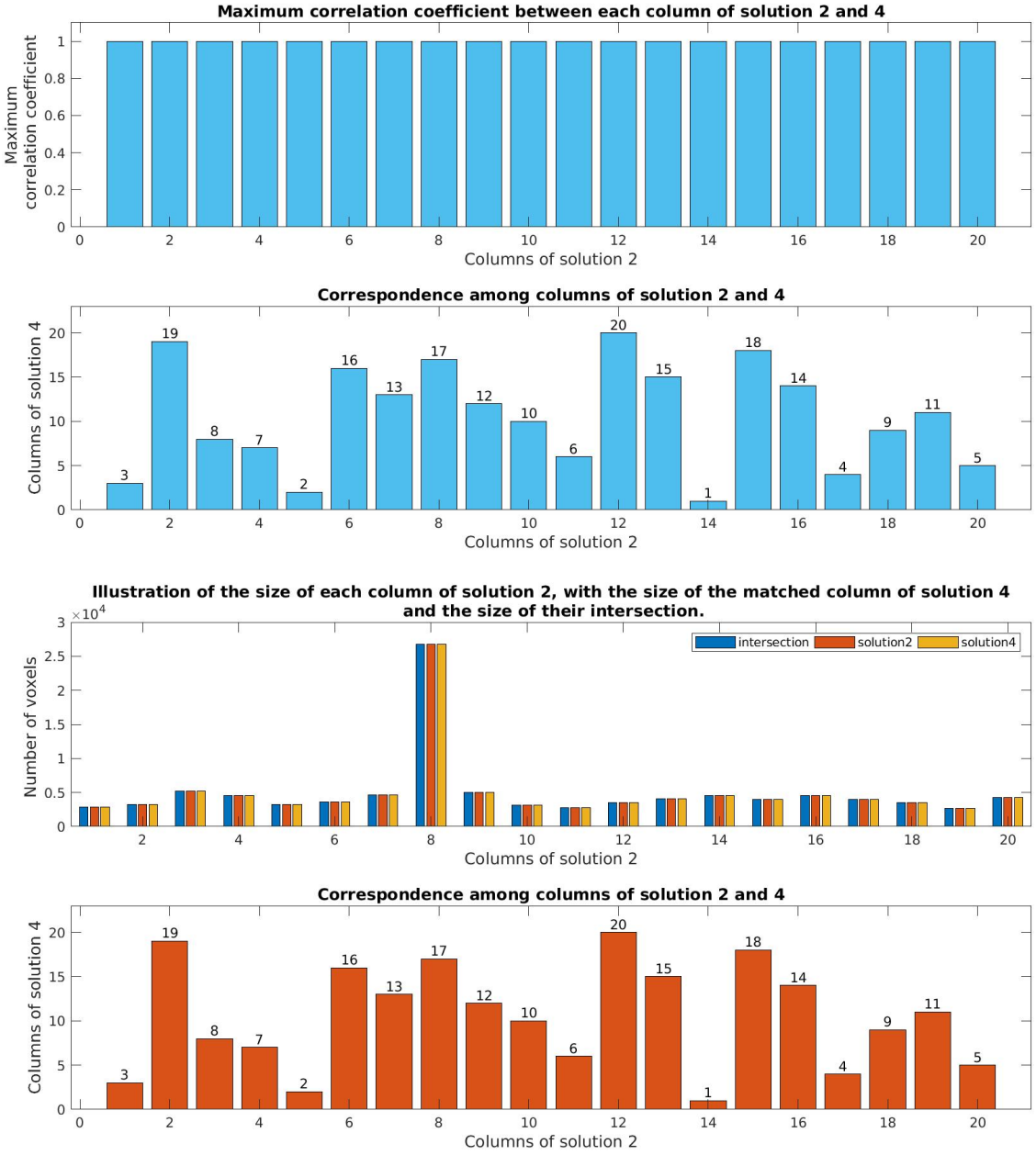


Figure A.230: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated, based on intersection. Solutions 2, 4 correspond to accelerated penalty method and SNR=0 dB.

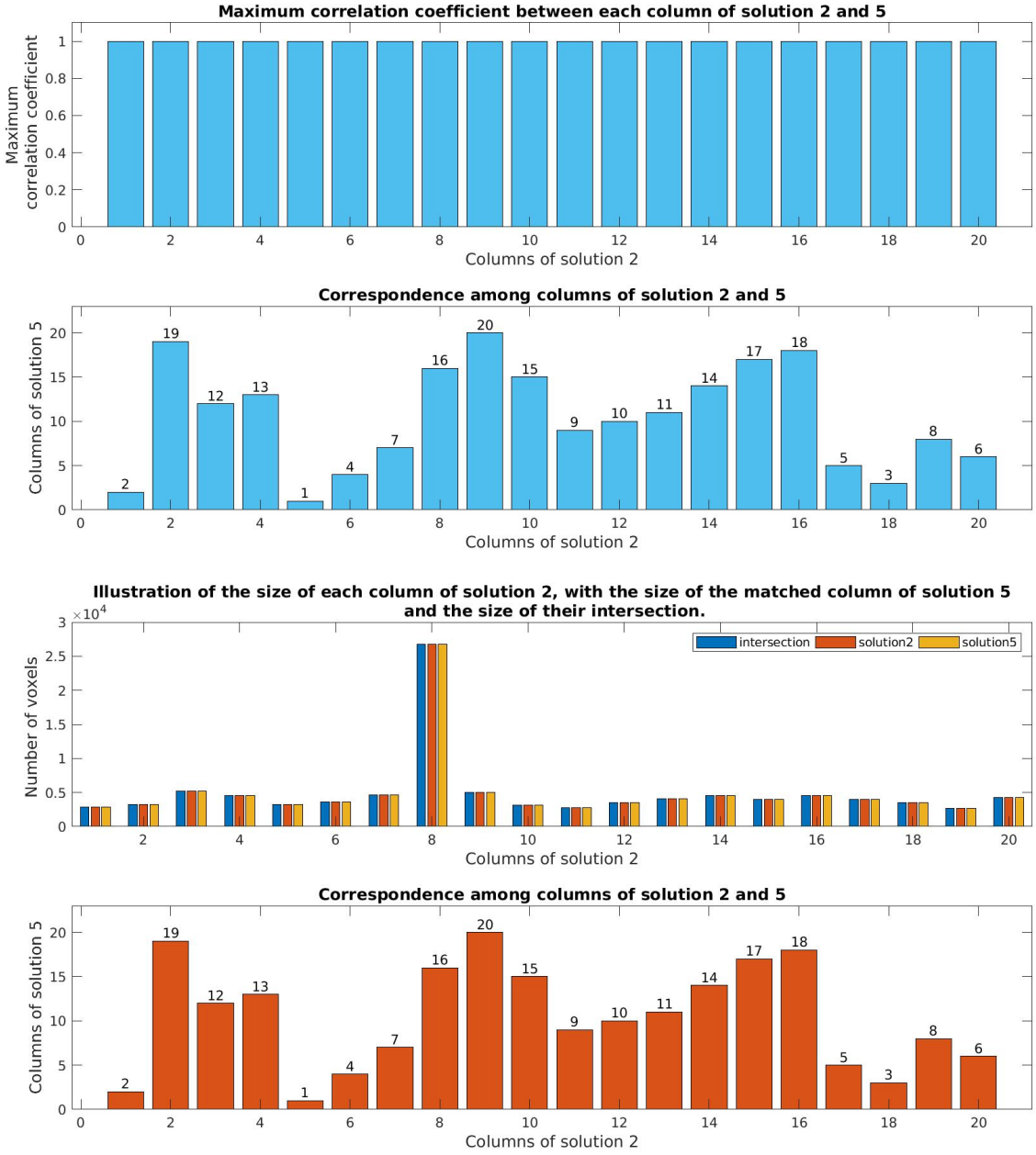


Figure A.231: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 5. We associate the column r of the solution 2 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 2 is associated, based on intersection. Solutions 2, 5 correspond to accelerated penalty method and SNR=0 dB.

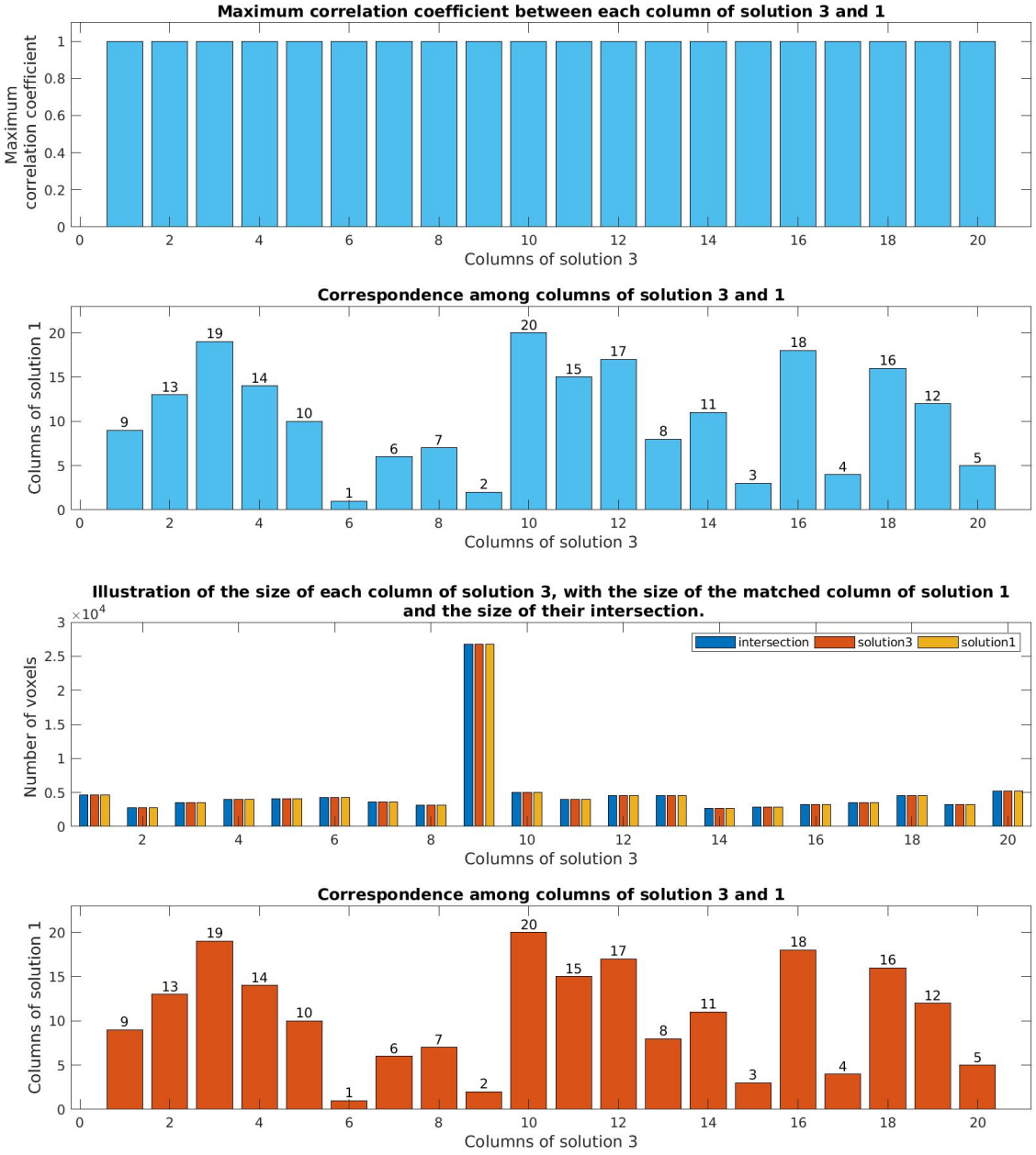


Figure A.232: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated, based on intersection. Solutions 3, 1 correspond to accelerated penalty method and SNR=0 dB.

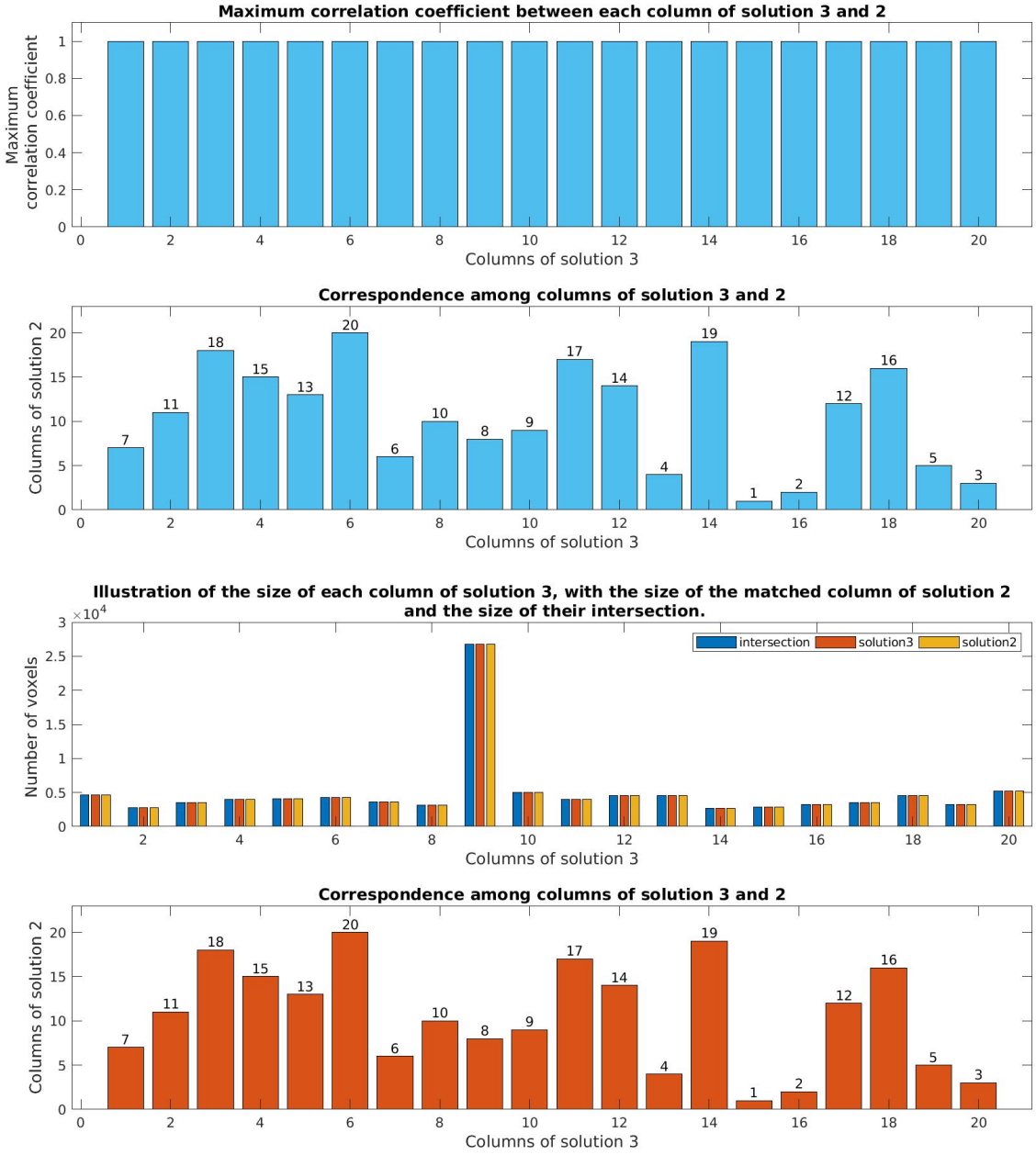


Figure A.233: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated, based on intersection. Solutions 3, 2 correspond to accelerated penalty method and SNR=0 dB.



Figure A.234: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated, based on intersection. Solutions 3, 4 correspond to accelerated penalty method and SNR=0 dB.

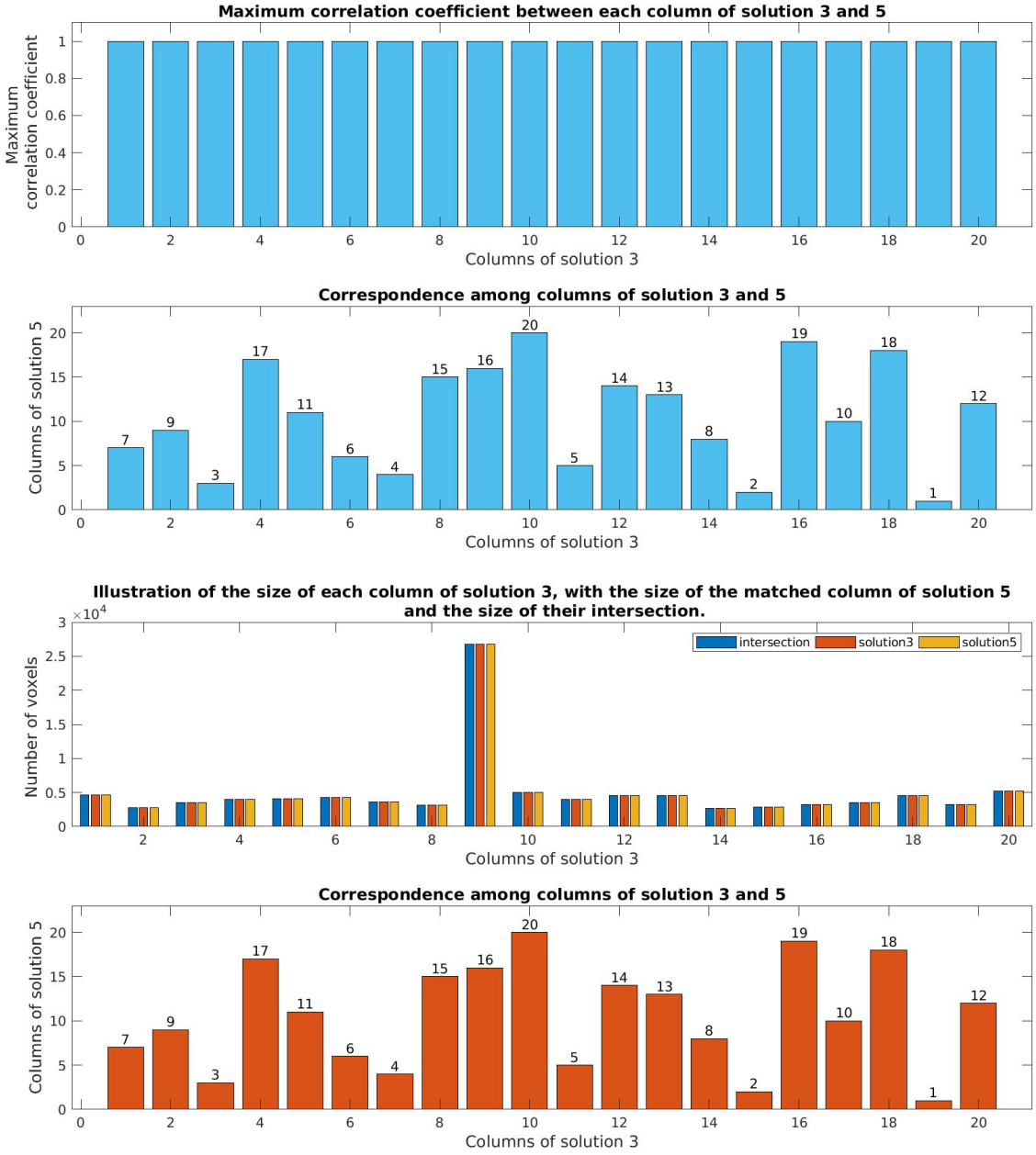


Figure A.235: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 5. We associate the column r of the solution 3 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 3 is associated, based on intersection. Solutions 3, 5 correspond to accelerated penalty method and SNR=0 dB.

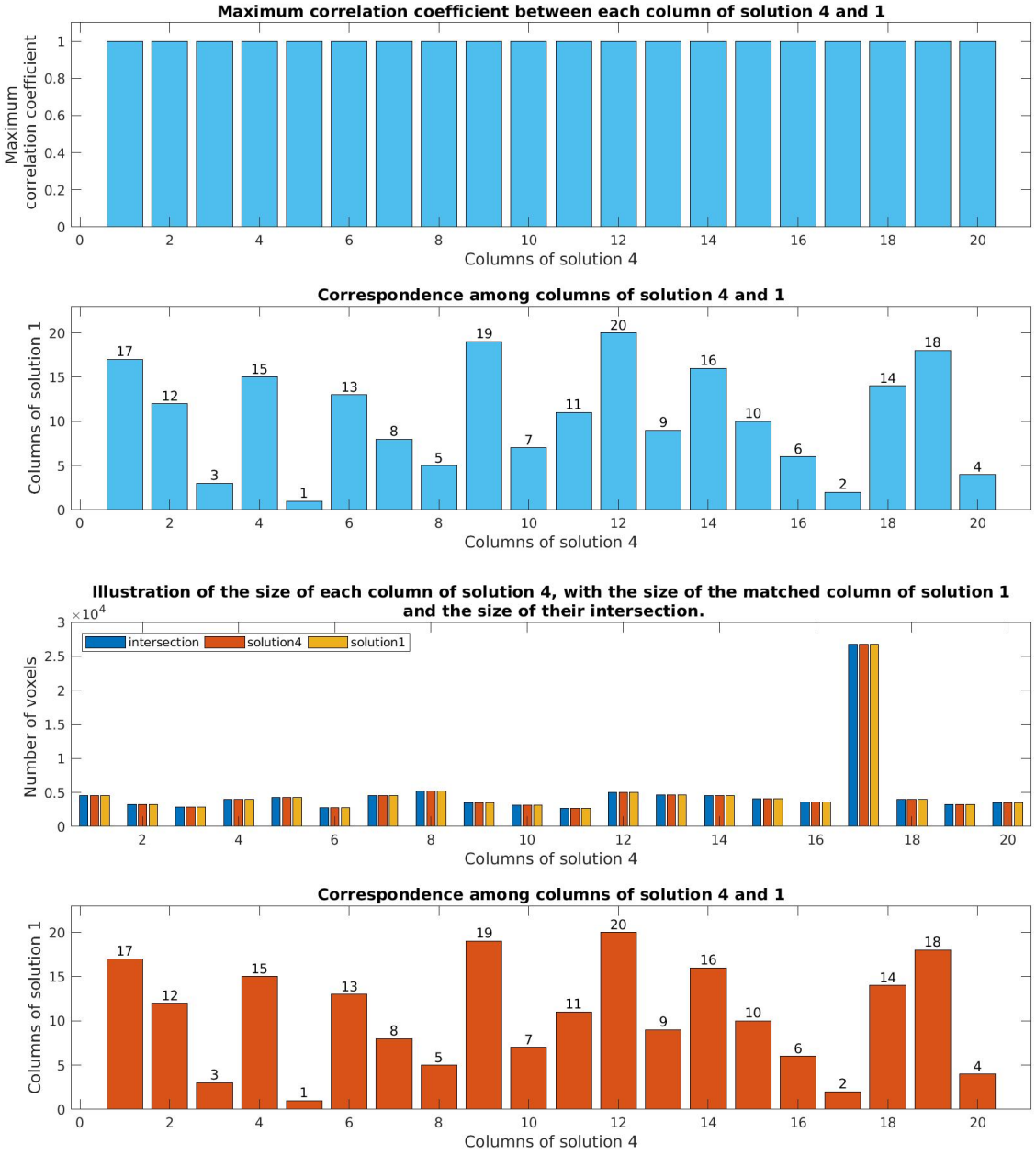


Figure A.236: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated, based on intersection. Solutions 4, 1 correspond to accelerated penalty method and SNR=0 dB.

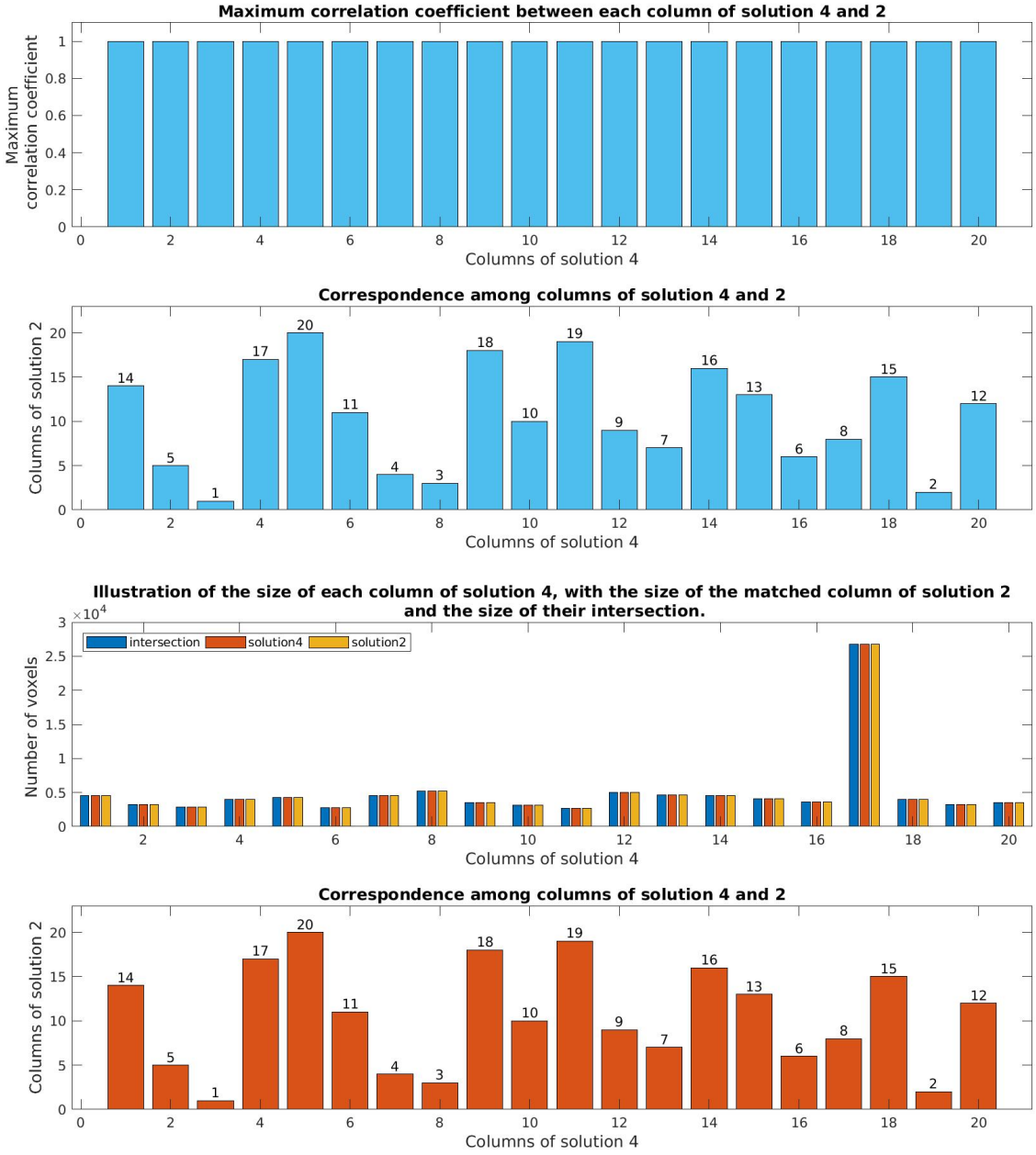


Figure A.237: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated, based on intersection. Solutions 4, 2 correspond to accelerated penalty method and SNR=0 dB.

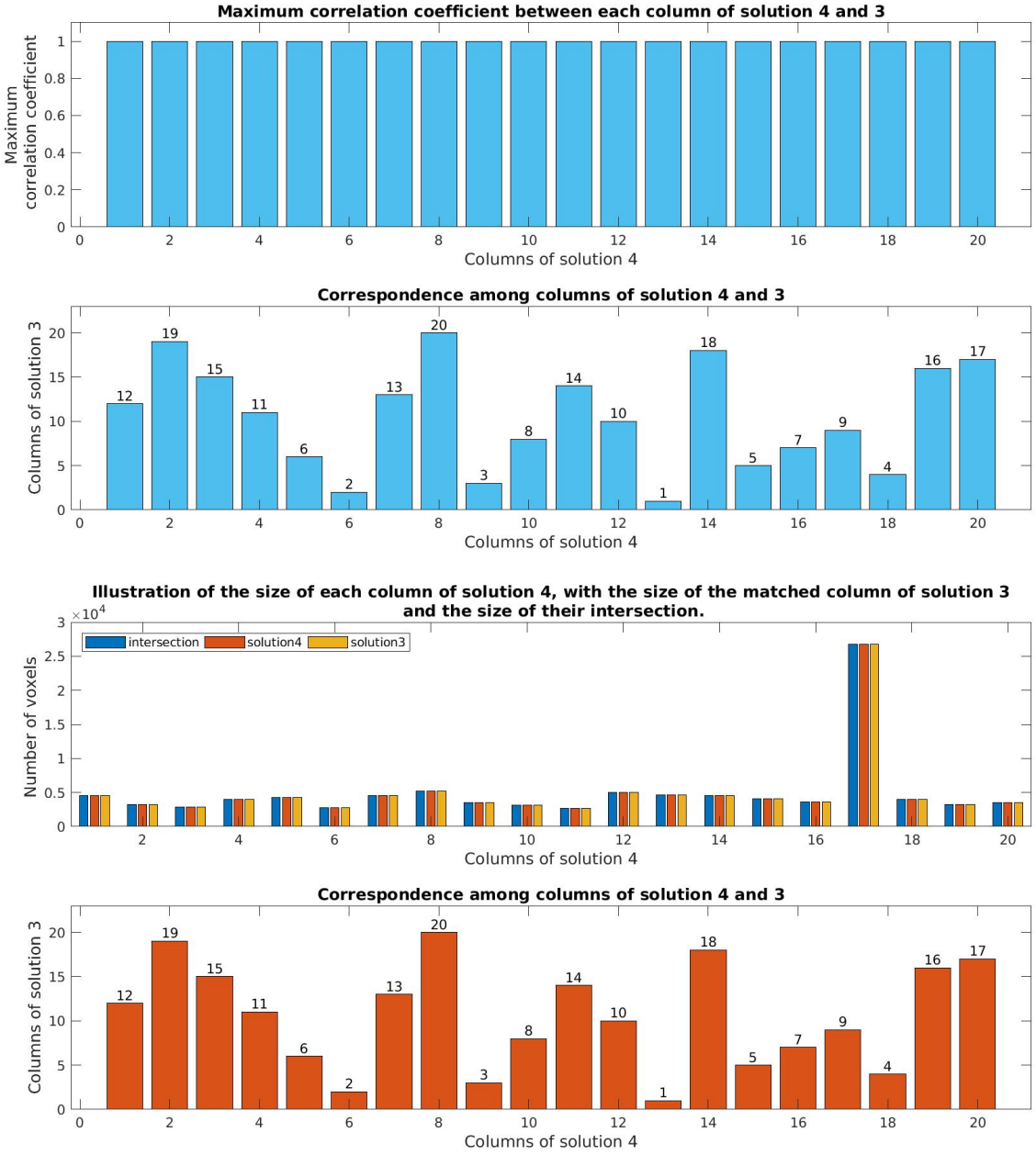


Figure A.238: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated, based on intersection. Solutions 4, 3 correspond to accelerated penalty method and SNR=0 dB.

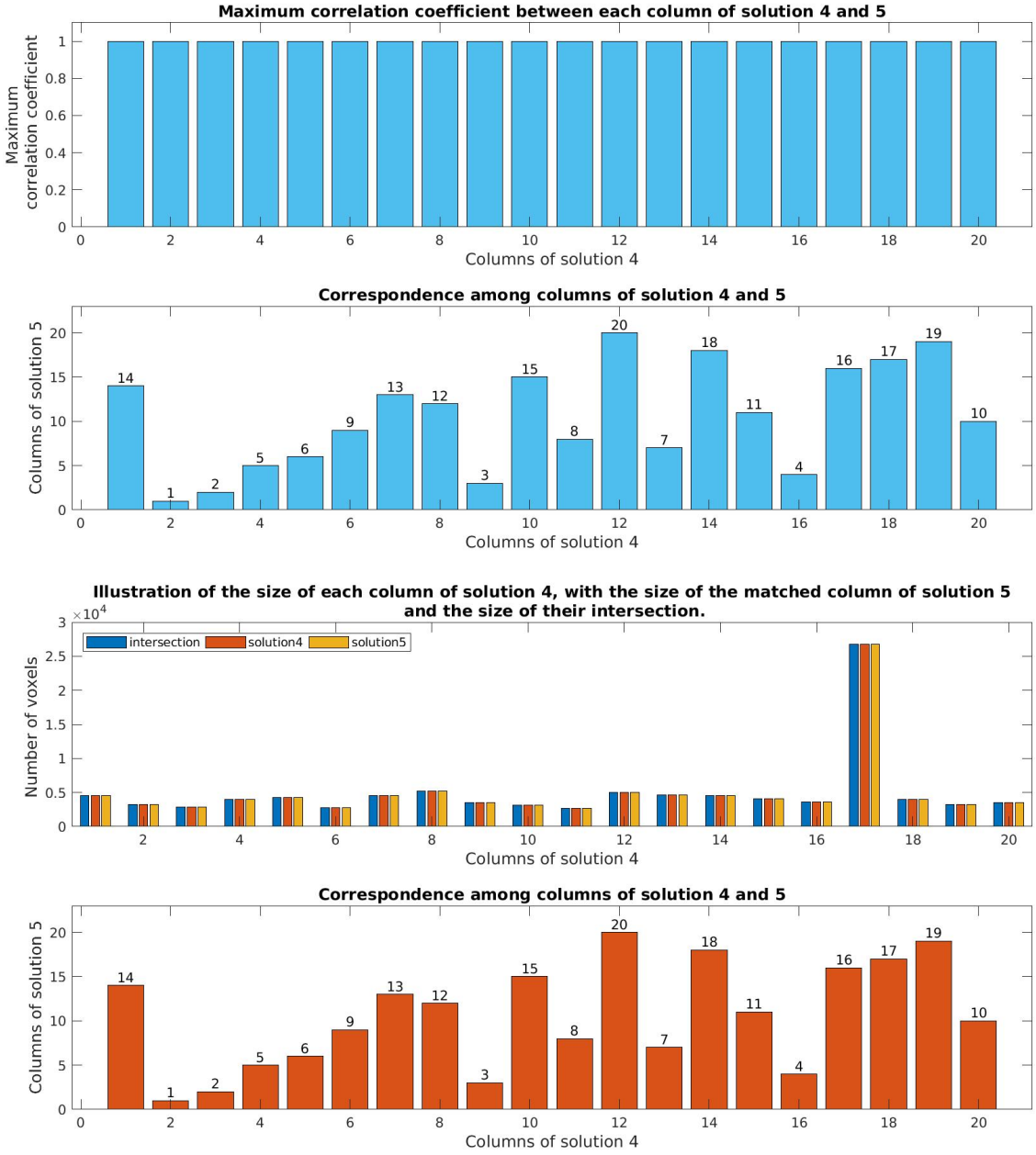


Figure A.239: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 5. We associate the column r of the solution 4 with the column of the solution 5 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 5. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 5 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 5 each column of the solution 4 is associated, based on intersection. Solutions 4, 5 correspond to accelerated penalty method and SNR=0 dB.

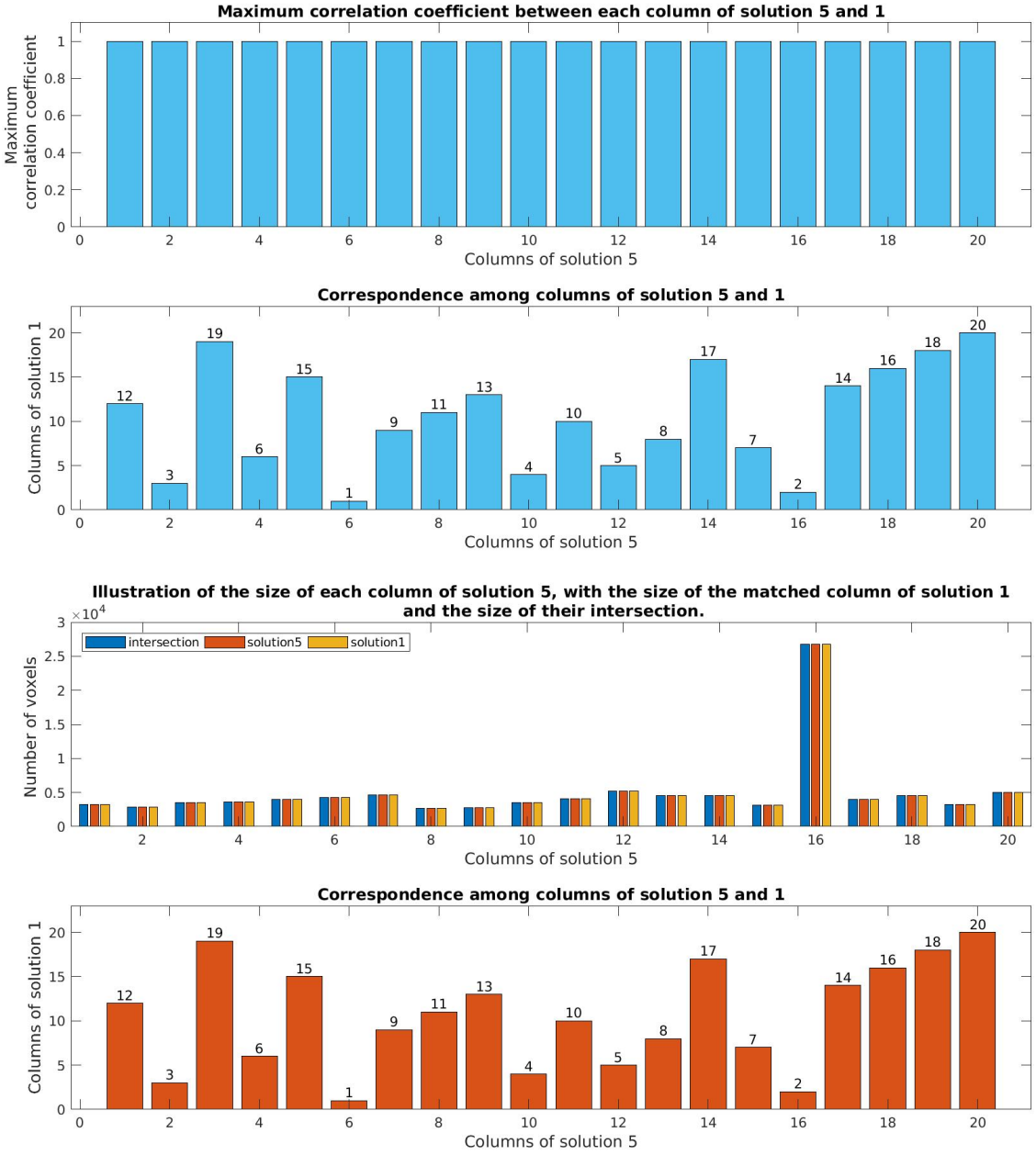


Figure A.240: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 1. We associate the column r of the solution 5 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 5 is associated, based on intersection. Solutions 5, 1 correspond to accelerated penalty method and SNR=0 dB.

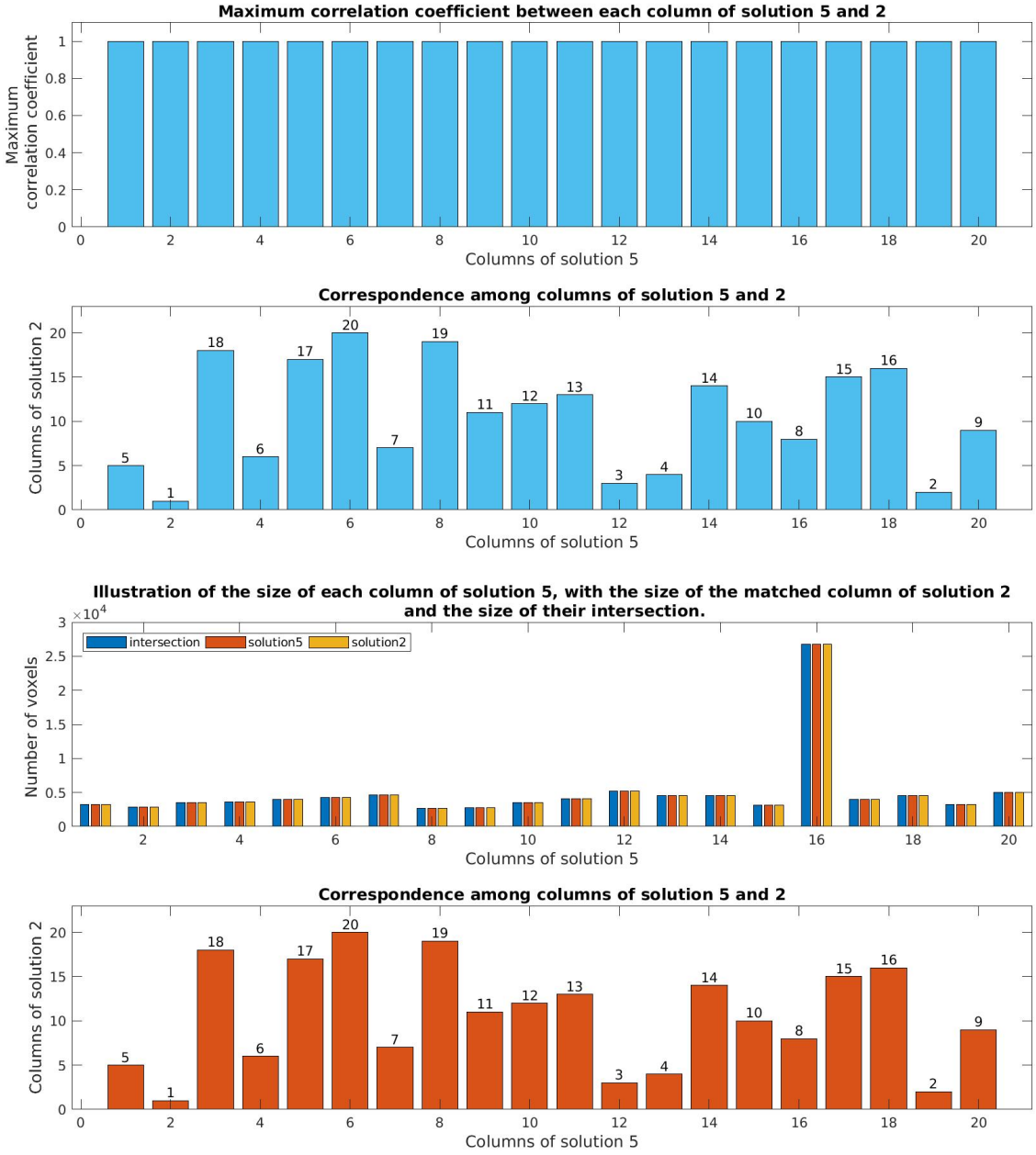


Figure A.241: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 2. We associate the column r of the solution 5 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 5 is associated, based on intersection. Solutions 5, 2 correspond to accelerated penalty method and SNR=0 dB.

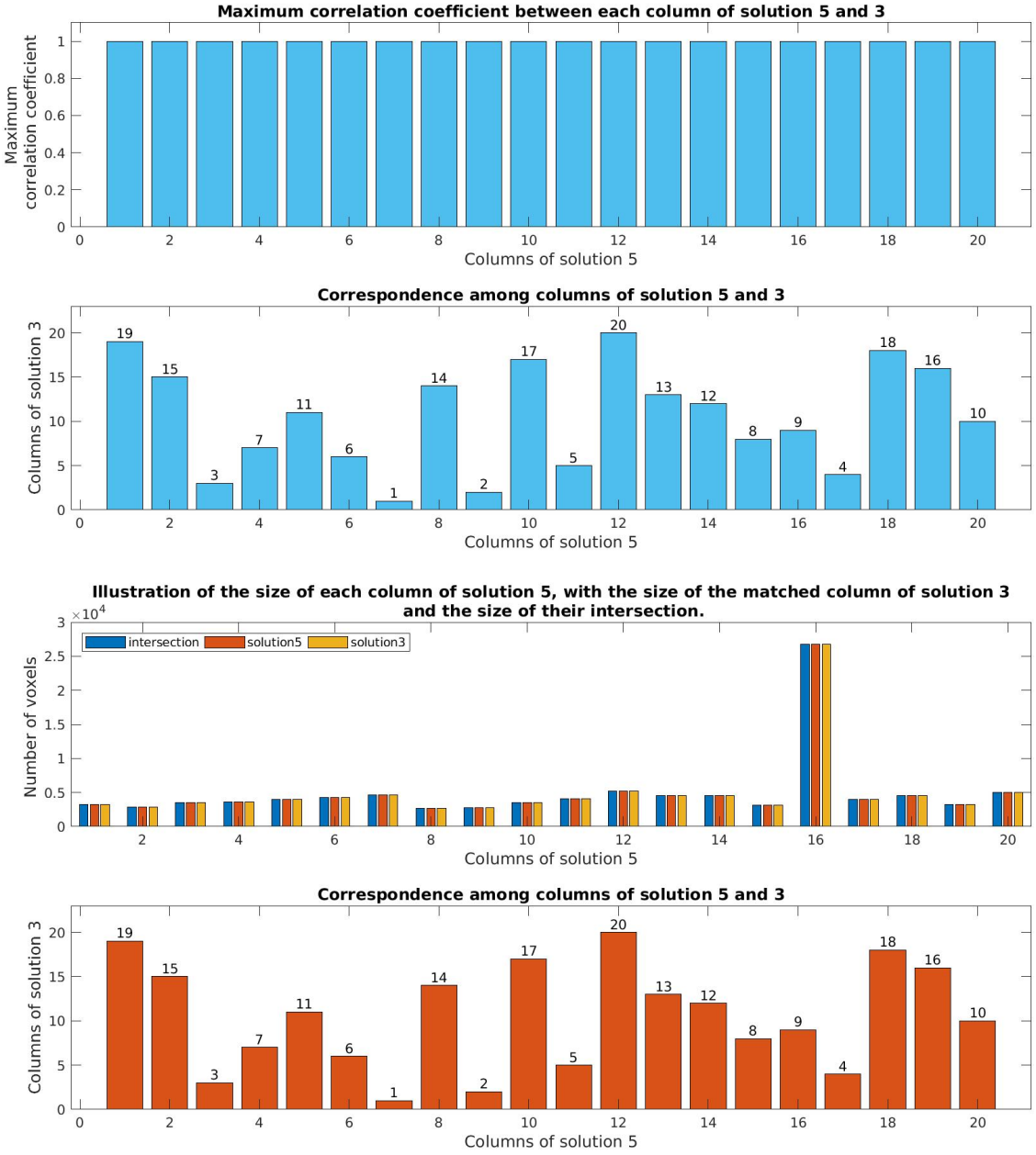


Figure A.242: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 3. We associate the column r of the solution 5 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 5 is associated, based on intersection. Solutions 5, 3 correspond to accelerated penalty method and SNR=0 dB.

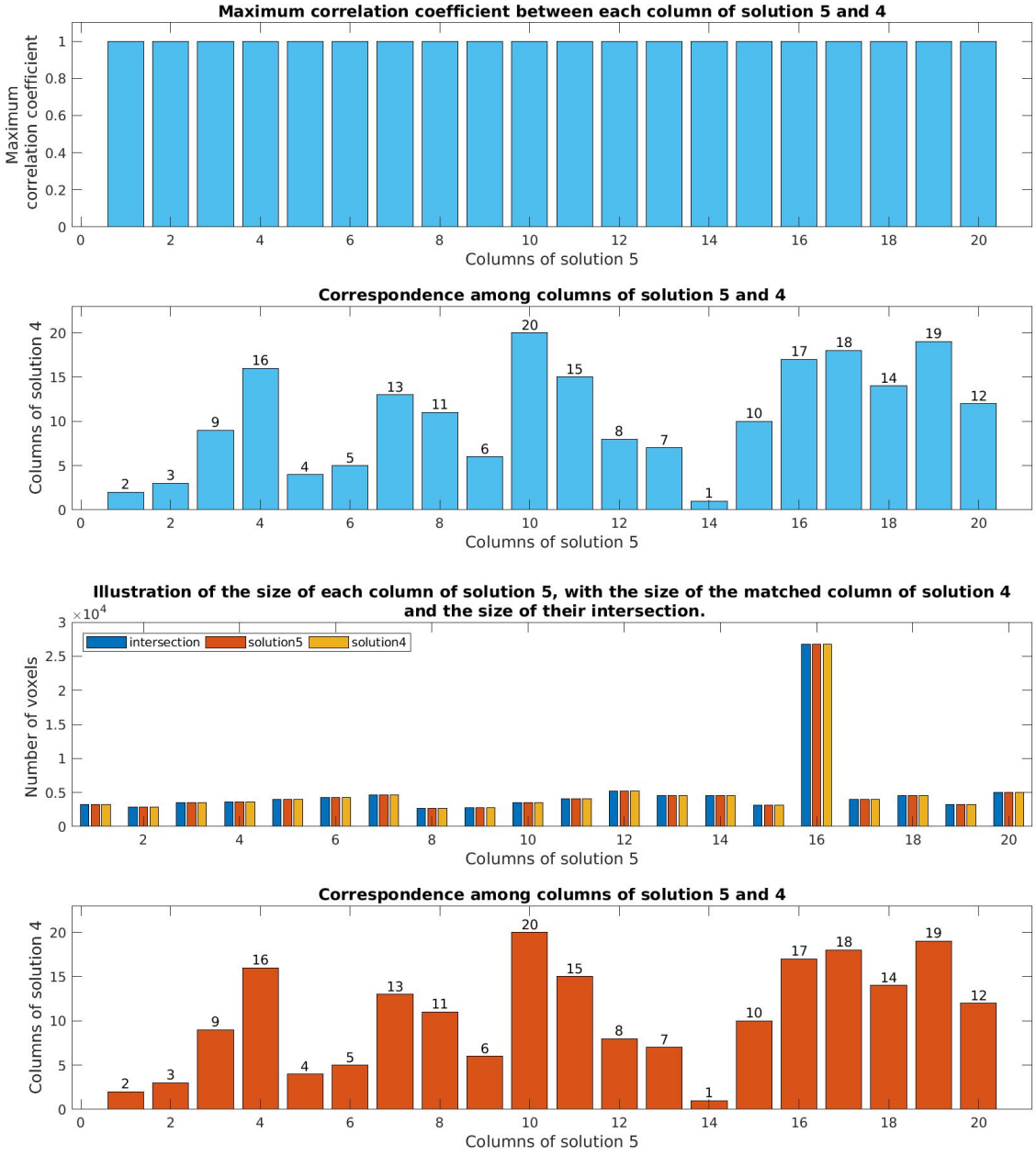


Figure A.243: **First figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 5 and each column of the solution 4. We associate the column r of the solution 5 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on correlation coefficients. **Third figure**: For each column r of the solution 5, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 5 is associated, based on intersection. Solutions 5, 4 correspond to accelerated penalty method and SNR=0 dB.

· **Comparison of Accelerated penalty method solutions to real factor A.**

· **Noiseless case**

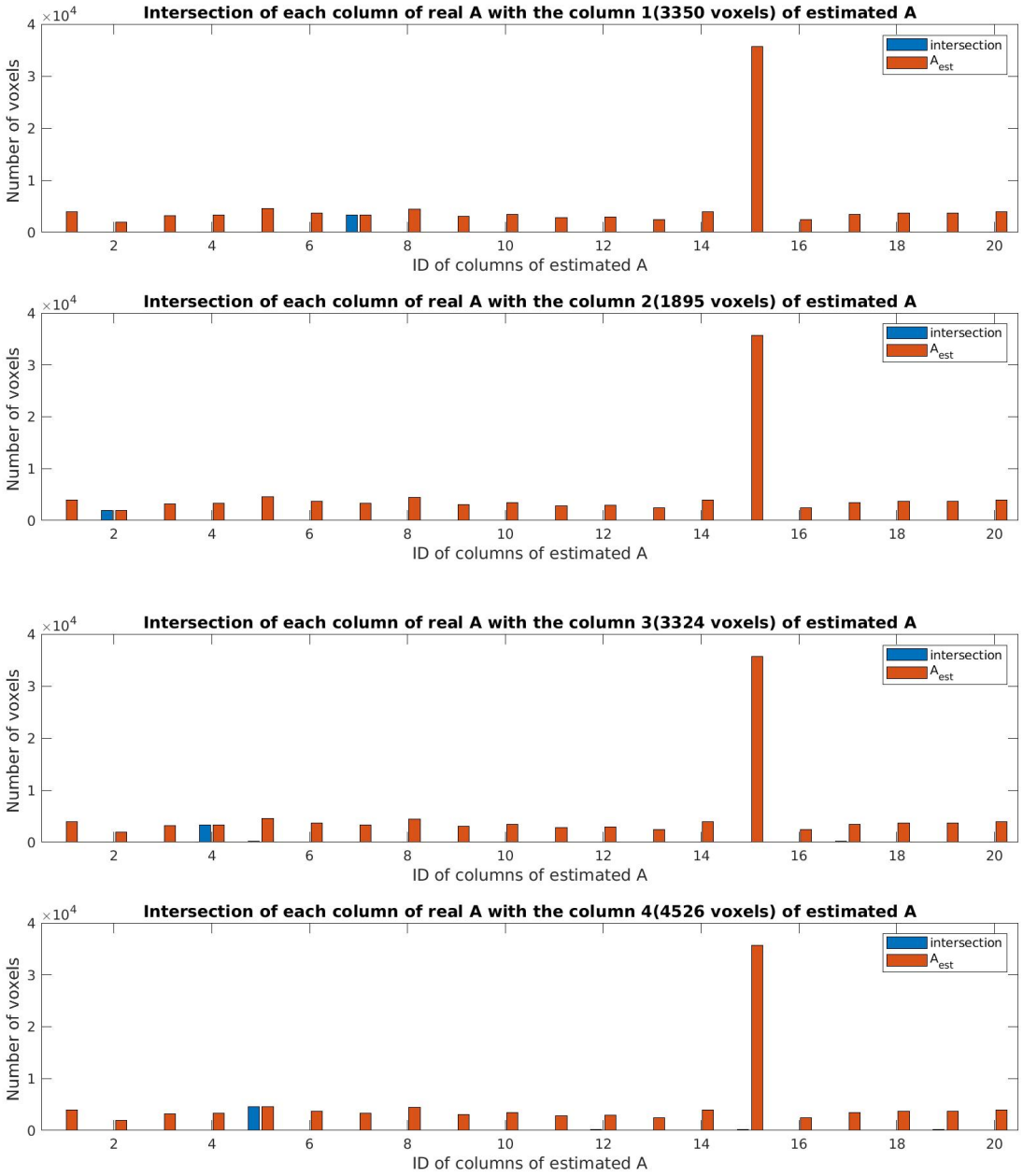


Figure A.244: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

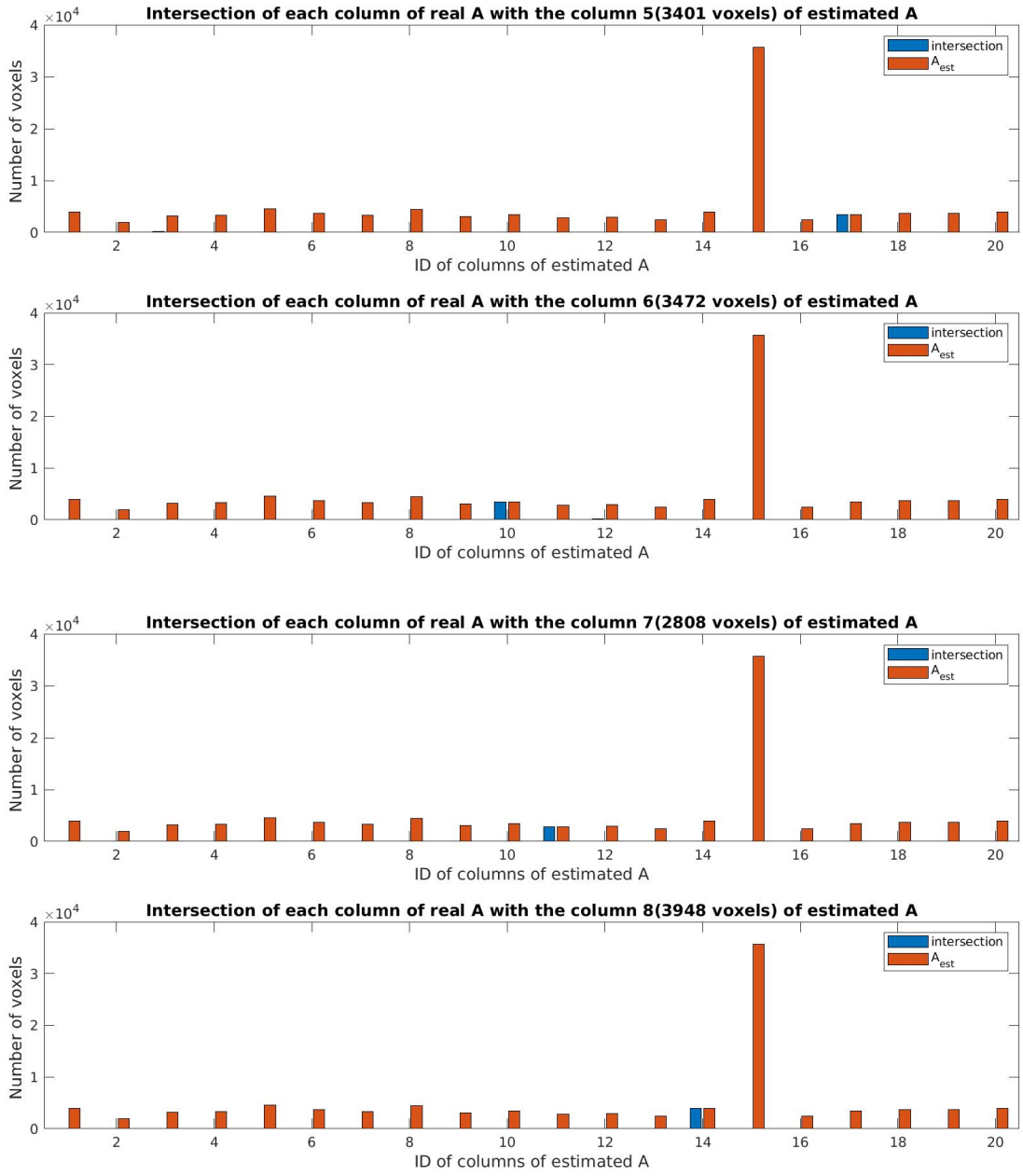


Figure A.245: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

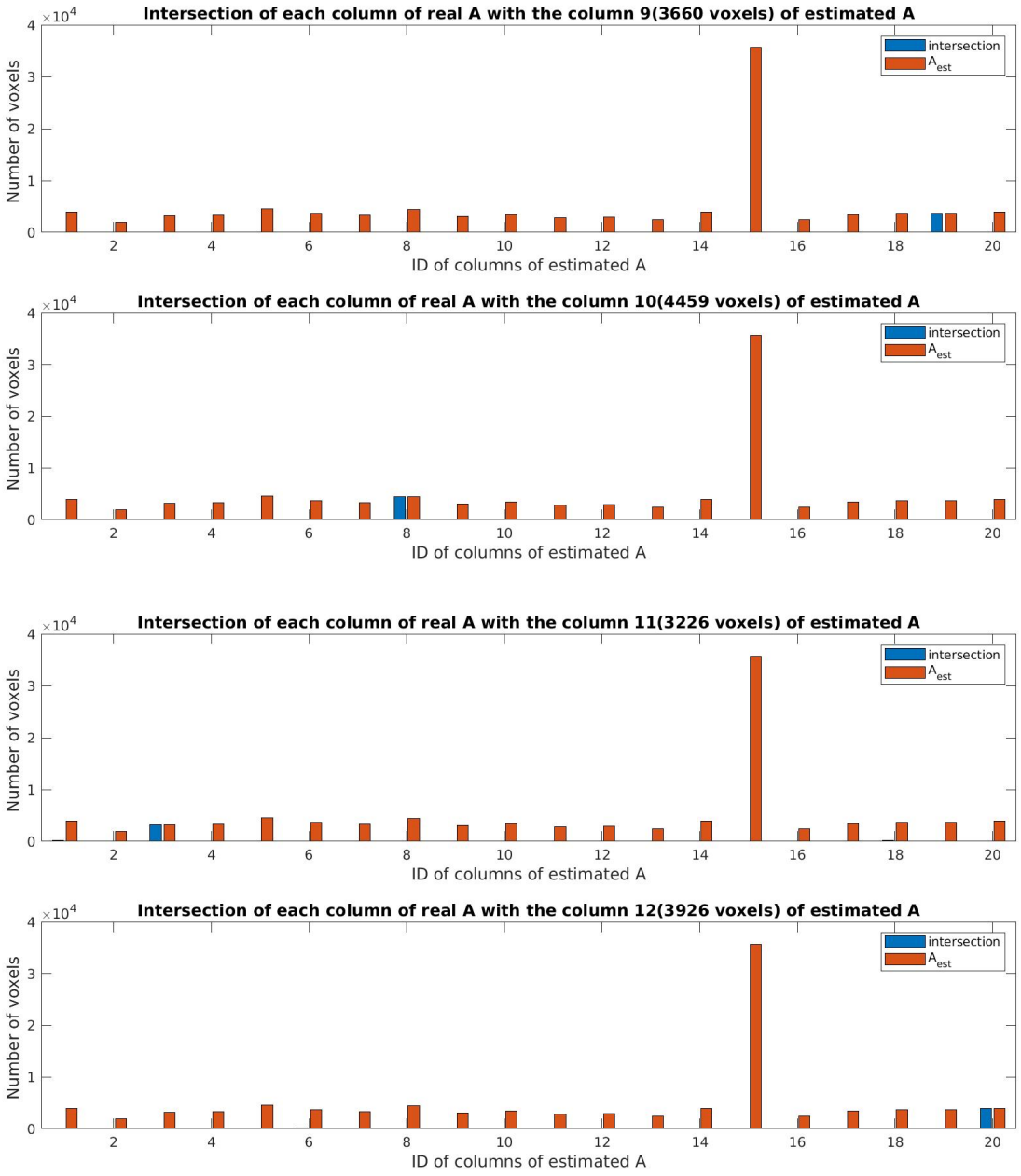


Figure A.246: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

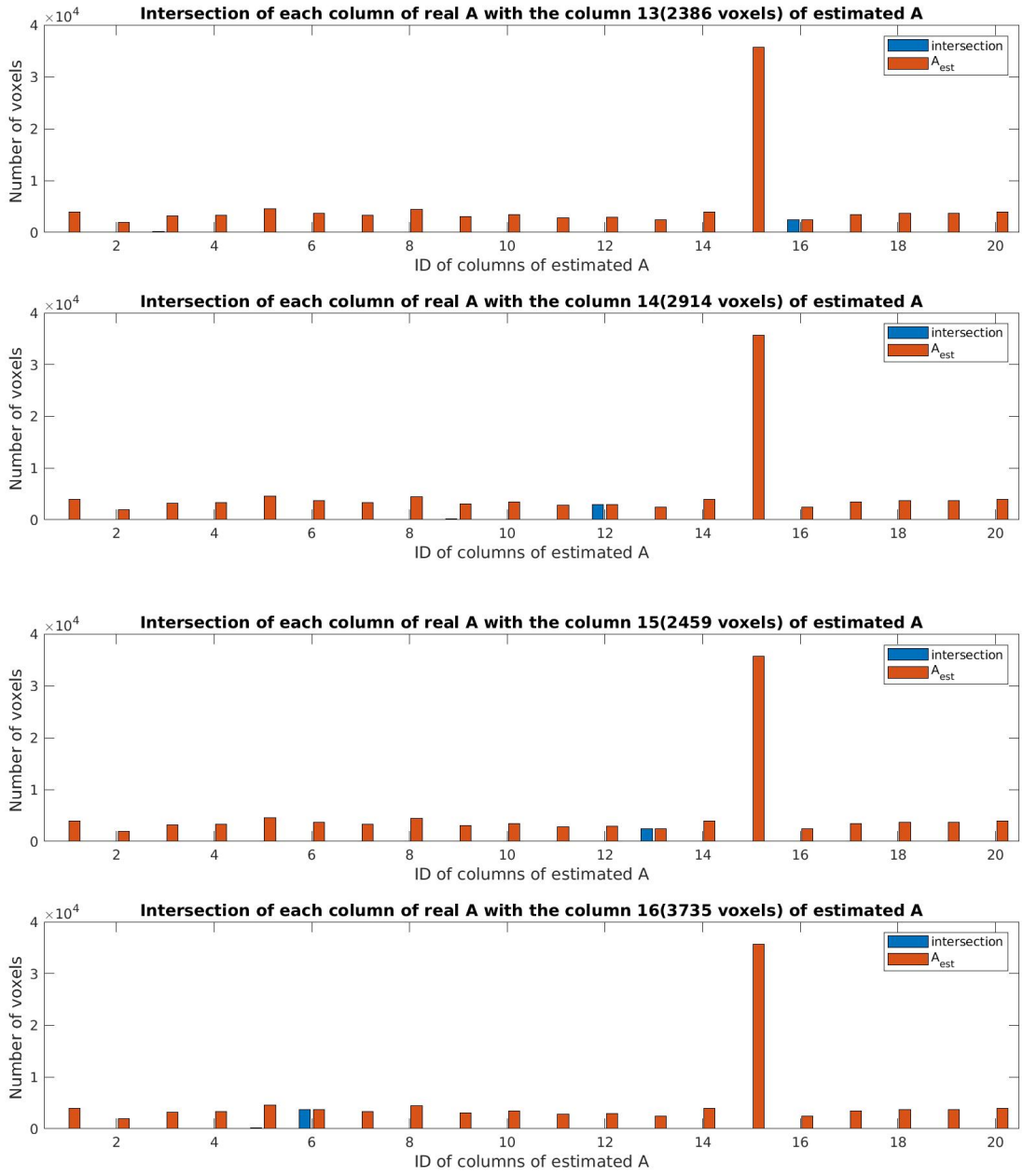


Figure A.247: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

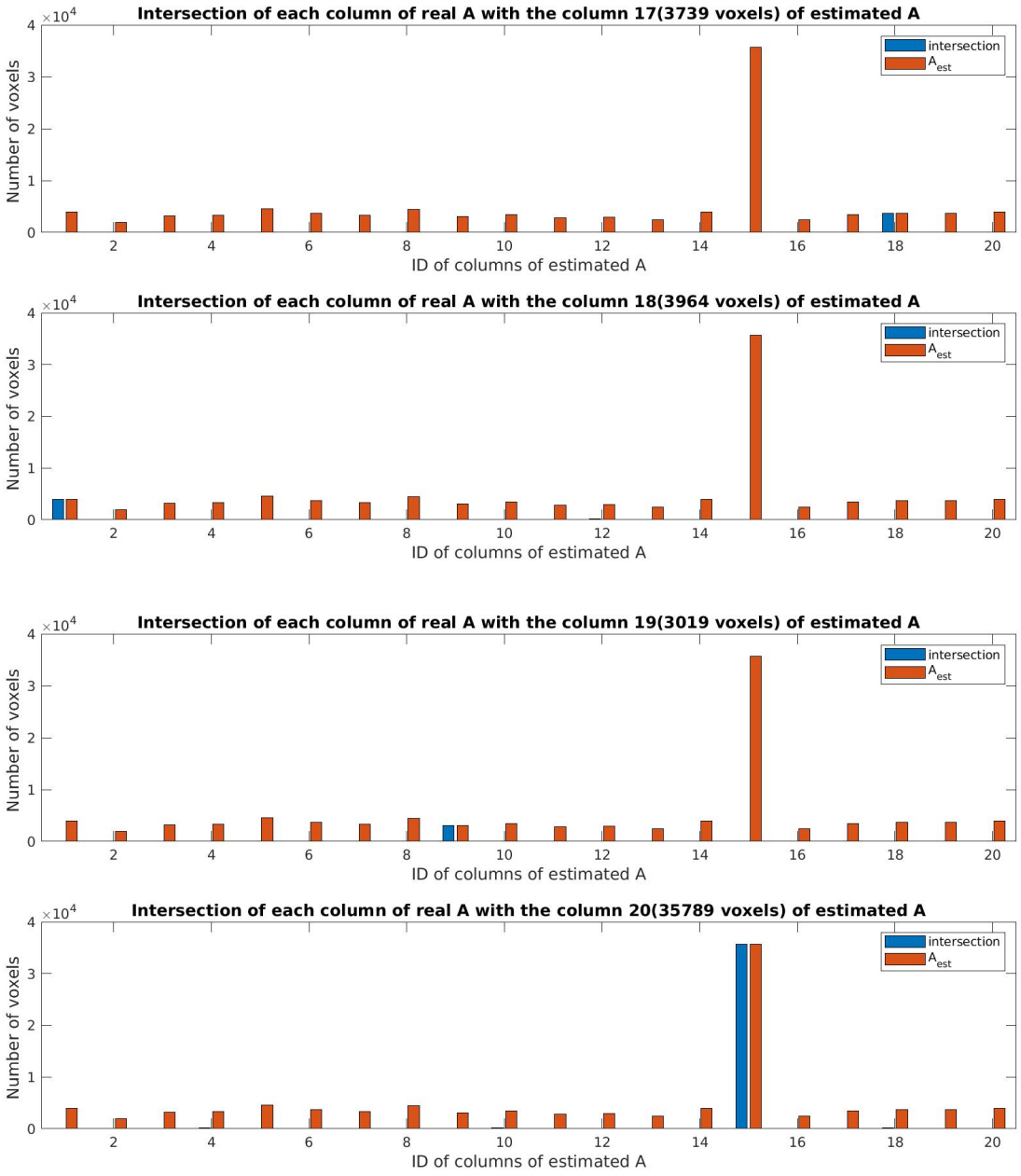


Figure A.248: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and noiseless case.

· **SNR= 0 dB**

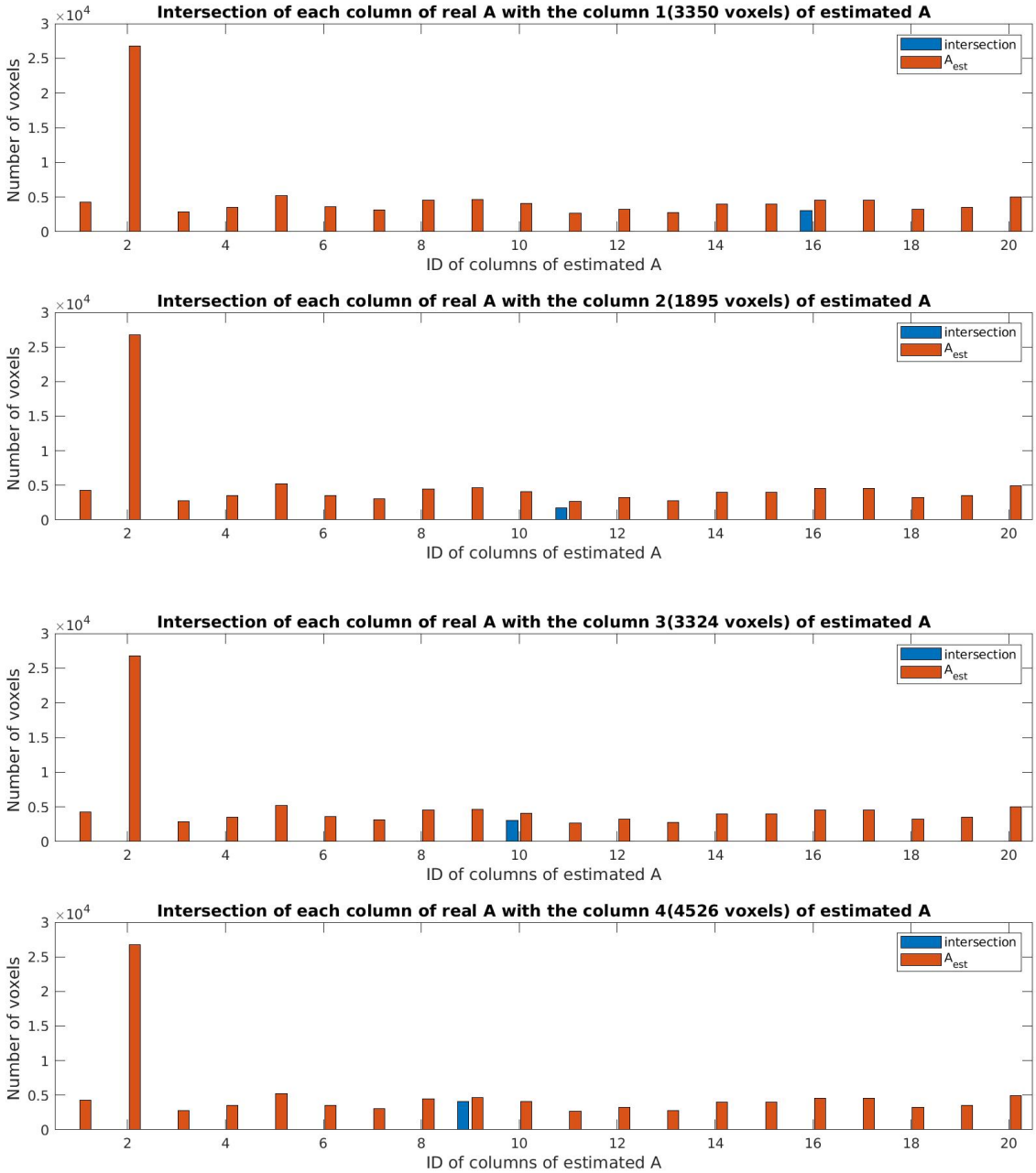


Figure A.249: For the column r, for $r = 1, \ldots, 4$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

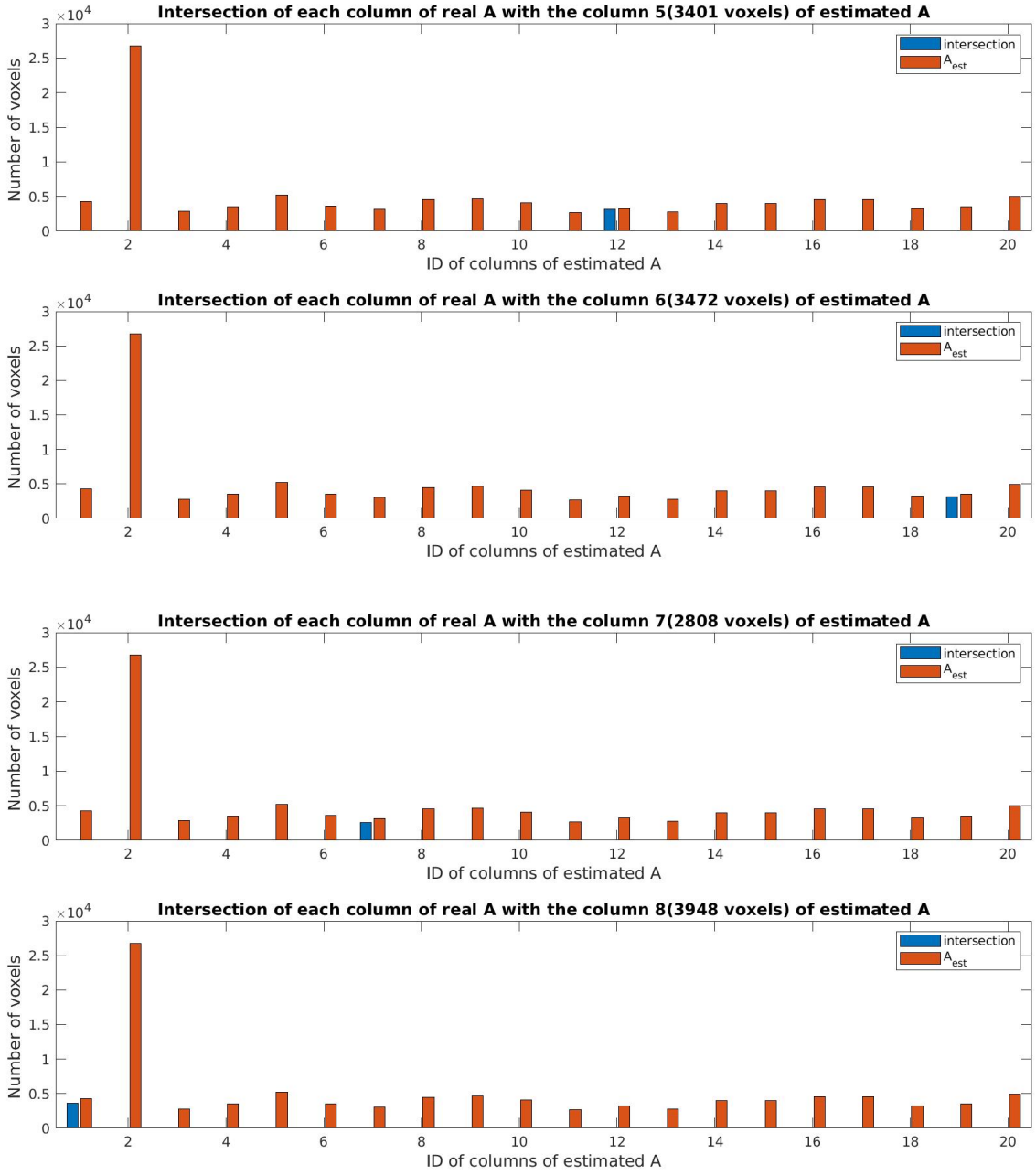


Figure A.250: For the column r, for $r = 5, \ldots, 8$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

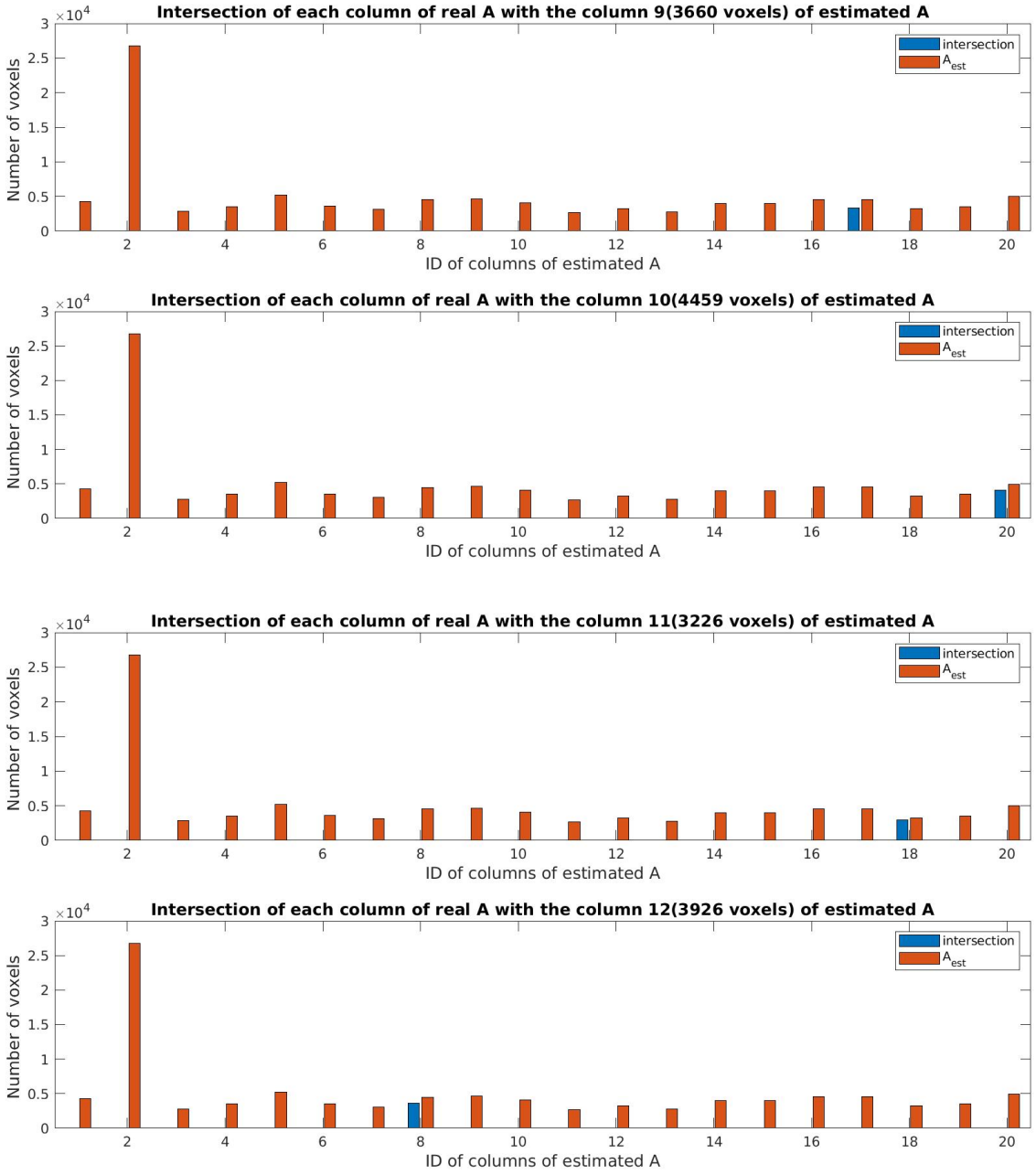


Figure A.251: For the column r, for $r = 9, \ldots, 12$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

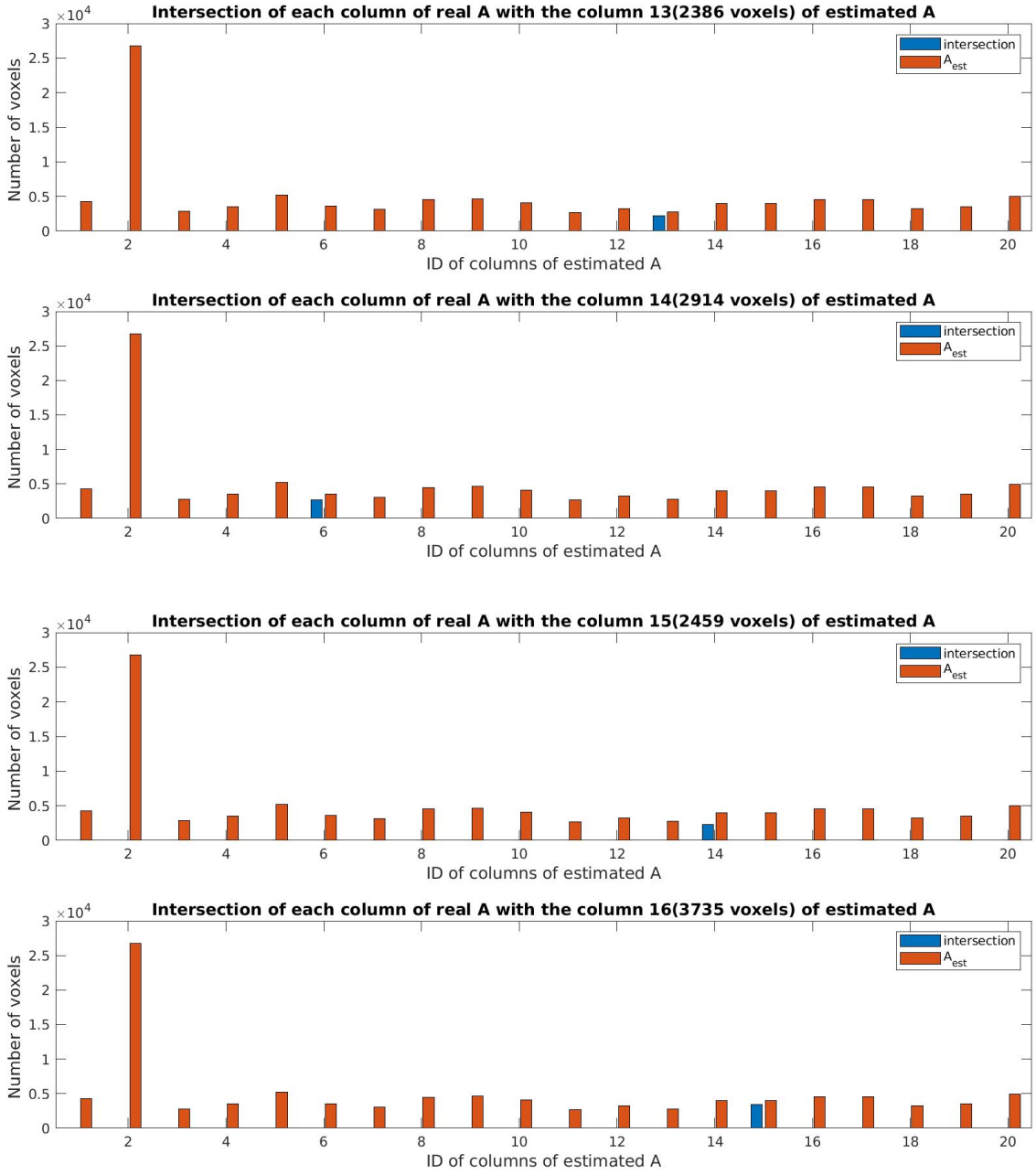


Figure A.252: For the column r, for $r = 13, \ldots, 16$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

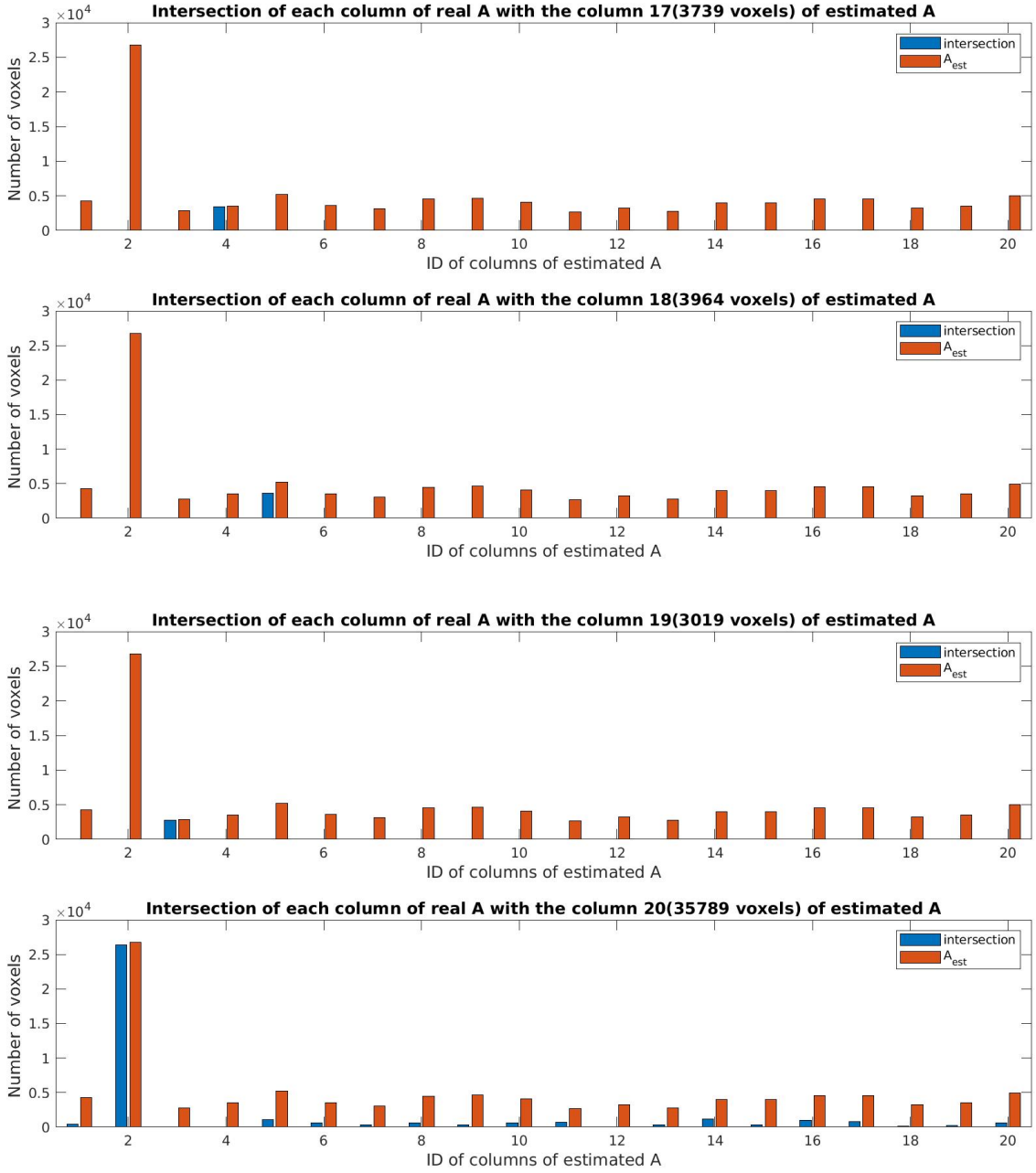


Figure A.253: For the column r, for $r = 17, \ldots, 20$, of the true $\mathbf{A}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the solution 1 (blue bar). We also illustrate the number of the non-zero elements of each column of the solution 1 (red bar). The solution 1 corresponds to accelerated penalty method and SNR=0 dB.

· **Histograms for SNR=0 dB**

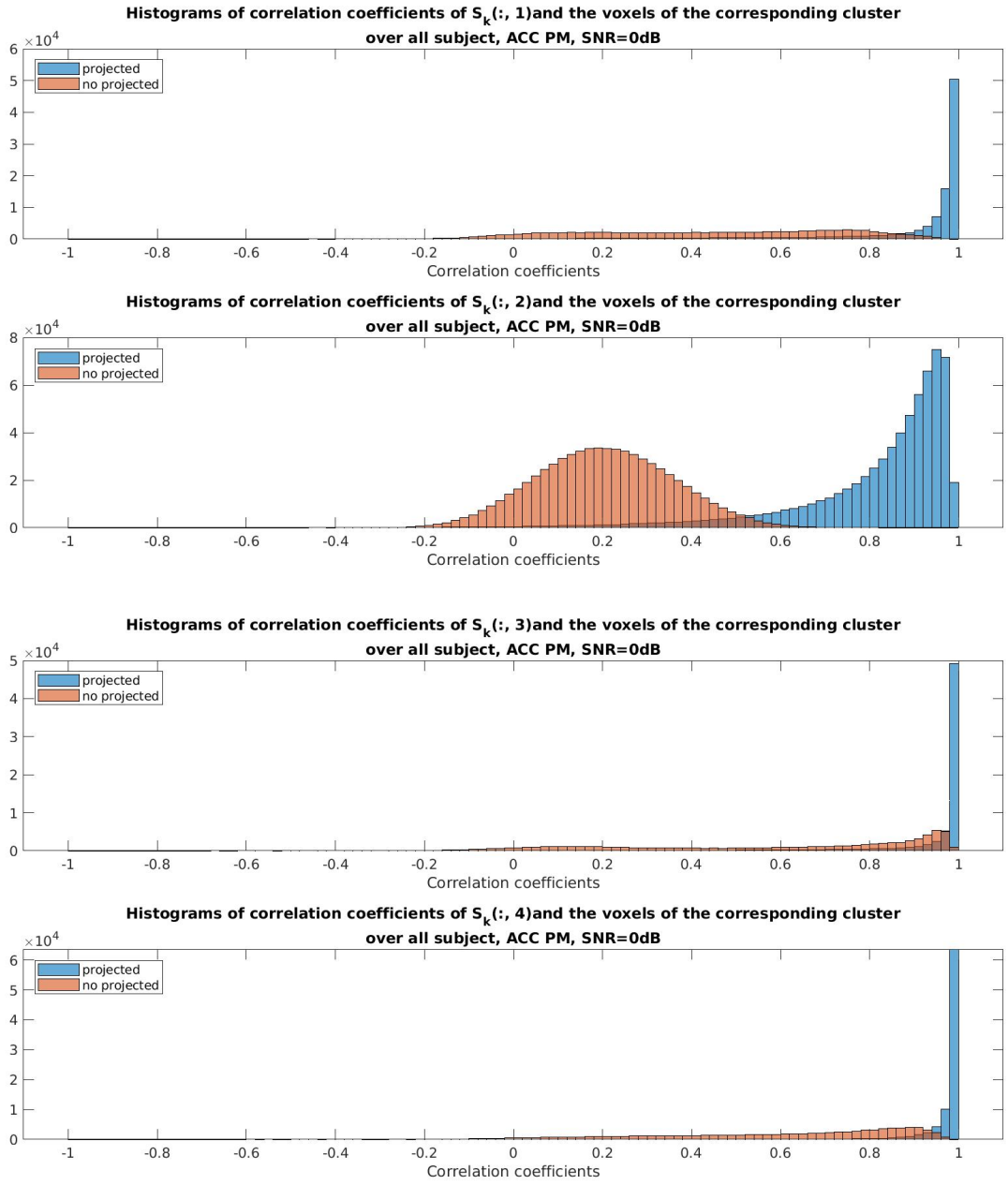


Figure A.254: For the column r and subject k, for $r = 1, \ldots, 4$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 1 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 2 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 3 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 4 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

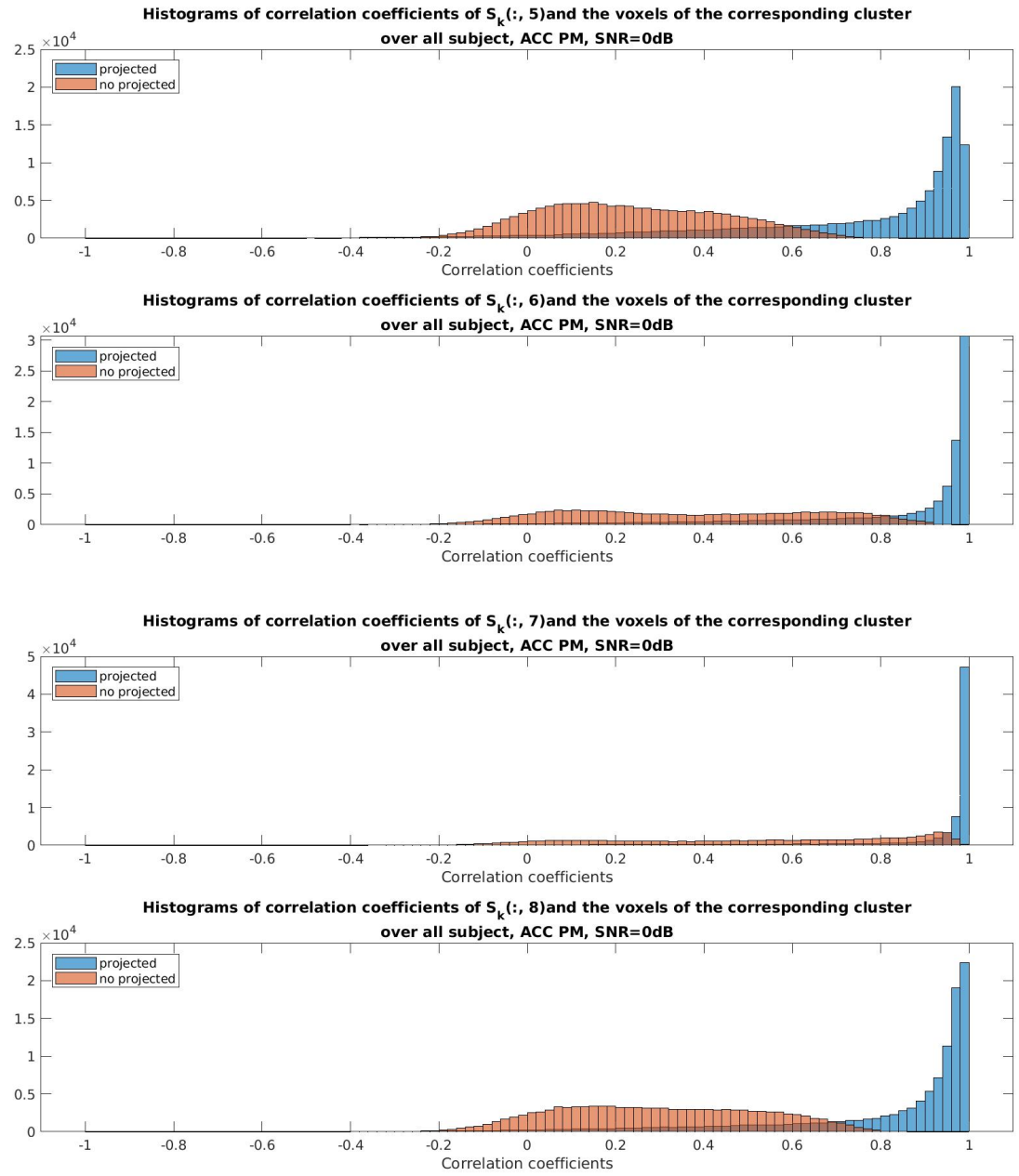


Figure A.255: For the column r and subject k, for $r = 5, \ldots, 8$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 5 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 6 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 7 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 8 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

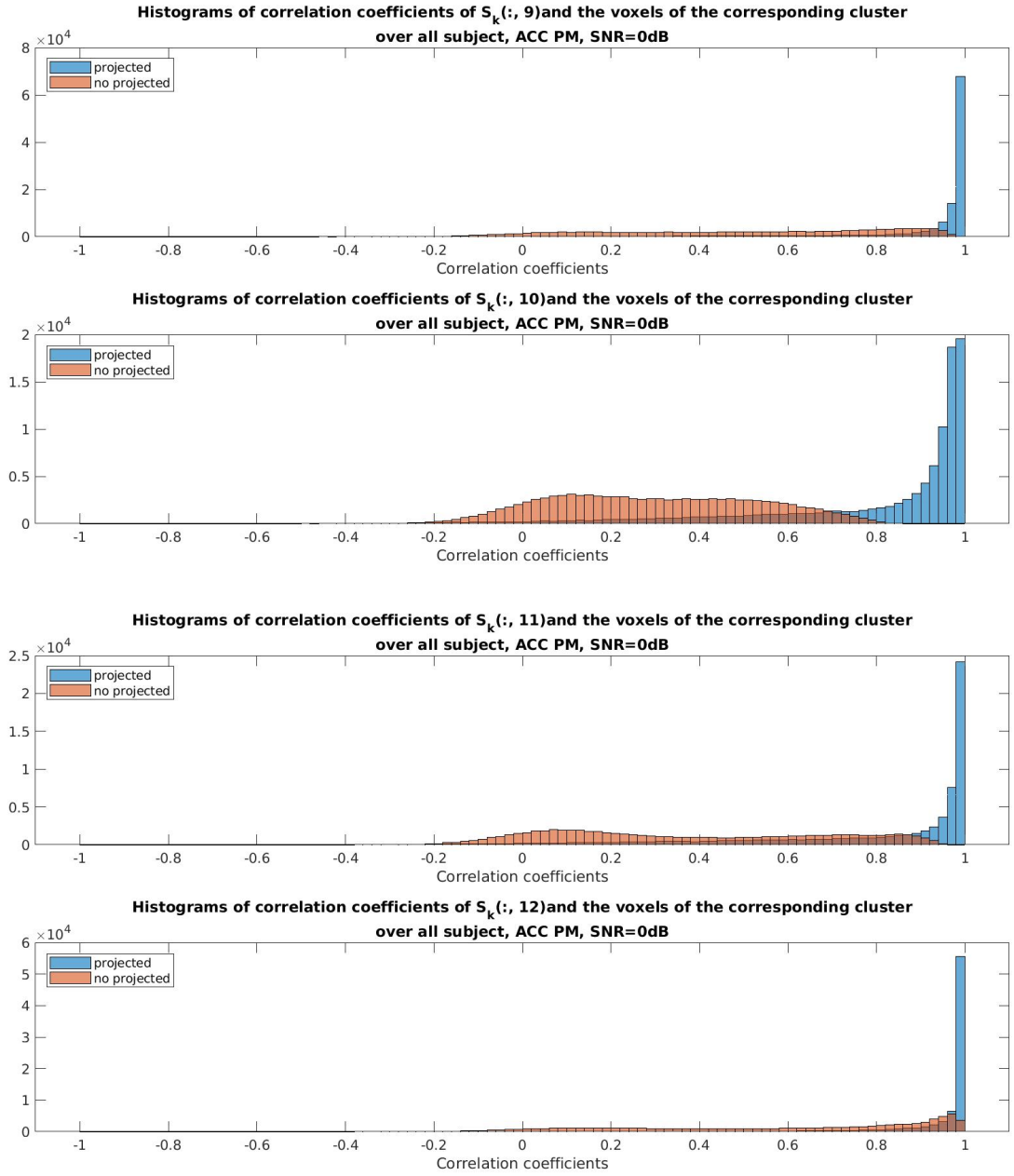


Figure A.256: For the column r and subject k, for $r = 9, \ldots, 12$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 9 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 10 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 11 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 12 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

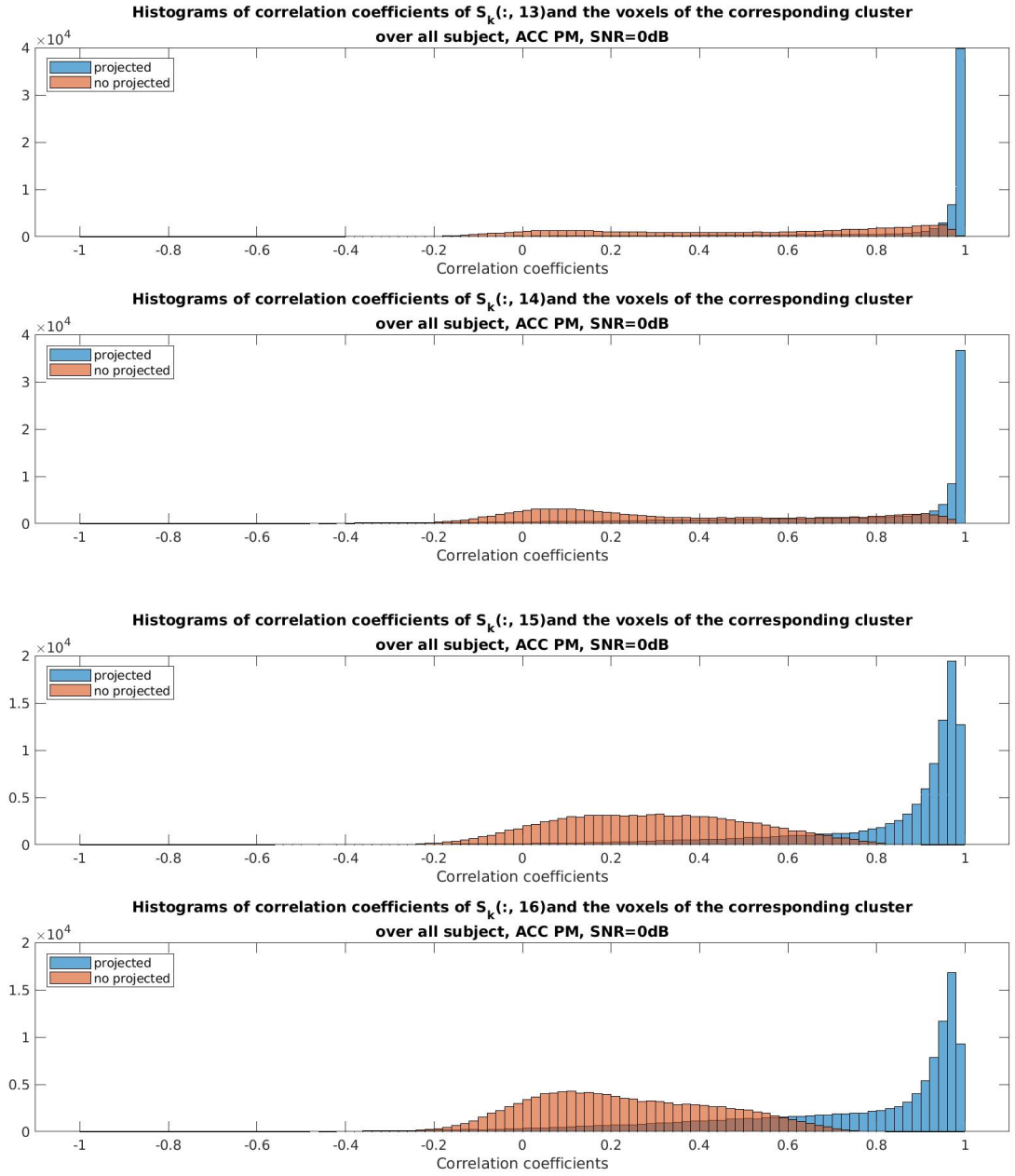


Figure A.257: For the column r and subject k, for $r = 13, \ldots, 16$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 13 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 14 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 15 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 16 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

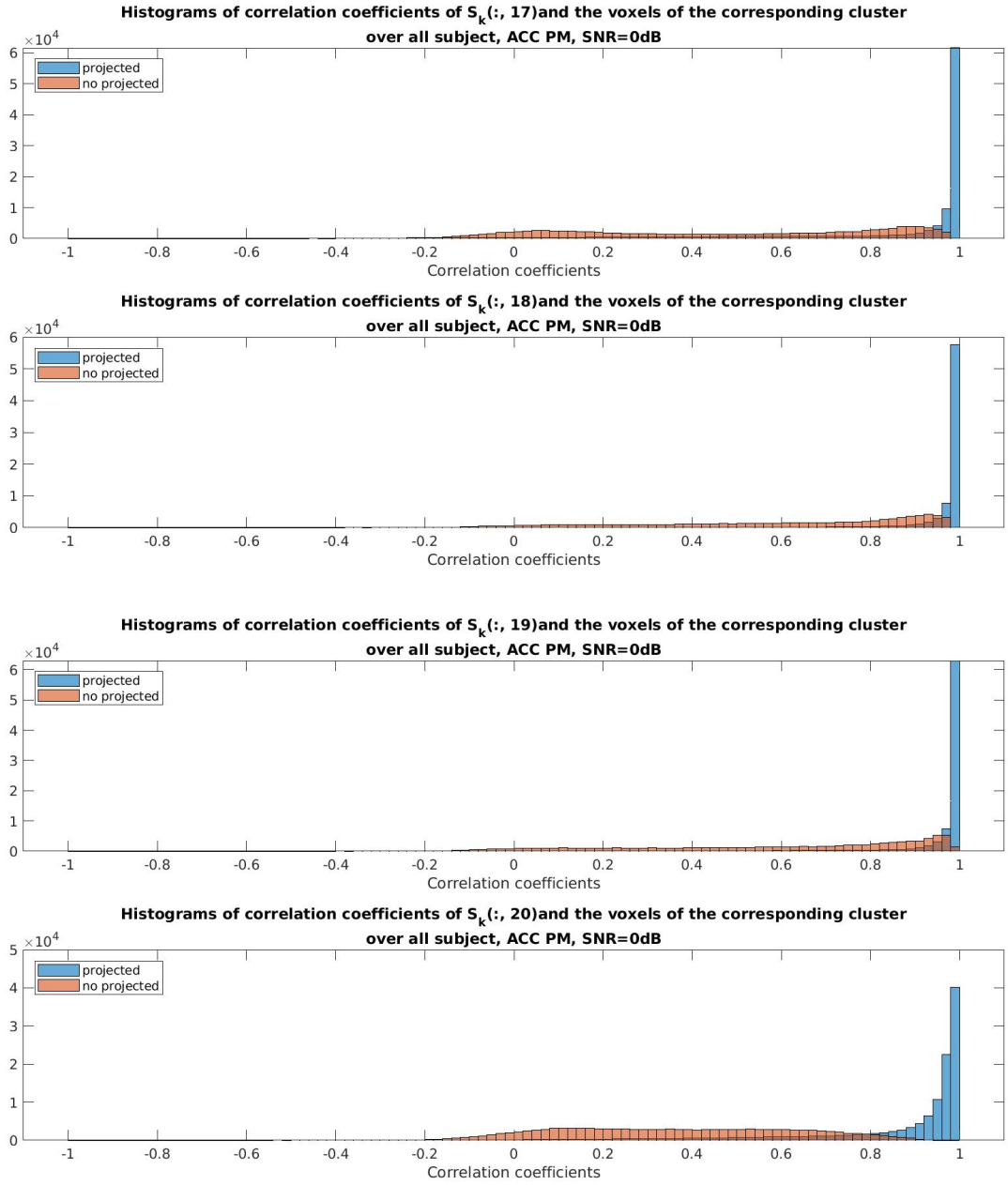


Figure A.258: For the column r and subject k, for $r = 17, \ldots, 20$ and $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, r)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color). The data correspond to data with SNR=0 dB. **First image**: It shows the histograms which correspond to column 17 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Second image**: It shows the histograms which correspond to column 18 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Third image**: It shows the histograms which correspond to column 19 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error. **Fourth image**: It shows the histograms which correspond to column 20 and the clusters are indicated by $\mathbf{A}_{est}$, which has been obtained by accelerated penalty method and corresponds to the best solution based on relative error.

Appendix B

Real Data

B.1 Estimation of common subspace dimension

		minimum gap	mean gap	maximum gap			minimum gap	mean gap	maximum gap
	1	0.16258	0.22154	0.28803		2	0.31692	0.46775	0.73569
	3	0.40105	0.52353	0.71246		4	0.45747	0.61435	0.75542
	5	0.46745	0.69169	0.86347		6	0.44763	0.52251	0.59657
	7	0.57349	0.65276	0.75586		8	0.85784	0.96116	0.99767
	9	0.82375	0.93735	0.99887		10	0.78742	0.90886	0.99072
	11	0.81419	0.89577	0.98005		12	0.78577	0.87961	0.9572
	13	0.91323	0.97391	0.99658		14	0.90624	0.97468	0.9983
	15	0.89852	0.9669	0.99125		16	0.94433	0.98931	0.99888
	17	0.85545	0.93268	0.99746		18	0.79914	0.88254	0.9834
C	19	0.97159	0.98993	0.9992	C	20	0.97851	0.99253	0.99958
O	21	0.89987	0.94883	0.99168	O	22	0.94849	0.97757	0.99788
M	23	0.97715	0.99033	0.9991	M	24	0.97773	0.99267	0.99992
M	25	0.97436	0.99273	0.99996	M	26	0.94681	0.97583	0.99839
O	27	0.9743	0.99044	0.99854	O	28	0.98706	0.99455	0.99926
N	29	0.98334	0.99496	0.99976	N	30	0.98369	0.99591	0.99977
	31	0.99073	0.9972	0.99983		32	0.99217	0.99757	0.99996
S	33	0.98444	0.9948	0.99971	S	34	0.98287	0.99558	0.99999
U	35	0.99442	0.99795	0.99995	U	36	0.98654	0.99578	0.99949
B	37	0.99414	0.99719	0.99985	B	38	0.98069	0.99356	0.99803
S	39	0.98898	0.99468	0.99988	S	40	0.987	0.99695	0.99981
P	41	0.99063	0.99693	0.99961	P	42	0.99646	0.99859	0.99994
A	43	0.99718	0.999	0.99997	A	44	0.99814	0.99899	0.99992
C	45	0.99892	0.99968	0.99998	C	46	0.99826	0.99965	0.99999
E	47	0.99869	0.99979	1	E	48	0.9983	0.99957	0.99998
	49	0.99914	0.99973	0.99999		50	0.99906	0.99979	1
D	51	0.99853	0.99959	0.99999	D	52	0.99934	0.99982	0.99997
I	53	0.99929	0.99971	1	I	54	0.99916	0.99982	1
M	55	0.99903	0.99961	0.99998	M	56	0.99903	0.99963	0.99997
E	57	0.99933	0.99977	0.99999	E	58	0.99839	0.99957	0.99994
N	59	0.99892	0.99969	0.99999	N	60	0.99912	0.99974	0.99999
S	61	0.99913	0.99972	0.99999	S	62	0.99962	0.99983	1
I	63	0.99935	0.99979	0.99999	I	64	0.99952	0.99982	0.99998
0	65	0.99955	0.99984	1	0	66	0.99969	0.99991	1
N	67	0.99966	0.99991	1	N	68	0.99972	0.9999	0.99999
	69	0.99953	0.99986	1		70	0.99971	0.99991	1
	71	0.99981	0.99992	1		72	0.99972	0.99992	0.99999
	73	0.99976	0.99992	0.99999		74	0.99986	0.99995	1
	75	0.99978	0.99993	0.99999		76	0.99976	0.99991	1
	77	0.99985	0.99996	1		78	0.99975	0.99993	0.99999
	79	0.99983	0.99994	1		80	0.99983	0.99995	1

		minimum gap	mean gap	maximum gap			minimum gap	mean gap	maximum gap
	81	0.99992	0.99998	1		82	0.99989	0.99996	1
C	83	0.99988	0.99995	0.99999	C	84	0.9999	0.99997	1
O	85	0.9997	0.99995	1	O	86	0.99993	0.99998	1
M	87	0.99992	0.99998	1	M	88	0.99994	0.99997	1
M	89	0.99991	0.99997	1	M	90	0.99997	0.99998	1
O	91	0.99993	0.99998	1	O	92	0.9999	0.99997	1
N	93	0.99993	0.99998	1	N	94	0.99984	0.99996	1
	95	0.99991	0.99997	0.99999		96	0.99991	0.99997	1
S	97	0.99993	0.99998	1	S	98	0.99995	0.99998	1
U	99	0.99995	0.99998	1	U	100	0.99995	0.99998	1
B	101	0.99996	0.99998	1	B	102	0.99995	0.99998	1
S	103	0.99995	0.99997	1	S	104	0.99995	0.99998	1
P	105	0.9999	0.99997	1	P	106	0.99991	0.99997	1
A	107	0.99993	0.99998	1	A	108	0.99991	0.99997	1
C	109	0.9999	0.99997	1	C	110	0.99994	0.99997	1
E	111	0.99995	0.99998	1	E	112	0.99994	0.99998	1
	113	0.99991	0.99999	1		114	0.99995	0.99998	1
D	115	0.99996	0.99998	1	D	116	0.99994	0.99999	1
I	117	0.99998	0.99999	1	I	118	0.99997	0.99999	1
M	119	0.99994	0.99999	1	M	120	0.99994	0.99998	1
E	121	0.99994	0.99998	1	E	122	0.99996	0.99998	1
N	123	0.99994	0.99998	1	N	124	0.99994	0.99998	1
S	125	0.99993	0.99998	1	S	126	0.99993	0.99997	1
I	127	0.99997	0.99998	1	I	128	0.99991	0.99998	1
0	129	0.99996	0.99999	1	0	130	0.99995	0.99999	1
N	131	0.99997	0.99999	1	N	132	0.99998	0.99999	1
	133	0.99994	0.99999	1		134	0.99998	0.99999	1
	135	0.99996	0.99999	1		136	0.99996	0.99999	1
	137	0.99996	0.99999	1		138	0.99998	0.99999	1
	139	0.99998	0.99999	1		140	0.99998	0.99999	1
	141	0.99998	0.99999	1		142	0.99998	0.99999	1
	143	0.99995	0.99999	1					

Table B.1: We apply the criterion to 25 random partitions of data. For each dimension of common subspace, we remove the 3 smallest and the 3 largest values of gap from 25 values and then we calculate the minimum, mean and maximum value for each dimension.

		minimum gap	mean gap	maximum gap			minimum gap	mean gap	maximum gap
	1	0.14116	0.21789	0.2981		2	0.26507	0.46772	0.7836
	3	0.33918	0.51655	0.83025		4	0.41754	0.62328	0.82677
	5	0.43038	0.66062	0.97595		6	0.42159	0.51178	0.59657
	7	0.54129	0.66608	0.90636		8	0.80388	0.94492	0.99981
	9	0.79556	0.93343	0.9999		10	0.71515	0.90418	1
	11	0.70802	0.88671	0.9894		12	0.70587	0.87333	0.99752
	13	0.85745	0.96529	0.99998		14	0.8366	0.95952	0.99977
	15	0.86687	0.95266	0.99903		16	0.87117	0.97967	0.99999
	17	0.77777	0.91934	0.99976		18	0.72387	0.87718	0.99696
C	19	0.9336	0.98542	0.99901	C	20	0.97851	0.99262	0.99976
O	21	0.84643	0.95083	0.99866	O	22	0.87917	0.97402	0.99964
M	23	0.96594	0.99014	1	M	24	0.96628	0.99133	0.99997
M	25	0.97095	0.991	0.99998	M	26	0.88788	0.97018	0.99966
O	27	0.95237	0.98528	1	O	28	0.96582	0.99261	0.99983
N	29	0.97827	0.99396	1	N	30	0.97419	0.99296	0.99998
	31	0.97896	0.99532	0.99995		32	0.98365	0.99702	1
S	33	0.97727	0.99361	0.99987	S	34	0.97702	0.99405	1
U	35	0.97617	0.99653	0.99997	U	36	0.96307	0.99341	1
B	37	0.98103	0.99611	0.99995	B	38	0.97492	0.99257	1
S	39	0.97968	0.99386	0.99995	S	40	0.98067	0.99617	1
P	41	0.99027	0.99685	0.99997	P	42	0.99214	0.99818	0.99999
A	43	0.99327	0.99859	1	A	44	0.99747	0.9989	0.99998
C	45	0.99485	0.9992	1	C	46	0.99782	0.99935	1
E	47	0.99695	0.99943	1	E	48	0.99373	0.99904	1
	49	0.99788	0.99959	1		50	0.99832	0.99966	1
D	51	0.99779	0.99935	1	D	52	0.99852	0.9997	1
I	53	0.99863	0.99961	1	I	54	0.99523	0.99941	1
M	55	0.9951	0.99931	1	M	56	0.99799	0.99954	0.99999
E	57	0.99906	0.9997	1	E	58	0.99729	0.99928	1
N	59	0.9979	0.99945	1	N	60	0.99845	0.99956	1
S	61	0.99847	0.99958	1	S	62	0.99914	0.99978	1
I	63	0.99877	0.99967	1	I	64	0.99871	0.9997	1
0	65	0.99941	0.9998	1	0	66	0.99868	0.99979	1
N	67	0.9991	0.99981	1	N	68	0.99939	0.99986	1
	69	0.99924	0.9998	1		70	0.99947	0.99986	1
	71	0.99964	0.99991	1		72	0.99951	0.9999	1
	73	0.99972	0.99991	1		74	0.99979	0.99994	1
	75	0.99939	0.99991	1		76	0.99926	0.99988	1
	77	0.99978	0.99995	1		78	0.9996	0.99991	1
	79	0.99978	0.99993	1		80	0.99977	0.99995	1

		minimum gap	mean gap	maximum gap			minimum gap	mean gap	maximum gap
	81	0.99986	0.99997	1		82	0.9998	0.99996	1
C	83	0.99972	0.99995	1	C	84	0.9998	0.99995	1
O	85	0.99937	0.99988	1	O	86	0.99967	0.99996	1
M	87	0.99985	0.99997	1	M	88	0.99972	0.99995	1
M	89	0.99991	0.99997	1	M	90	0.99989	0.99998	1
O	91	0.99981	0.99997	1	O	92	0.99974	0.99995	1
N	93	0.99988	0.99998	1	N	94	0.99979	0.99995	1
	95	0.99959	0.99995	0.99999		96	0.9998	0.99995	1
S	97	0.9998	0.99996	1	S	98	0.99988	0.99998	1
U	99	0.99988	0.99997	1	U	100	0.99989	0.99997	1
B	101	0.9999	0.99998	1	B	102	0.99994	0.99998	1
S	103	0.99989	0.99997	1	S	104	0.99981	0.99996	1
P	105	0.9999	0.99998	1	P	106	0.99979	0.99997	1
A	107	0.99974	0.99997	1	A	108	0.99987	0.99997	1
C	109	0.99989	0.99997	1	C	110	0.99989	0.99996	1
E	111	0.99989	0.99998	1	E	112	0.9999	0.99998	1
	113	0.99988	0.99998	1		114	0.99991	0.99998	1
D	115	0.99978	0.99997	1	D	116	0.99993	0.99998	1
I	117	0.99994	0.99999	1	I	118	0.99989	0.99999	1
M	119	0.99984	0.99998	1	M	120	0.99991	0.99998	1
E	121	0.9999	0.99998	1	E	122	0.99989	0.99998	1
N	123	0.99978	0.99996	1	N	124	0.9999	0.99998	1
S	125	0.9999	0.99998	1	S	126	0.99993	0.99997	1
I	127	0.99995	0.99998	1	I	128	0.99989	0.99998	1
0	129	0.99995	0.99999	1	0	130	0.99994	0.99998	1
N	131	0.99996	0.99999	1	N	132	0.99996	0.99999	1
	133	0.99992	0.99998	1		134	0.99993	0.99998	1
	135	0.99995	0.99999	1		136	0.9999	0.99998	1
	137	0.99989	0.99998	1		138	0.99995	0.99999	1
	139	0.99994	0.99999	1		140	0.99995	0.99999	1
	141	0.99994	0.99999	1		142	0.99998	0.99999	1
	143	0.99992	0.99999	1					

Table B.2: We apply the criterion to 25 random partitions of data. For each partition, we estimate the common subspace dimension. To end this, we consider that the common subspace dimension is the one after which the values of gap are greater than a threshold equal to 0.99. We remove the 3 partitions which have the smallest common subspace dimension and the 3 partitions which have the largest one and then we calculate the minimum, mean and maximum value for each dimension.

		COMMON SUBSPACE DIMENSION														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	1	0.27579	0.66889	0.42344	0.54508	0.69009	0.58117	0.61879	0.92832	0.89756	0.89277	0.86883	0.85673	0.95824	0.99836	0.94568
	2	0.28731	0.50958	0.45481	0.49052	0.55718	0.59246	0.6191	0.85784	0.91371	0.98004	0.732	0.90157	0.91202	0.88141	0.89414
	3	0.15583	0.26507	0.41979	0.44163	0.81657	0.44763	0.57349	0.8263	0.79556	0.79659	0.94777	0.90775	0.91323	0.90624	0.86687
	4	0.2981	0.7449	0.71246	0.75884	0.69198	0.5327	0.54129	0.91584	0.83341	0.9959	0.88837	0.90425	0.97752	0.98325	0.9389
	5	0.19431	0.579	0.61538	0.75542	0.70618	0.57148	0.72398	0.98101	0.98428	0.96485	0.87402	0.90721	0.98571	0.87983	0.88666
P	6	0.28803	0.47411	0.52503	0.45708	0.46141	0.48941	0.63323	0.90259	0.9999	0.94561	0.97871	0.84026	0.98742	0.99435	0.9889
A	7	0.2682	0.75973	0.69862	0.82649	0.81515	0.51009	0.69322	0.99499	0.99887	0.98045	0.9894	0.9933	0.9995	0.91716	0.97121
R	8	0.20659	0.46971	0.37759	0.45747	0.93178	0.47184	0.56169	0.99833	0.81779	0.78742	0.85299	0.83219	0.97224	0.99826	0.99253
T	9	0.18237	0.55278	0.7188	0.82677	0.56298	0.56668	0.85359	0.9929	0.99286	0.99072	0.93918	0.92546	0.99998	0.99354	0.96942
I	10	0.26216	0.7836	0.83025	0.7255	0.97595	0.59657	0.71772	0.99365	0.98804	1	0.98177	0.99752	0.99582	0.99752	0.98876
T	11	0.30955	0.66851	0.67676	0.58464	0.86347	0.58679	0.75586	0.99561	0.99909	0.9768	0.99685	0.96006	0.99044	0.98667	0.98556
I	12	0.16258	0.29692	0.42552	0.65732	0.44522	0.47573	0.57568	0.8724	0.8415	0.717	0.83067	0.78153	0.97845	0.97118	0.99903
0	13	0.34155	0.73569	0.79209	0.6387	0.72324	0.68759	0.81747	0.99767	0.99003	0.99959	0.93957	0.9572	0.99274	0.99926	0.98864
N	14	0.2081	0.36326	0.4669	0.55583	0.89278	0.51827	0.62375	0.97252	0.96657	0.91956	0.81419	0.91973	0.95924	0.9983	0.98484
	15	0.2313	0.44402	0.50278	0.56593	0.83868	0.61596	0.5866	0.70723	0.70993	0.77987	0.81475	0.86844	0.95349	0.98122	0.97859
	16	0.20427	0.32548	0.40105	0.74107	0.43038	0.45666	0.66456	0.97993	0.99286	0.83003	0.92456	0.9195	0.99696	0.8366	0.9226
	17	0.20887	0.52884	0.5172	0.58758	0.60162	0.48341	0.68944	0.99054	0.96284	0.99054	0.89262	0.87229	0.85745	0.96075	0.89852
	18	0.19108	0.29346	0.48074	0.70541	0.70954	0.42159	0.61475	0.80388	0.90267	0.93284	0.8537	0.78577	0.90124	0.99232	0.98851
	19	0.22673	0.50126	0.65667	0.65752	0.71862	0.65266	0.70565	0.9658	0.96646	0.93736	0.81011	0.89059	0.94293	0.97374	0.97639
	20	0.21005	0.31692	0.51311	0.48972	0.72043	0.59352	0.70252	0.96067	0.97397	0.8353	0.83265	0.77968	0.94695	0.99483	0.92247
	21	0.20222	0.34049	0.45621	0.70182	0.46745	0.53493	0.69409	0.99862	0.82375	0.82411	0.88715	0.80399	0.99367	0.99977	0.9615
	22	0.22407	0.4097	0.45131	0.41754	0.52854	0.46599	0.90636	0.99564	0.99942	0.97852	0.98005	0.88285	0.99461	0.98514	0.99899
	23	0.13956	0.32507	0.54934	0.65429	0.65194	0.41844	0.4789	0.9782	0.91277	0.81539	0.97601	0.82881	0.97914	0.94399	0.98705
	24	0.14116	0.33025	0.36454	0.59814	0.70757	0.43313	0.58205	0.98598	0.98241	0.71515	0.70802	0.70587	0.99658	0.95089	0.98225
	25	0.17532	0.34371	0.33918	0.56063	0.77079	0.45236	0.62794	0.99981	0.88503	0.88945	0.92382	0.90796	0.98586	0.98948	0.99125

Table B.3: We provide the gap value for each and every of 25 partitions and each common subspace dimension from 1 to 15.

		COMMON SUBSPACE DIMENSION														
		16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	1	0.99072	0.98453	0.87576	0.97099	0.9973	0.96603	0.96632	1	0.99992	0.99361	0.9632	0.98096	0.99517	0.99782	0.99977
	2	0.95465	0.87377	0.82868	0.97159	0.98529	0.90227	0.99788	0.98423	0.97379	0.97436	0.94681	0.98286	0.99941	0.99958	0.99943
	3	0.99175	0.86504	0.72387	0.97637	0.99957	0.86244	0.98107	0.99995	0.97773	0.97333	0.9557	0.95237	0.98391	0.97827	0.99998
	4	0.98896	0.88575	0.8642	0.99465	0.99976	0.97862	0.98274	0.99969	0.99286	0.99694	0.9598	0.96142	0.99049	0.99653	0.99582
	5	0.87117	0.85545	0.86343	0.99498	0.98835	0.99324	0.99543	0.99568	0.98995	0.99548	0.99405	0.99854	0.99775	0.99987	0.99993
P	6	0.99989	0.99891	0.94429	0.98375	0.98603	0.95176	0.97476	0.99702	0.98685	0.99998	0.97857	0.99179	0.96582	0.98334	0.97419
A	7	0.99047	0.97954	0.90183	0.99815	0.99958	0.99546	0.96234	0.97715	0.99993	0.99339	0.99547	0.99662	0.99978	0.98701	0.99178
R	8	0.99846	0.95871	0.98353	0.9992	0.99889	0.90344	0.88179	0.97899	0.99667	0.99978	0.98624	0.99505	0.99834	0.98054	0.99868
T	9	0.99878	0.99976	0.99696	0.99686	0.99945	0.97626	0.99488	0.9991	0.9992	0.99122	0.99839	0.98834	0.99535	0.99609	0.98028
I	10	0.99893	0.99746	0.97653	0.99901	0.99974	0.99168	0.98593	0.9861	0.99728	0.99982	0.97213	0.9743	0.99315	0.99954	0.99957
T	11	0.99592	0.97716	0.9834	0.99999	0.97361	0.93772	0.9993	0.99547	0.99645	1	0.99916	0.9991	0.9947	0.99819	0.99862
I	12	0.9953	0.94816	0.85381	0.9988	0.99429	0.99866	0.9925	0.98335	0.99675	0.99998	0.97896	0.95426	0.98808	0.99624	0.99995
0	13	0.99604	0.99303	0.93323	0.99109	0.96204	0.9459	0.97848	0.97382	0.9482	0.98981	0.99995	0.99621	0.9988	0.99844	0.9958
N	14	0.99473	0.99087	0.97206	0.98555	0.98731	0.91848	0.9948	0.99361	0.99997	0.98736	0.98553	1	0.99707	0.9986	0.98726
	15	0.94433	0.88601	0.79914	0.96693	0.95048	0.89987	0.9324	0.9834	0.99984	0.99016	0.9873	0.99986	0.98012	0.99409	0.99716
	16	0.99999	0.84806	0.8175	0.99704	0.99752	0.94984	0.9597	0.98779	0.99997	0.99409	0.96556	0.97987	0.99602	0.99432	0.99739
	17	0.99588	0.83344	0.75699	0.9336	0.97991	0.98823	0.95464	0.99414	0.99936	0.97672	0.93596	0.99088	0.99349	0.9999	0.99677
	18	0.94015	0.91225	0.85347	0.99625	0.97851	0.84643	0.98299	0.96594	0.96628	0.99996	0.97809	0.99732	0.99843	0.98476	0.99964
	19	0.99103	0.99927	0.85658	0.99988	0.99994	0.97138	0.96777	0.99269	0.97825	0.9518	0.93827	0.9966	0.99446	0.99957	0.99937
	20	0.93129	0.77777	0.85424	0.97168	0.9813	0.95343	0.98785	0.99704	0.99639	0.99956	0.9947	0.99726	0.98706	0.9946	0.99414
	21	0.99268	0.98176	0.99261	0.9936	0.99328	0.98019	0.99964	0.99638	0.97817	0.98754	0.95967	0.99565	0.99108	0.99797	0.99681
	22	0.99885	0.88359	0.80848	0.99517	0.99766	0.93982	0.96527	0.98818	0.99852	0.99964	0.99966	0.98895	0.99215	1	0.98369
	23	0.99888	0.89649	0.79822	0.99941	0.99899	0.83791	0.99861	0.99426	0.99412	0.99725	0.95725	0.97822	0.99926	0.98788	0.99954
	24	0.98619	0.95026	0.93411	0.97407	0.99841	0.92143	0.87917	0.97551	0.99985	0.99508	0.88788	0.99276	0.99563	0.99976	0.97885
	25	0.9933	0.9011	0.8476	0.99087	0.99644	0.95149	0.94849	0.99175	0.98257	0.97095	0.98333	0.99608	0.99983	0.98108	0.991

Table B.4: We provide the gap value for each and every of 25 partitions and each common subspace dimension from 16 to 30.

		COMMON SUBSPACE DIMENSION														
		31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	1	0.99995	0.99973	0.99971	0.97702	0.97617	0.98925	0.99524	0.99022	0.99995	0.9975	0.99793	0.99995	0.99999	0.99836	0.99717
	2	0.99927	0.99897	0.99896	0.99997	0.99954	0.99656	0.99681	0.98021	0.97968	0.98067	0.99144	0.99947	0.99997	0.99814	1
	3	0.9936	0.99996	0.99987	0.99963	0.9991	0.99993	0.99453	0.99079	0.99316	0.99796	0.99821	0.99646	0.99937	0.99822	0.99982
	4	0.99776	0.9951	0.998	0.99862	0.99995	0.99826	0.99967	0.97492	0.98898	0.995	0.99997	0.99868	0.99989	0.99998	0.99958
	5	0.99523	0.99739	0.99841	0.99476	0.99872	0.96307	0.99854	0.99999	0.9859	0.99986	0.9996	0.9966	0.99725	0.99868	0.99965
P	6	0.99934	0.99932	0.98847	0.9938	0.99545	0.99933	0.98103	0.98069	0.99876	0.99981	0.99959	0.99935	1	0.99896	0.99808
A	7	0.99073	0.99217	0.99516	0.9959	0.99513	0.99809	0.99934	0.99999	0.99977	0.99999	0.99961	0.99999	0.99933	0.99926	0.9999
R	8	0.99588	0.98199	0.99731	0.97531	0.97484	0.98677	0.96353	0.96147	0.98401	0.98512	0.97743	0.99981	0.99996	0.9987	1
T	9	0.99452	0.9996	0.9899	0.99777	0.99818	0.99934	0.98843	0.99384	0.99624	0.99823	0.99891	0.99971	0.99875	0.9999	0.99991
I	10	0.98856	0.99849	0.99973	1	0.99997	1	0.99554	0.9939	0.99993	0.99918	0.99879	0.99798	0.99718	0.99808	0.99991
T	11	0.99984	0.99889	0.99056	0.9966	0.99659	0.9992	0.99716	0.99715	0.99825	0.99776	0.99929	0.99988	0.9965	0.99757	0.99988
I	12	0.99933	0.99969	0.98929	0.98866	0.9974	0.99331	0.99414	0.99381	0.98943	0.99668	0.99809	0.99873	0.99706	0.99985	0.99915
0	13	0.9984	0.99996	0.99618	0.99578	0.99844	0.99949	0.9989	0.99421	0.99708	0.99841	0.99999	0.99915	0.99997	0.99951	0.99961
N	14	0.99925	0.99999	0.99191	0.99999	0.99977	0.99583	0.99453	0.99315	0.98983	0.98388	0.99898	0.99986	0.99986	0.99994	0.99998
	15	0.99474	0.99218	0.99252	1	0.99999	0.99501	0.99755	0.99768	0.99997	0.99647	0.9995	0.99804	0.99987	0.99992	0.99981
	16	0.99807	0.99517	0.98152	0.98286	0.99681	0.98654	0.99991	0.99638	0.99035	1	0.99027	0.99789	0.99942	0.99936	0.99985
	17	0.99987	1	0.97727	0.99808	0.99729	0.99756	0.99985	0.98512	0.99938	0.99979	0.99396	0.99772	0.99327	0.99747	0.99485
	18	0.99661	0.98365	0.9793	0.98429	0.99442	0.99801	0.99995	0.99803	0.99136	0.99923	0.99063	0.9973	0.9982	0.9982	0.9993
	19	0.99898	0.99241	0.99905	0.99817	0.99702	0.9992	0.99695	0.99133	0.99075	0.987	0.98824	0.99724	0.99847	0.99908	0.99988
	20	0.99909	0.99824	0.99945	1	0.9999	0.99748	0.99579	1	0.99988	0.99198	0.99555	0.99429	0.99859	0.99997	0.99985
	21	0.99983	0.99972	0.99824	0.99327	0.99995	0.99967	0.9991	0.99757	0.98994	0.98949	0.99997	0.99995	0.99957	0.99867	0.99995
	22	0.9837	0.98753	0.99801	0.98287	0.98887	0.9858	0.99956	0.99734	0.99552	0.99913	0.99178	0.99214	0.99868	0.99828	0.99901
	23	0.9987	0.99827	0.98444	0.99844	0.99987	0.99838	0.99909	0.99351	0.9948	0.9995	0.99288	0.98444	0.99972	0.99982	0.99999
	24	0.99737	0.99868	0.99986	0.99953	0.99939	0.99227	0.99991	0.99578	0.99851	0.9998	0.99835	0.99947	0.9991	0.99898	0.99988
	25	0.97896	0.99994	0.99563	0.99996	0.99803	0.9845	0.99429	0.99717	0.99686	0.99907	0.9986	0.99994	0.99782	0.99882	0.99892

Table B.5: We provide the gap value for each and every of 25 partitions and each common subspace dimension from 31 to 45. For common subspace dimension greater than 42, all values of gap are larger than 0.99.

	COMMON SUBSPACE DIMENSION																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
common subspace dimension	36	40	29	39	39	38	29	41	37	31	21	39	25	40	28	36	38	34	41	28	40	36	42	30	36

Table B.6: We provide the gap value for each and every of 25 partitions and each common subspace dimension from 31 to 45. For common subspace dimension greater than 42, all values of gap are larger than 0.99.

B.2 Estimation of the common spatial factor A

1. Common subspace dimension $R = 27$

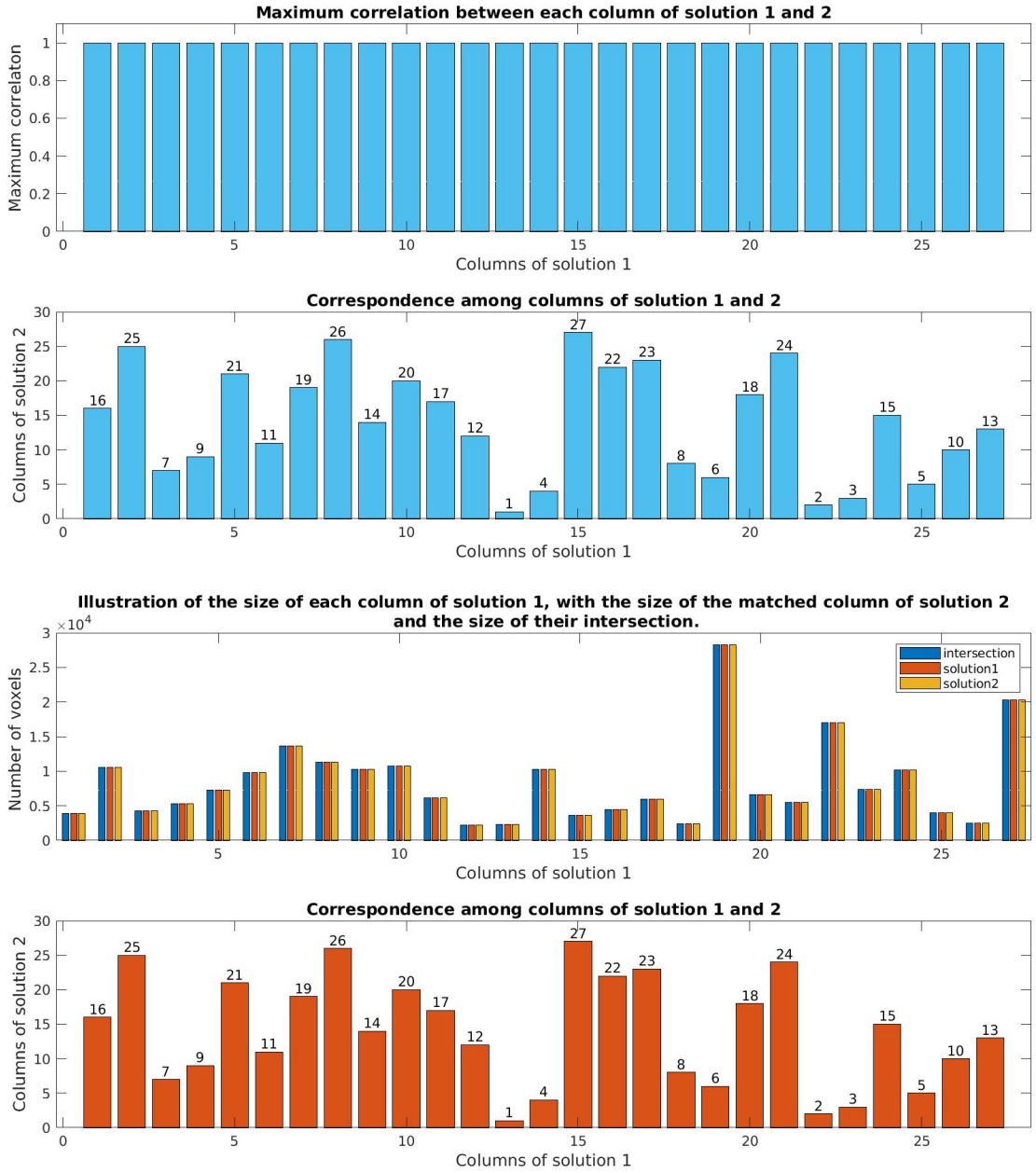


Figure B.1: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated based on intersection. Solutions 1, 2 correspond to accelerated penalty method and common subspace dimension equal to 27.

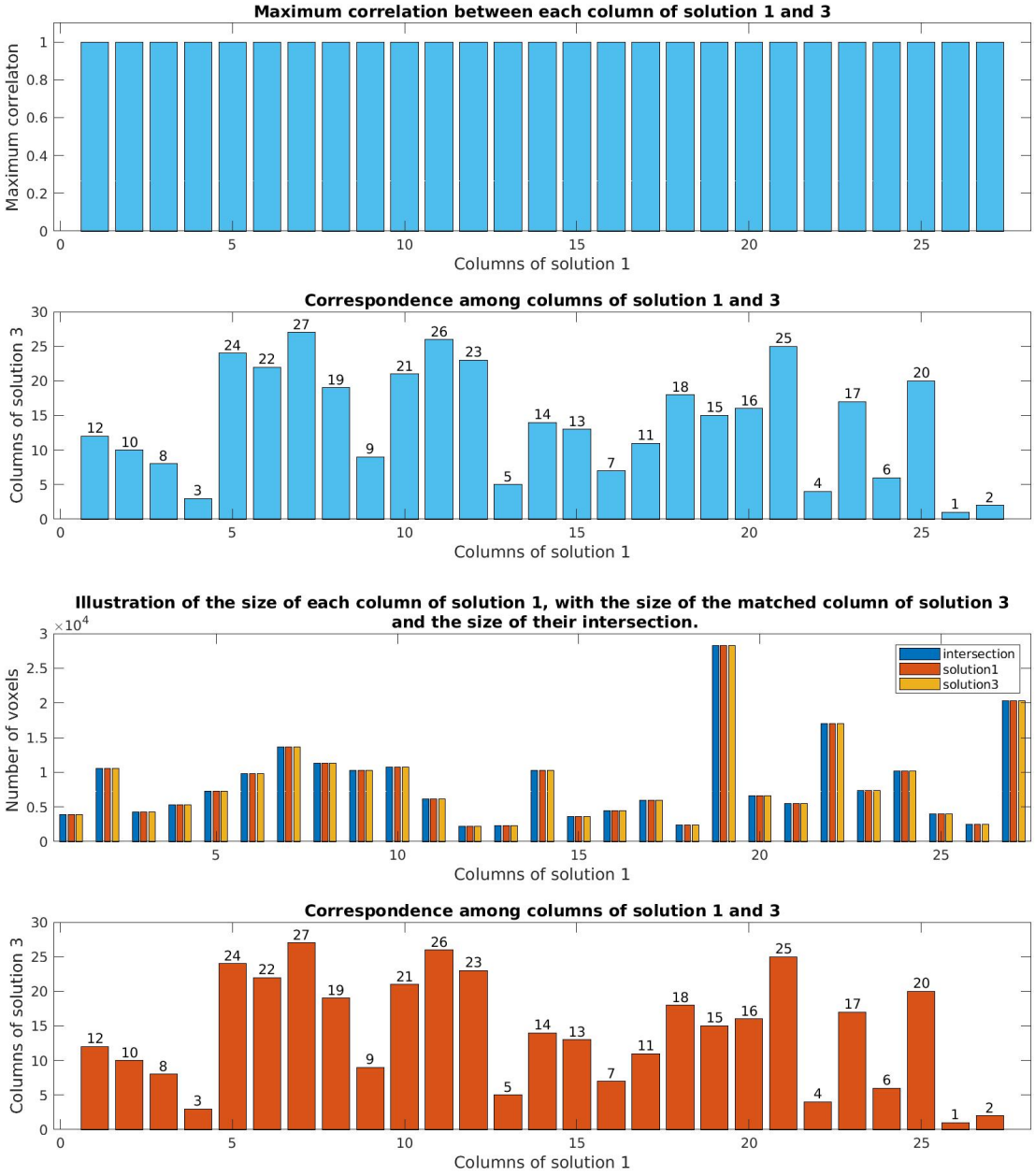


Figure B.2: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated based on intersection. Solutions 1, 3 correspond to accelerated penalty method and common subspace dimension equal to 27.

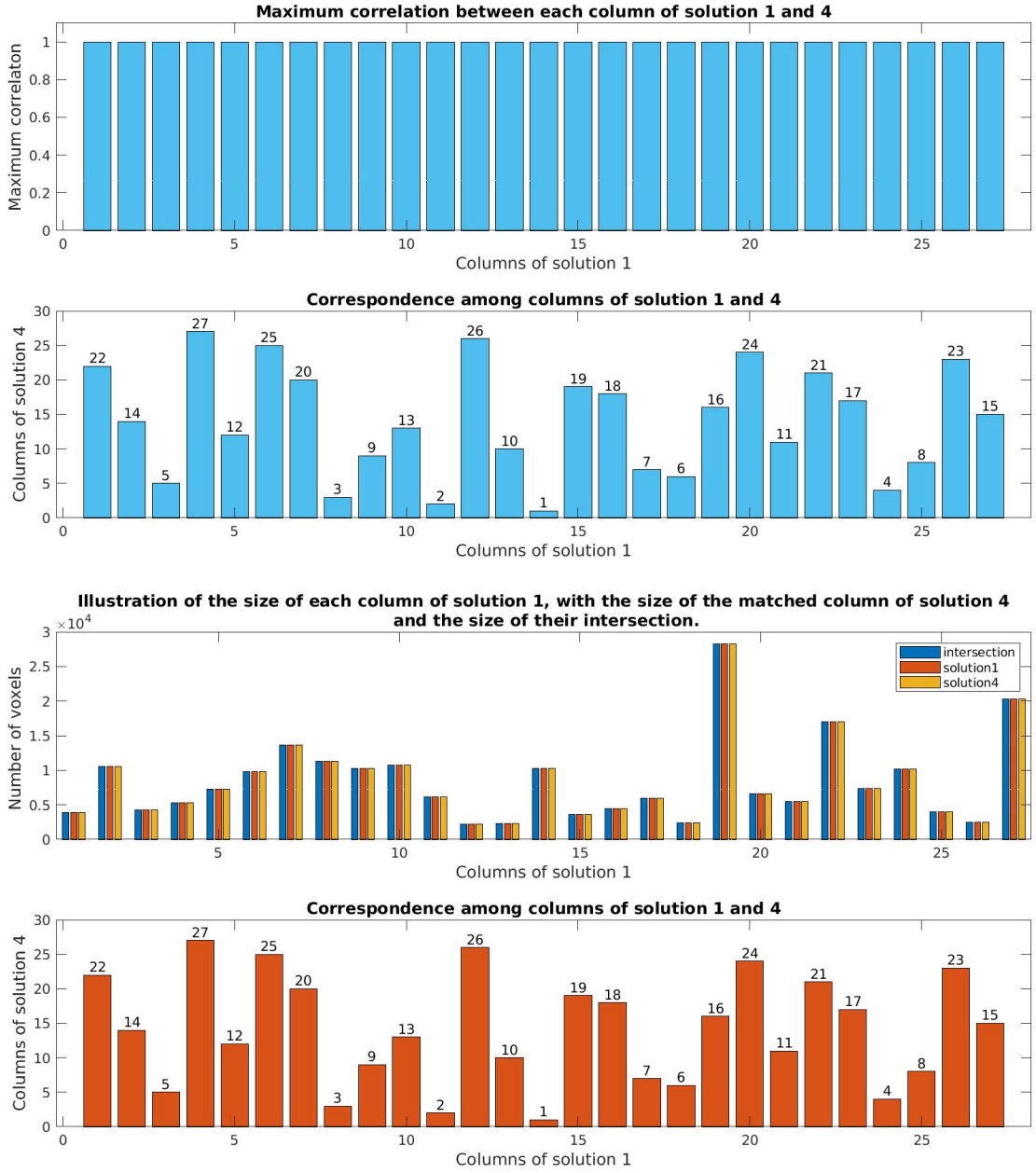


Figure B.3: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated based on intersection. Solutions 1, 4 correspond to accelerated penalty method and common subspace dimension equal to 27.

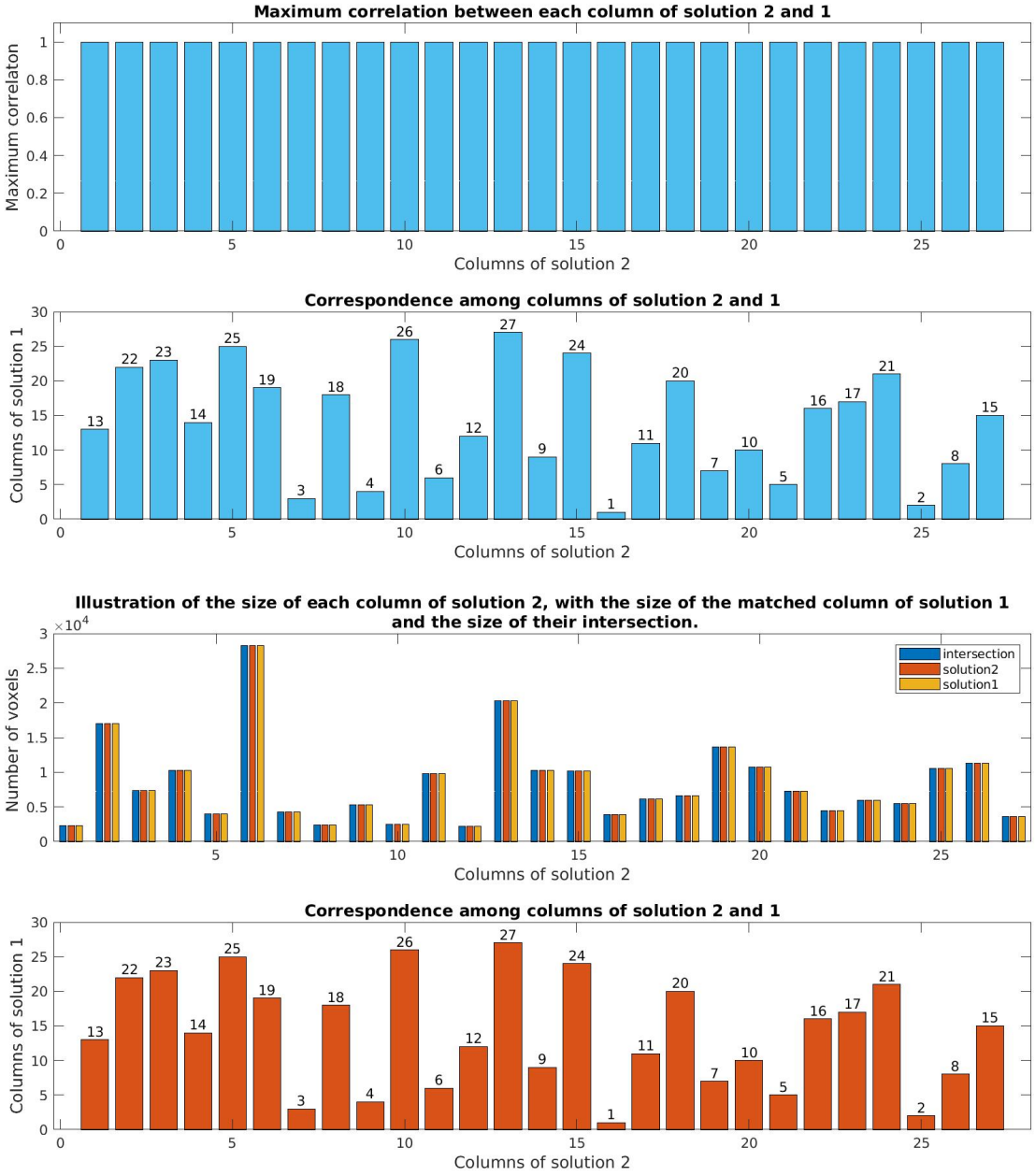


Figure B.4: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated based on intersection. Solutions 2, 1 correspond to accelerated penalty method and common subspace dimension equal to 27.

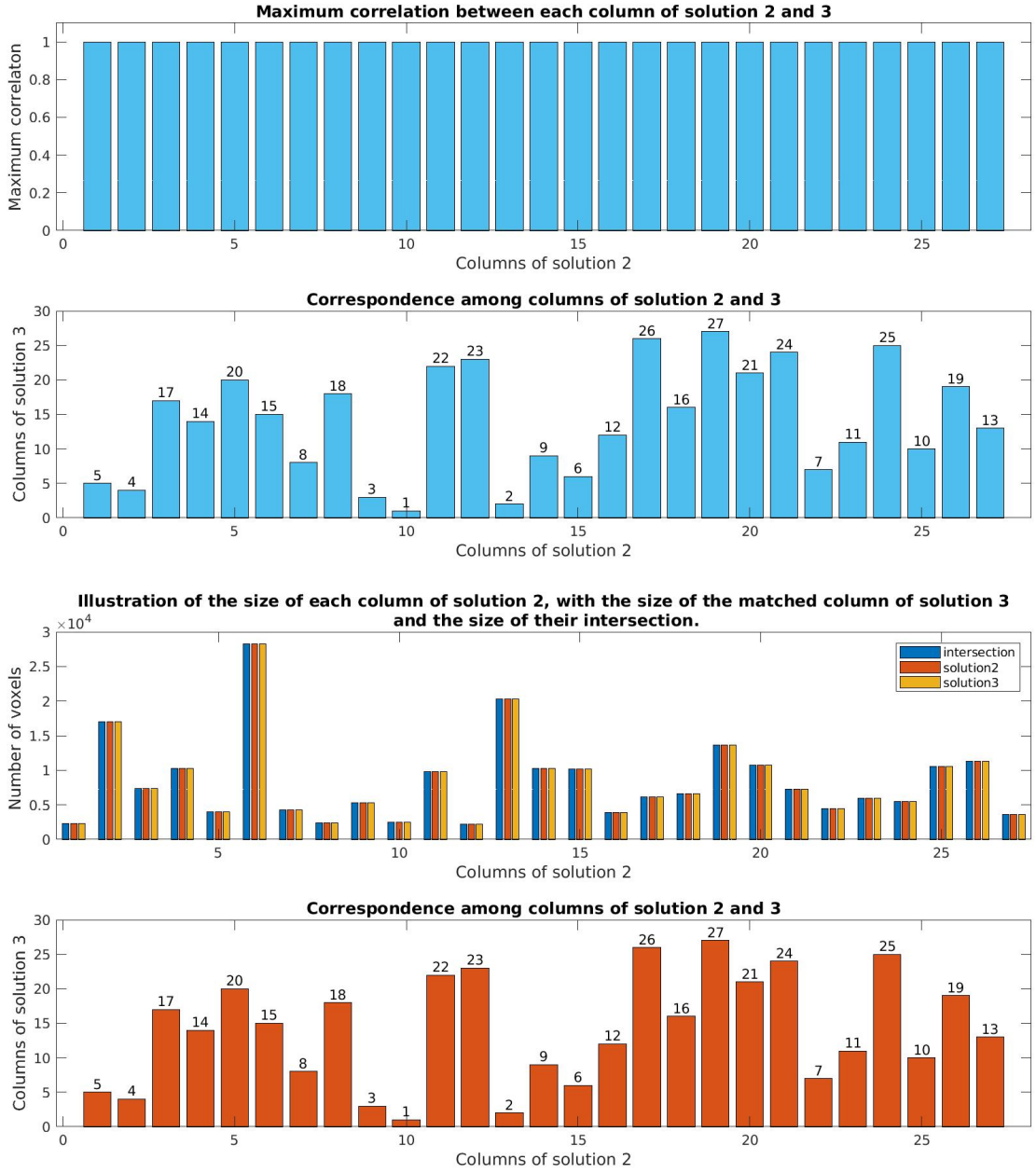


Figure B.5: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated based on intersection. Solutions 2, 3 correspond to accelerated penalty method and common subspace dimension equal to 27.

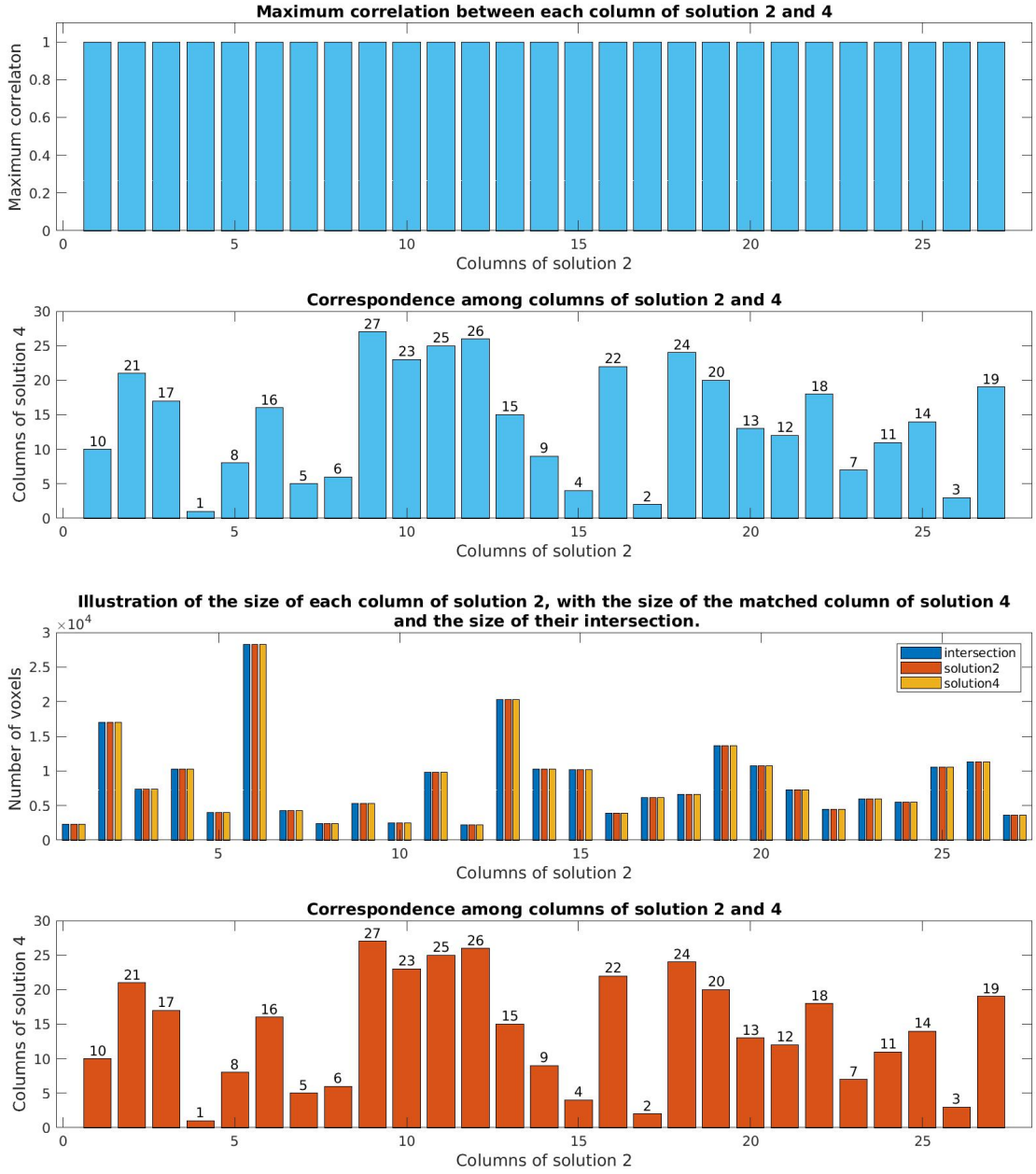


Figure B.6: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated based on intersection. Solutions 2, 4 correspond to accelerated penalty method and common subspace dimension equal to 27.

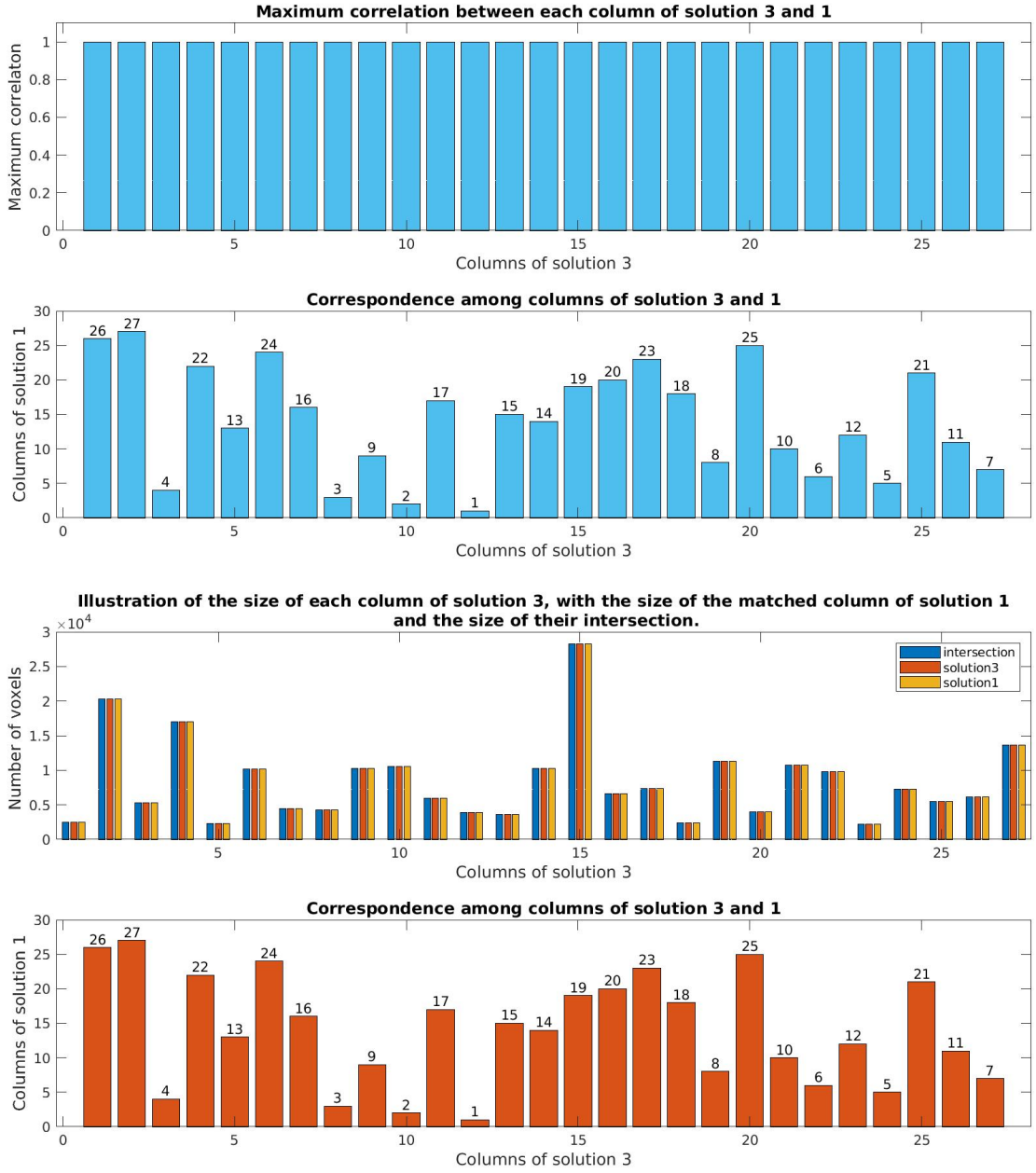


Figure B.7: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated based on intersection. Solutions 3, 1 correspond to accelerated penalty method and common subspace dimension equal to 27.

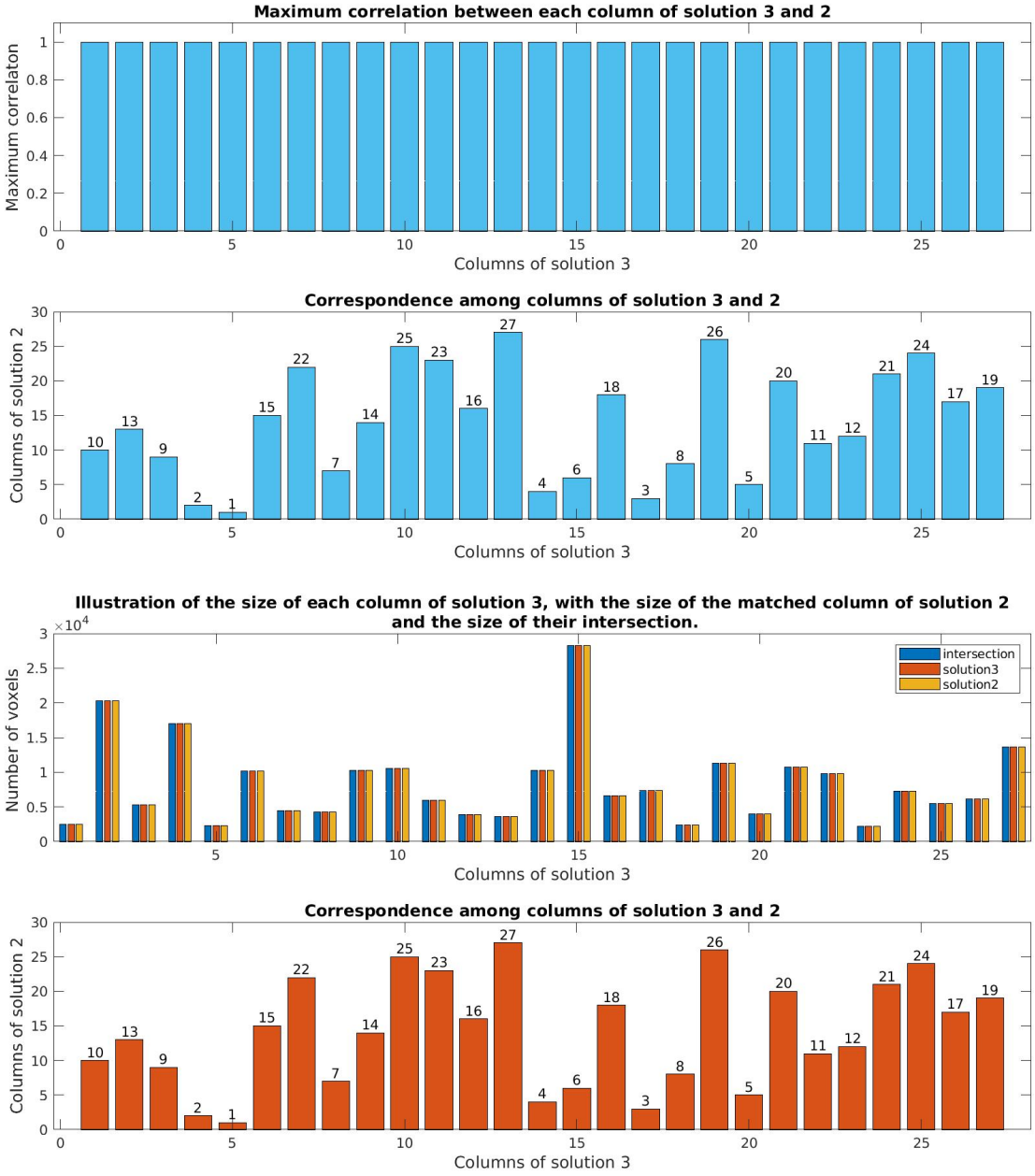


Figure B.8: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated based on intersection. Solutions 3, 2 correspond to accelerated penalty method and common subspace dimension equal to 27.

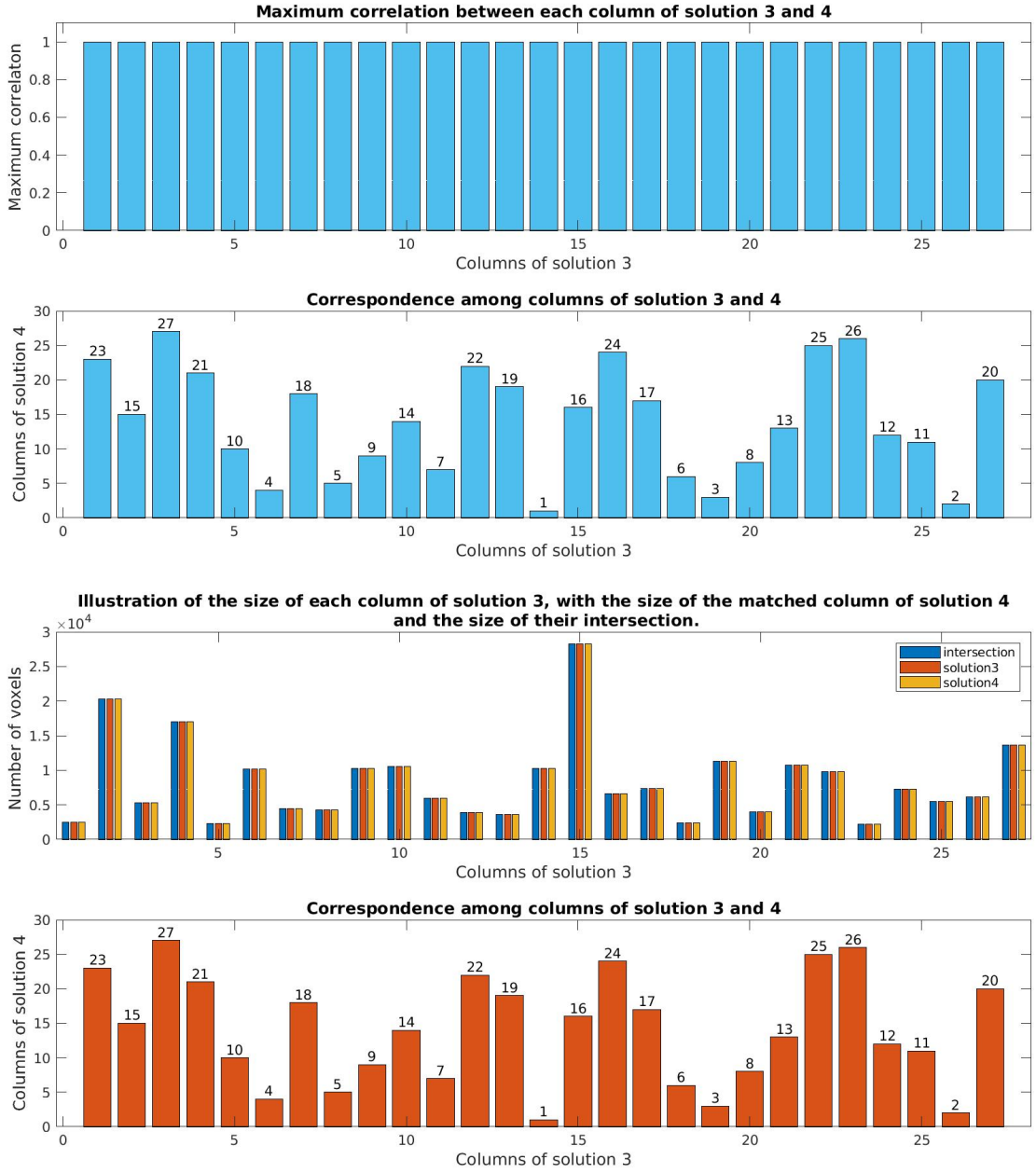


Figure B.9: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated based on intersection. Solutions 3, 4 correspond to accelerated penalty method and common subspace dimension equal to 27.

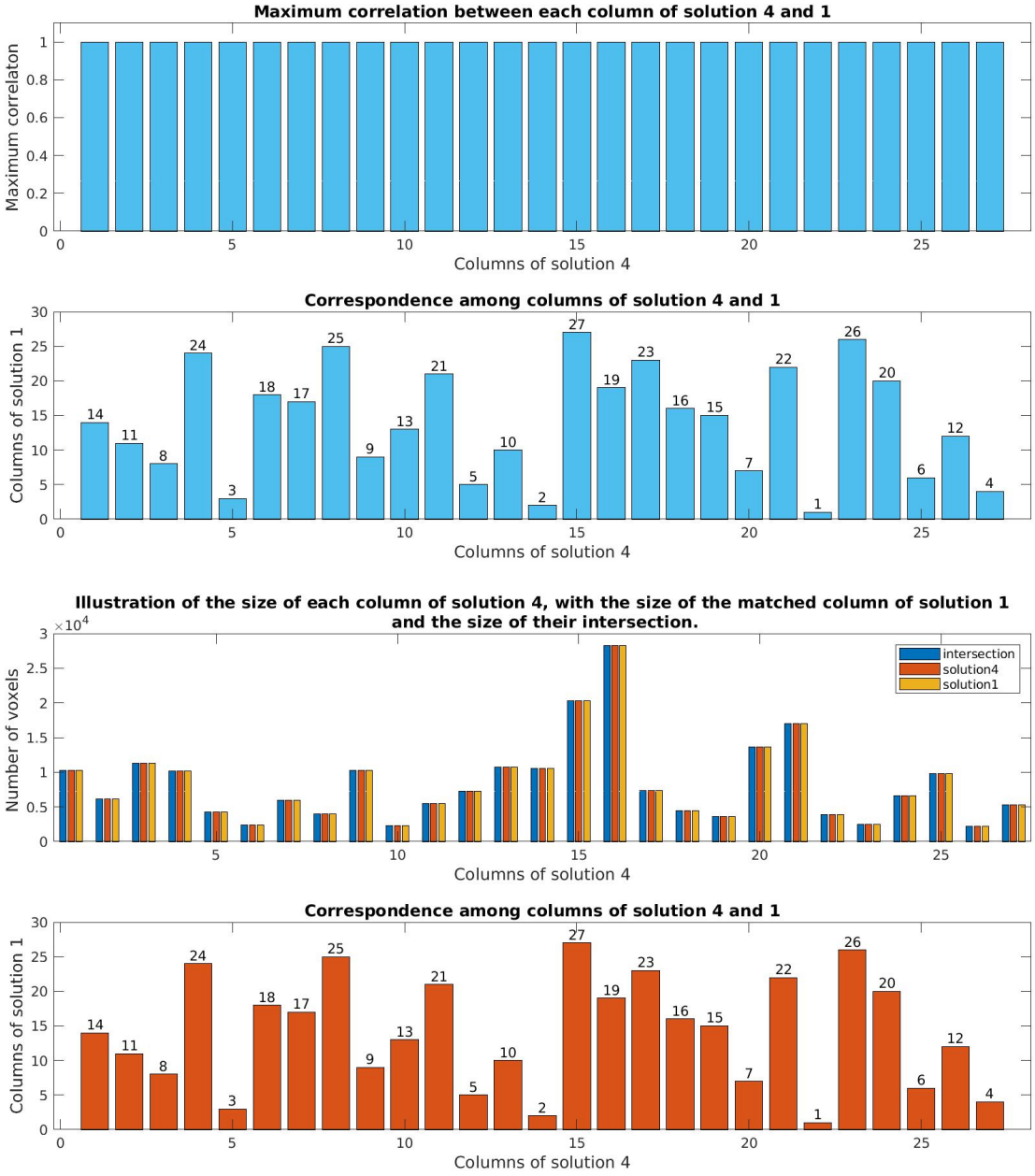


Figure B.10: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated based on intersection. Solutions 4, 1 correspond to accelerated penalty method and common subspace dimension equal to 27.

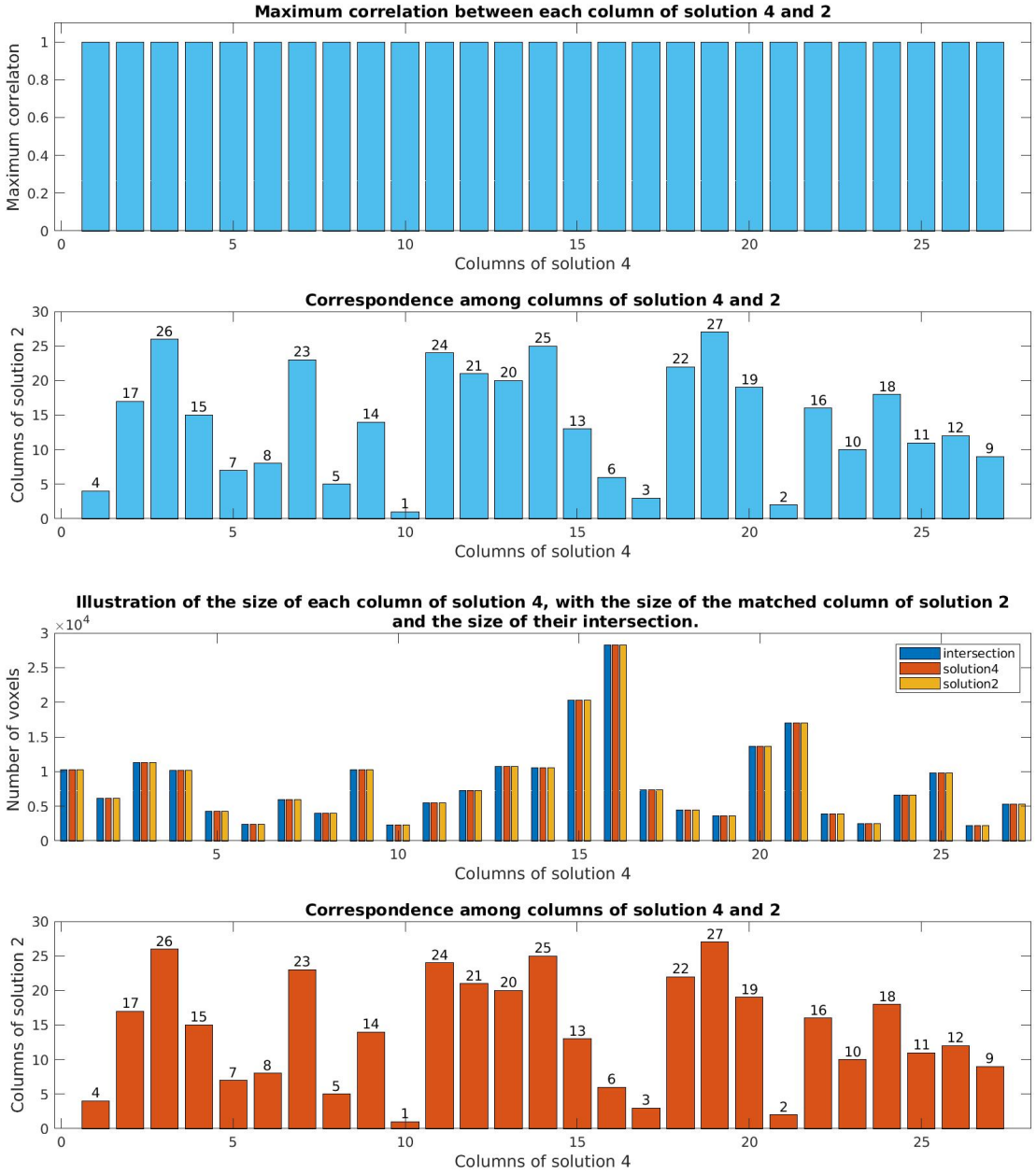


Figure B.11: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated based on intersection. Solutions 4, 2 correspond to accelerated penalty method and common subspace dimension equal to 27.

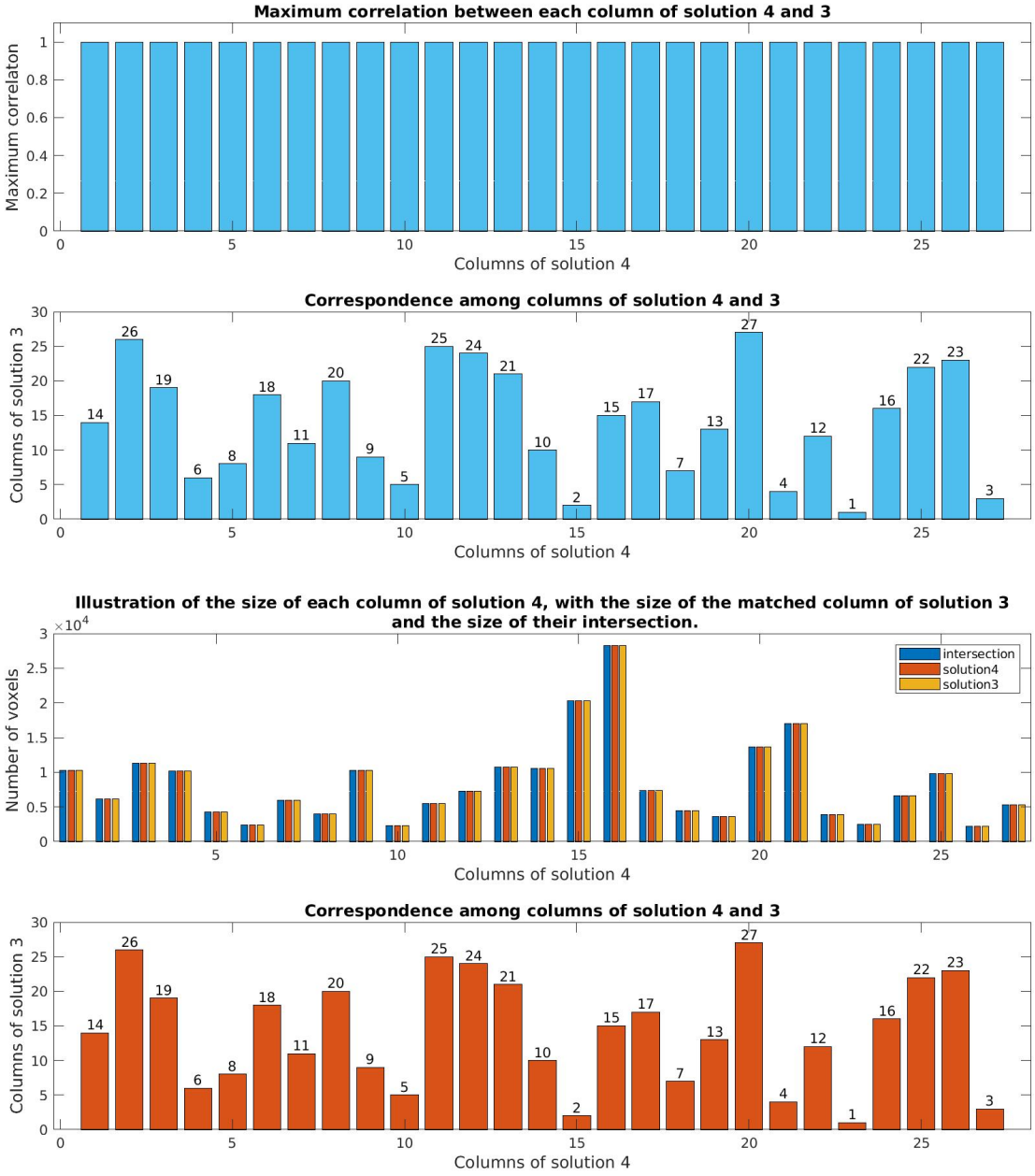


Figure B.12: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated based on intersection. Solutions 4, 3 correspond to accelerated penalty method and common subspace dimension equal to 27.

2. Common subspace dimension $R = 42$

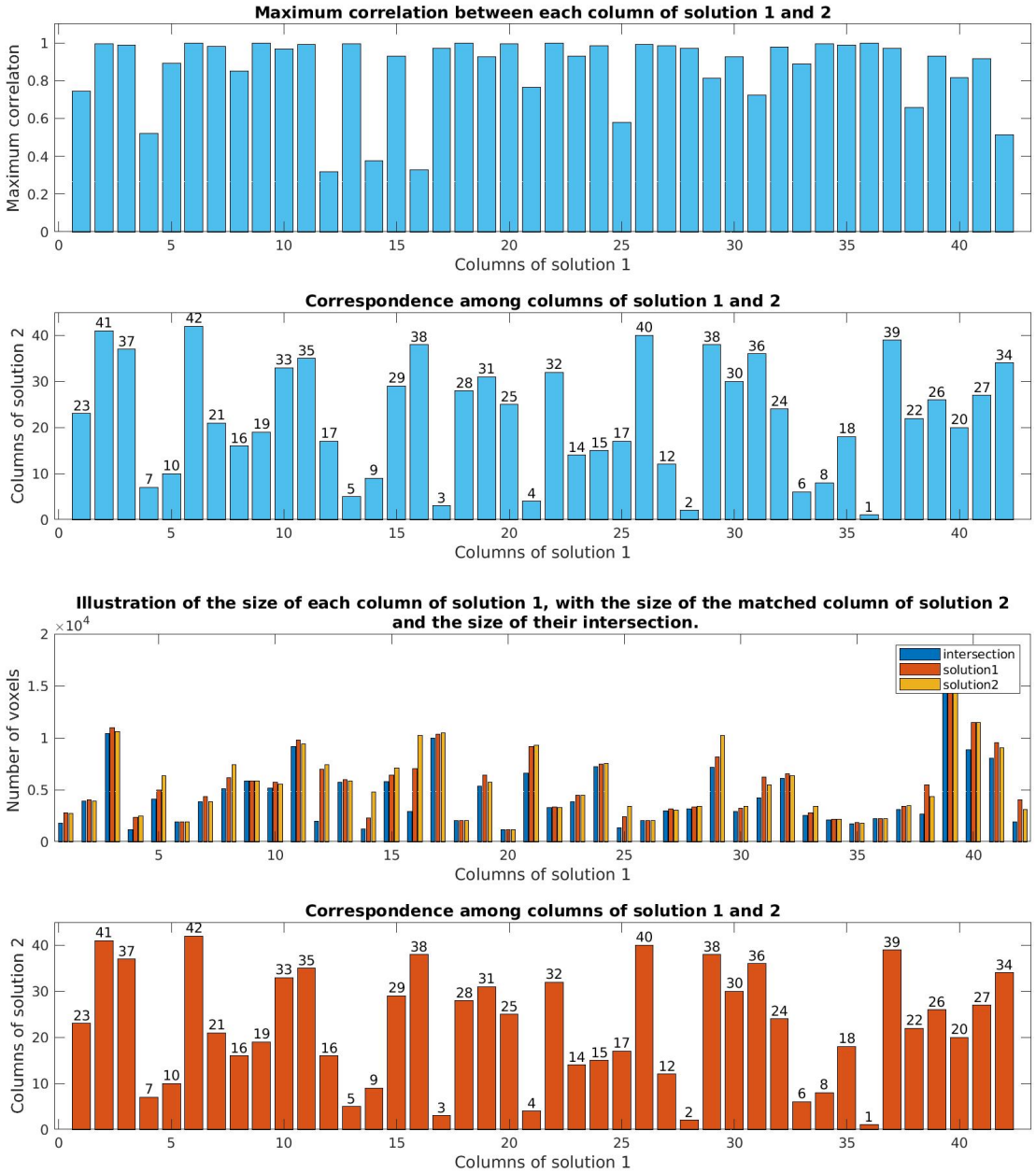


Figure B.13: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 2. We associate the column r of the solution 1 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 1 is associated based on intersection. Solutions 1, 2 correspond to accelerated penalty method and common subspace dimension equal to 42.

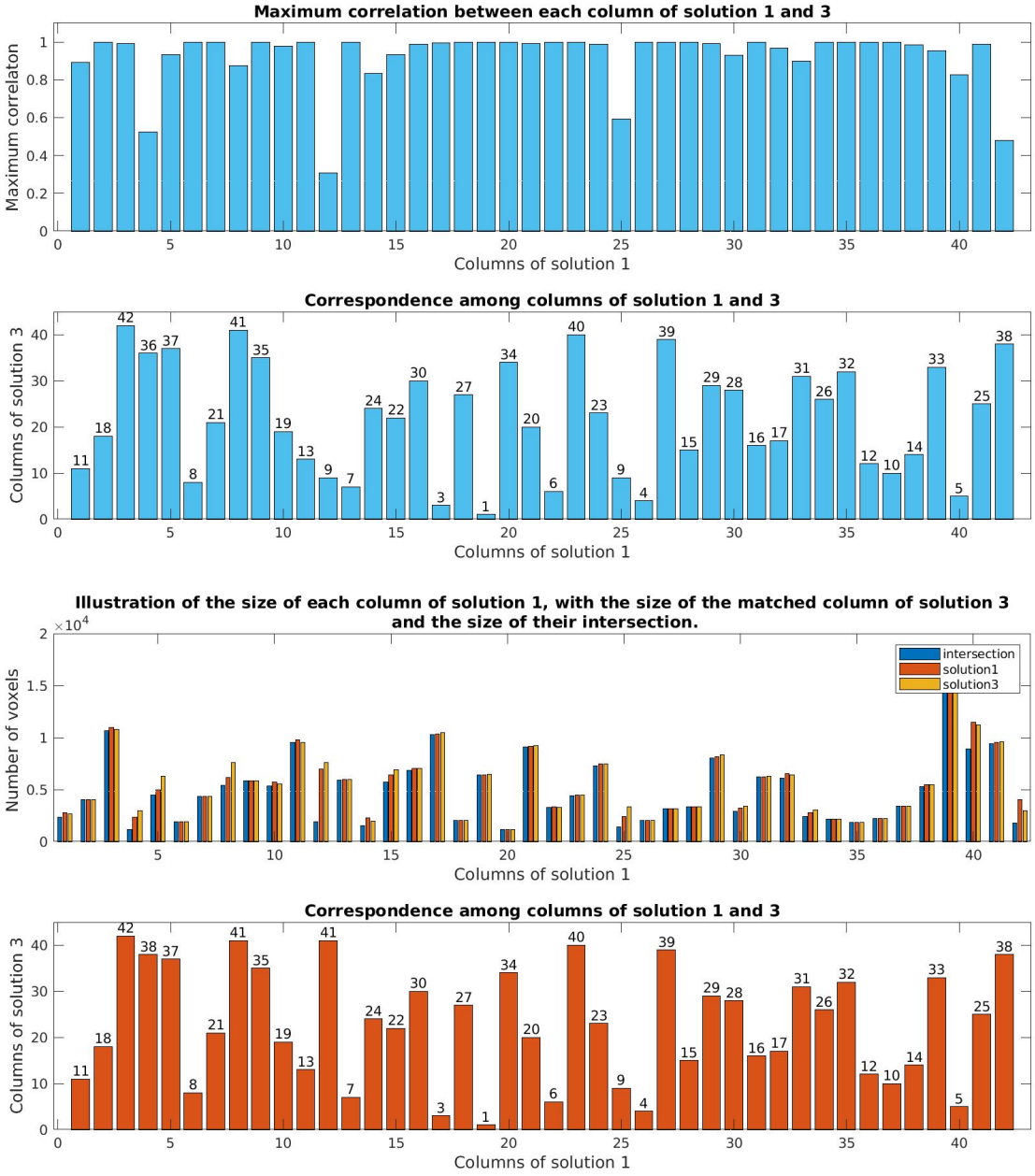


Figure B.14: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 3. We associate the column r of the solution 1 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 1 is associated based on intersection. Solutions 1, 3 correspond to accelerated penalty method and common subspace dimension equal to 42.

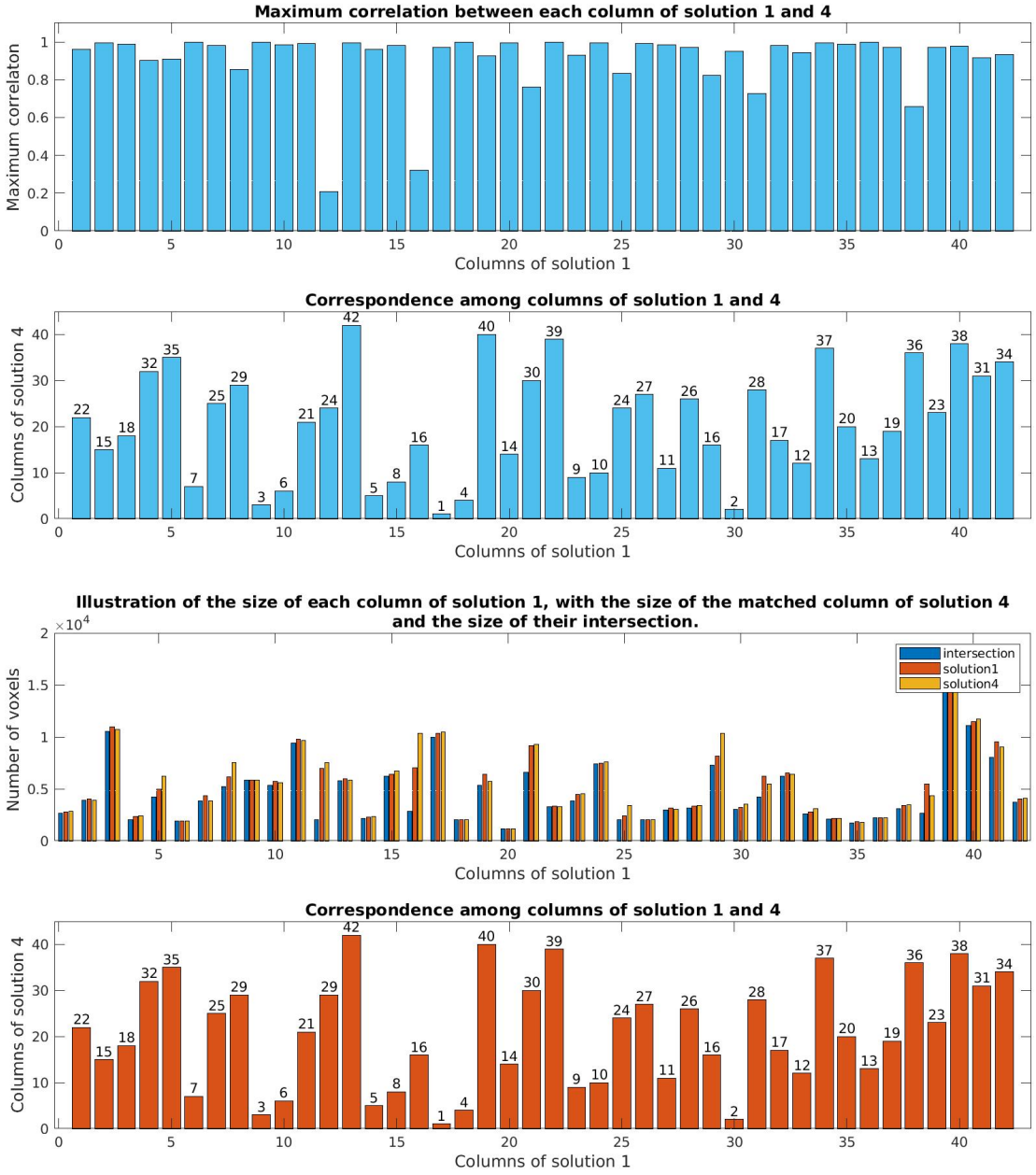


Figure B.15: **First figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 1 and each column of the solution 4. We associate the column r of the solution 1 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 1, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 1 is associated based on intersection. Solutions 1, 4 correspond to accelerated penalty method and common subspace dimension equal to 42.

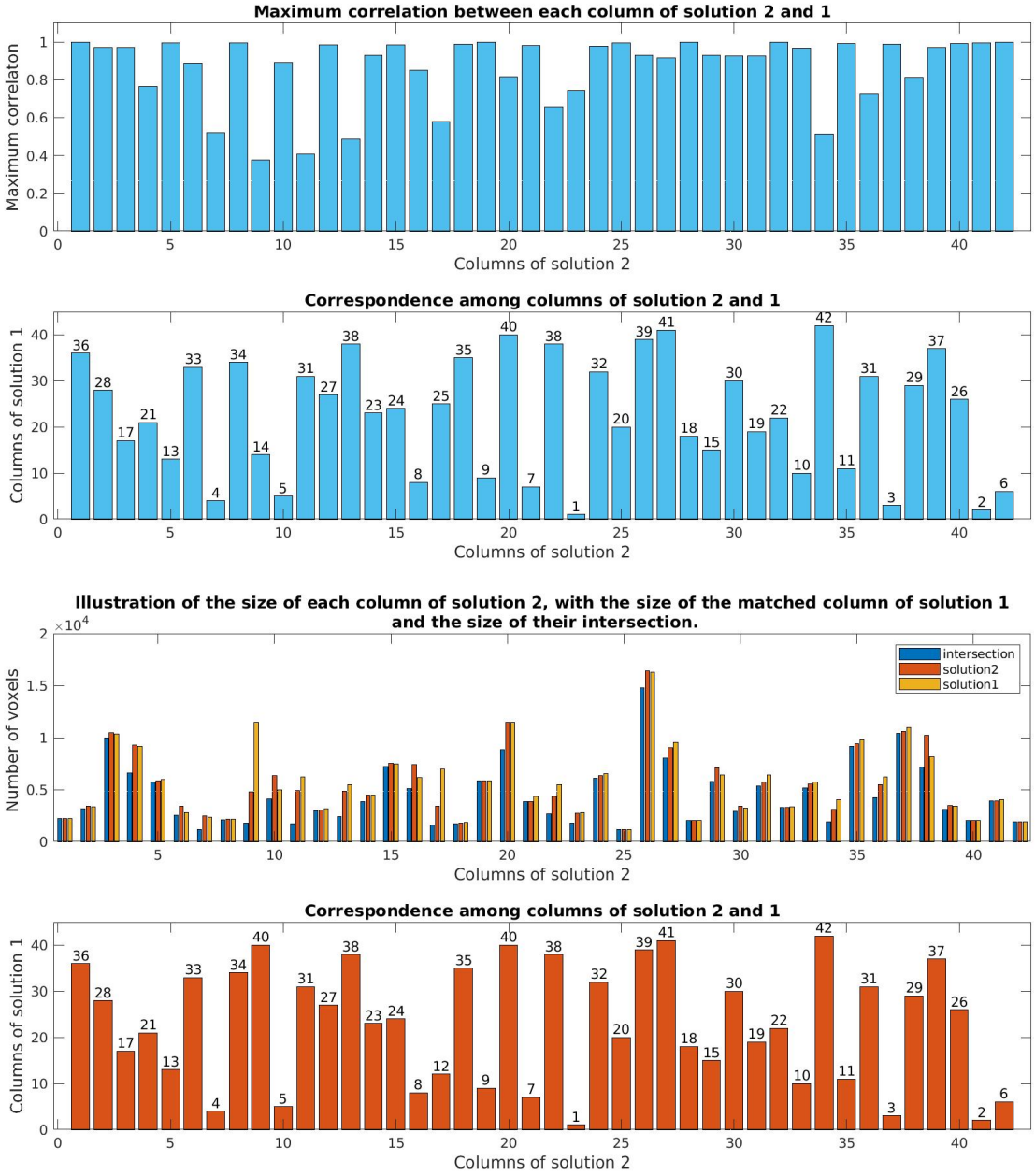


Figure B.16: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 1. We associate the column r of the solution 2 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 2 is associated based on intersection. Solutions 2, 1 correspond to accelerated penalty method and common subspace dimension equal to 42.

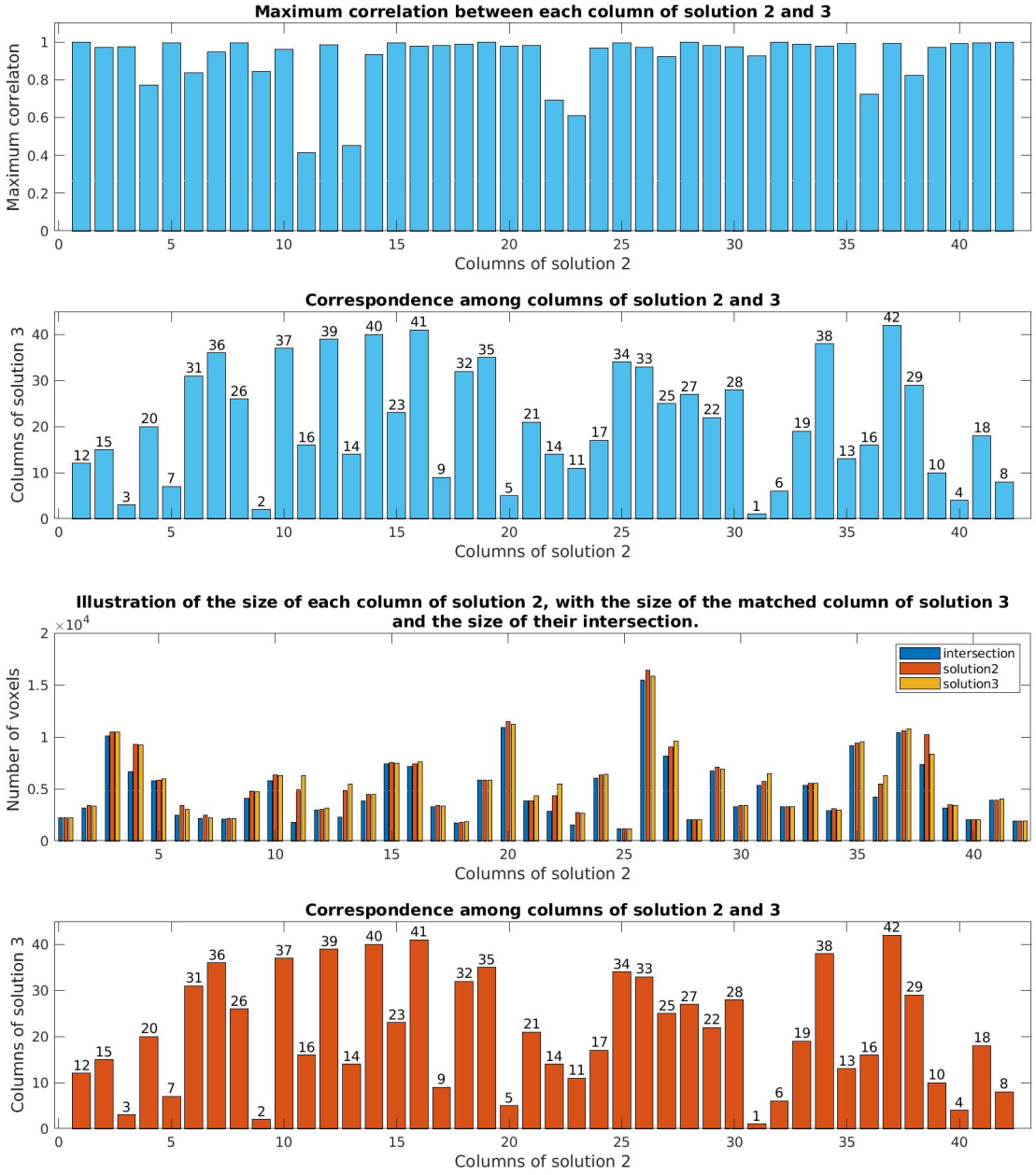


Figure B.17: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 3. We associate the column r of the solution 2 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 2 is associated based on intersection. Solutions 2, 3 correspond to accelerated penalty method and common subspace dimension equal to 42.

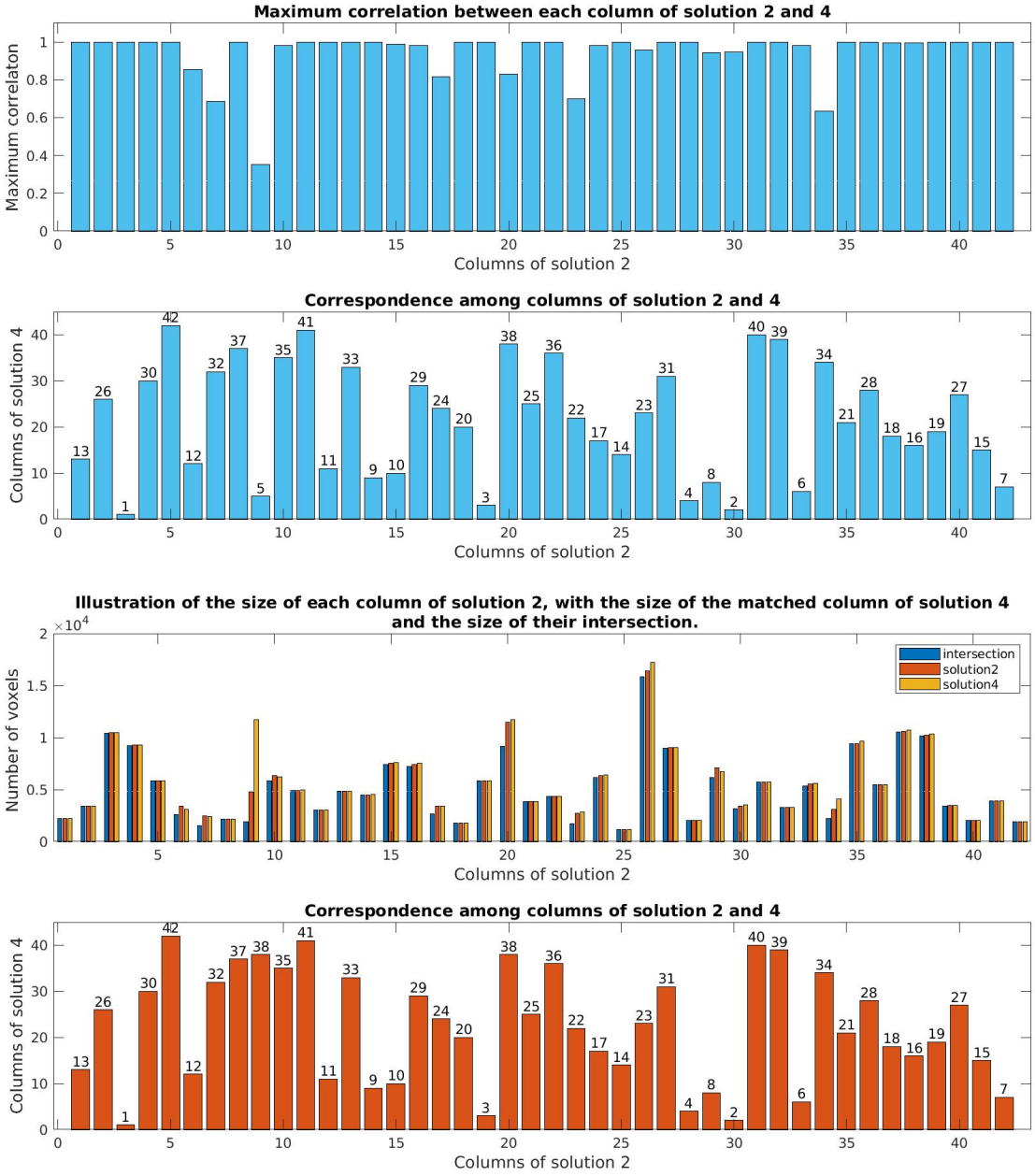


Figure B.18: **First figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 2 and each column of the solution 4. We associate the column r of the solution 2 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 2, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 2 is associated based on intersection. Solutions 2, 4 correspond to accelerated penalty method and common subspace dimension equal to 42.

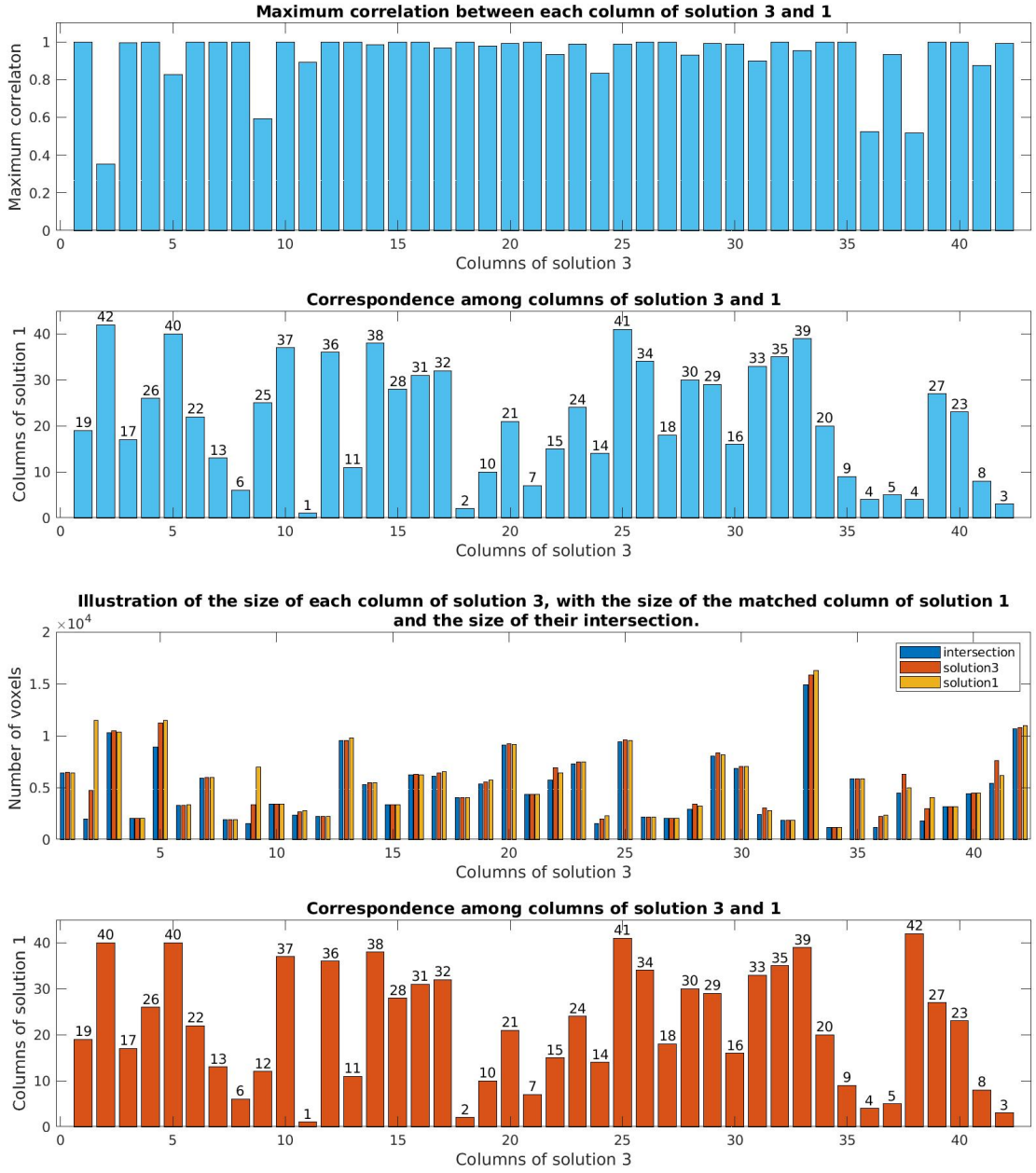


Figure B.19: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 1. We associate the column r of the solution 3 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 3 is associated based on intersection. Solutions 3, 1 correspond to accelerated penalty method and common subspace dimension equal to 42.

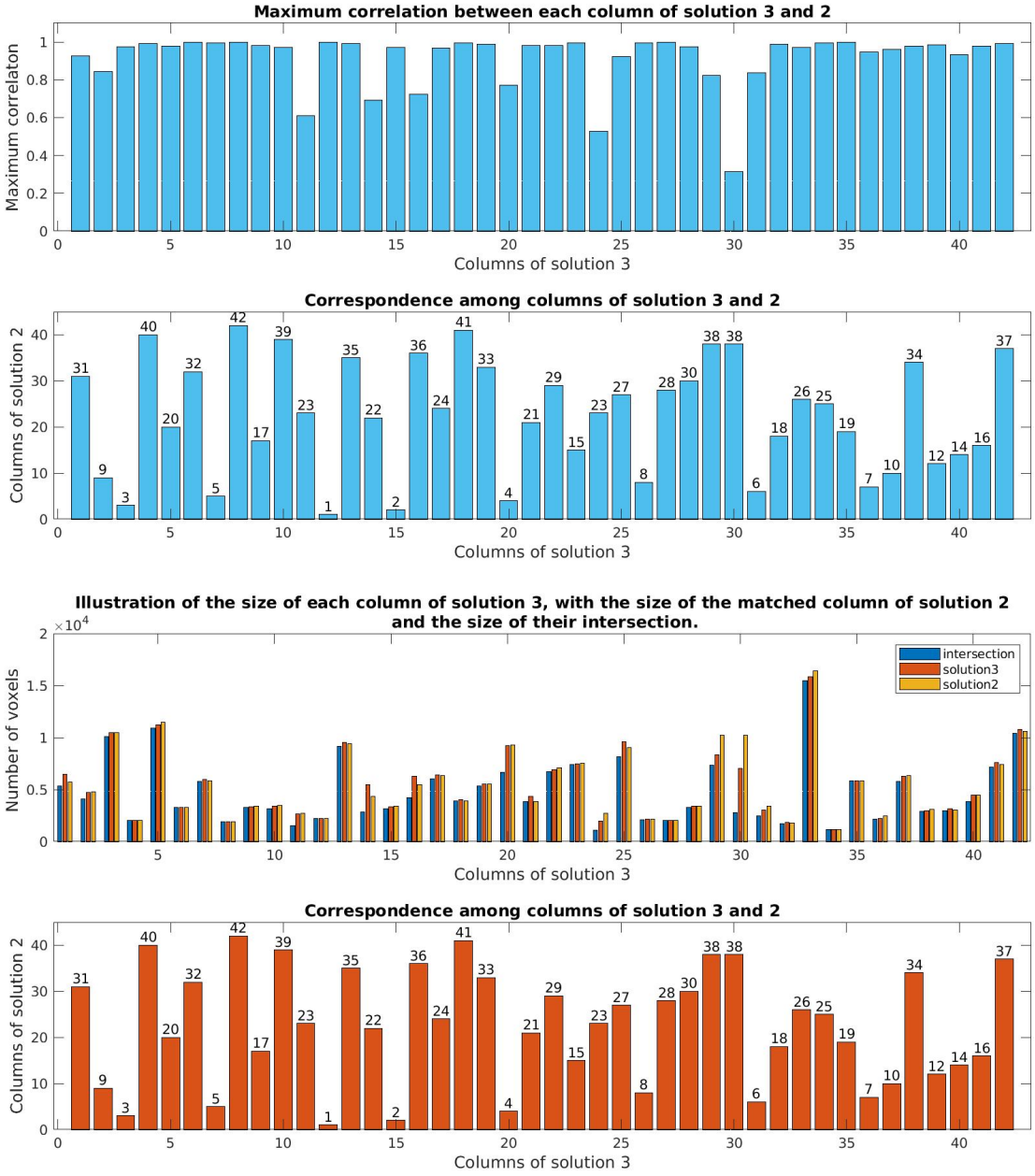


Figure B.20: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 2. We associate the column r of the solution 3 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 3 is associated based on intersection. Solutions 3, 2 correspond to accelerated penalty method and common subspace dimension equal to 42.

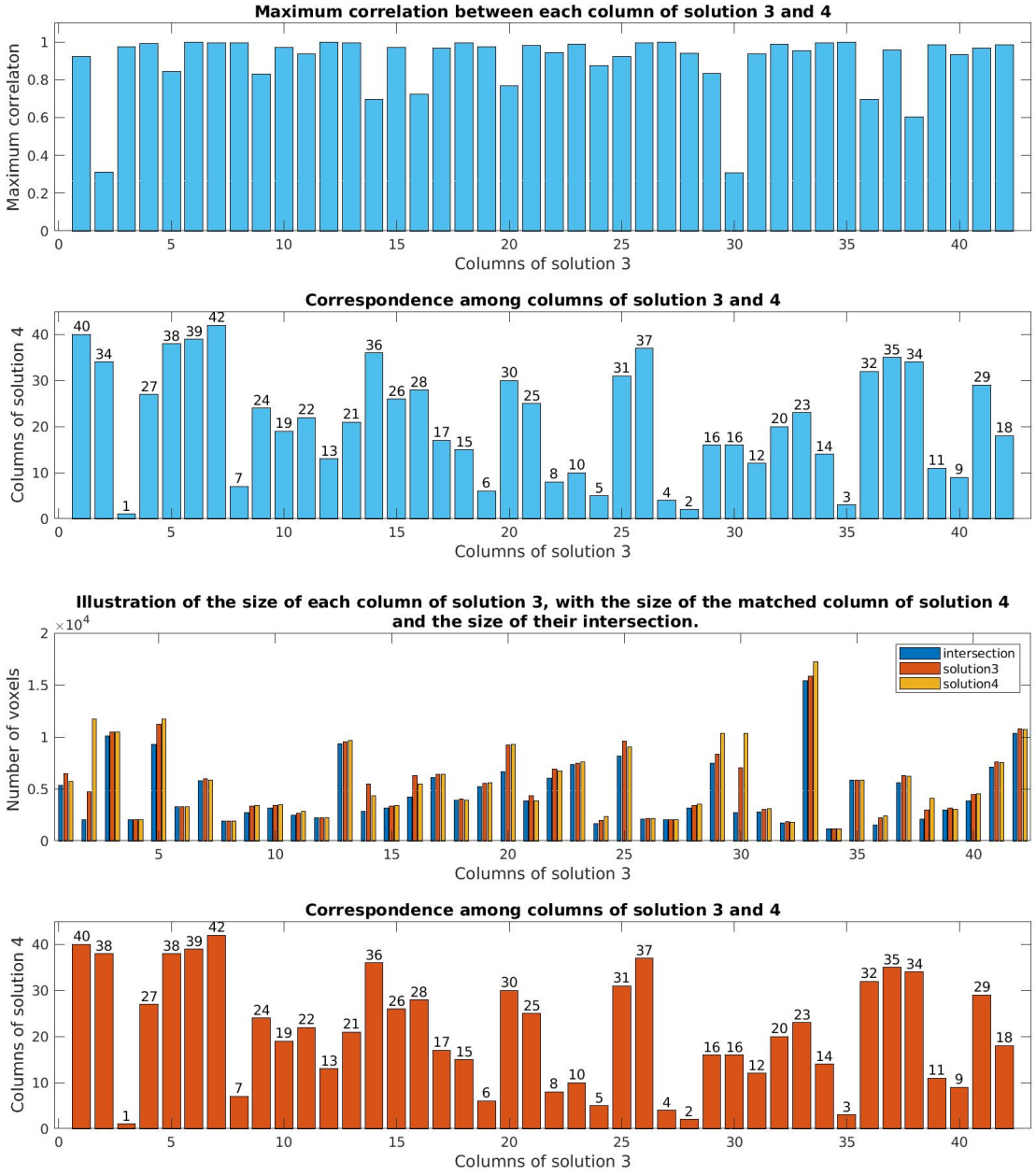


Figure B.21: **First figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 3 and each column of the solution 4. We associate the column r of the solution 3 with the column of the solution 4 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 3, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 4. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 4 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 4 each column of the solution 3 is associated based on intersection. Solutions 3, 4 correspond to accelerated penalty method and common subspace dimension equal to 42.

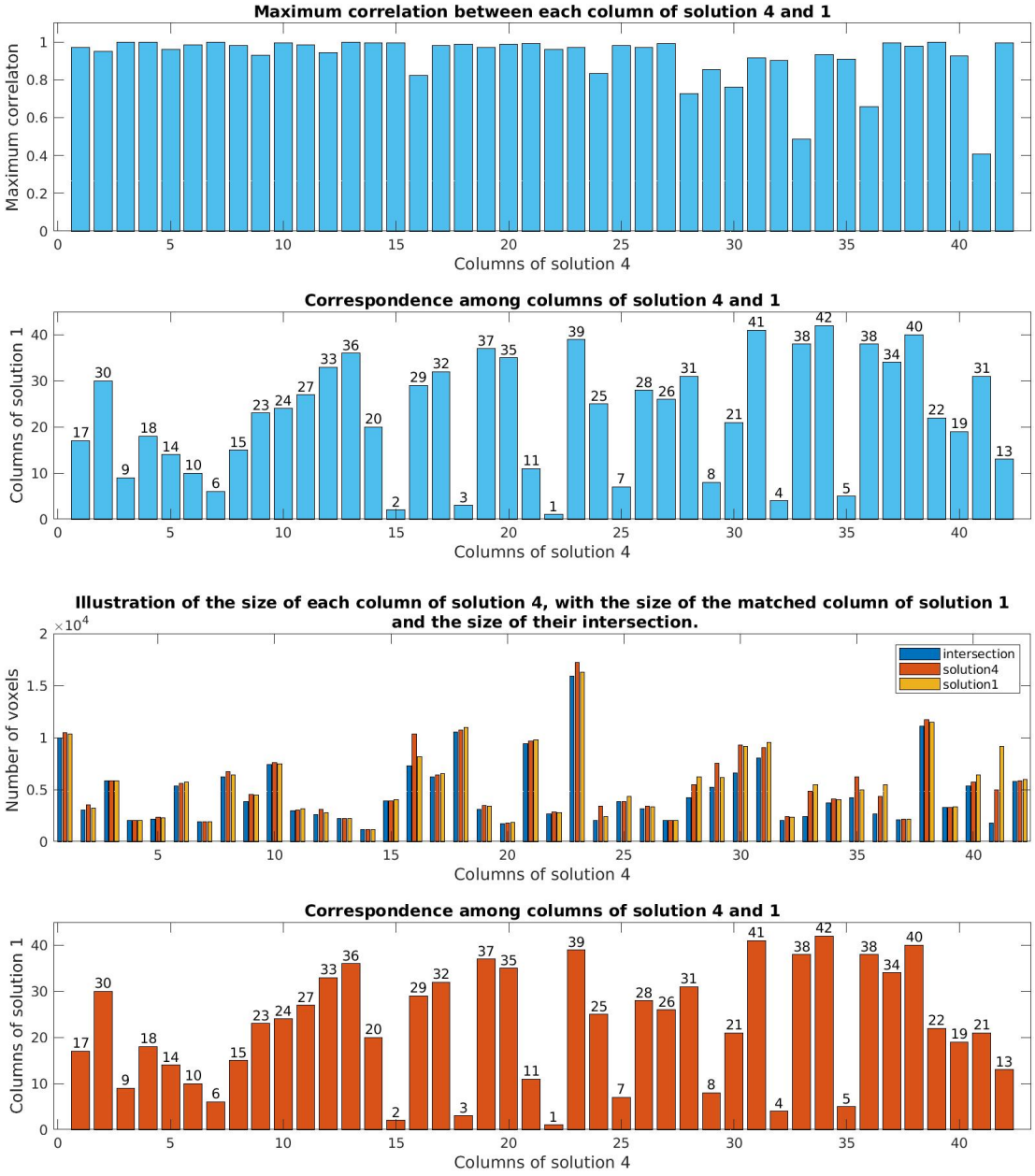


Figure B.22: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 1. We associate the column r of the solution 4 with the column of the solution 1 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 1. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 1 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 1 each column of the solution 4 is associated based on intersection. Solutions 4, 1 correspond to accelerated penalty method and common subspace dimension equal to 42.

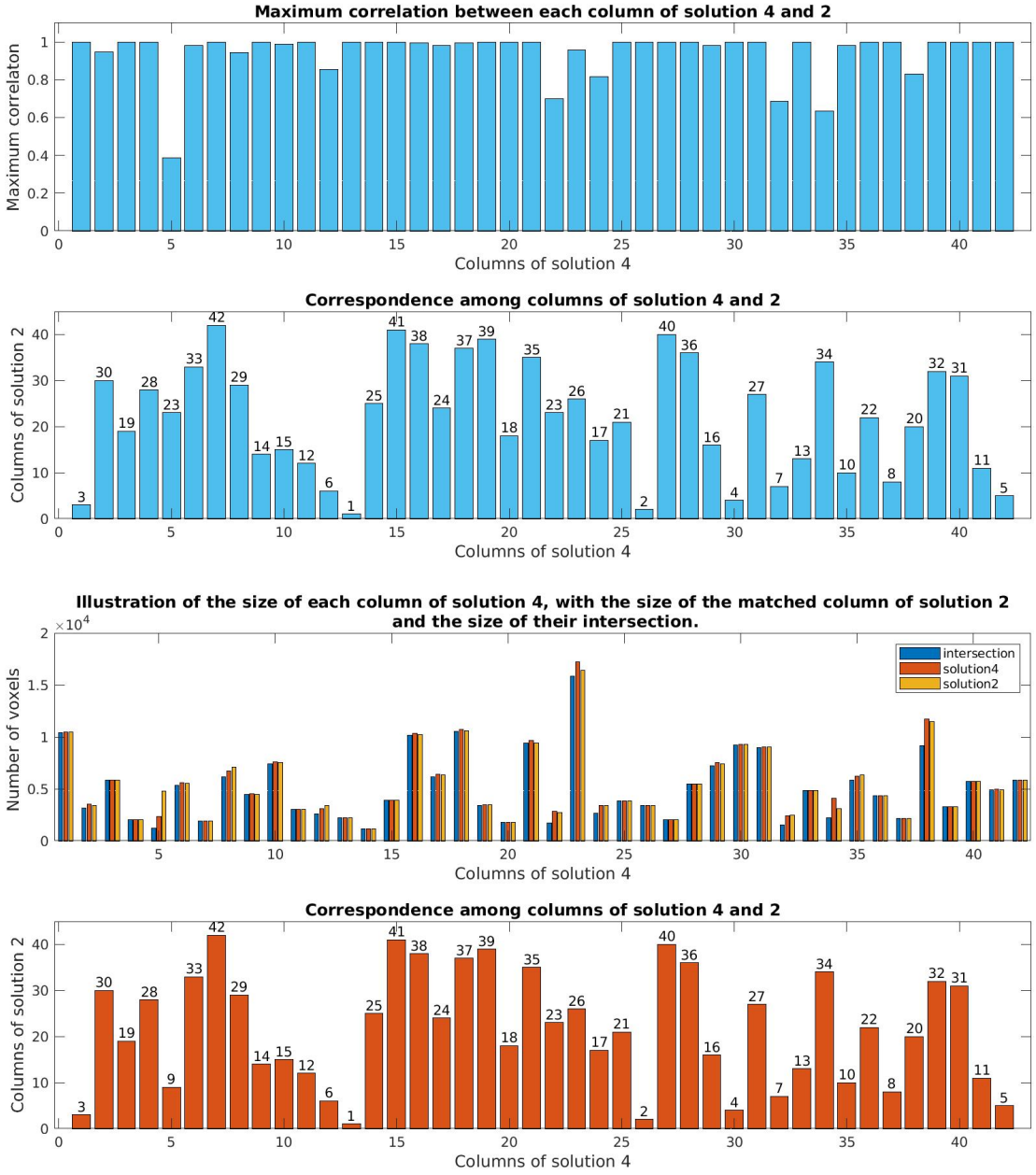


Figure B.23: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 2. We associate the column r of the solution 4 with the column of the solution 2 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 2. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 2 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 2 each column of the solution 4 is associated based on intersection. Solutions 4, 2 correspond to accelerated penalty method and common subspace dimension equal to 42.

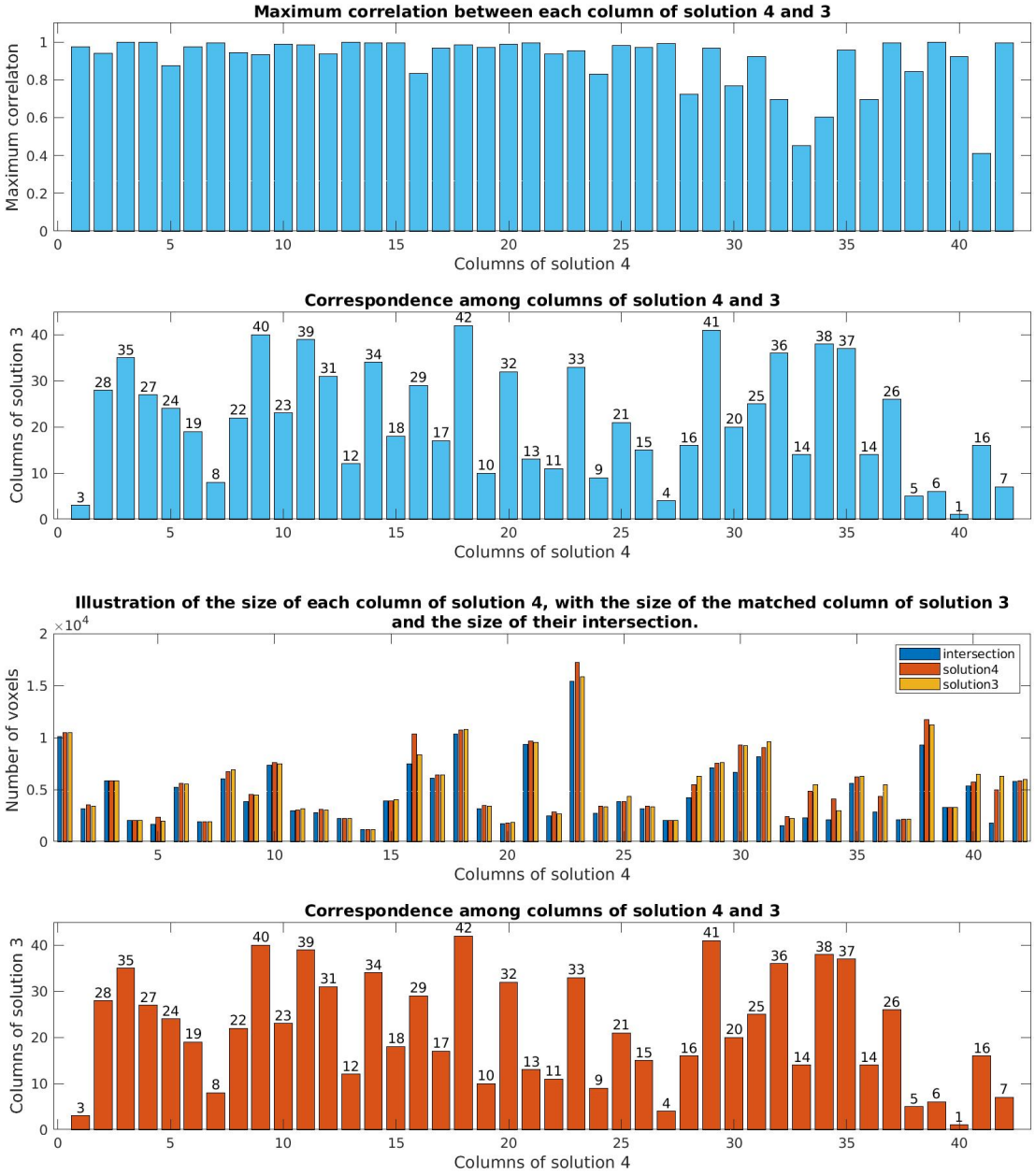


Figure B.24: **First figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the correlation coefficients between the column r of the solution 4 and each column of the solution 3. We associate the column r of the solution 4 with the column of the solution 3 with which the column r has the maximum correlation coefficient. We illustrate, for each column r, the maximum correlation coefficient. **Second figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated based on correlation coefficients. **Third figure**: For each column r of the solution 4, for $r = 1, \ldots, R$, we calculate the intersection of non-zero elements between the column r and each column of the solution 3. For each column r, we illustrate the number of non-zero elements of column r (red bar), the number of non-zero elements of column of solution 3 with which column r has the largest intersection (yellow bar) and the number of elements of the intersection (blue bar). **Fourth figure**: It illustrates with which column of the solution 3 each column of the solution 4 is associated based on intersection. Solutions 4, 3 correspond to accelerated penalty method and common subspace dimension equal to 42.

B.3 Clusters for common subspace dimension equal to 27

1. **Cluster 19**
   Cluster 19 contains 28245 voxels, which correspond to 12.500664% of the whole brain (background, soft tissue, bone are excluded).

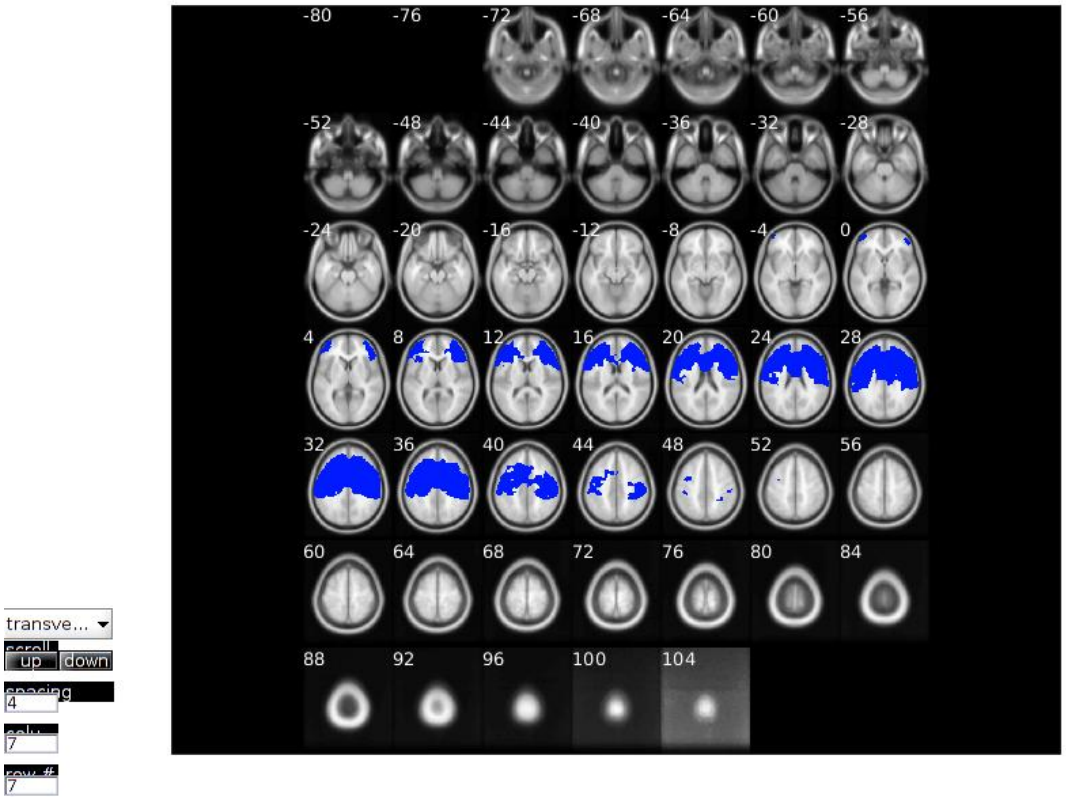


Figure B.25: Cluster 19. Blue areas indicate the voxels which belong to cluster 19.

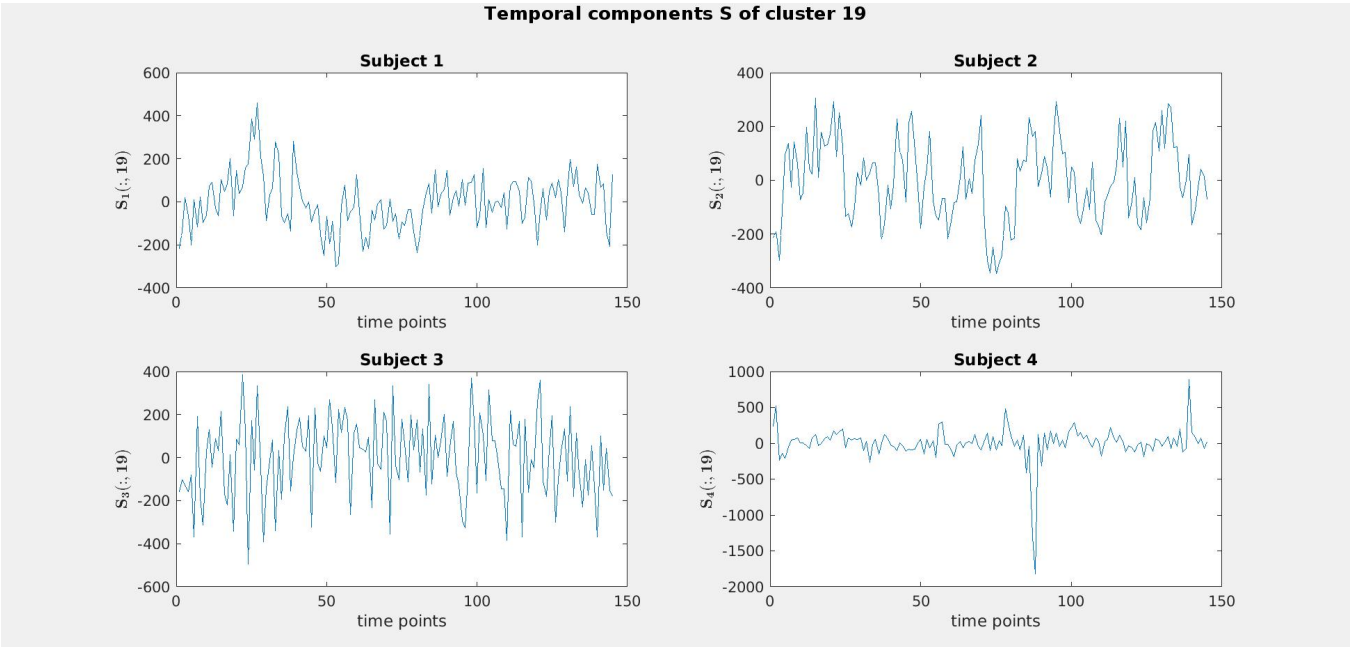


Figure B.26: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 1, \ldots, 4$ and R=27.

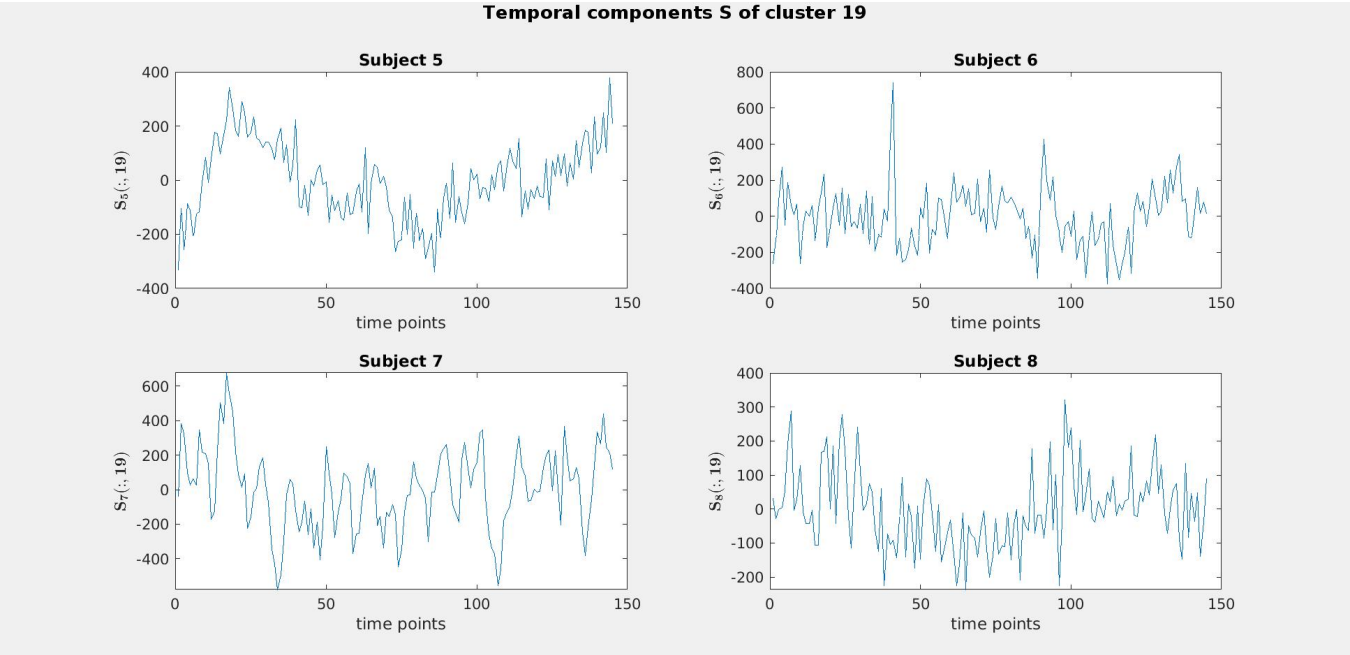


Figure B.27: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 5, \ldots, 8$ and R=27.

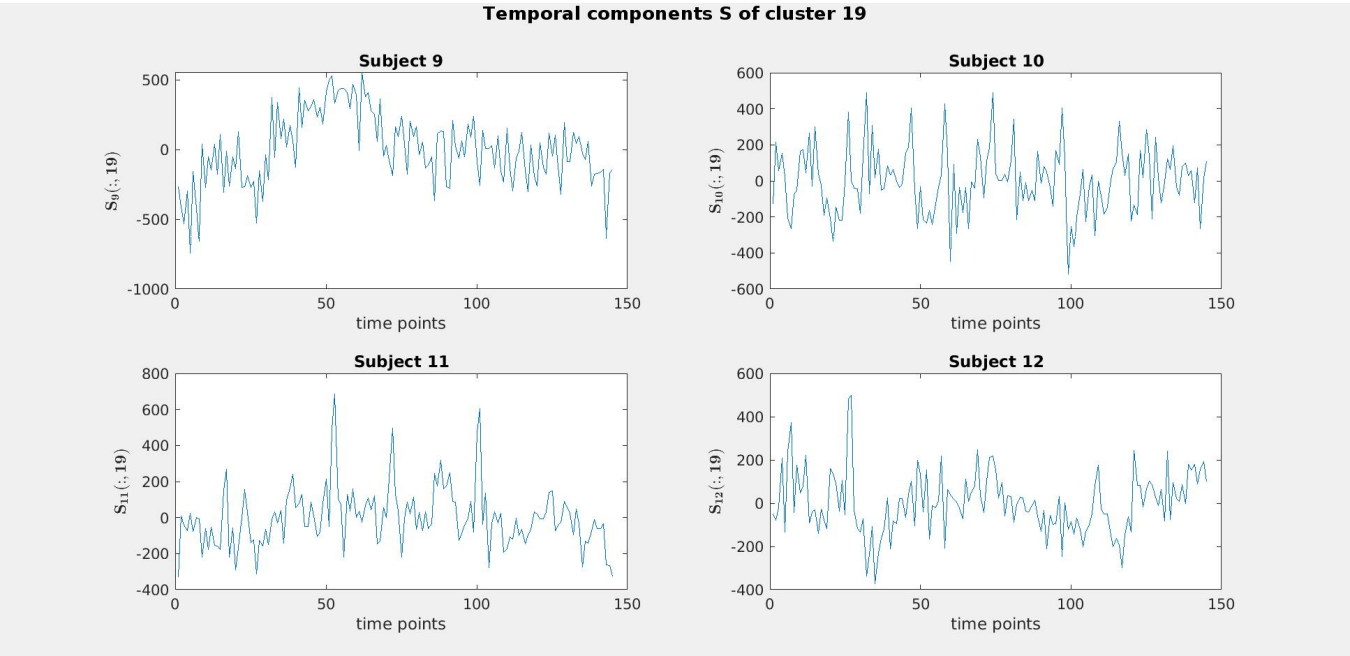


Figure B.28: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 9, \ldots, 12$ and R=27.

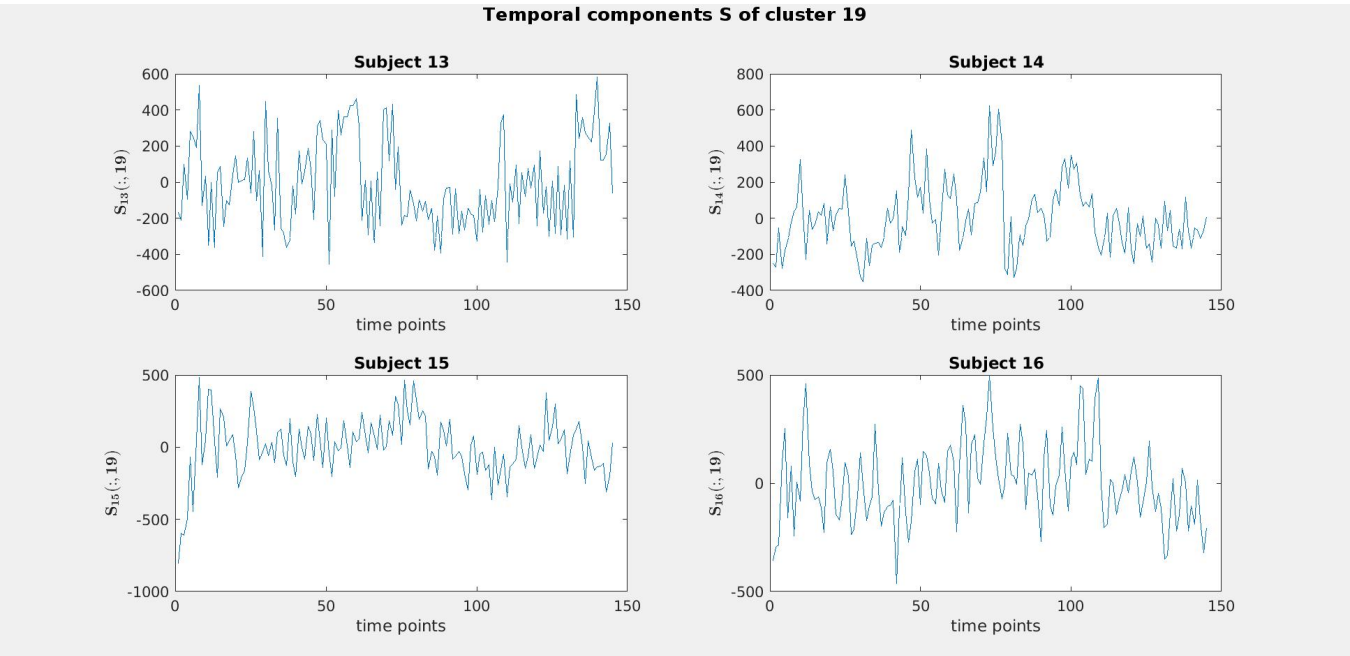


Figure B.29: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 13, \ldots, 16$ and R=27.

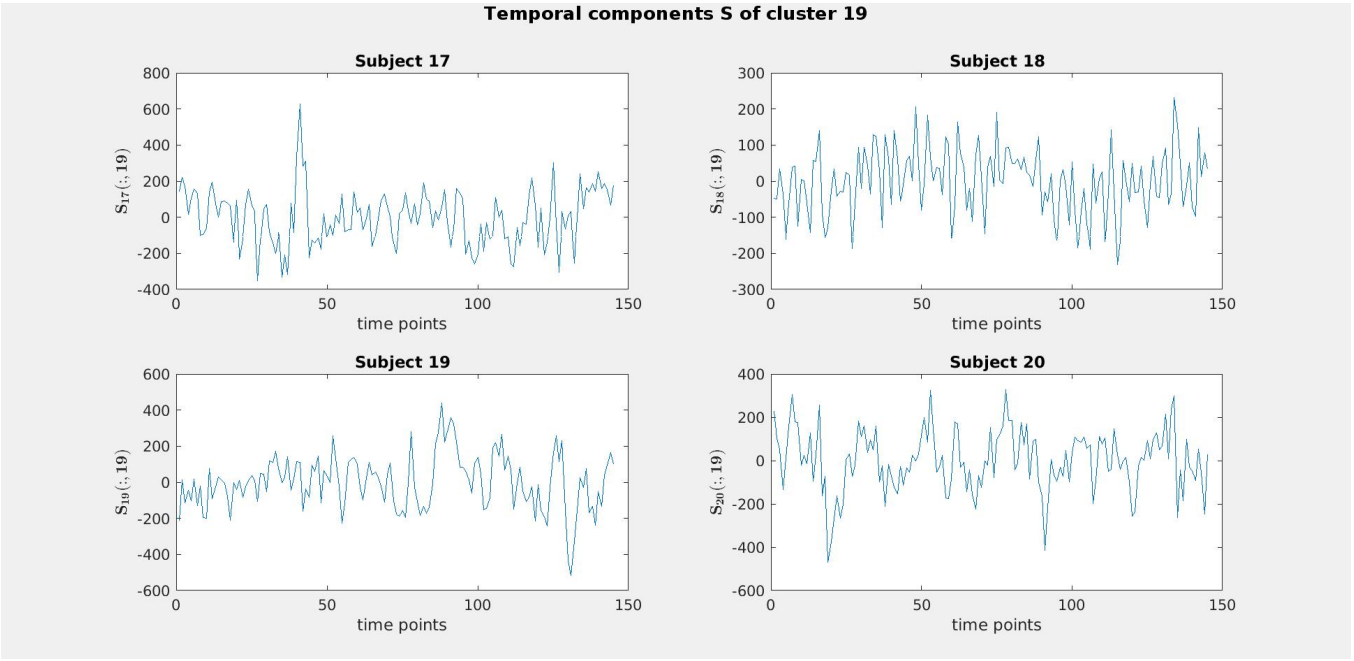


Figure B.30: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 17, \ldots, 20$ and R=27.

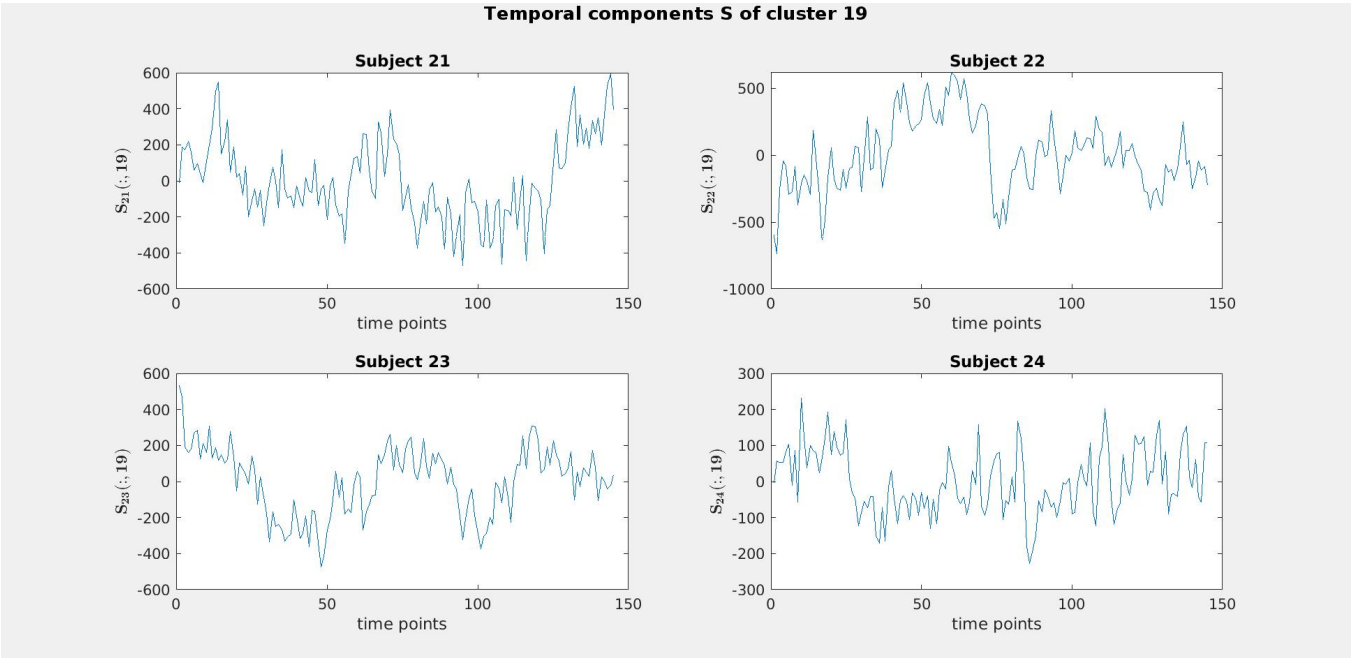


Figure B.31: Waveform of temporal components $\mathbf{S}_k(:, 119)$ for $k = 21, \ldots, 24$ and R=27.

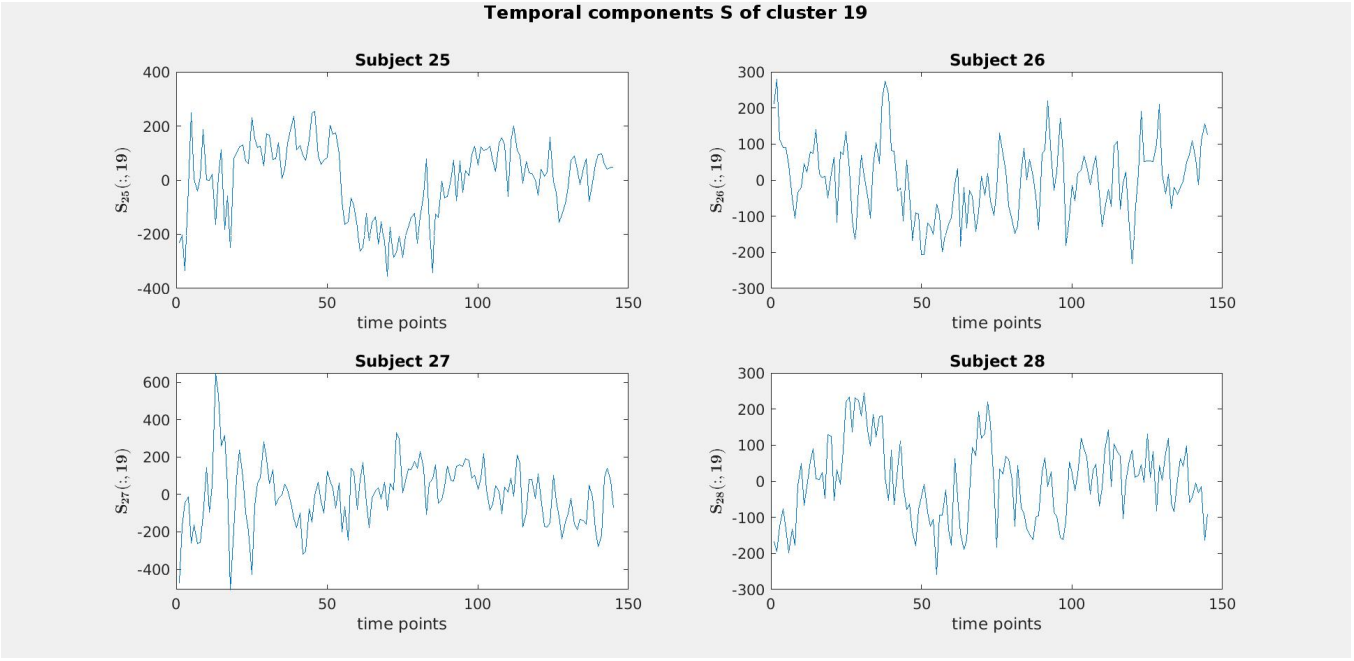


Figure B.32: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 25, \ldots, 28$ and R=27.

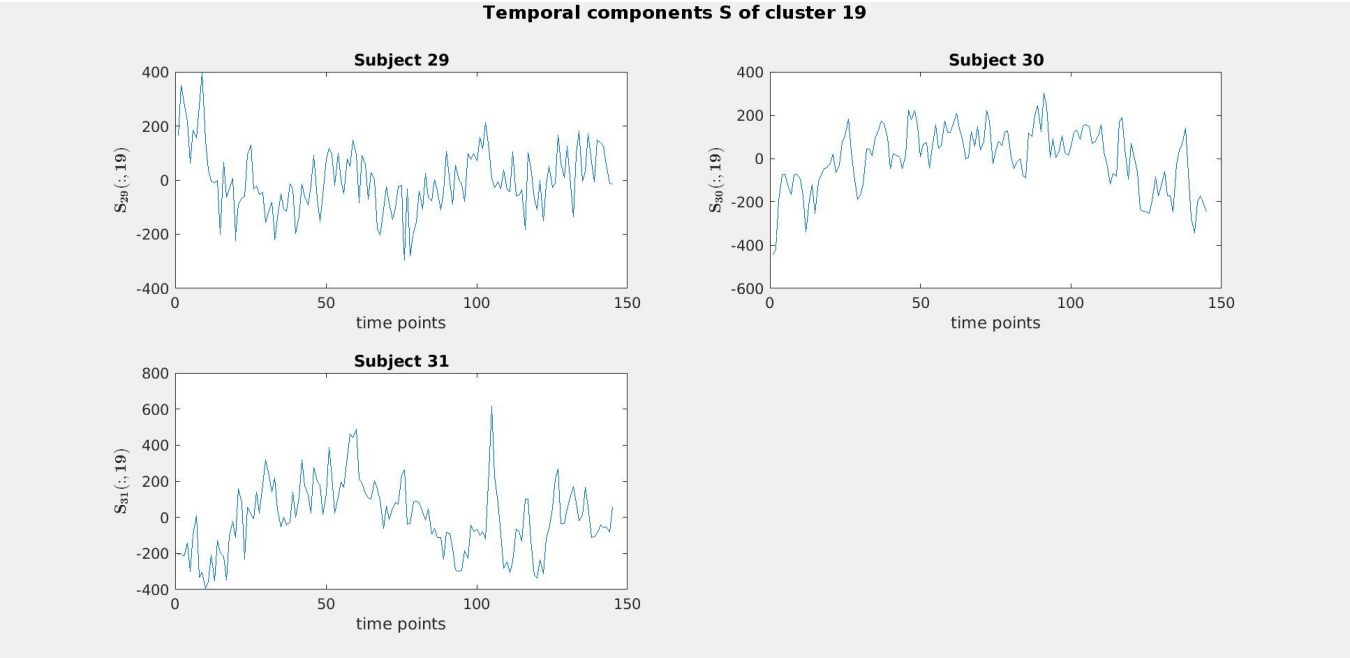


Figure B.33: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 29, \ldots, 31$ and R=27.

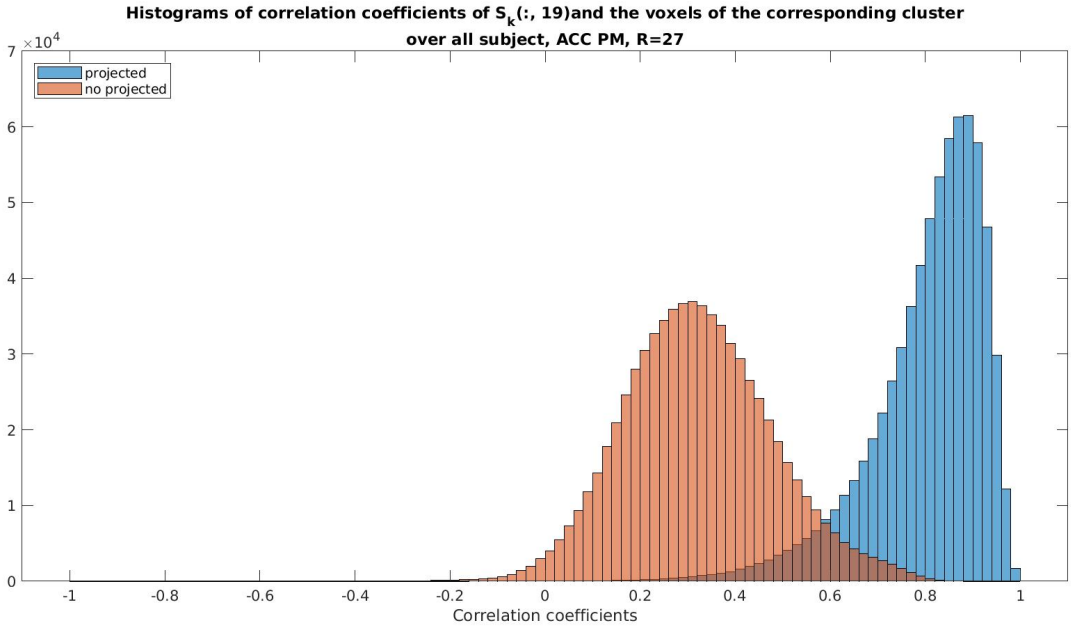


Figure B.34: For the column 19 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 19)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

2. **Cluster 27**
   Cluster 27 contains 20290 voxels, which correspond to 8.979942% of the whole brain (background, soft tissue, bone are excluded).

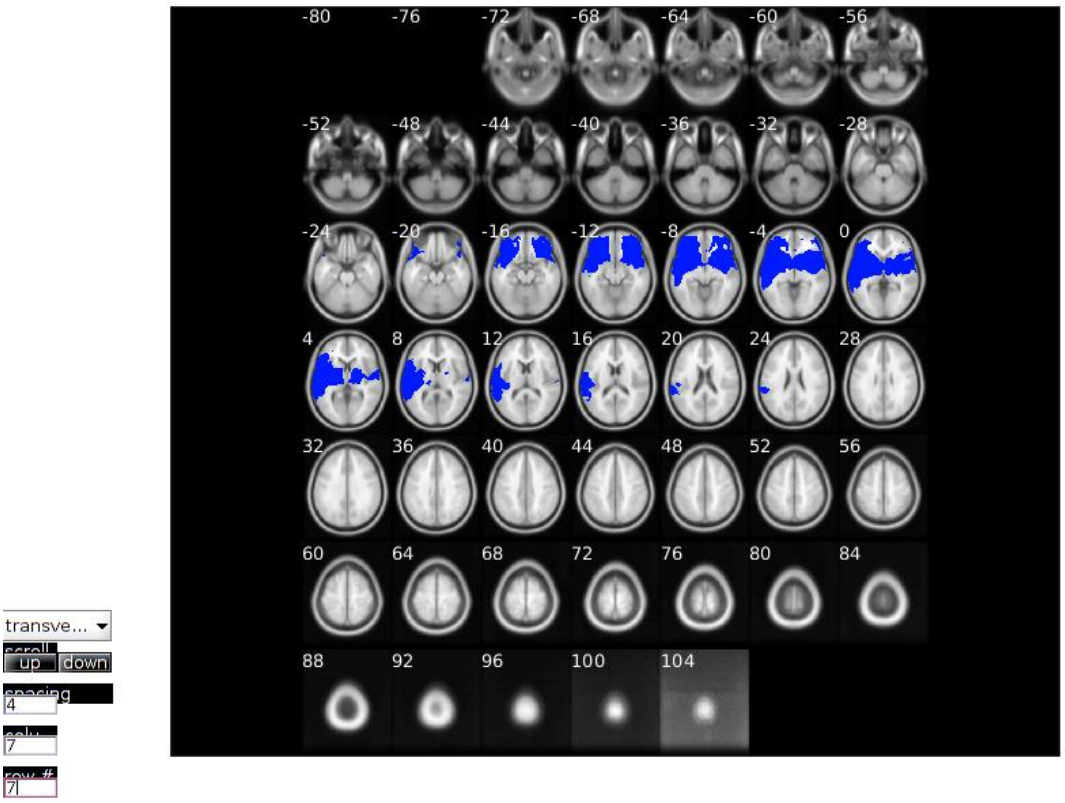


Figure B.35: Cluster 27. Blue areas indicate the voxels which belong to cluster 27.

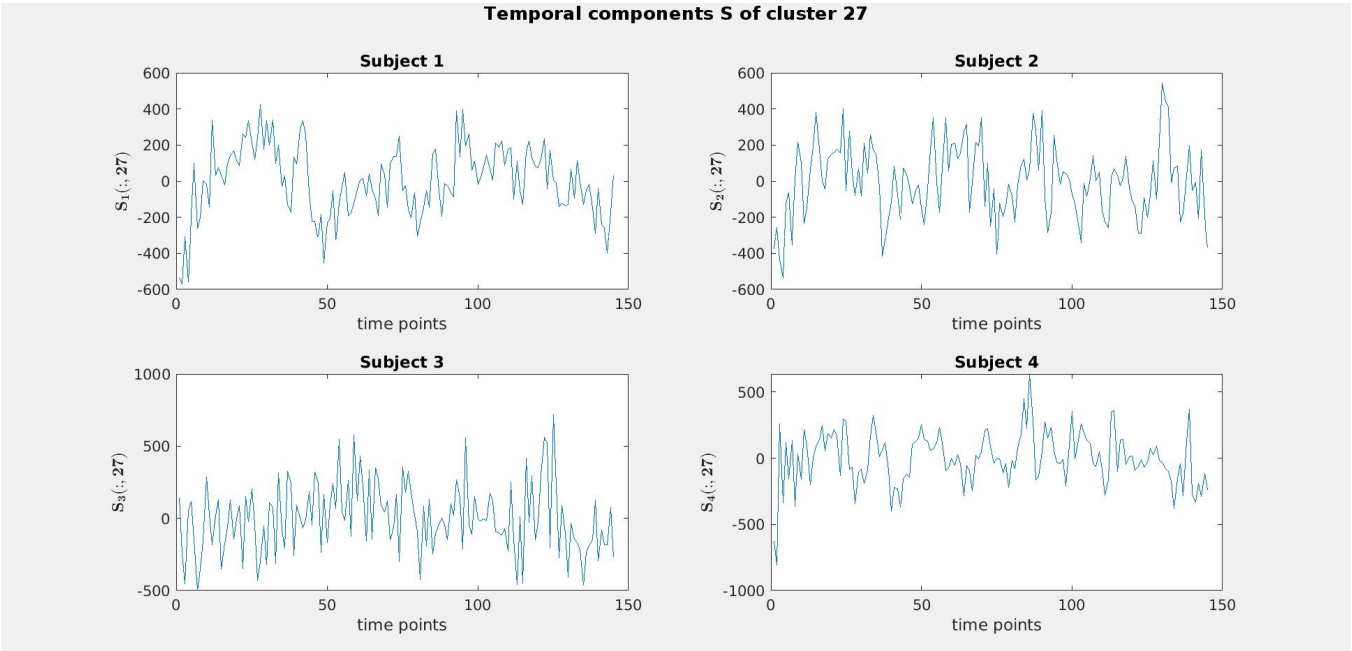


Figure B.36: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 1, \ldots, 4$ and R=27.

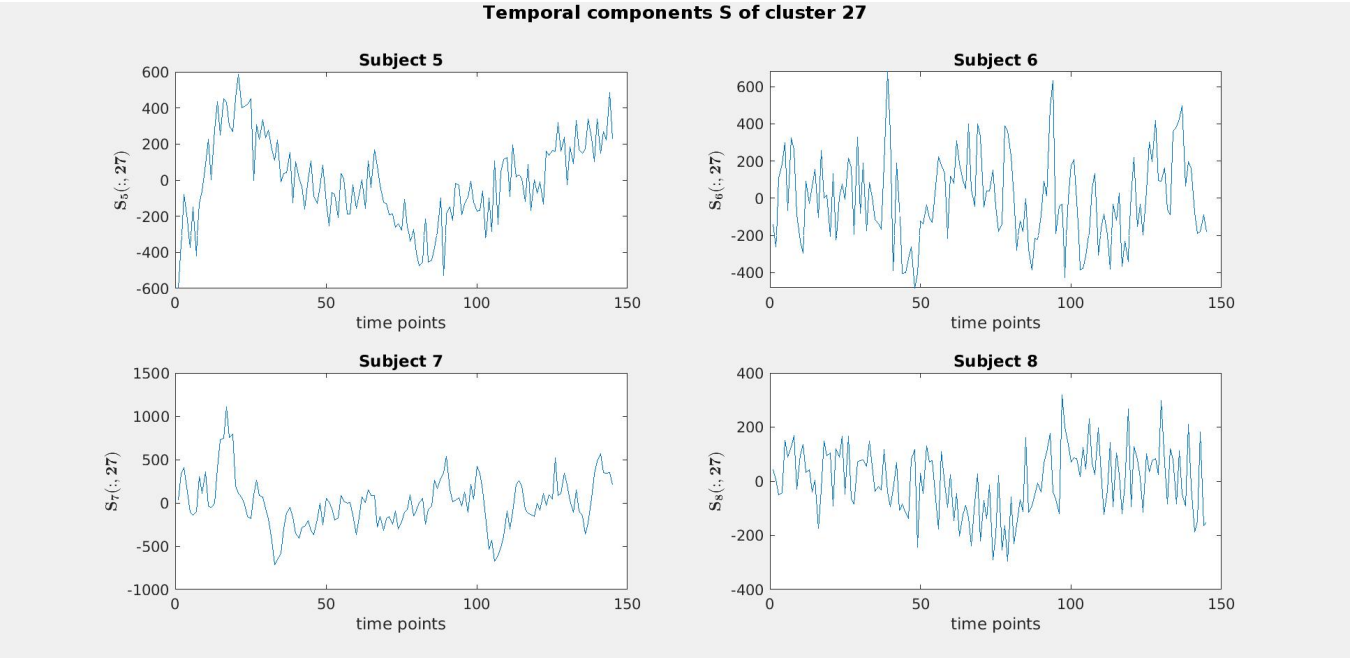


Figure B.37: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 5, \ldots, 8$ and R=27.

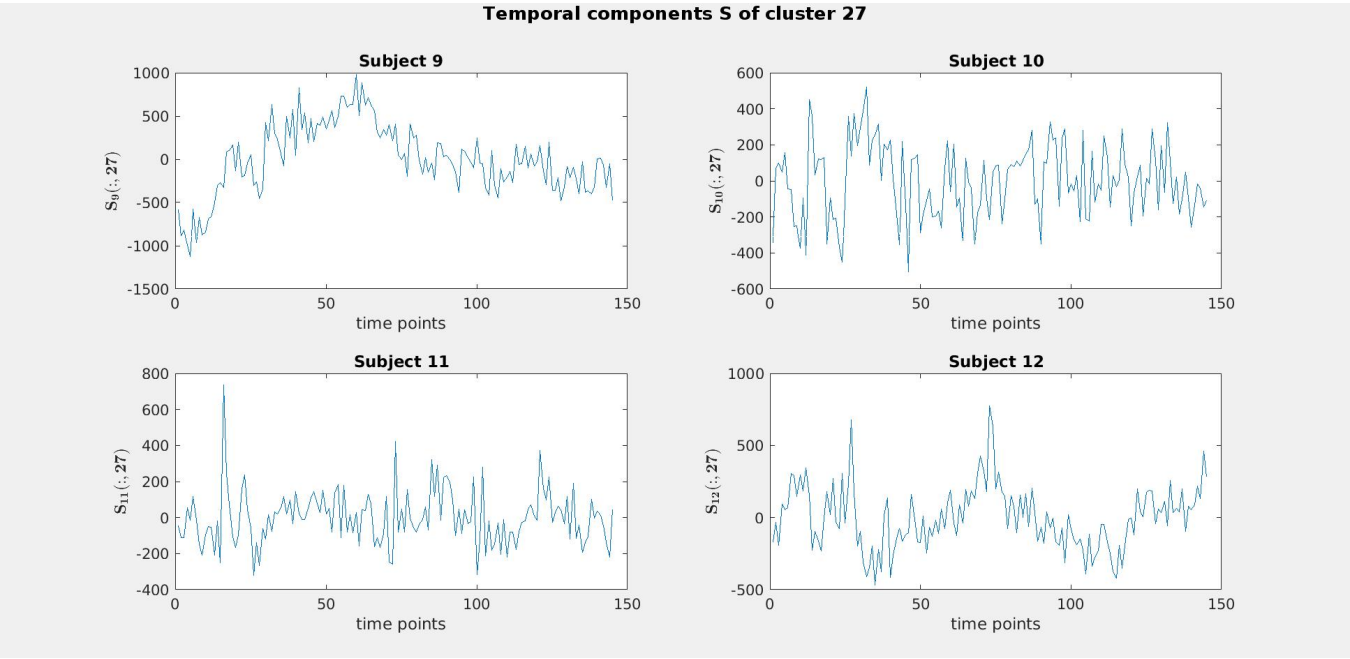


Figure B.38: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 9, \ldots, 12$ and R=27.

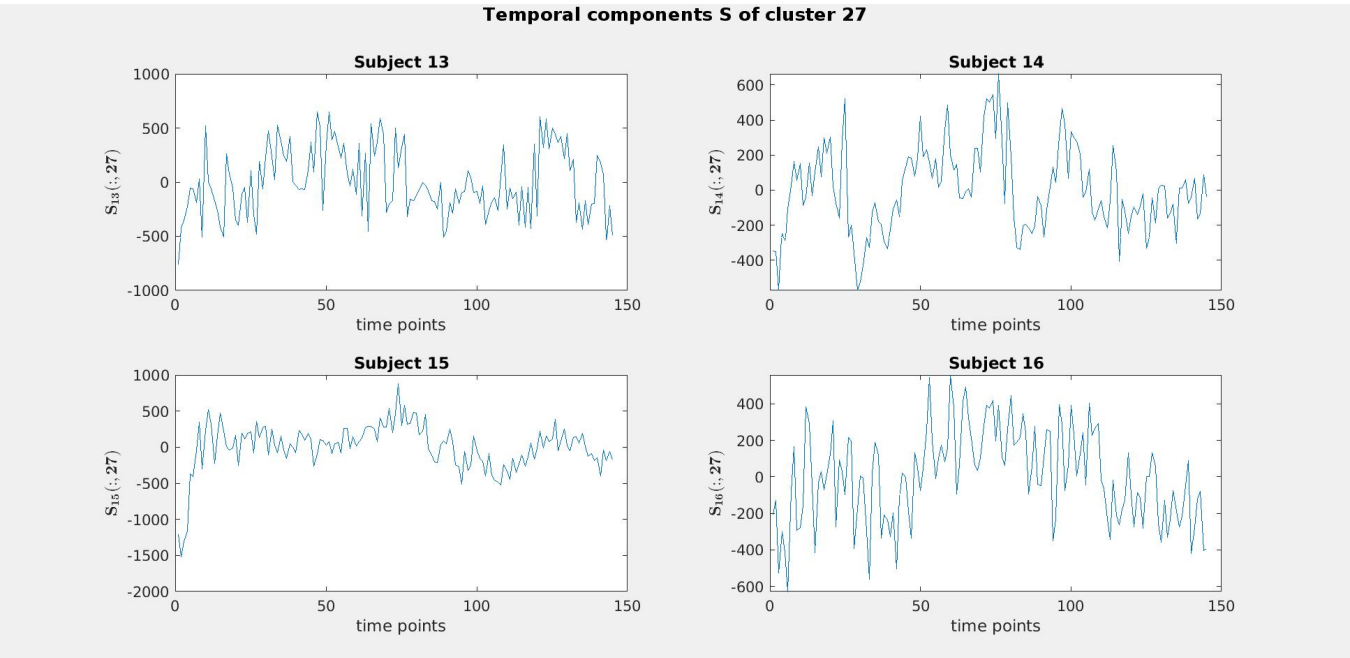


Figure B.39: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 13, \ldots, 16$ and R=27.

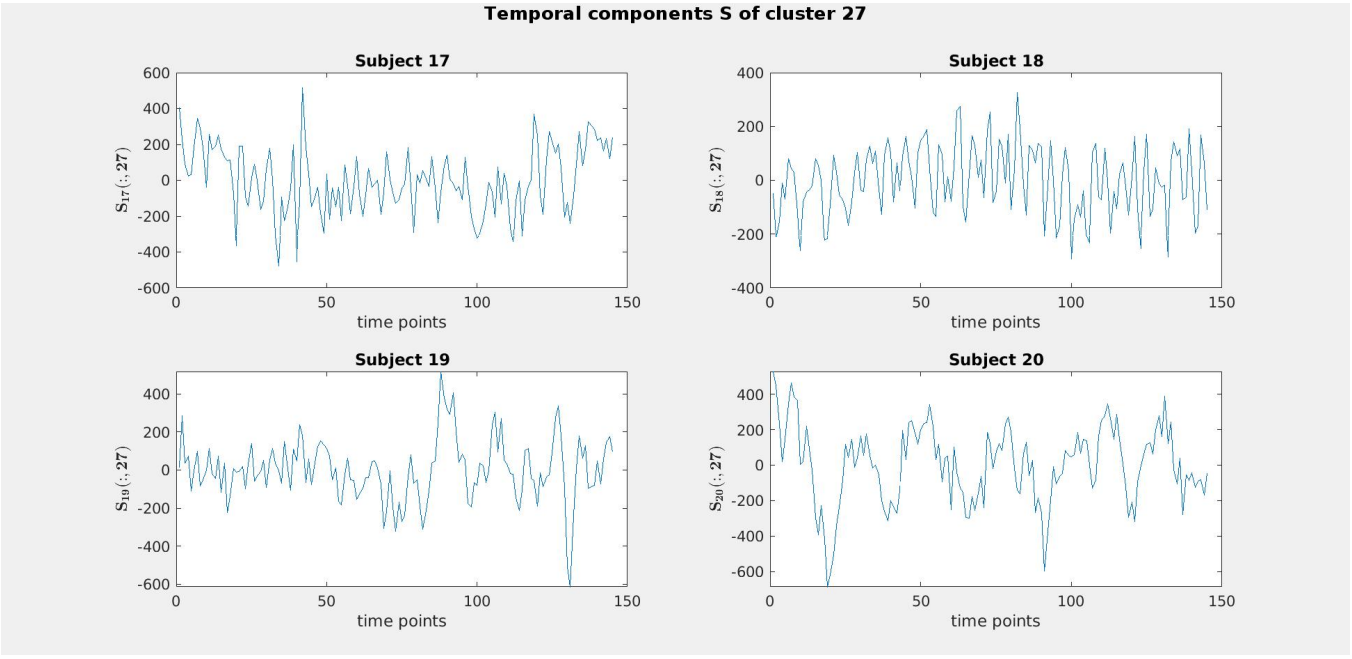


Figure B.40: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 17, \ldots, 20$ and R=27.

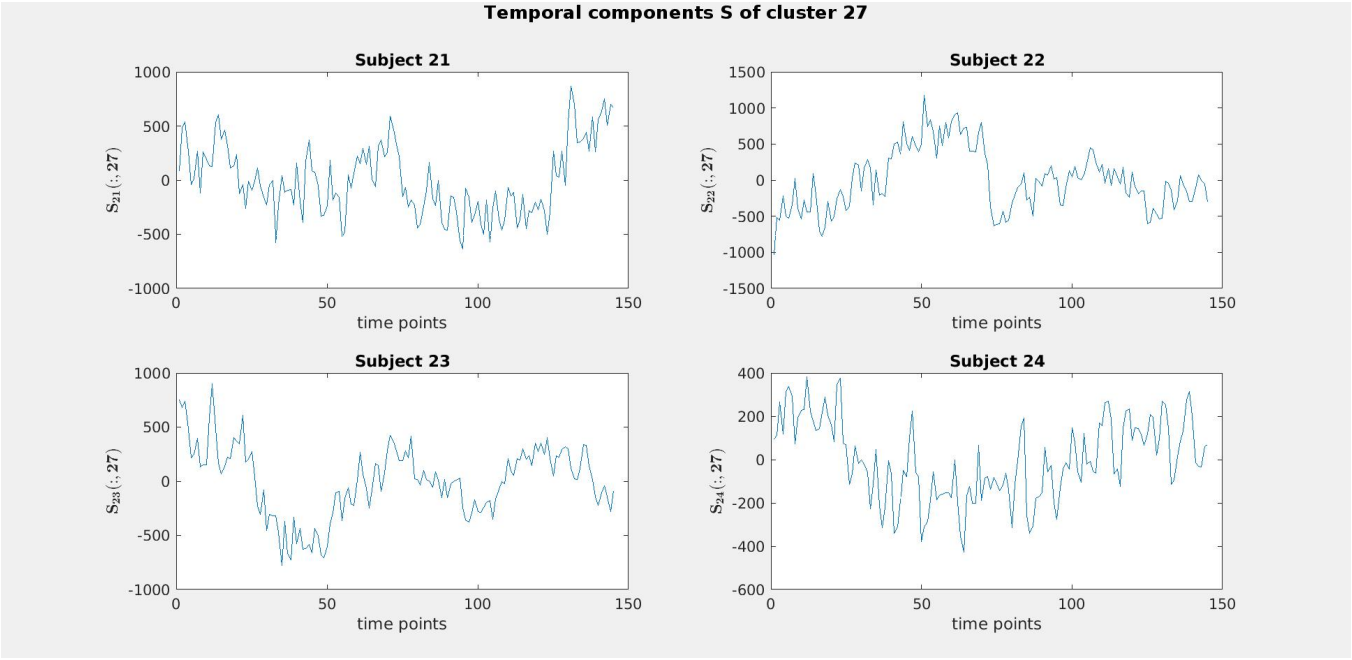


Figure B.41: Waveform of temporal components $\mathbf{S}_k(:, 127)$ for $k = 21, \ldots, 24$ and R=27.

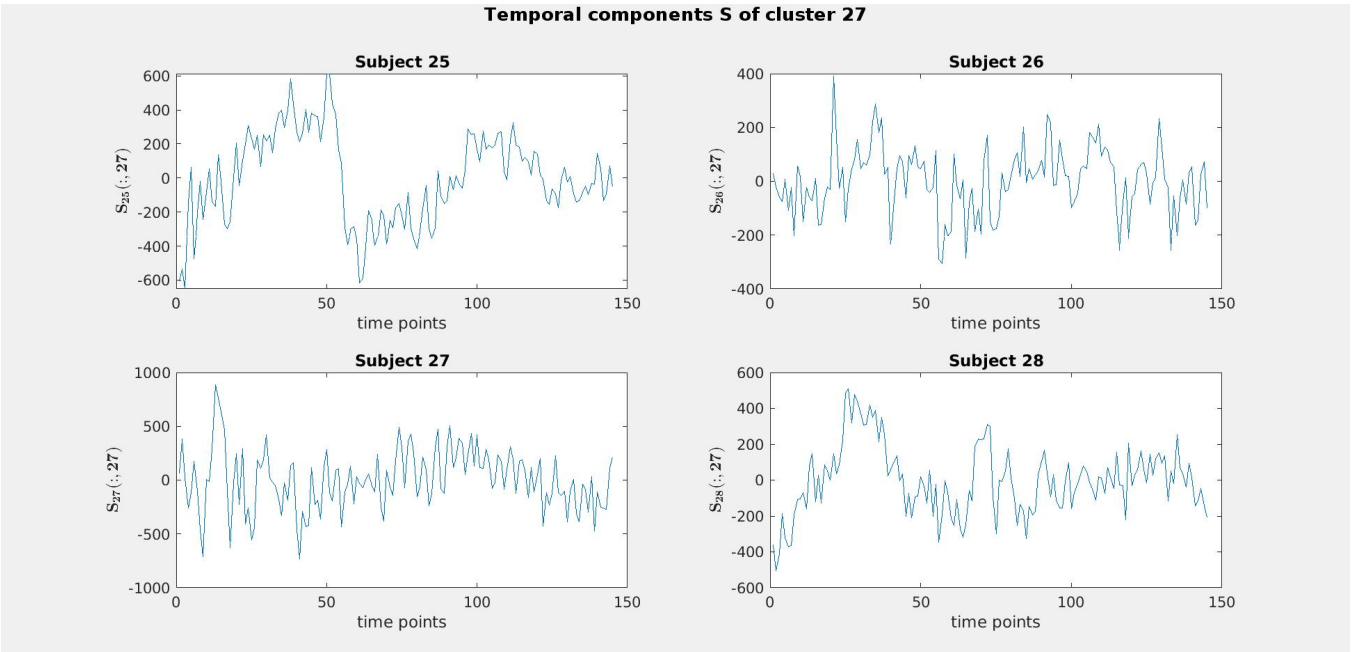


Figure B.42: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 25, \ldots, 28$ and R=27.

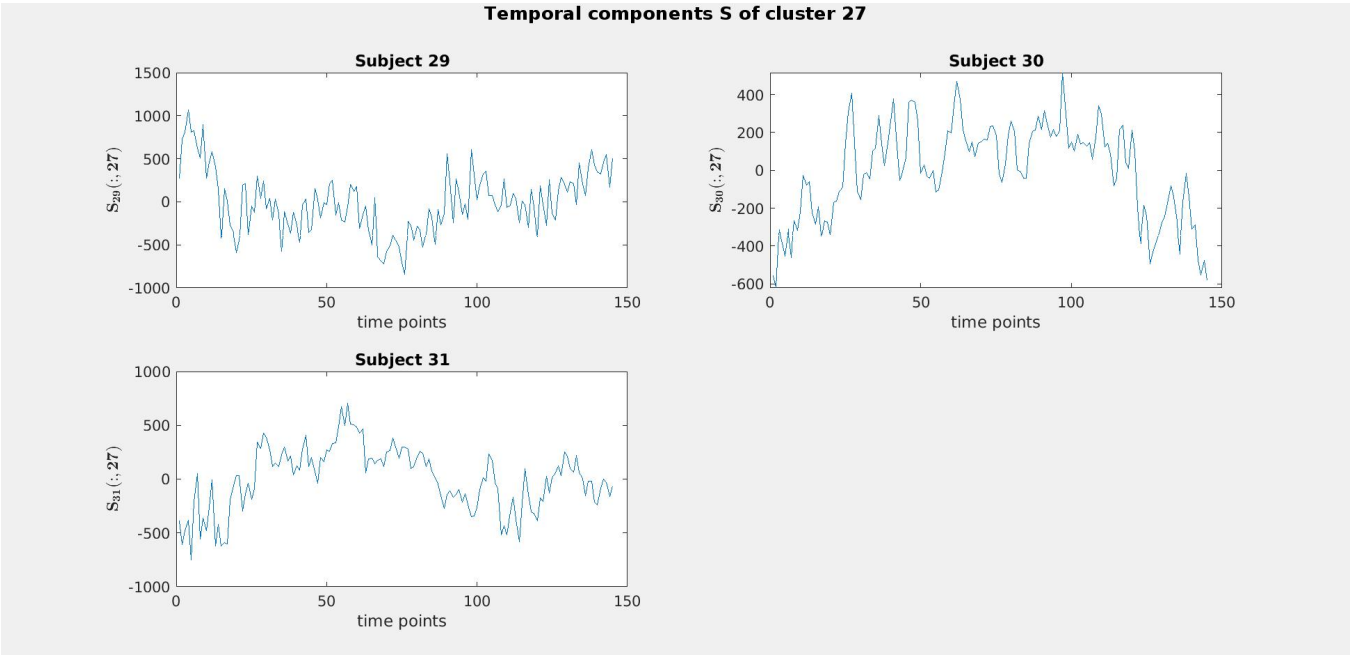


Figure B.43: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 29, \ldots, 31$ and R=27.

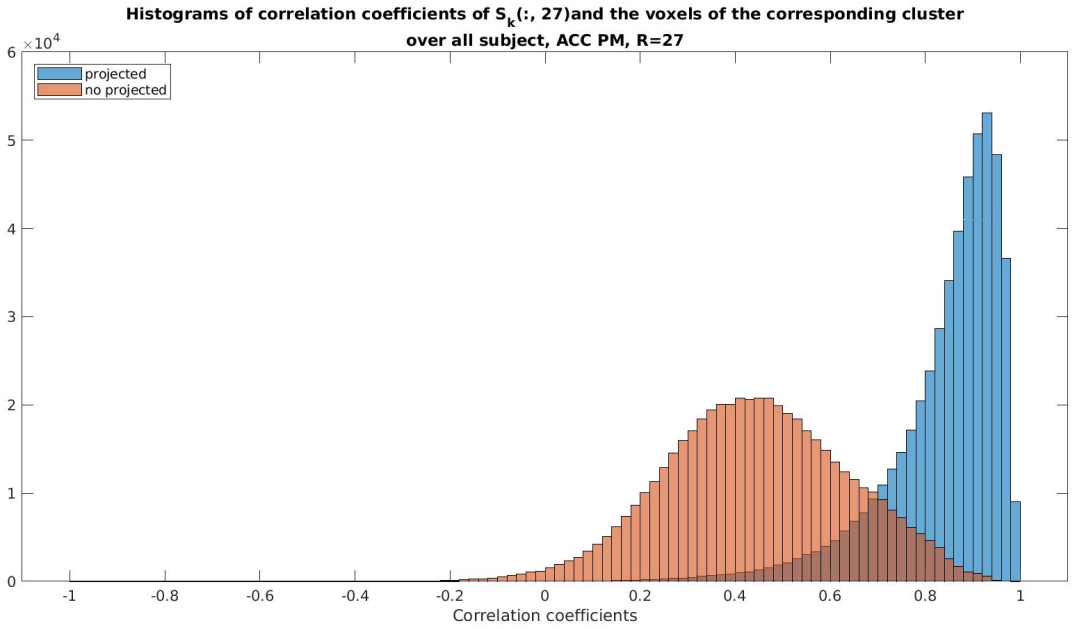


Figure B.44: For the column 27 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 27)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

3. **Cluster 22**
   Cluster 22 contains 17039 voxels, which correspond to 7.541116% of the whole brain (background, soft tissue, bone are excluded).

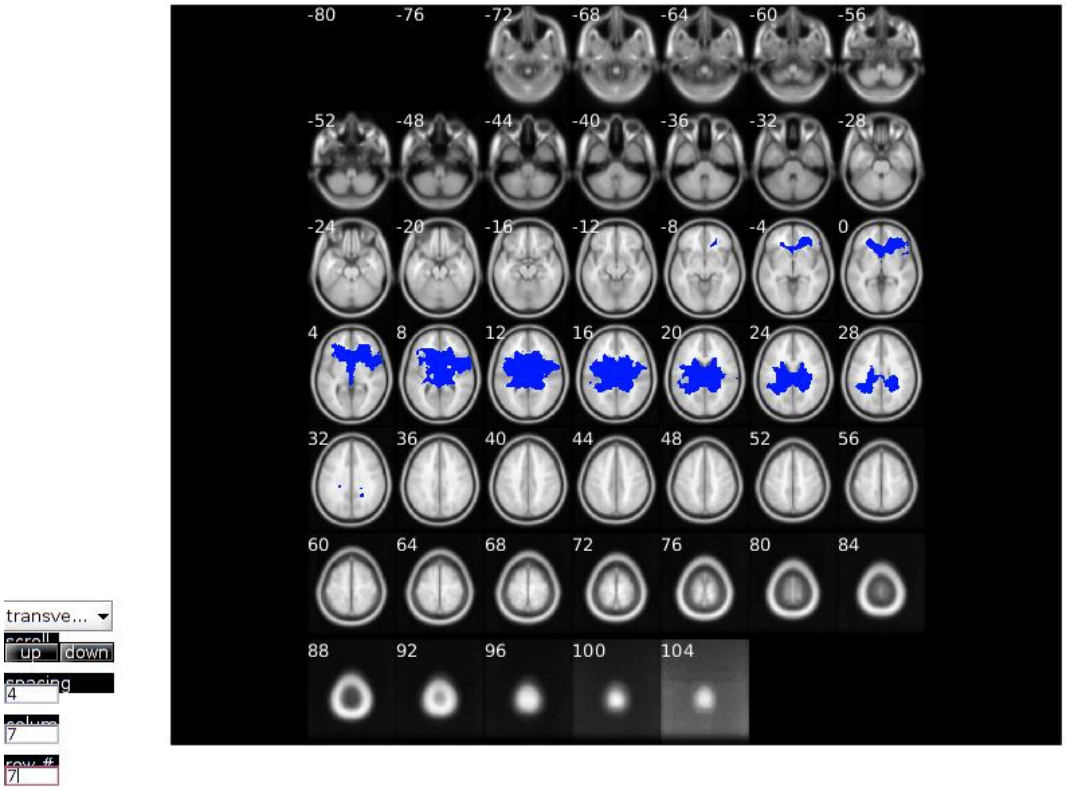


Figure B.45: Cluster 22. Blue areas indicate the voxels which belong to cluster 22.

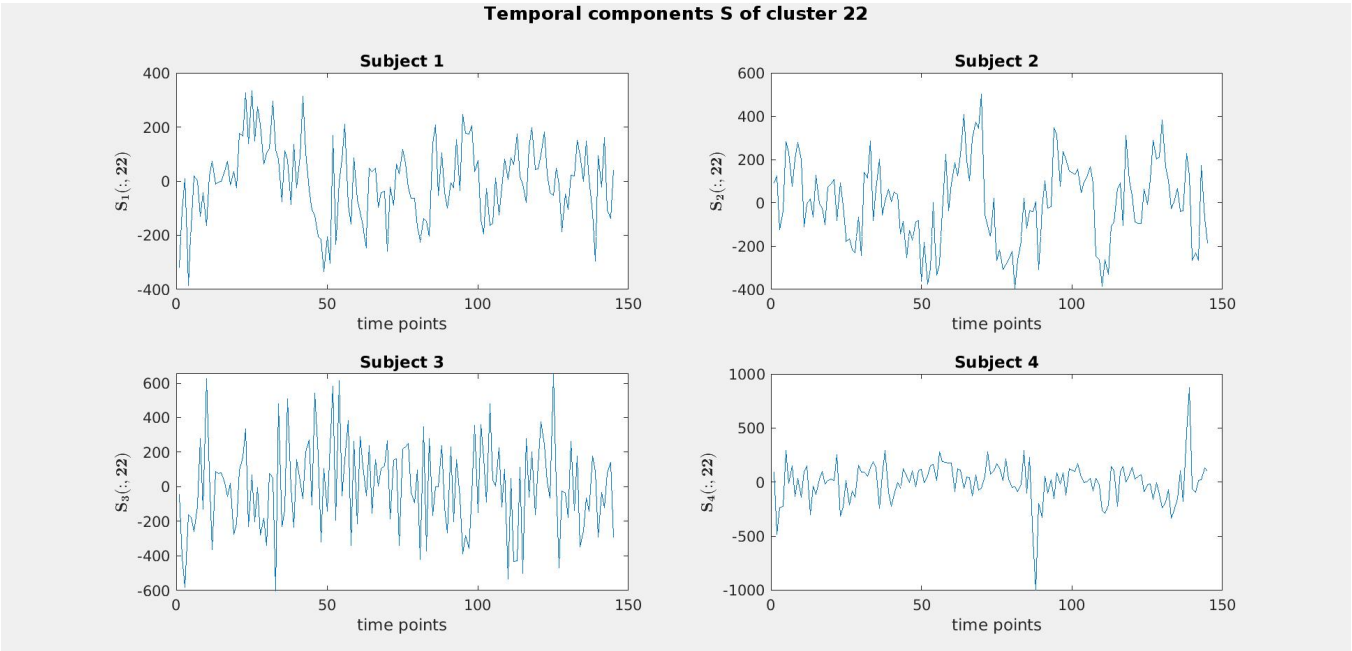


Figure B.46: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 1, \ldots, 4$ and R=27.

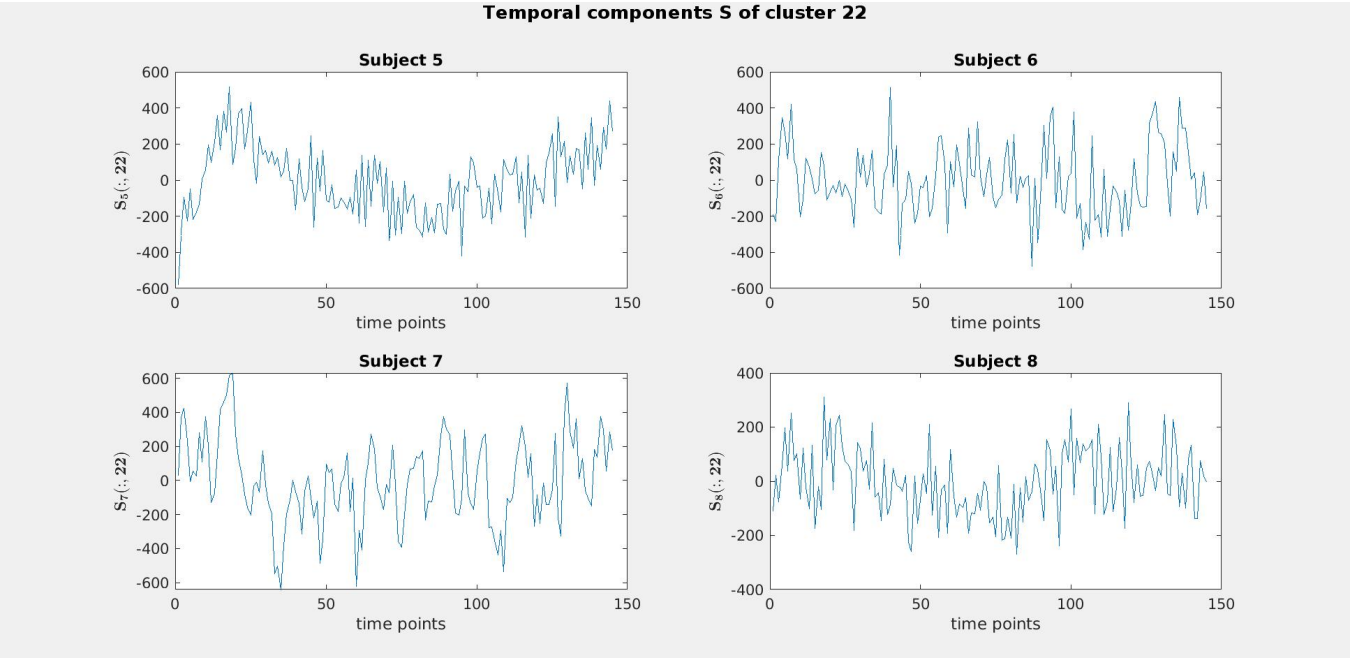


Figure B.47: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 5, \ldots, 8$ and R=27.

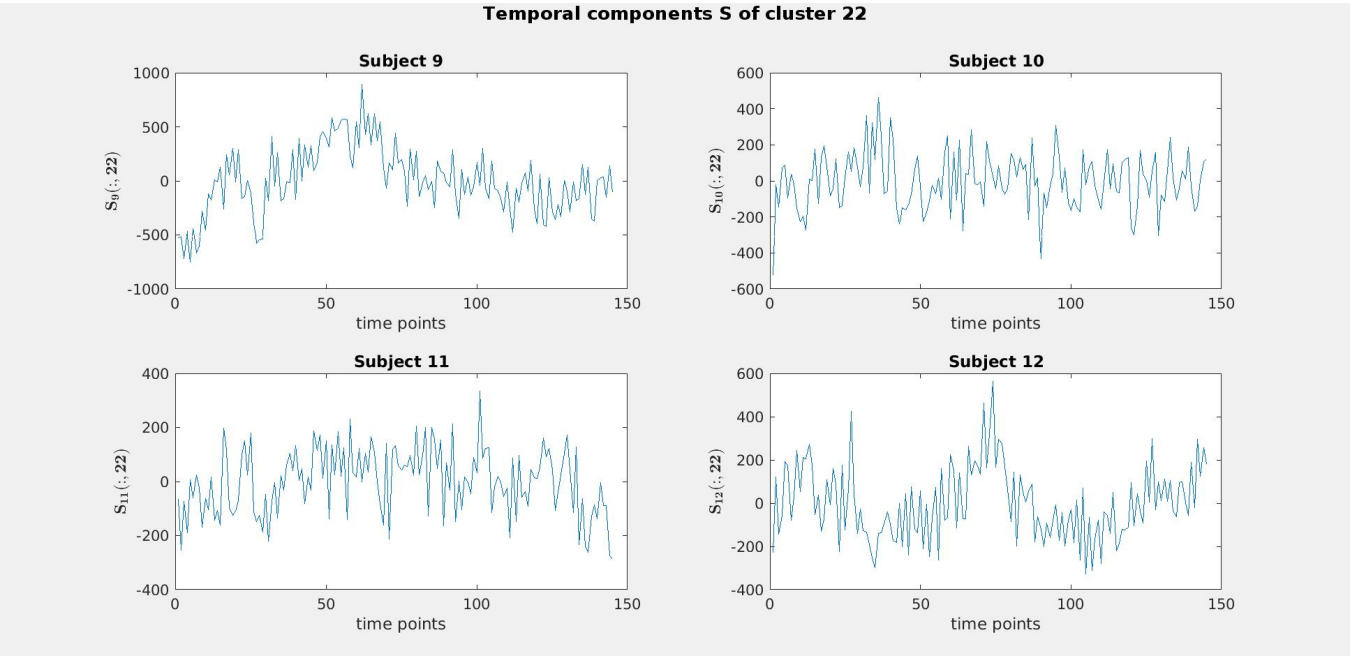


Figure B.48: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 9, \ldots, 12$ and R=27.

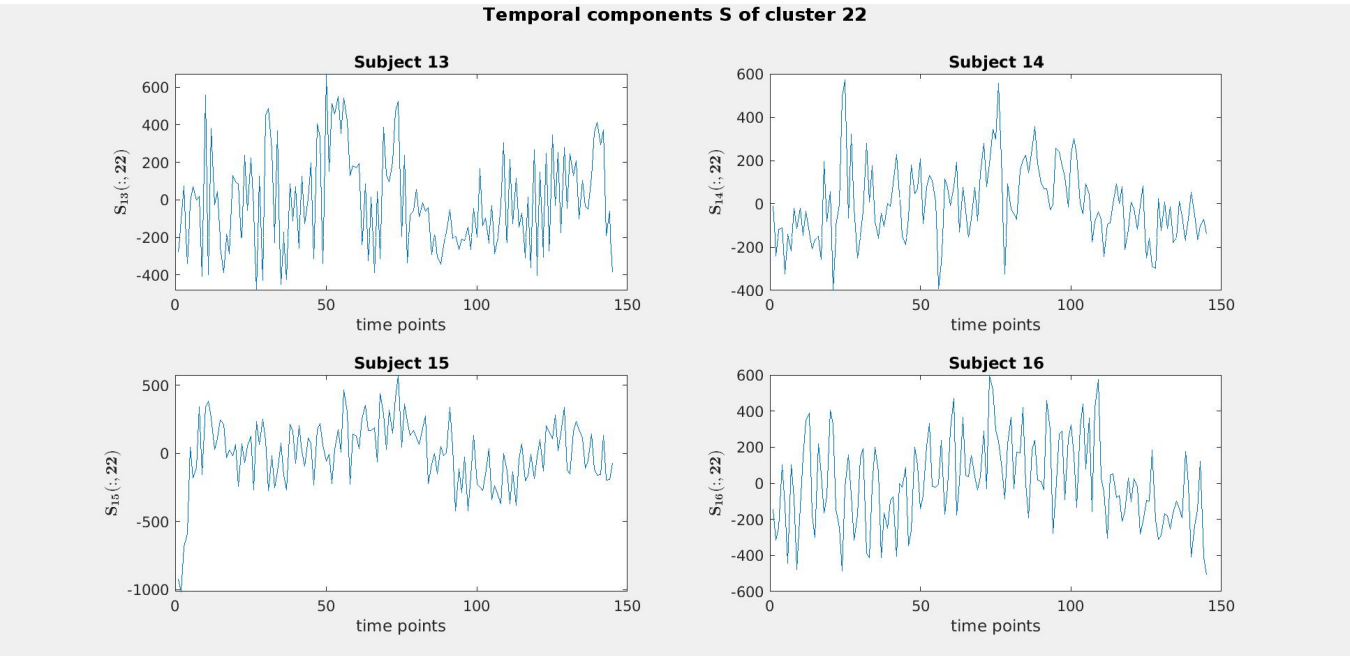


Figure B.49: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 13, \ldots, 16$ and R=27.

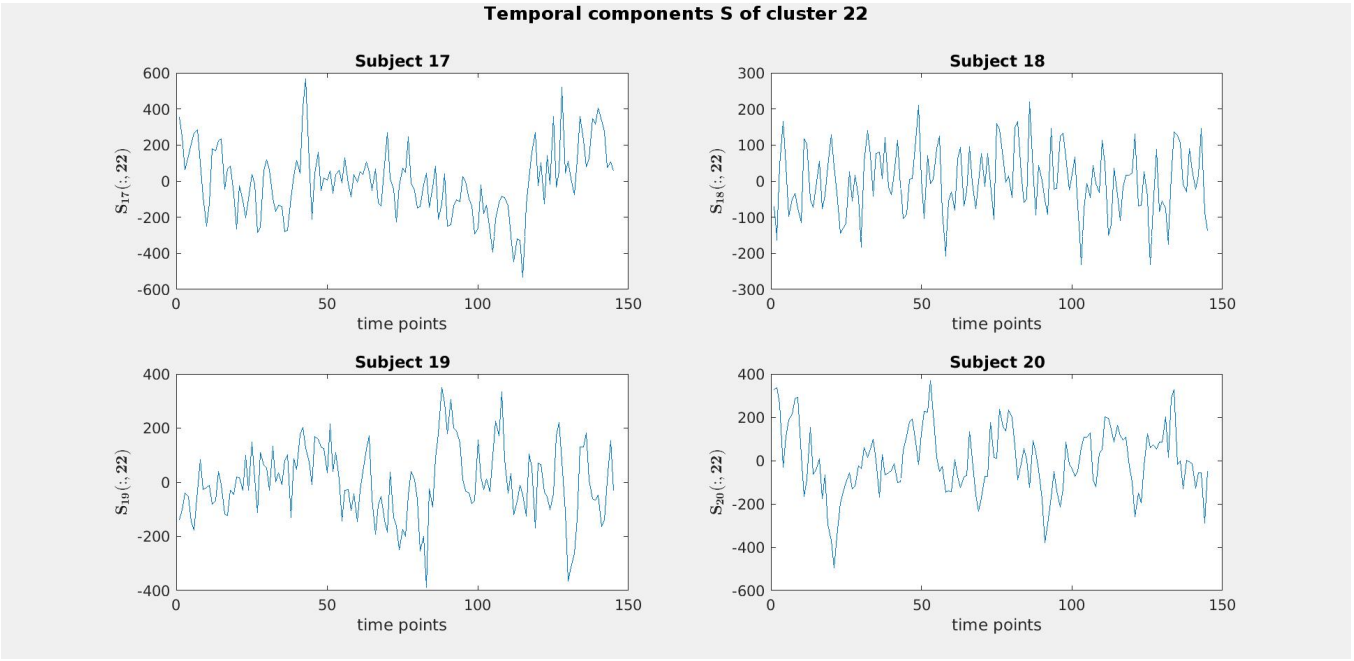


Figure B.50: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 17, \ldots, 20$ and R=27.

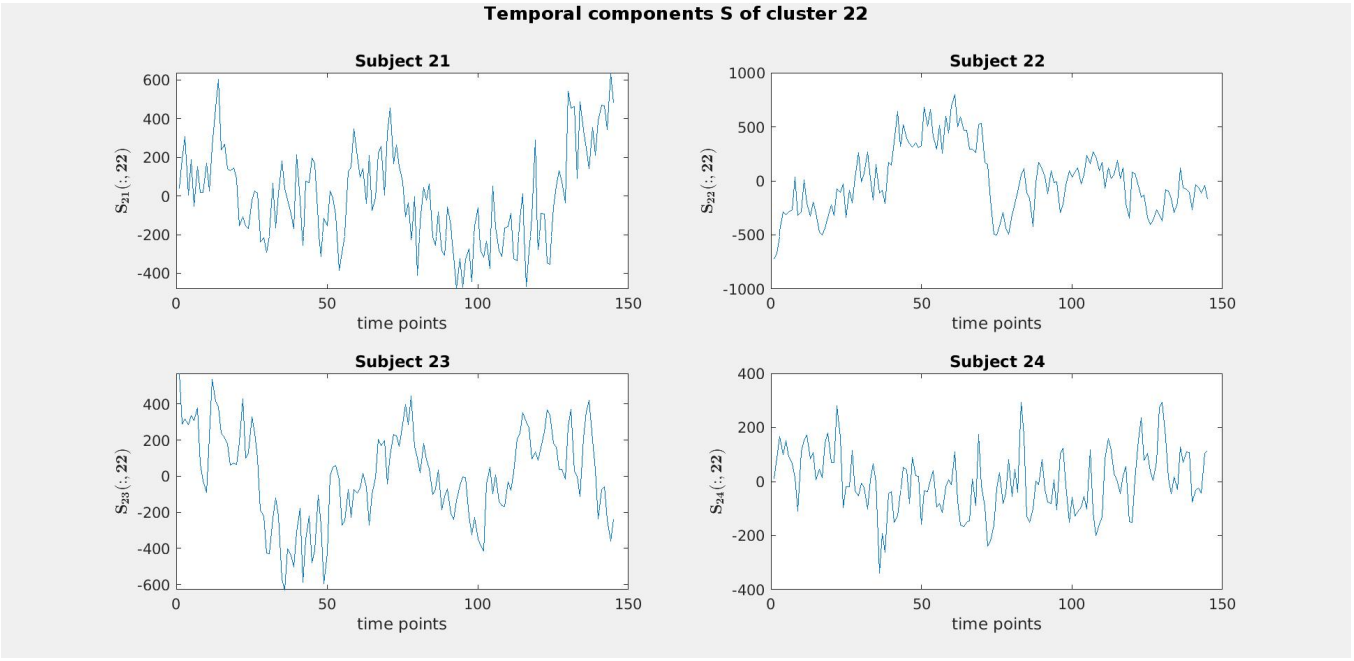


Figure B.51: Waveform of temporal components $\mathbf{S}_k(:, 122)$ for $k = 21, \ldots, 24$ and R=27.

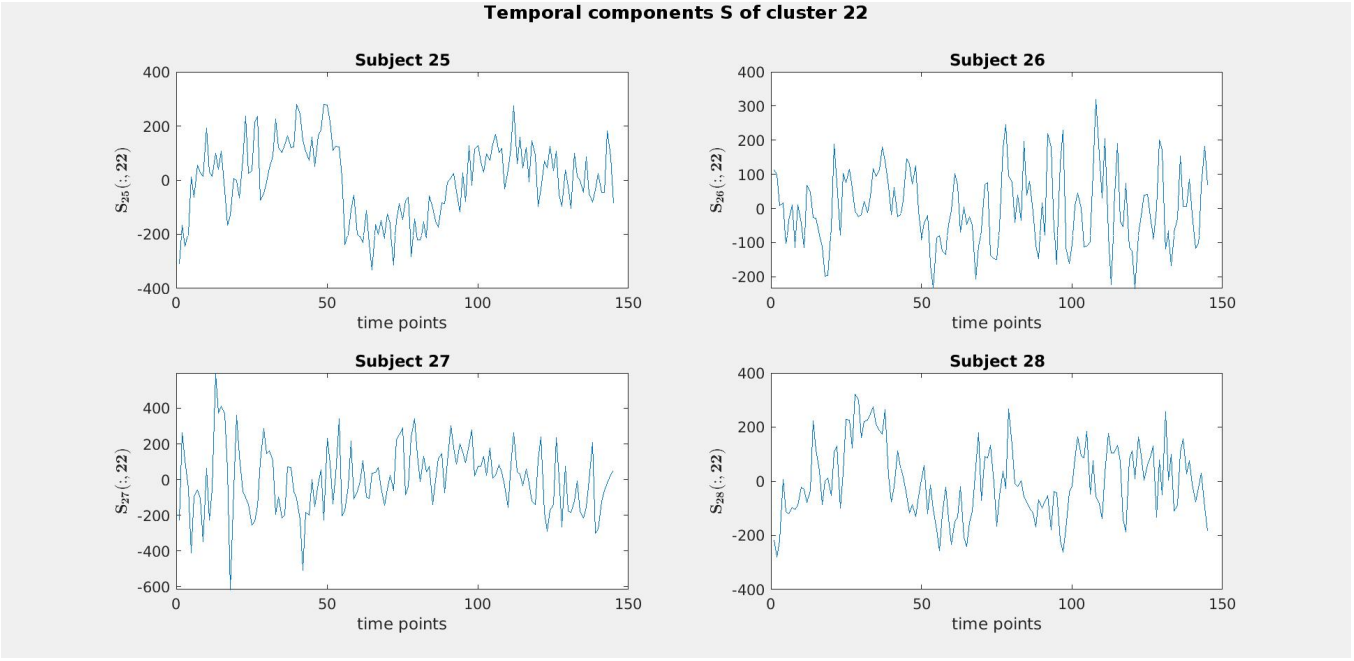


Figure B.52: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 25, \ldots, 28$ and R=27.

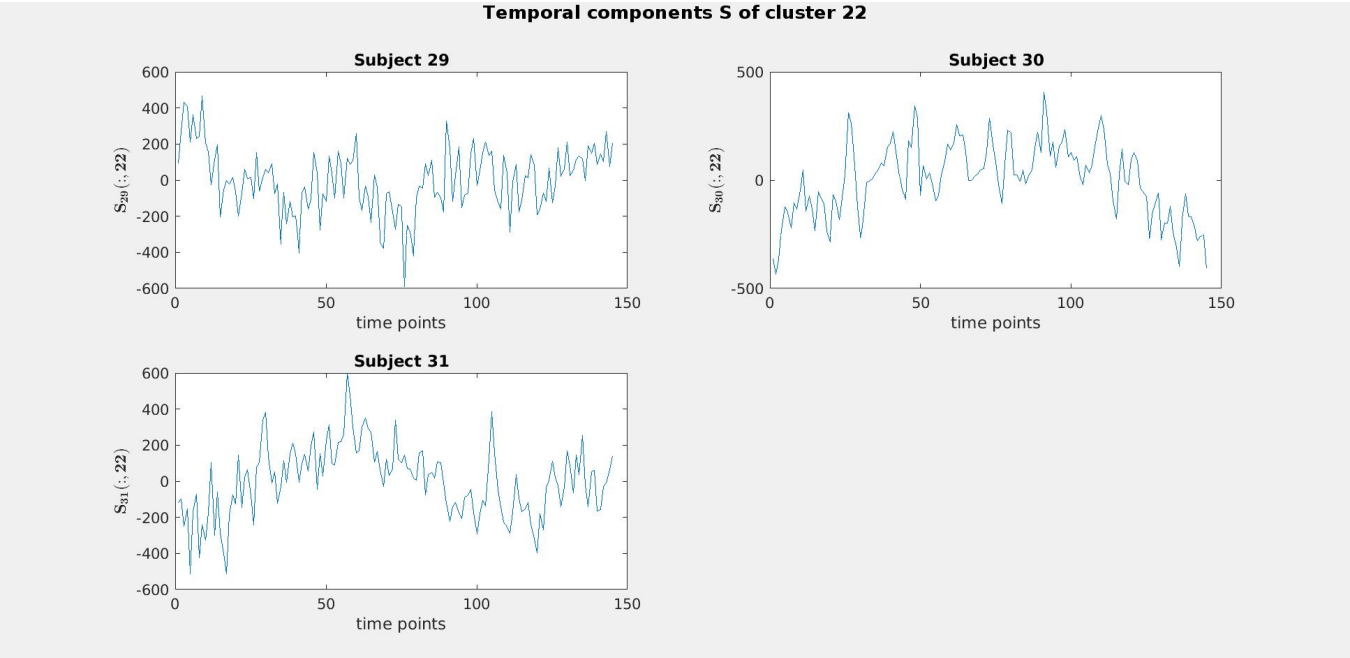


Figure B.53: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 29, \ldots, 31$ and R=27.

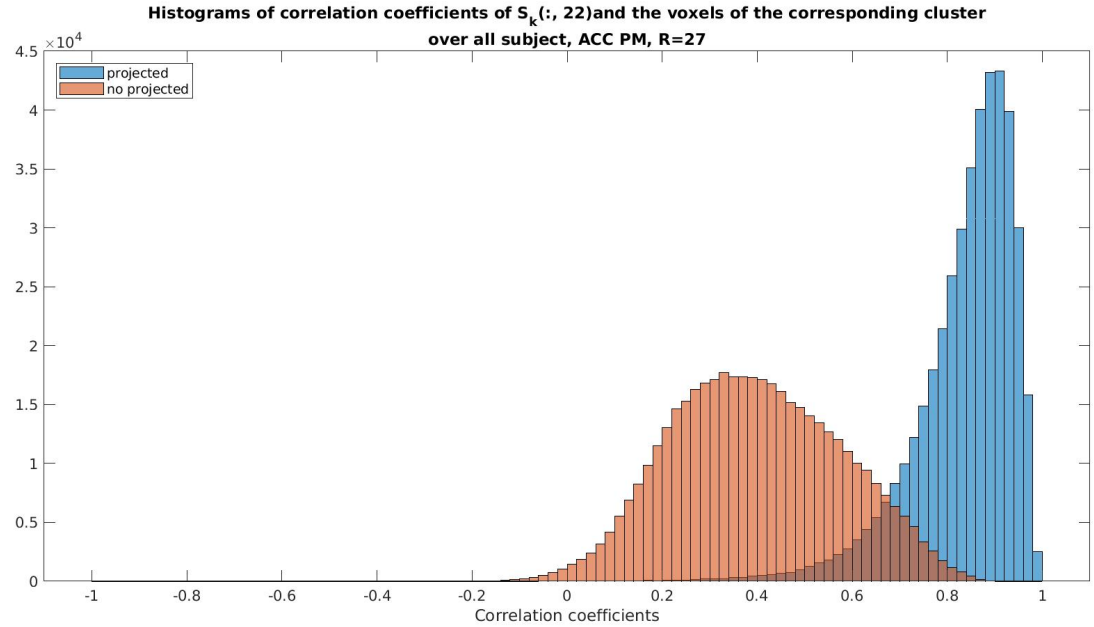


Figure B.54: For the column 22 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 22)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

4. **Cluster 7**
   Cluster 7 contains 13626 voxels, which correspond to 6.030591% of the whole brain (background, soft tissue, bone are excluded).

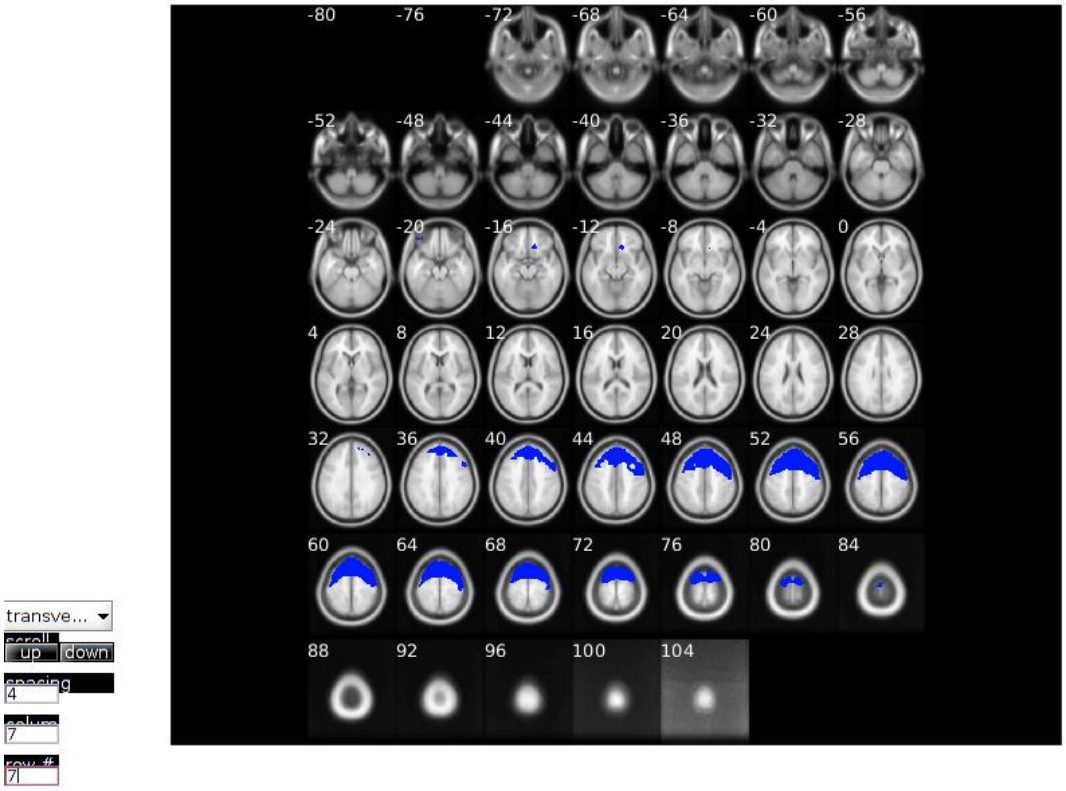


Figure B.55: Cluster 7. Blue areas indicate the voxels which belong to cluster 7.

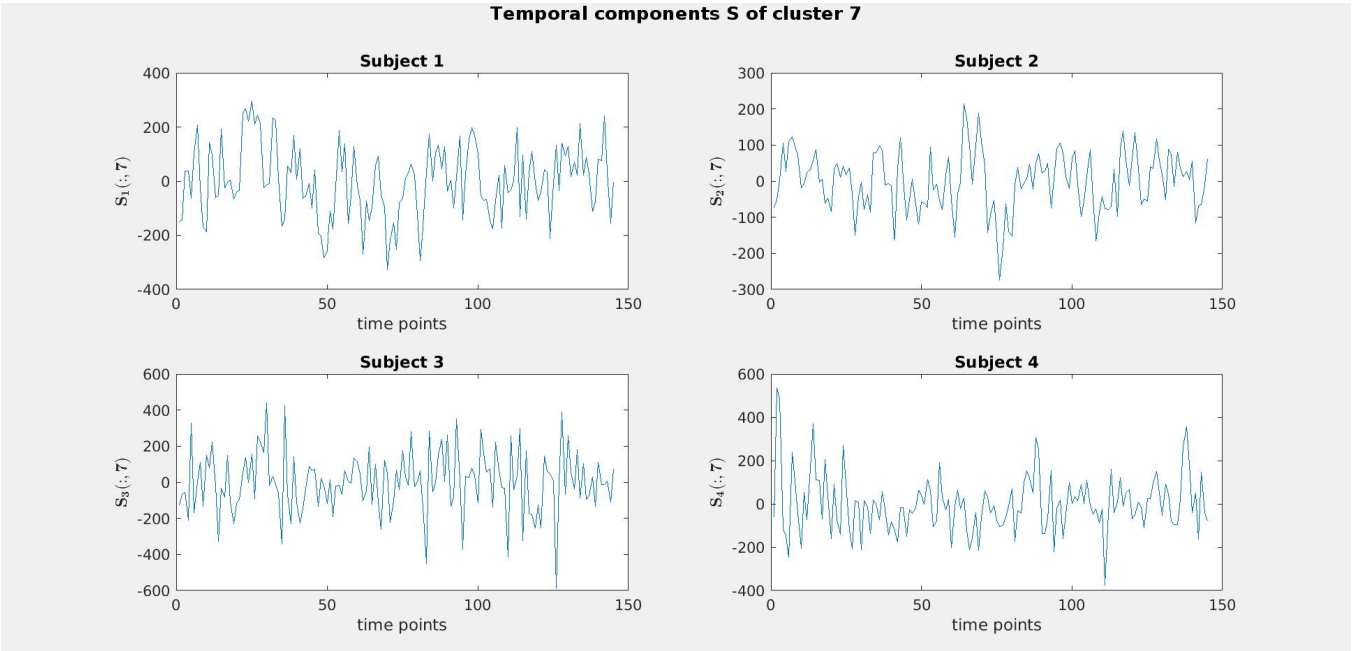


Figure B.56: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 1,\ldots,4$ and R=27.

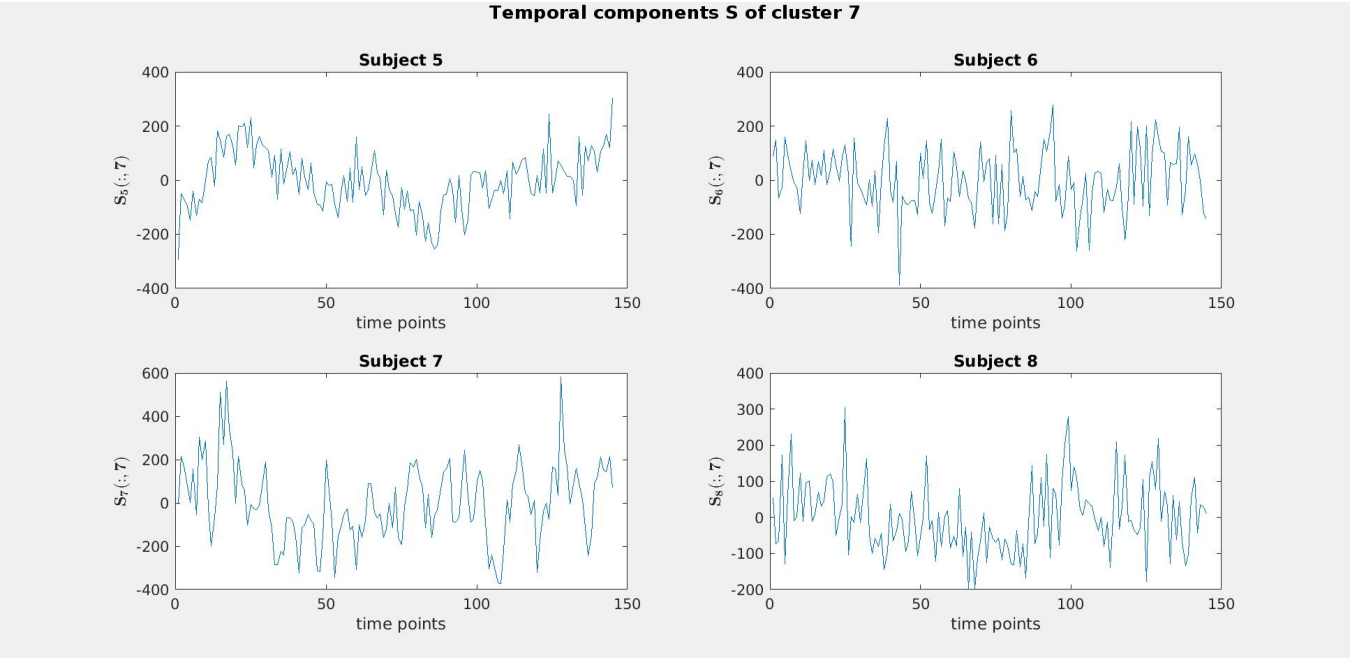


Figure B.57: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 5,\ldots,8$ and R=27.

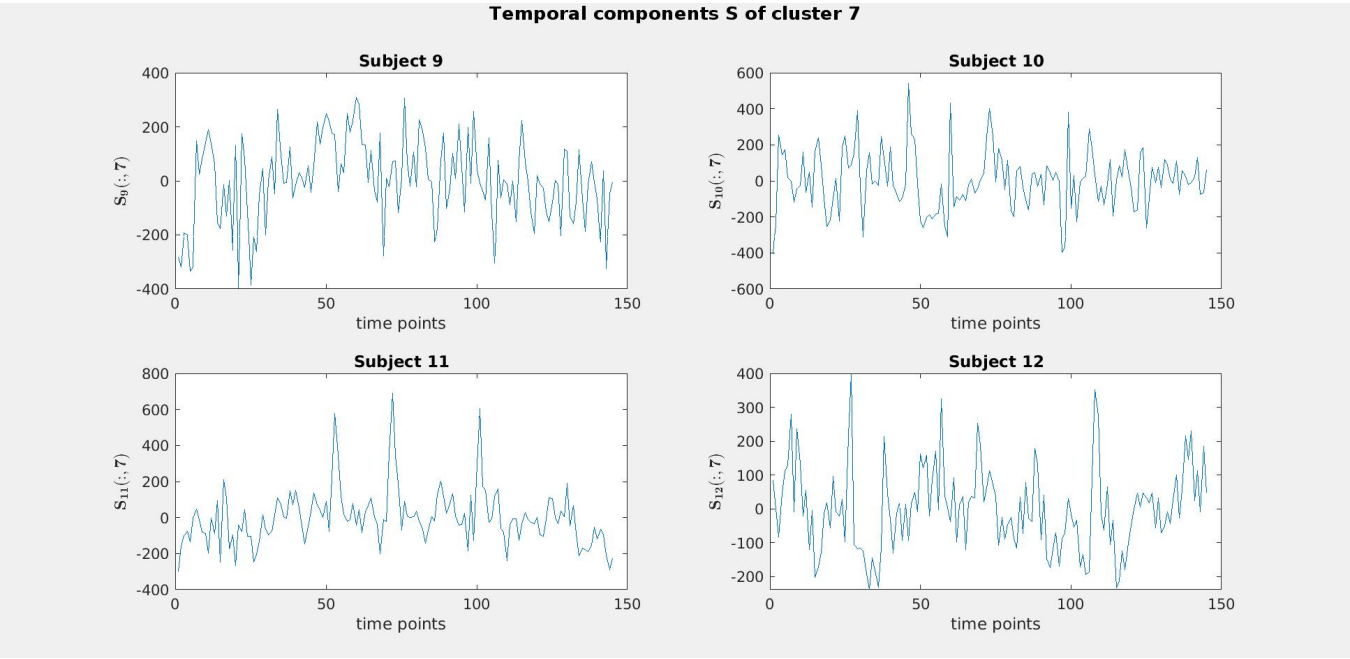


Figure B.58: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 9,\ldots,12$ and R=27.

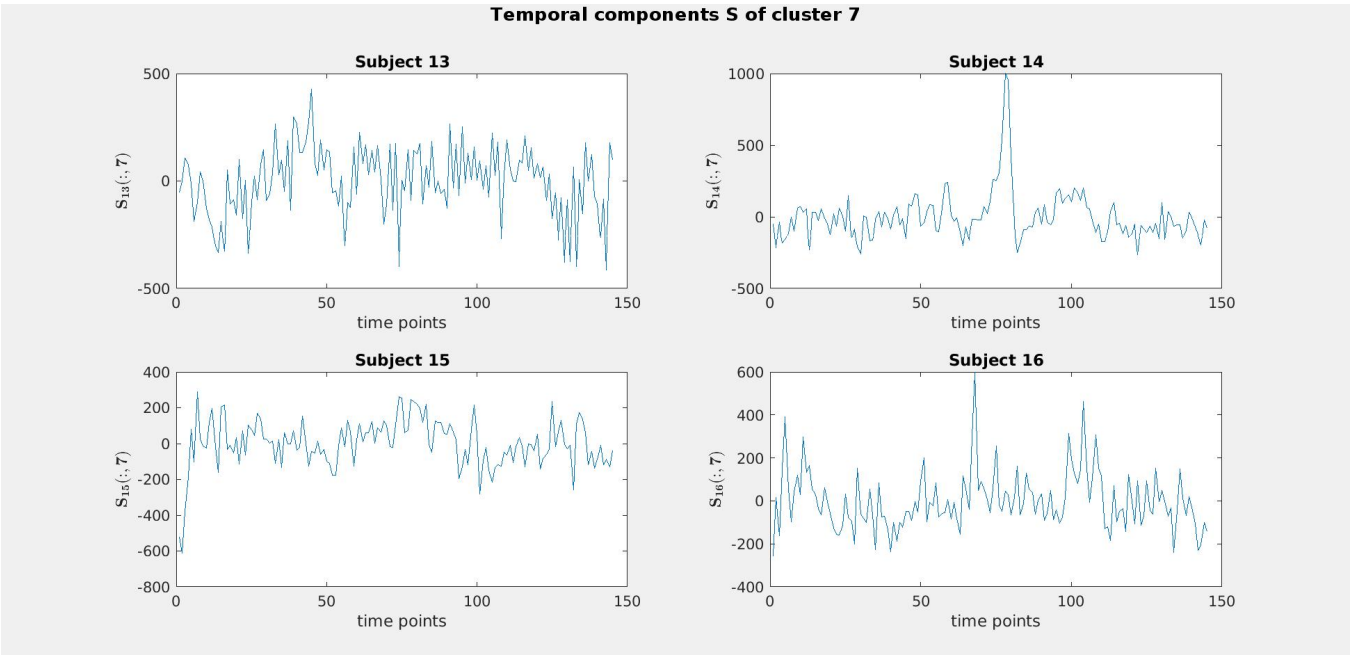


Figure B.59: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 13,\ldots,16$ and R=27.

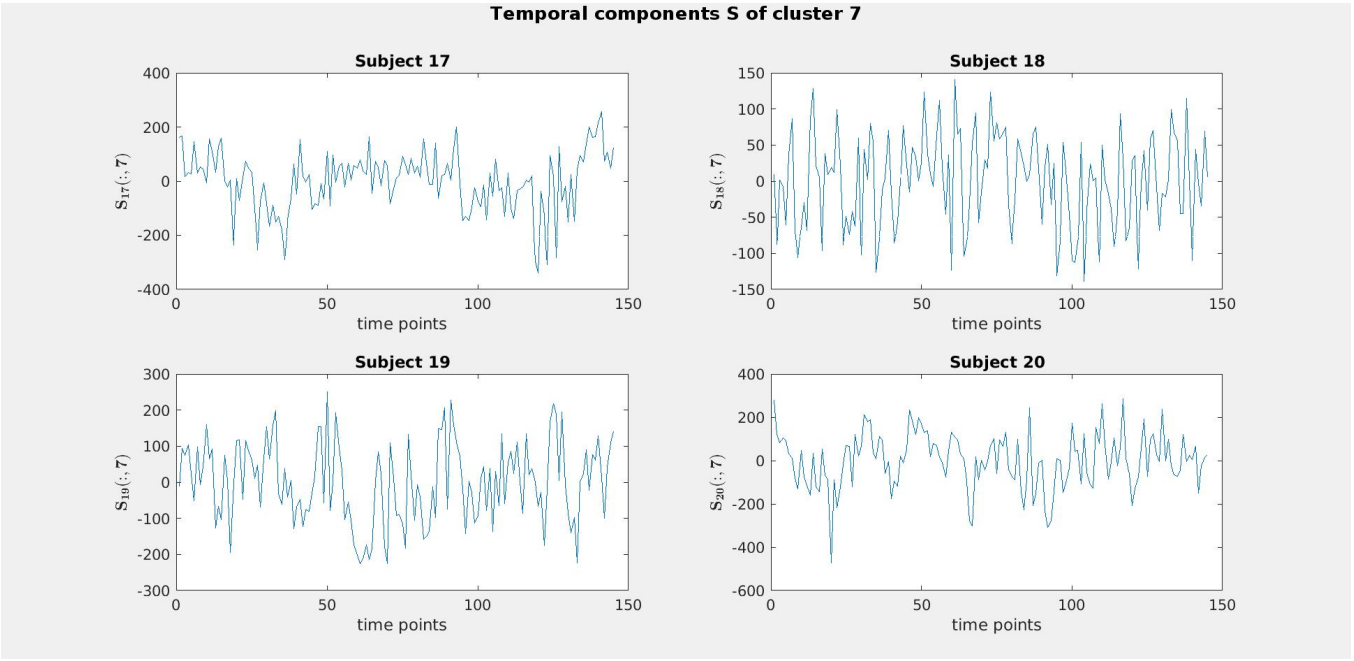


Figure B.60: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 17,\ldots,20$ and R=27.

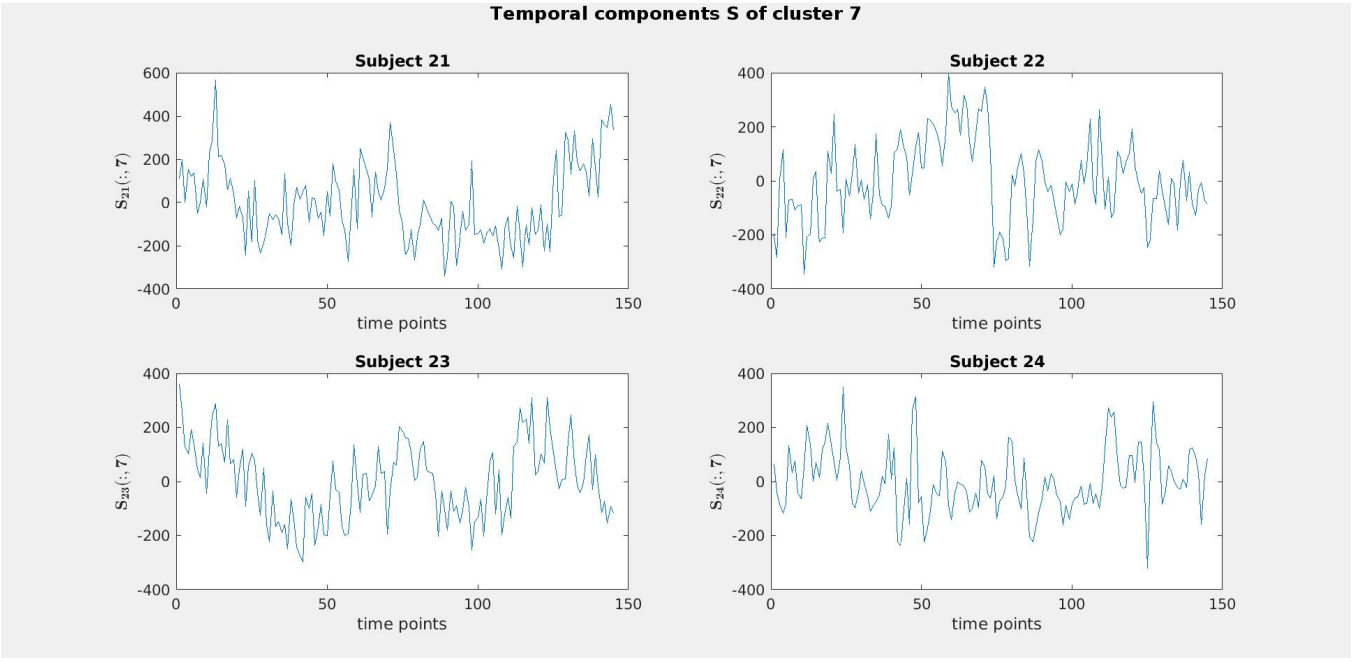


Figure B.61: Waveform of temporal components $\mathbf{S}_k(:,17)$ for $k = 21,\ldots,24$ and R=27.

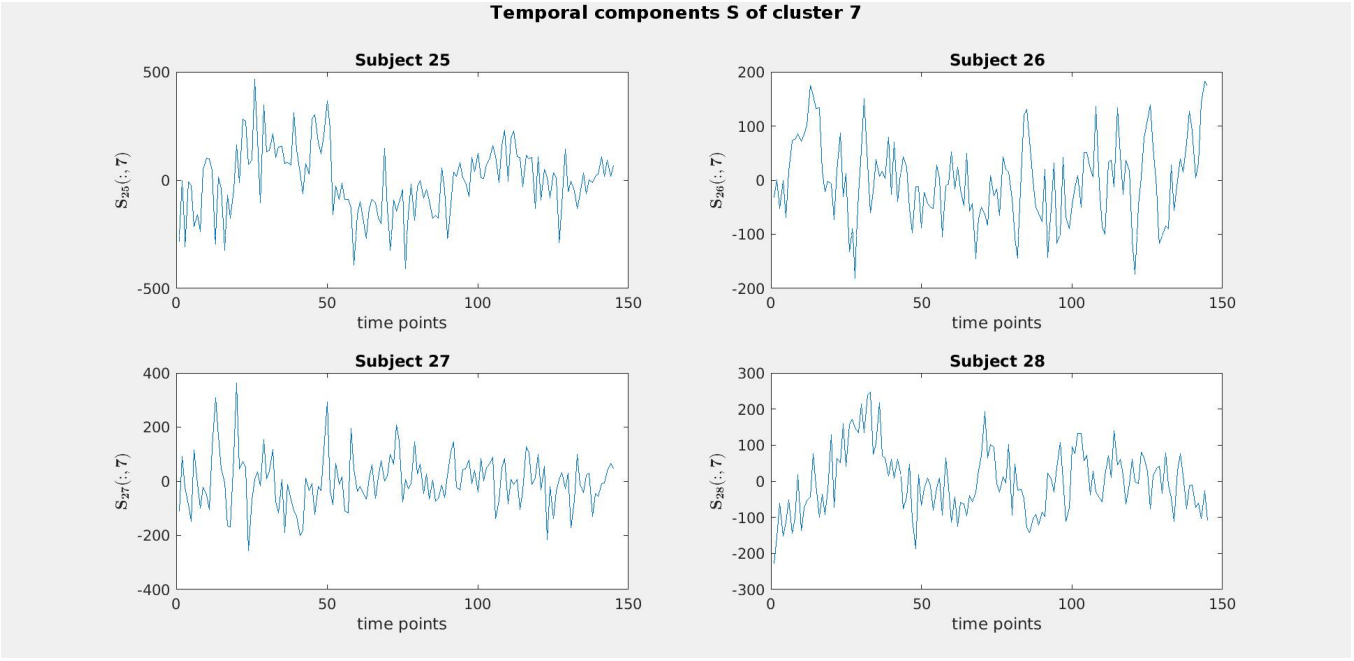


Figure B.62: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 25,\ldots,28$ and R=27.

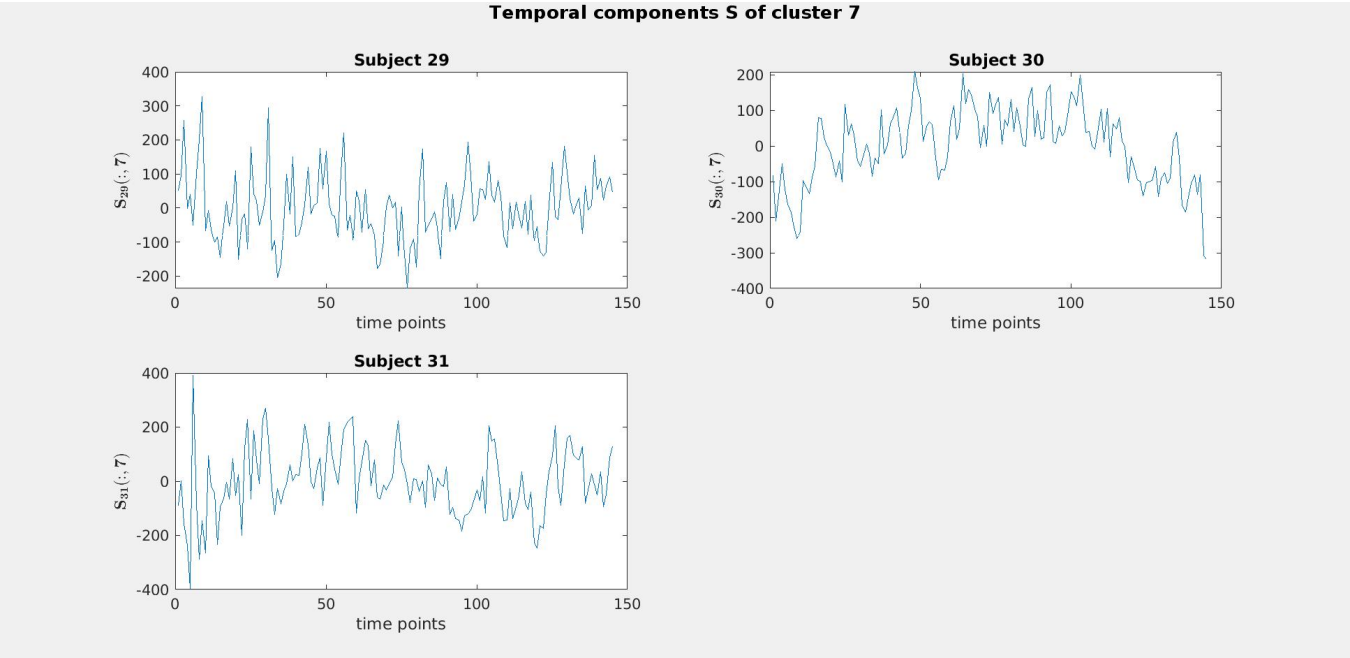


Figure B.63: Waveform of temporal components $\mathbf{S}_k(:, 7)$ for $k = 29, \ldots, 31$ and R=27.

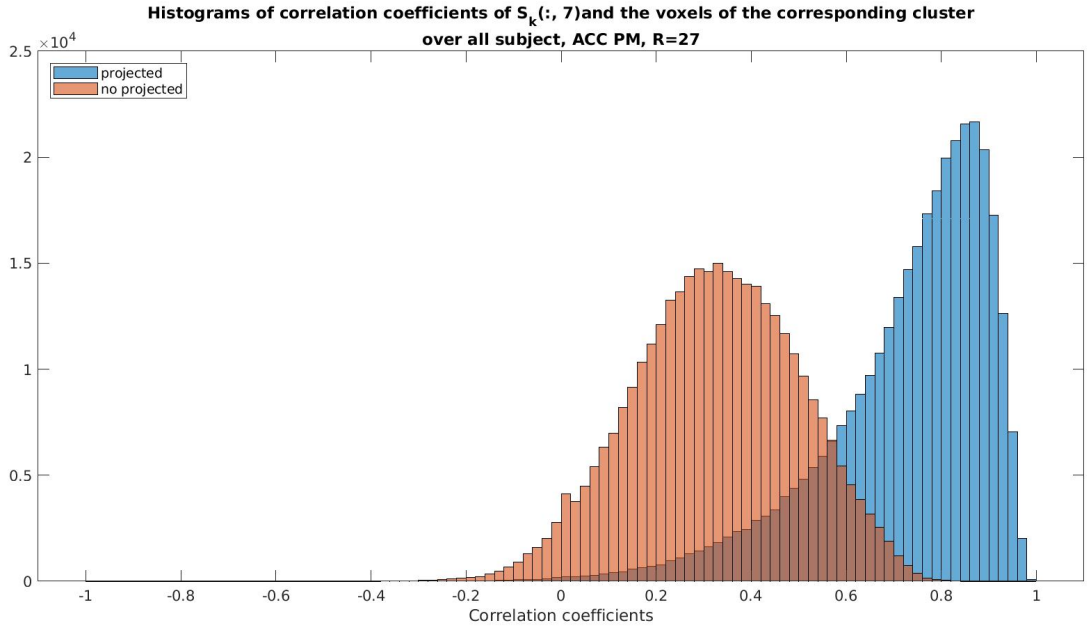


Figure B.64: For the column 7 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 7)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

5. **Cluster 8**
   Cluster 8 contains 11279 voxels, which correspond to 4.991857% of the whole brain (background, soft tissue, bone are excluded).

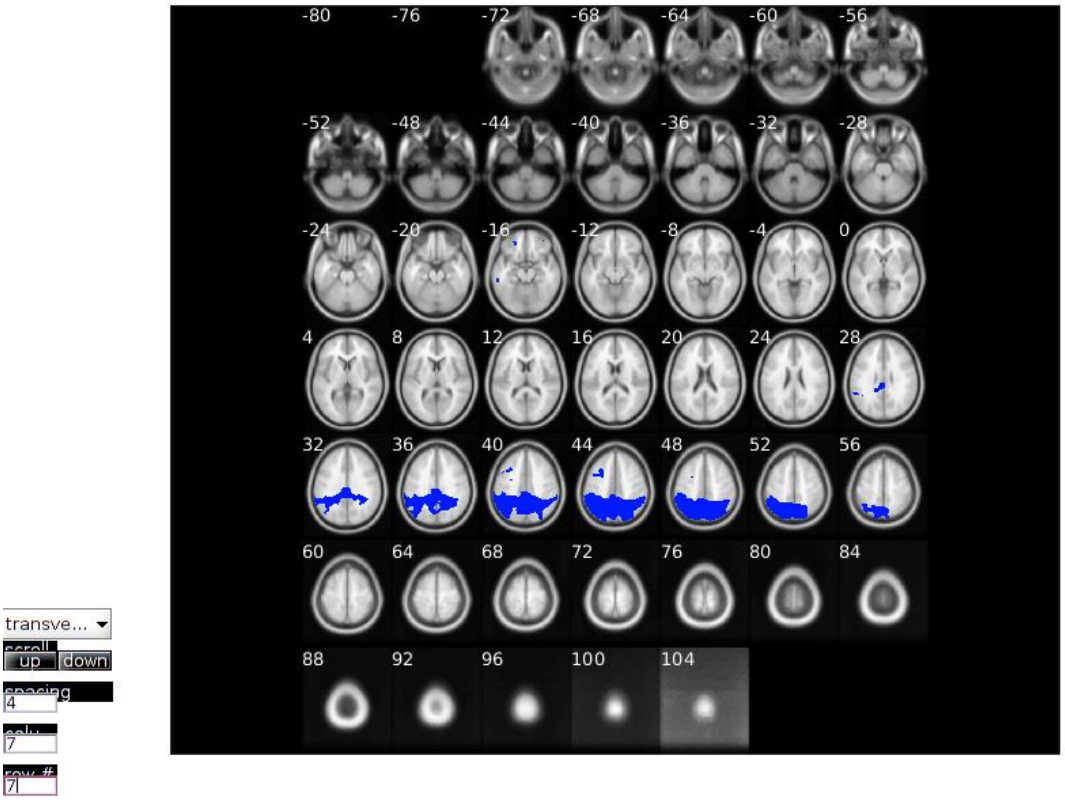


Figure B.65: Cluster 8. Blue areas indicate the voxels which belong to cluster 8.

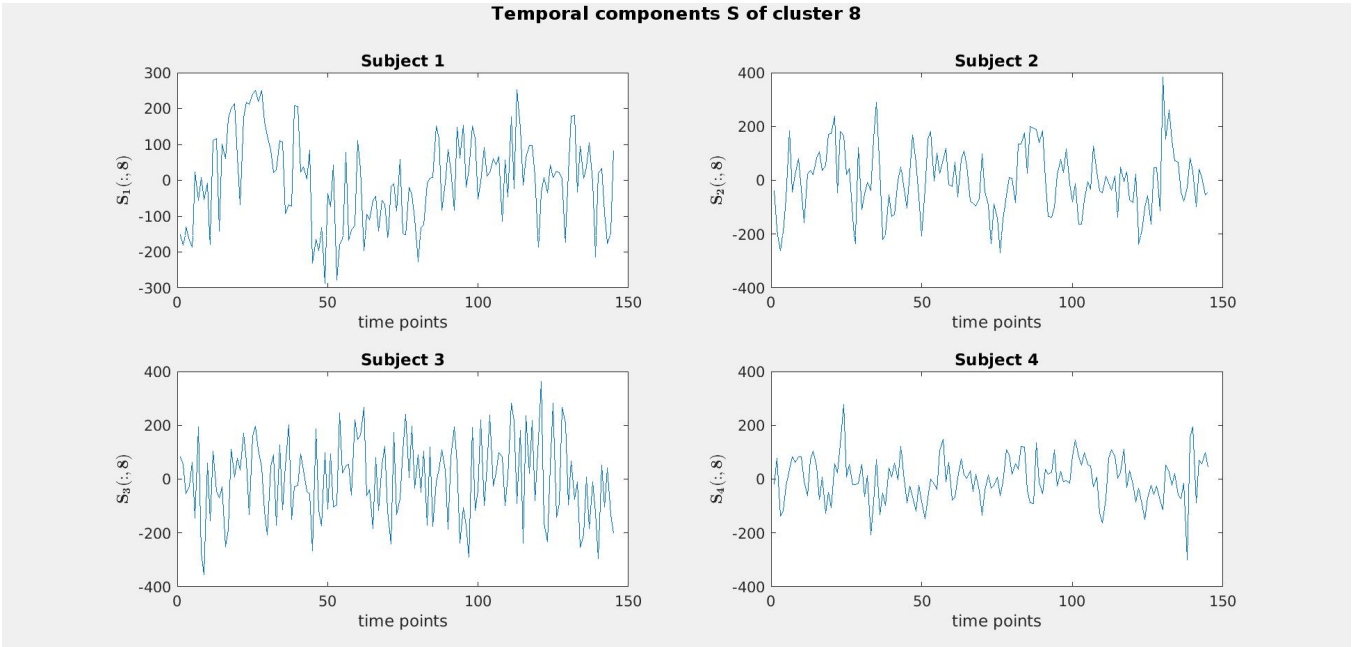


Figure B.66: Waveform of temporal components $\mathbf{S}_k(:,8)$ for $k = 1,\ldots,4$ and R=27.

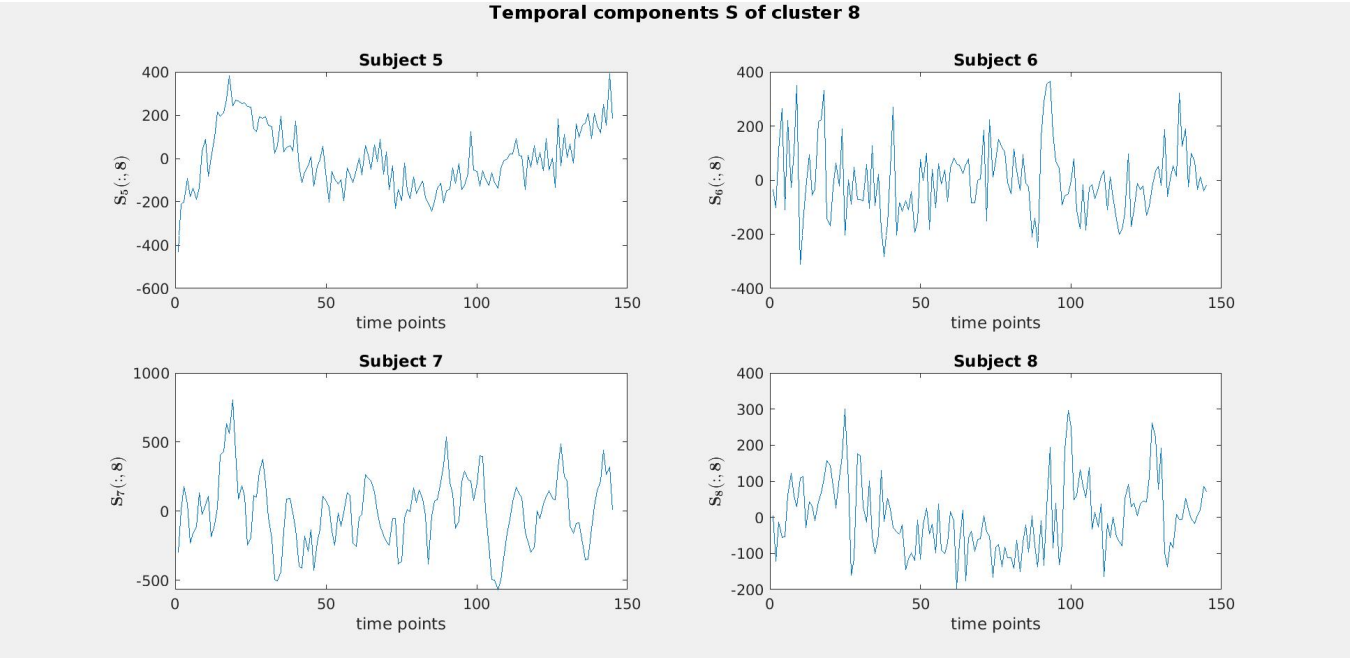


Figure B.67: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 5, \ldots, 8$ and R=27.

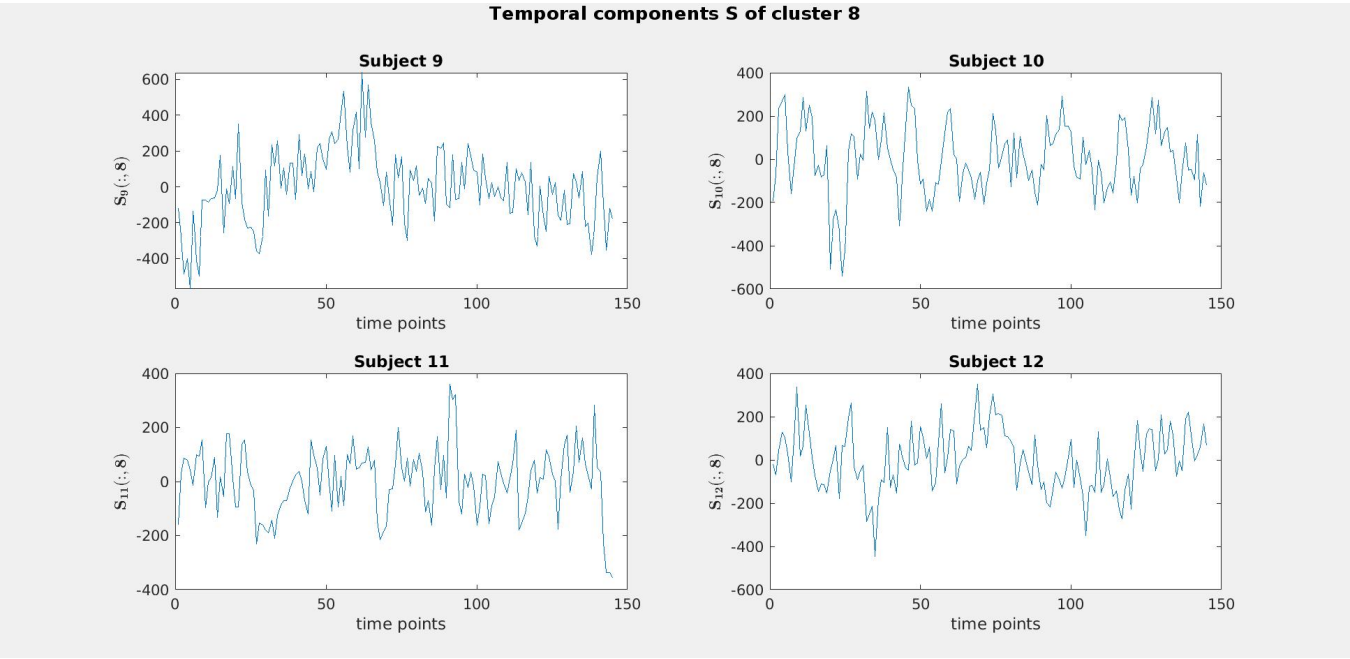


Figure B.68: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 9, \ldots, 12$ and R=27.

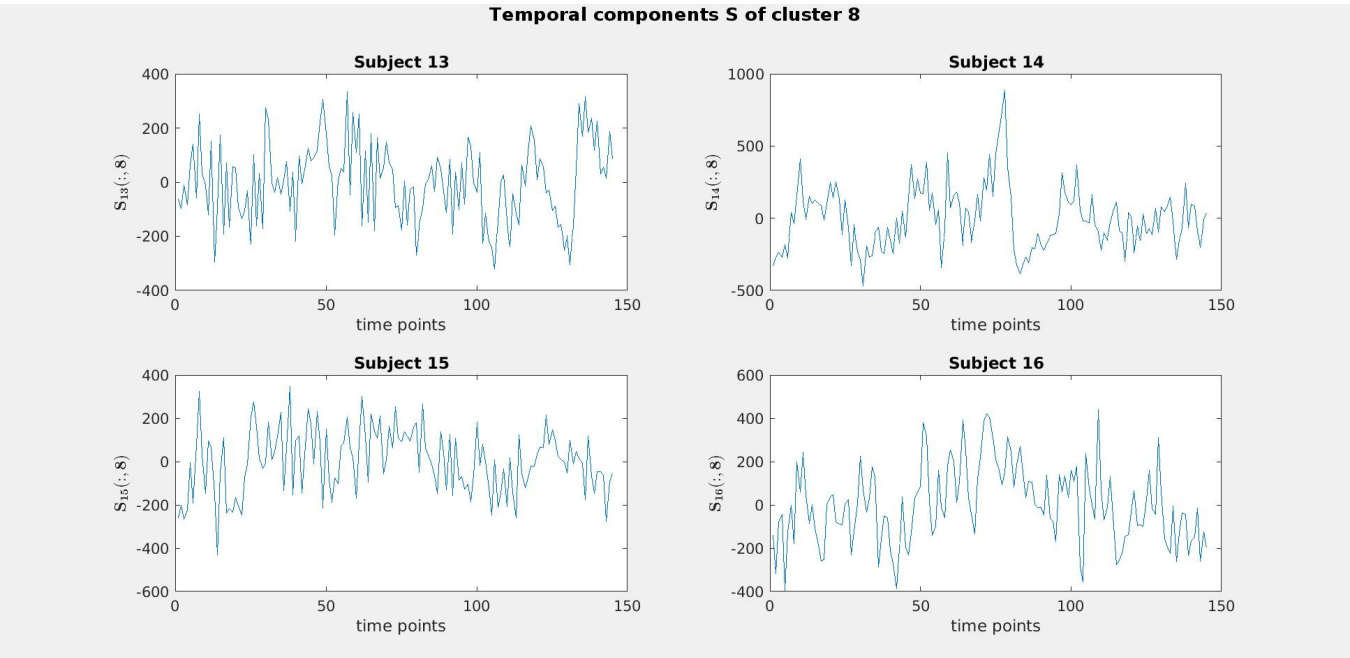


Figure B.69: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 13, \ldots, 16$ and R=27.

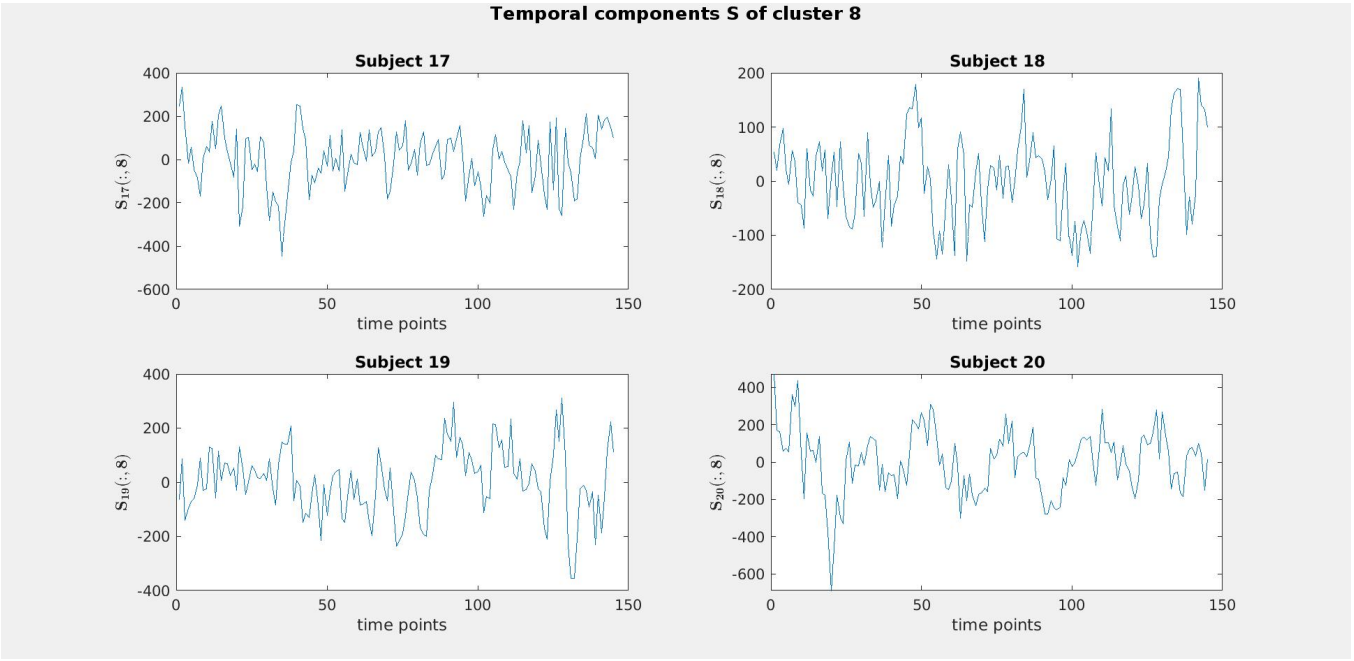


Figure B.70: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 17, \ldots, 20$ and R=27.

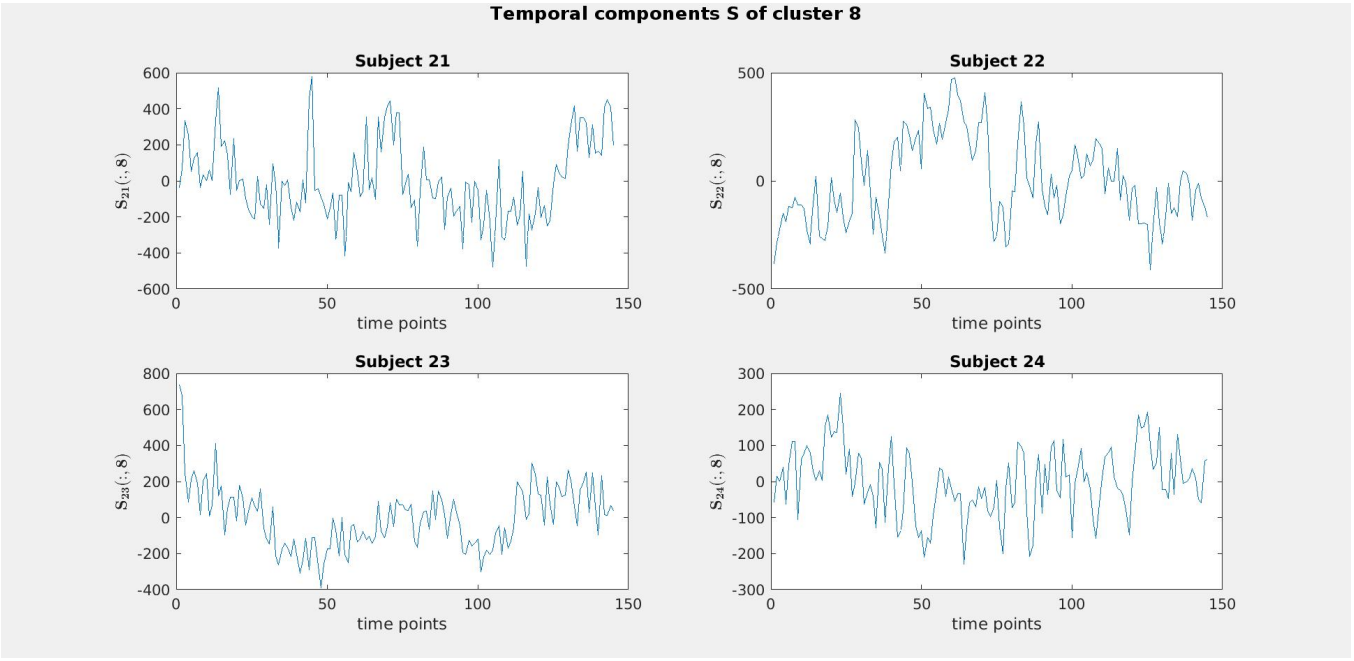


Figure B.71: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 21, \ldots, 24$ and R=27.

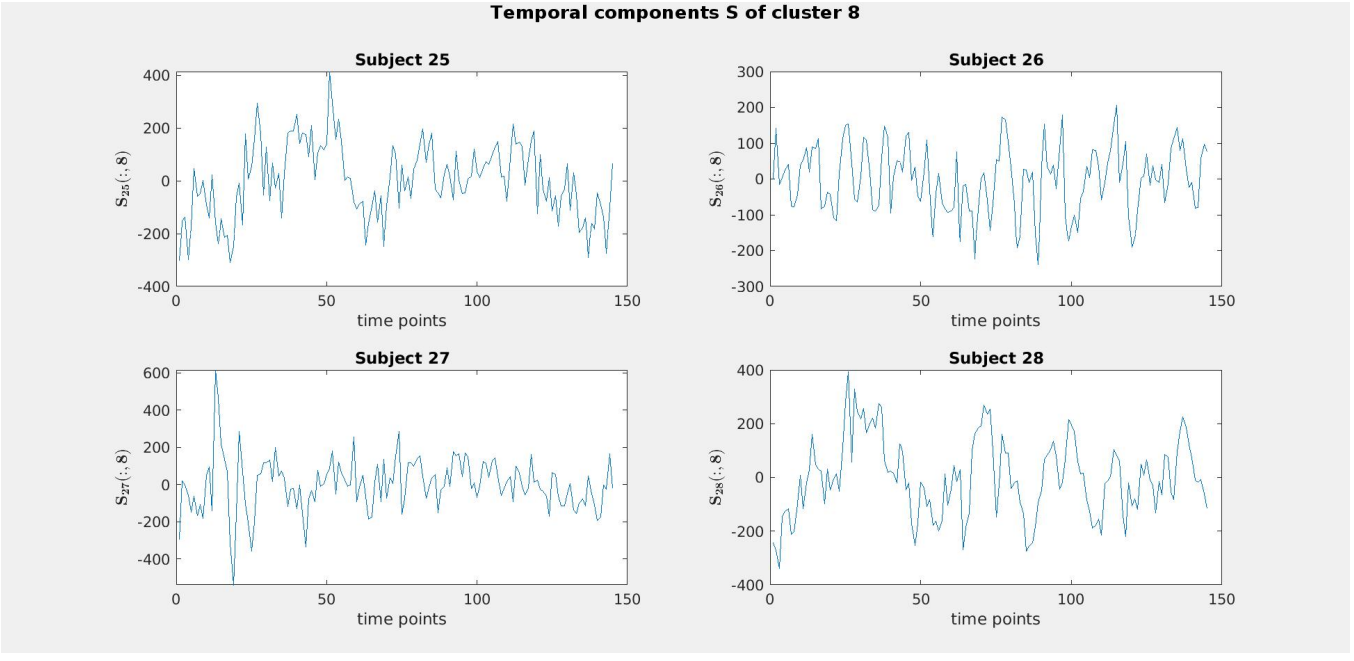


Figure B.72: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 25, \ldots, 28$ and R=27.

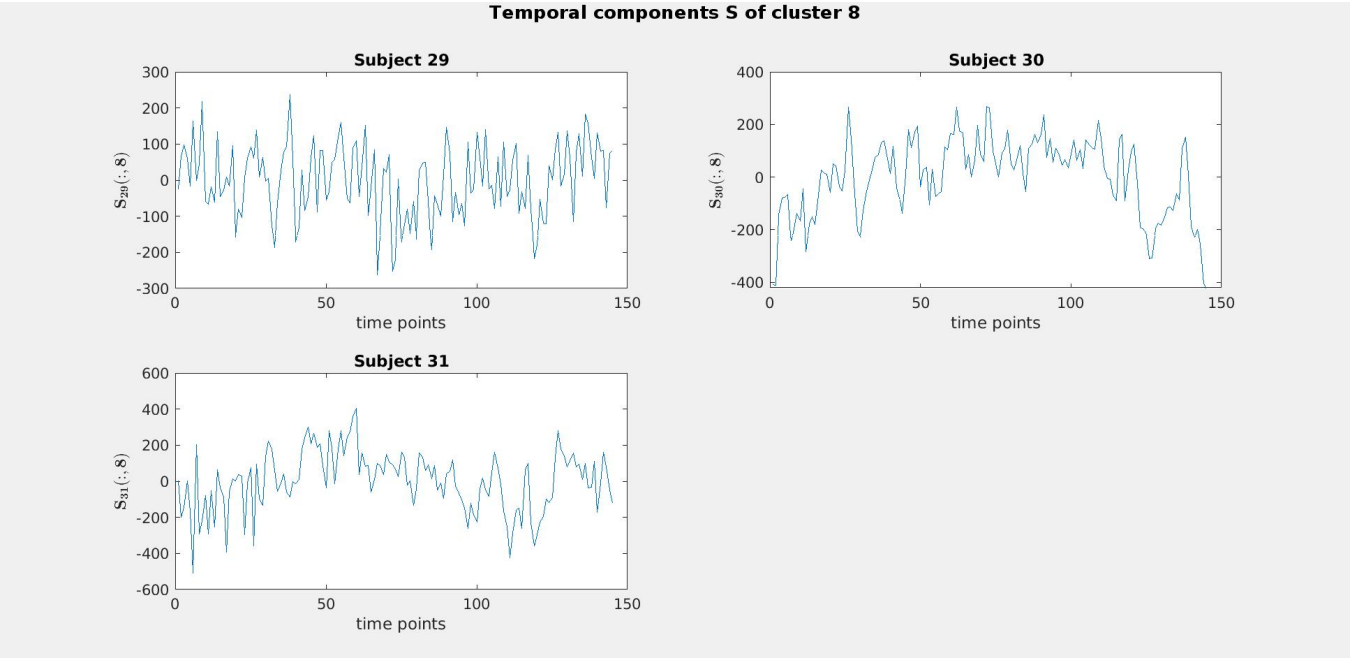


Figure B.73: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 29, \ldots, 31$ and R=27.

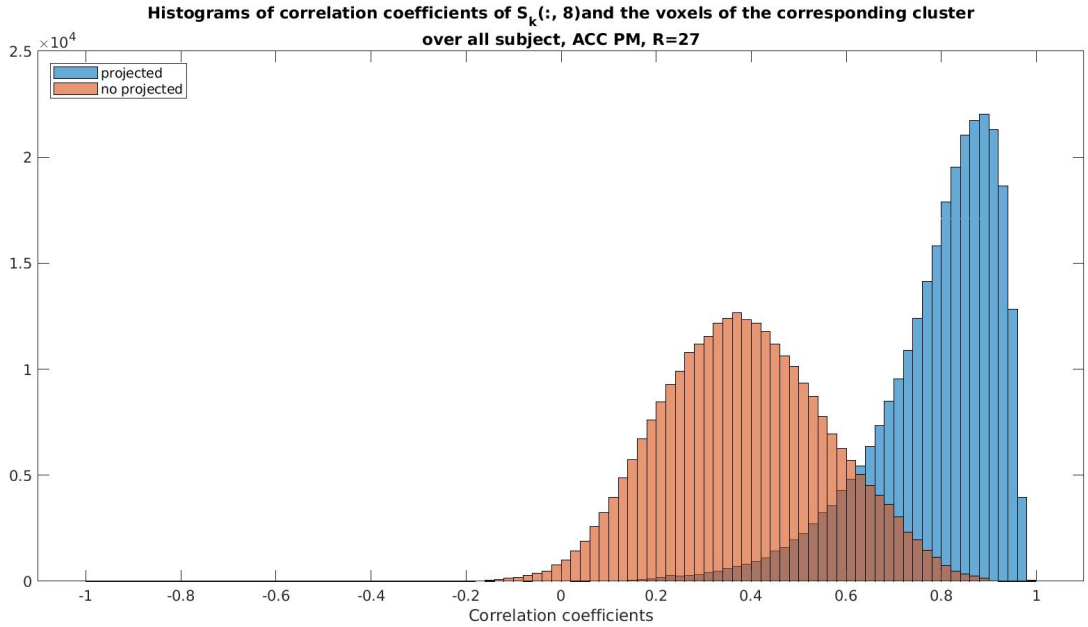


Figure B.74: For the column 8 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 8)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

6. **Cluster 10**
   Cluster 10 contains 10781 voxels, which correspond to 4.771452% of the whole brain (background, soft tissue, bone are excluded).

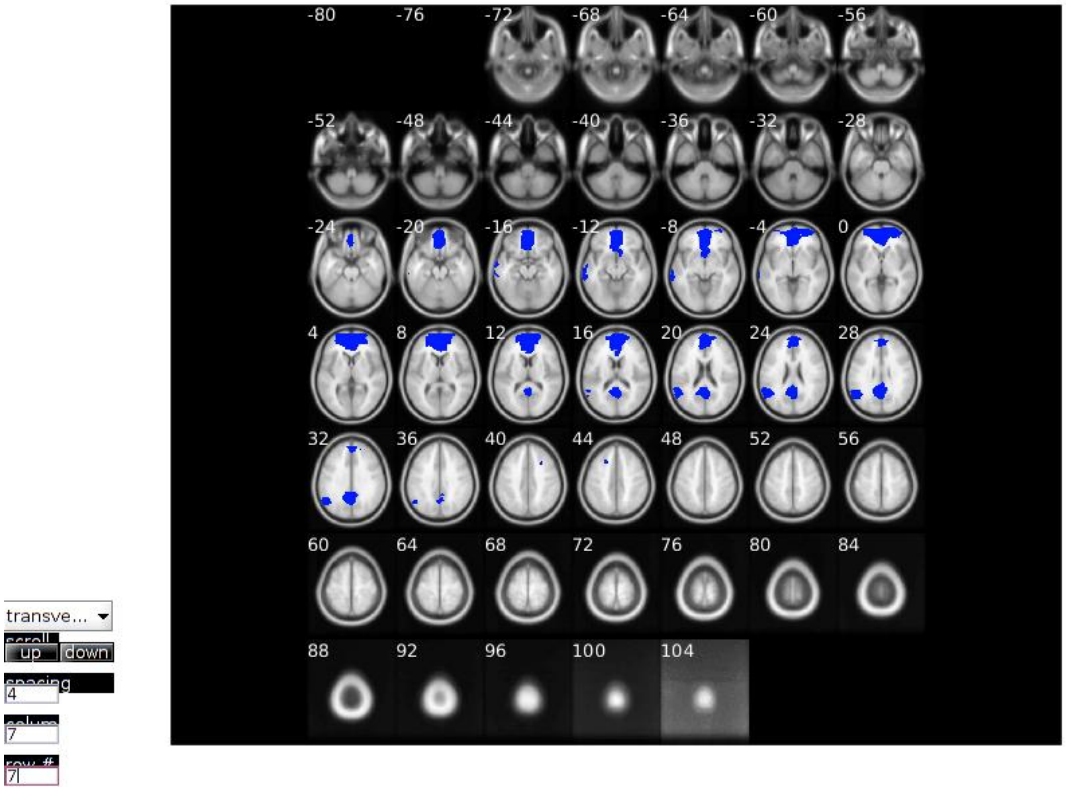


Figure B.75: Cluster 10. Blue areas indicate the voxels which belong to cluster 10.

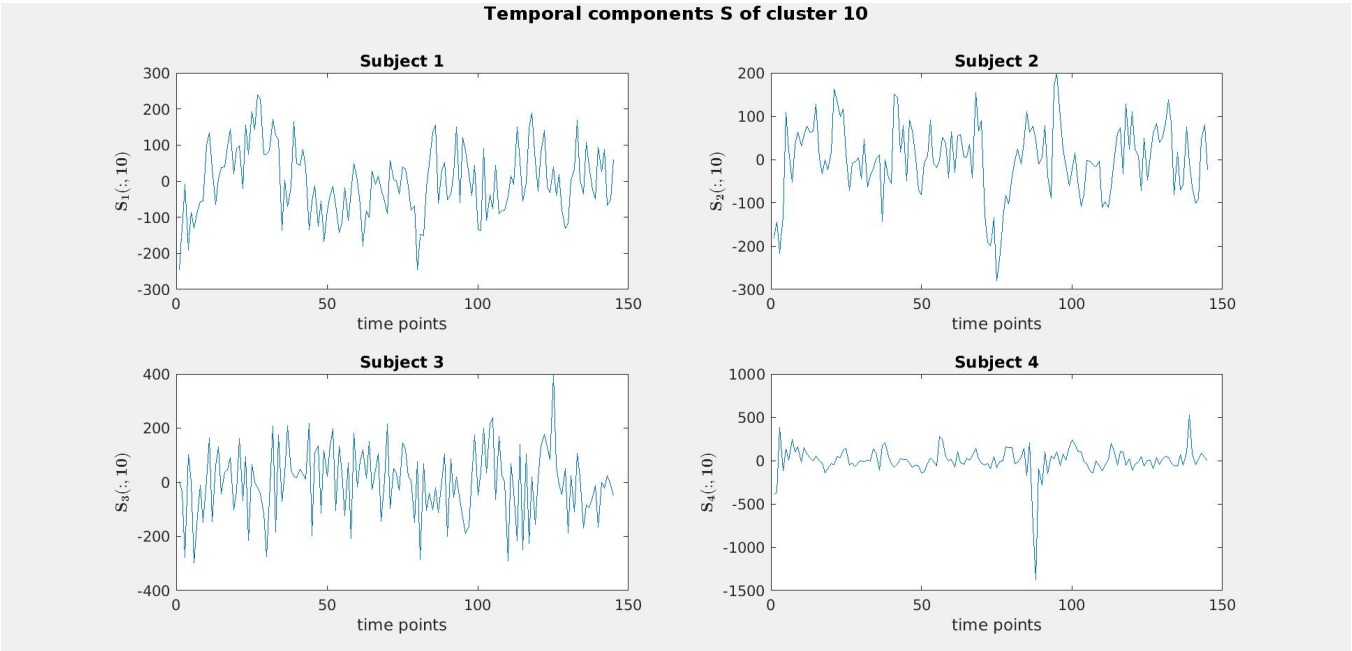


Figure B.76: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 1, \ldots, 4$ and R=27.

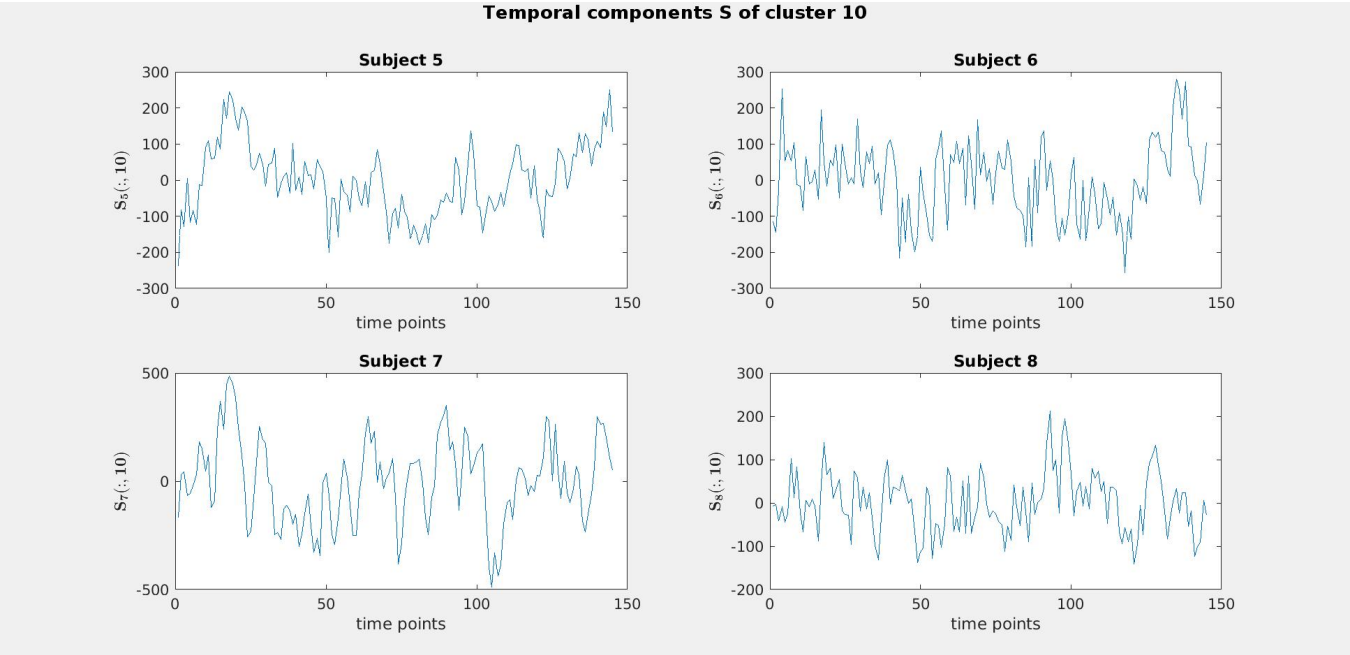


Figure B.77: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 5, \ldots, 8$ and R=27.

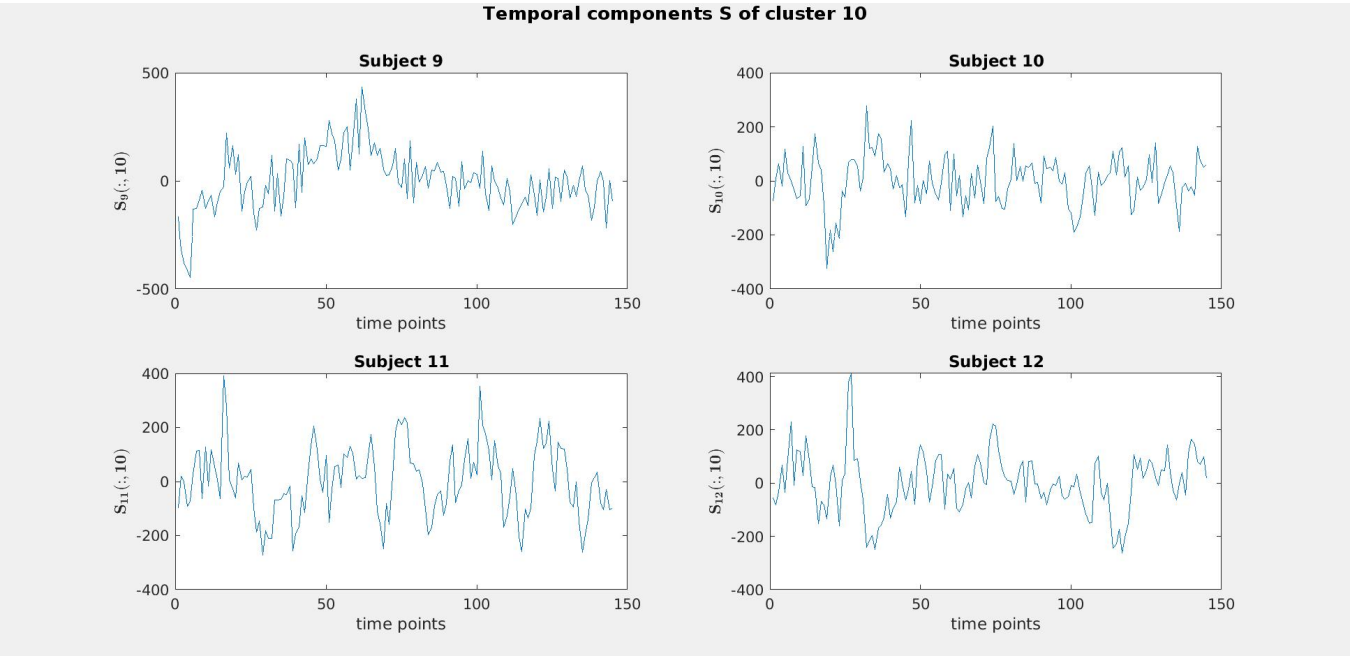


Figure B.78: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 9, \ldots, 12$ and R=27.

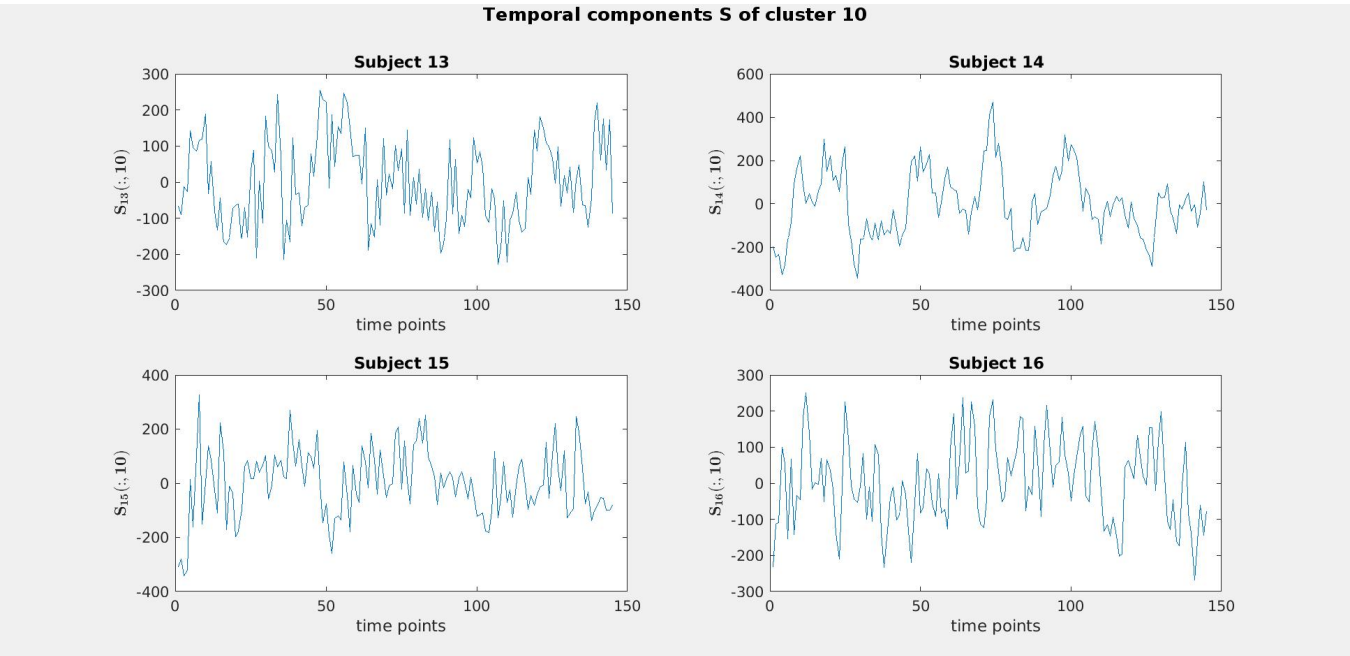


Figure B.79: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 13, \ldots, 16$ and R=27.

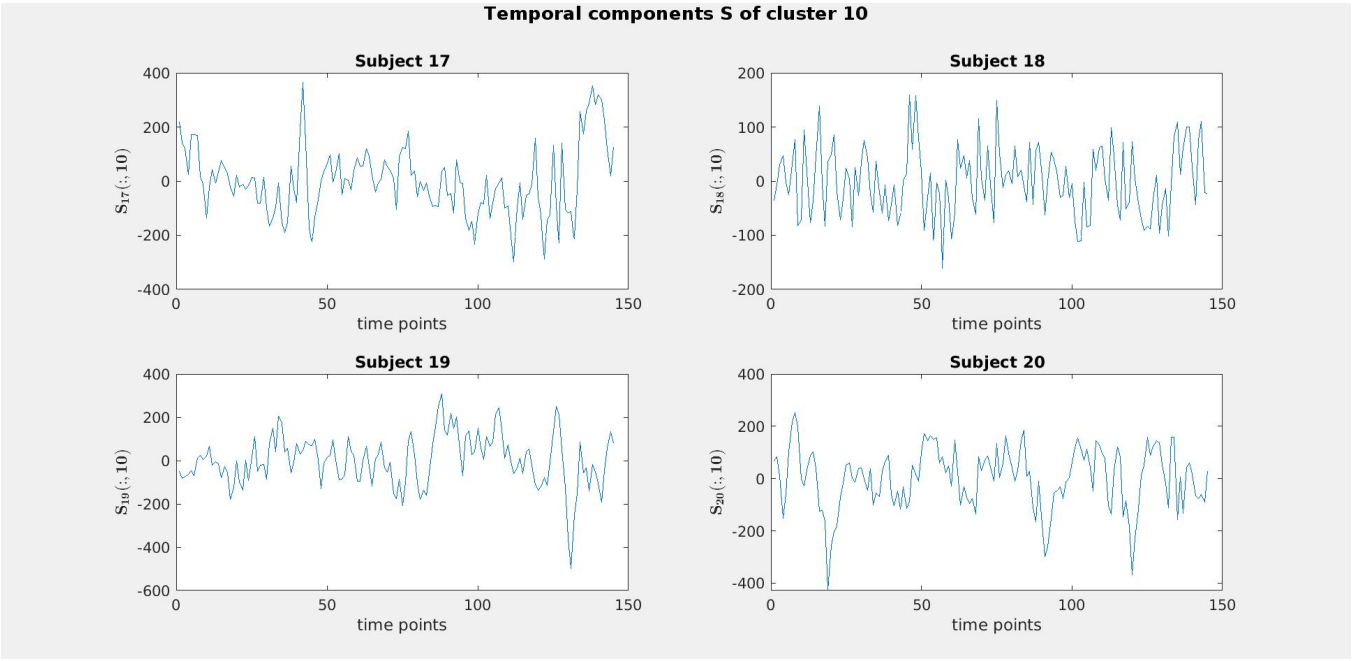


Figure B.80: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 17, \ldots, 20$ and R=27.

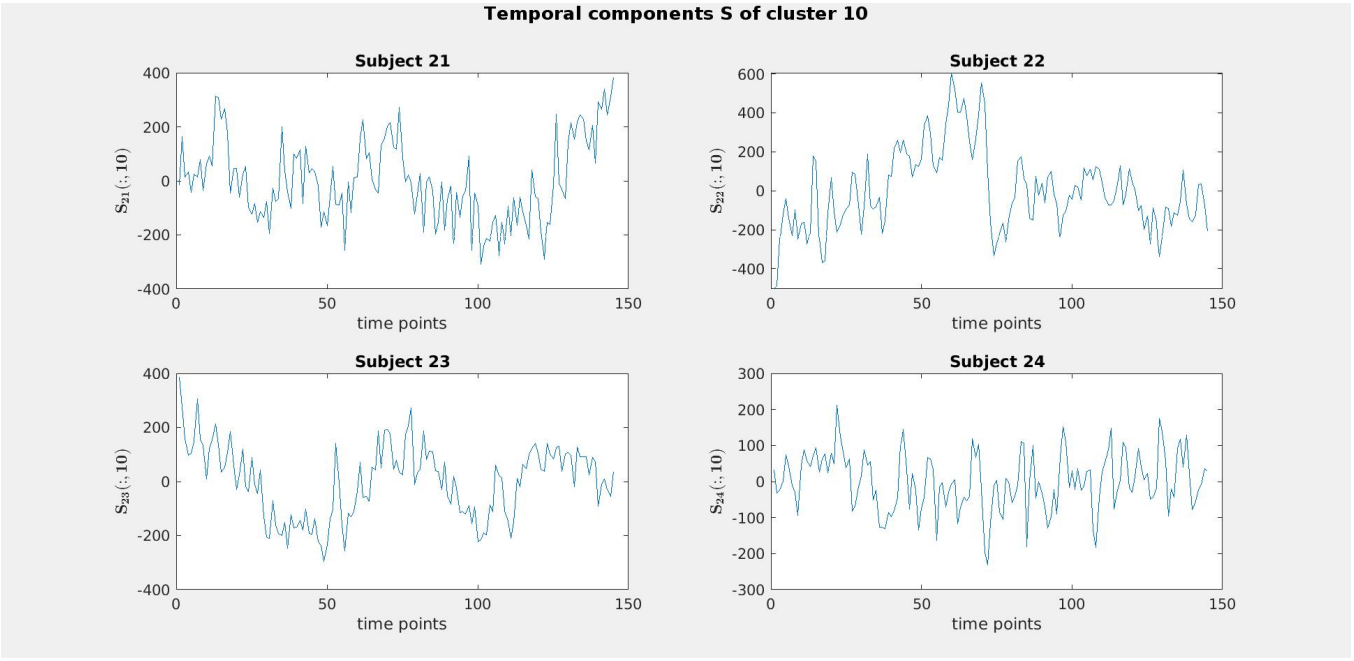


Figure B.81: Waveform of temporal components $\mathbf{S}_k(:, 110)$ for $k = 21, \ldots, 24$ and R=27.

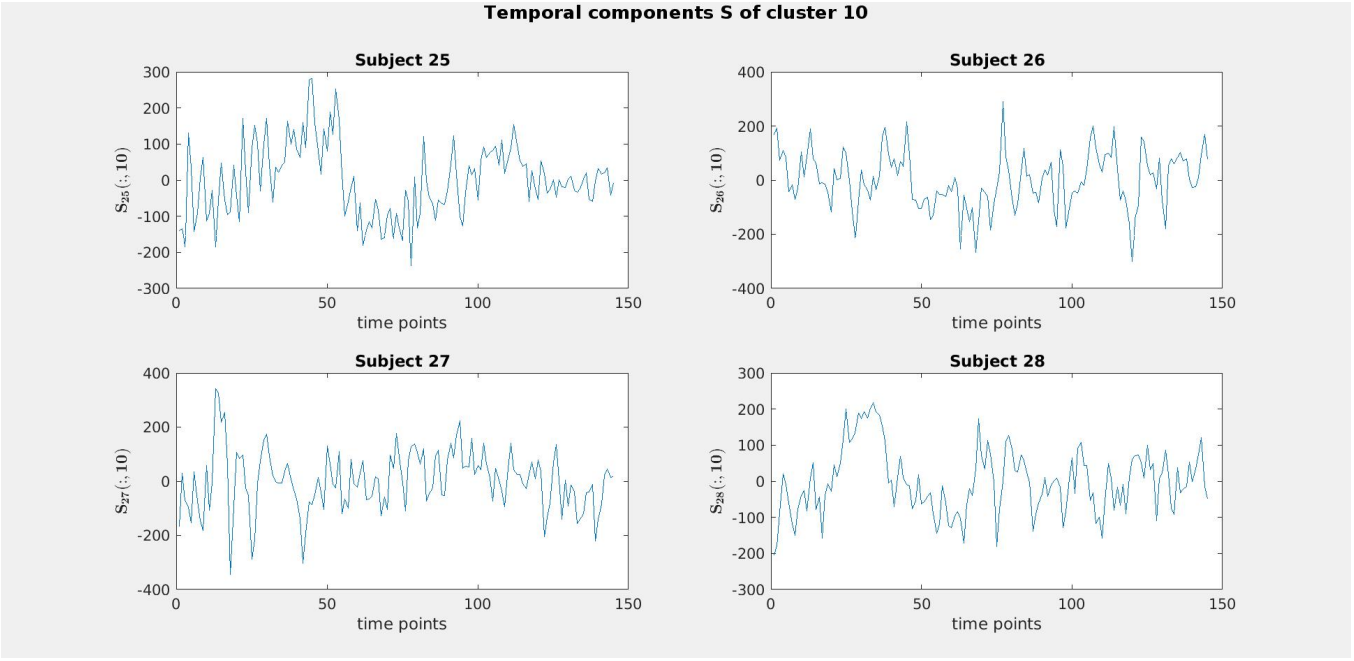


Figure B.82: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 25, \ldots, 28$ and R=27.

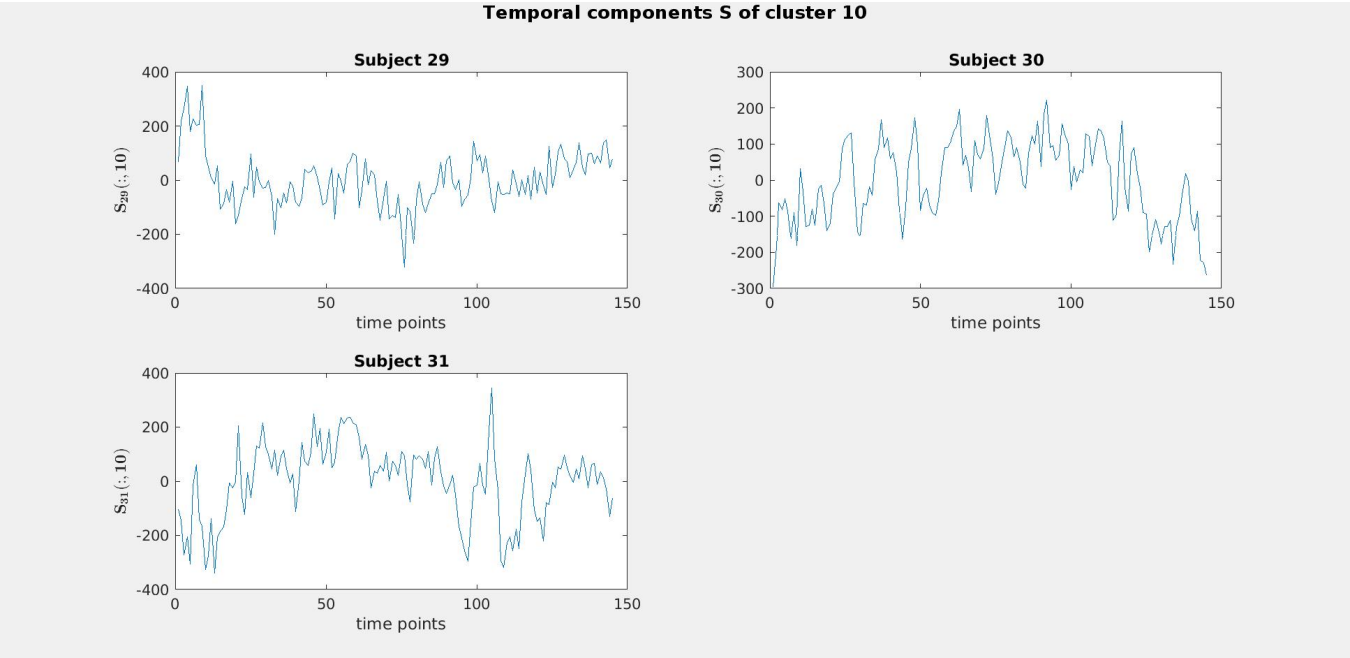


Figure B.83: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 29, \ldots, 31$ and R=27.

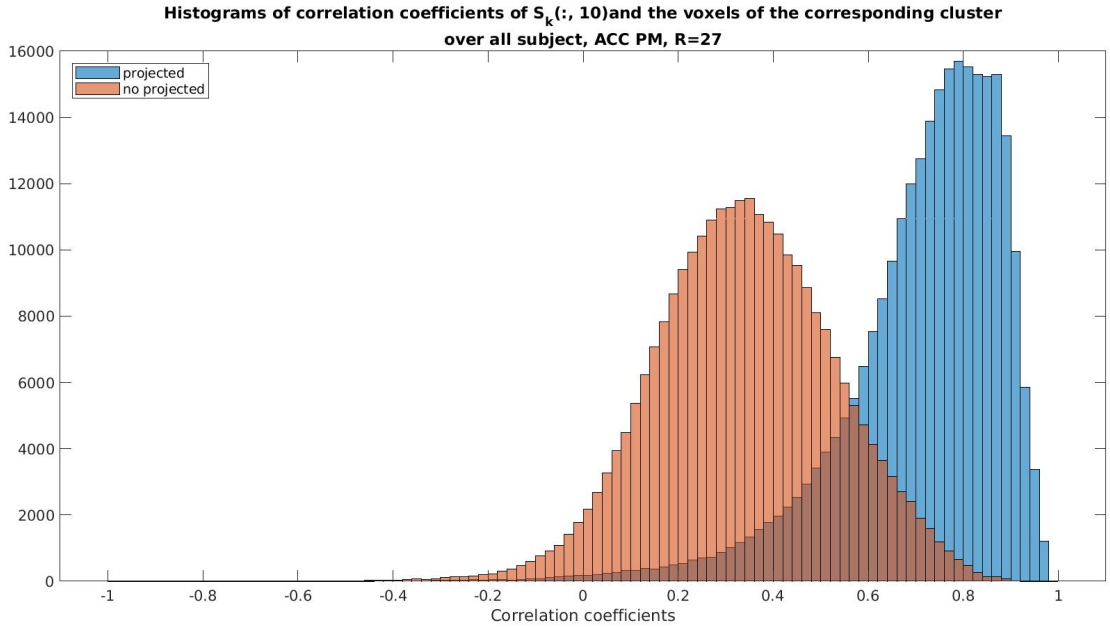


Figure B.84: For the column 10 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 10)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

7. **Cluster 2**
   Cluster 2 contains 10514 voxels, which correspond to 4.653283% of the whole brain (background, soft tissue, bone are excluded).

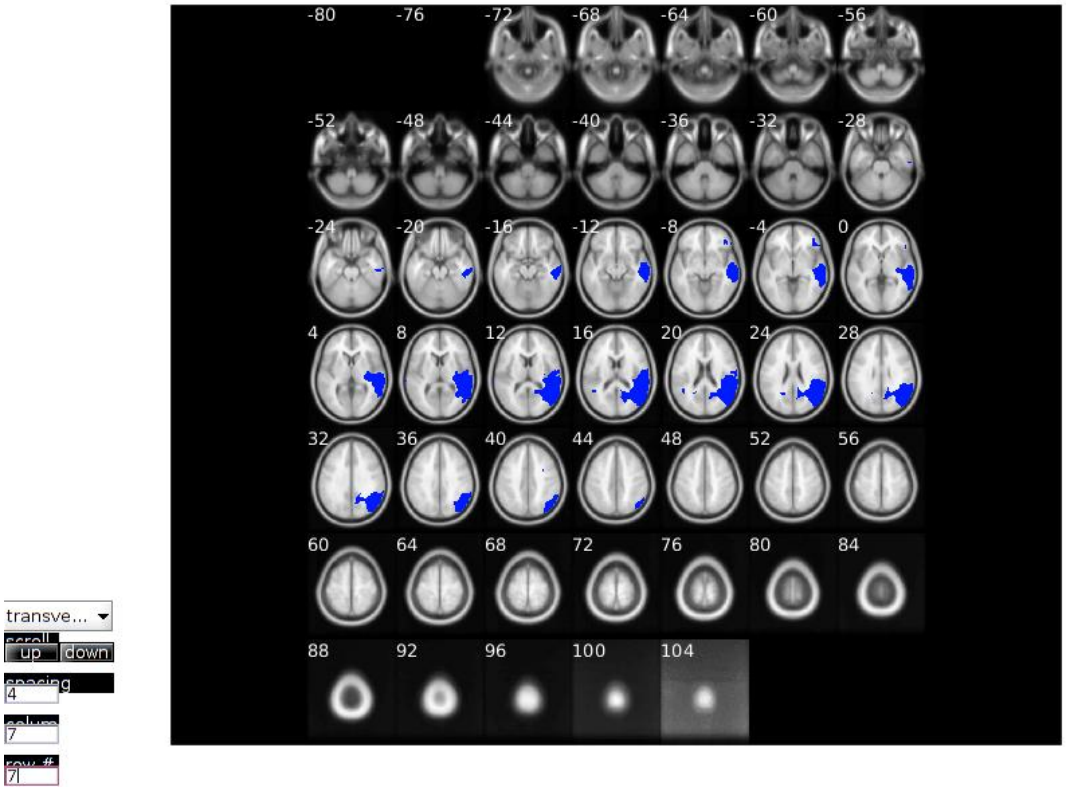


Figure B.85: Cluster 2. Blue areas indicate the voxels which belong to cluster 2.

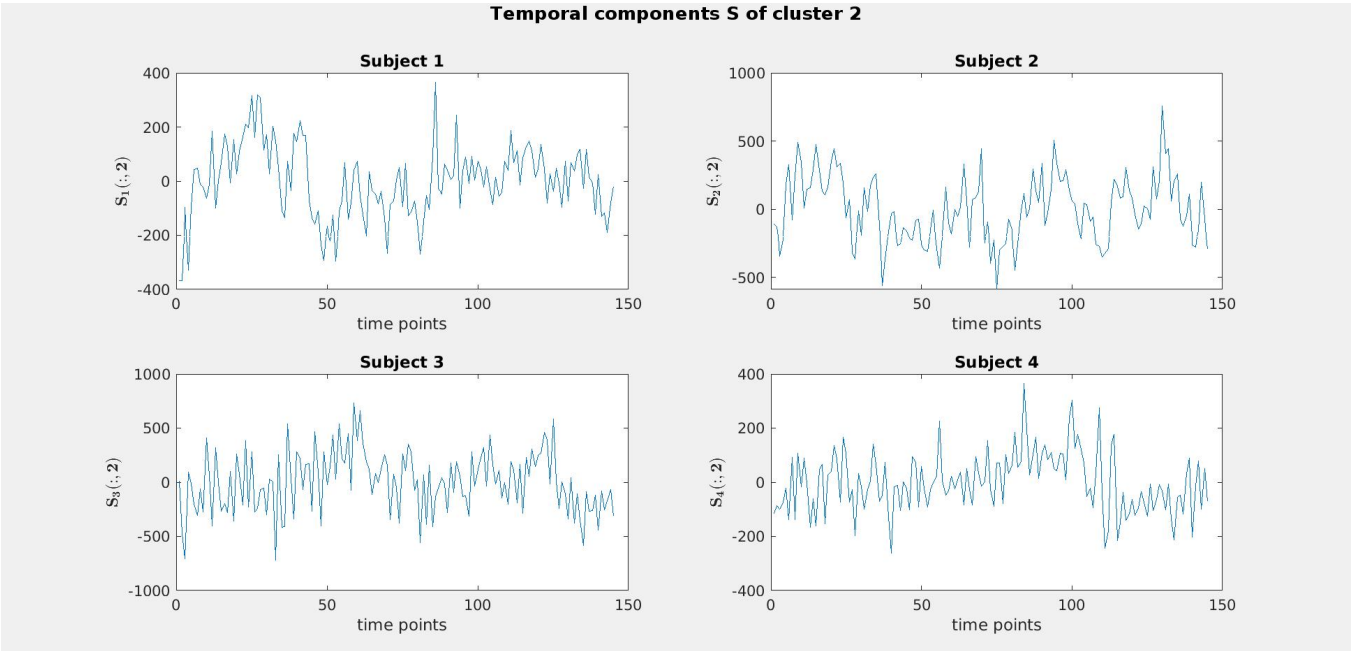


Figure B.86: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 1, \ldots, 4$ and R=27.

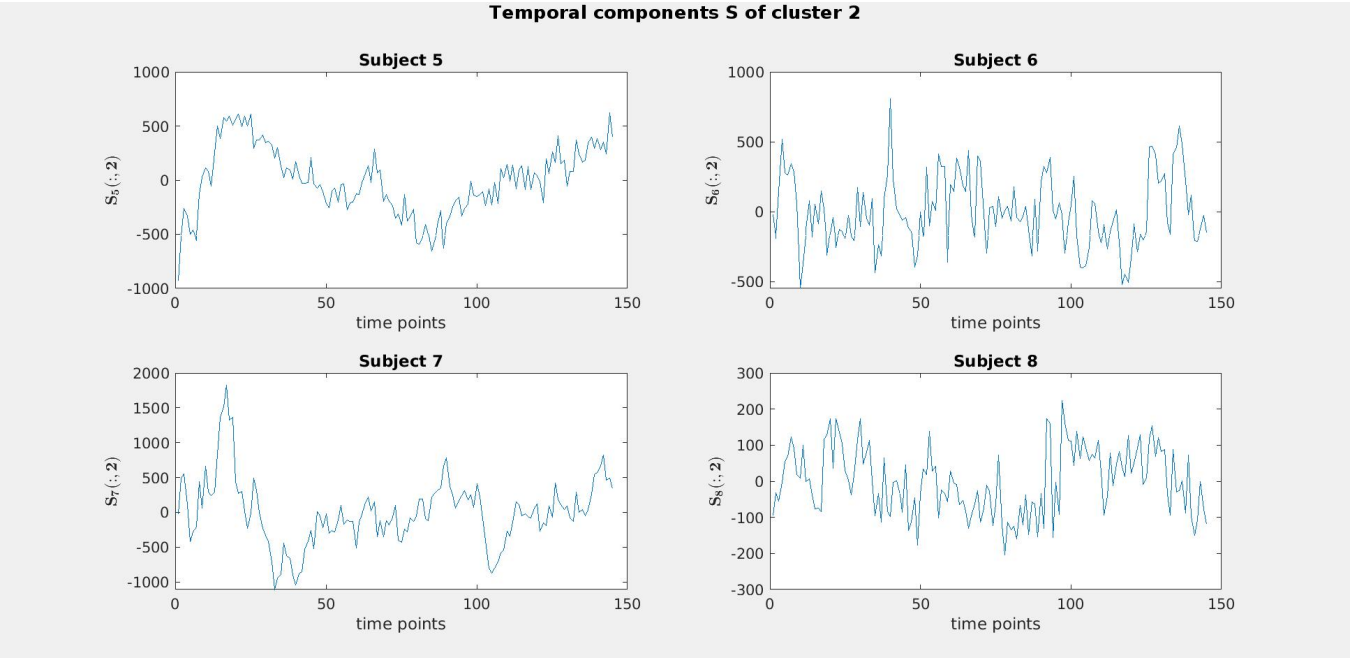


Figure B.87: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 5,\ldots,8$ and R=27.

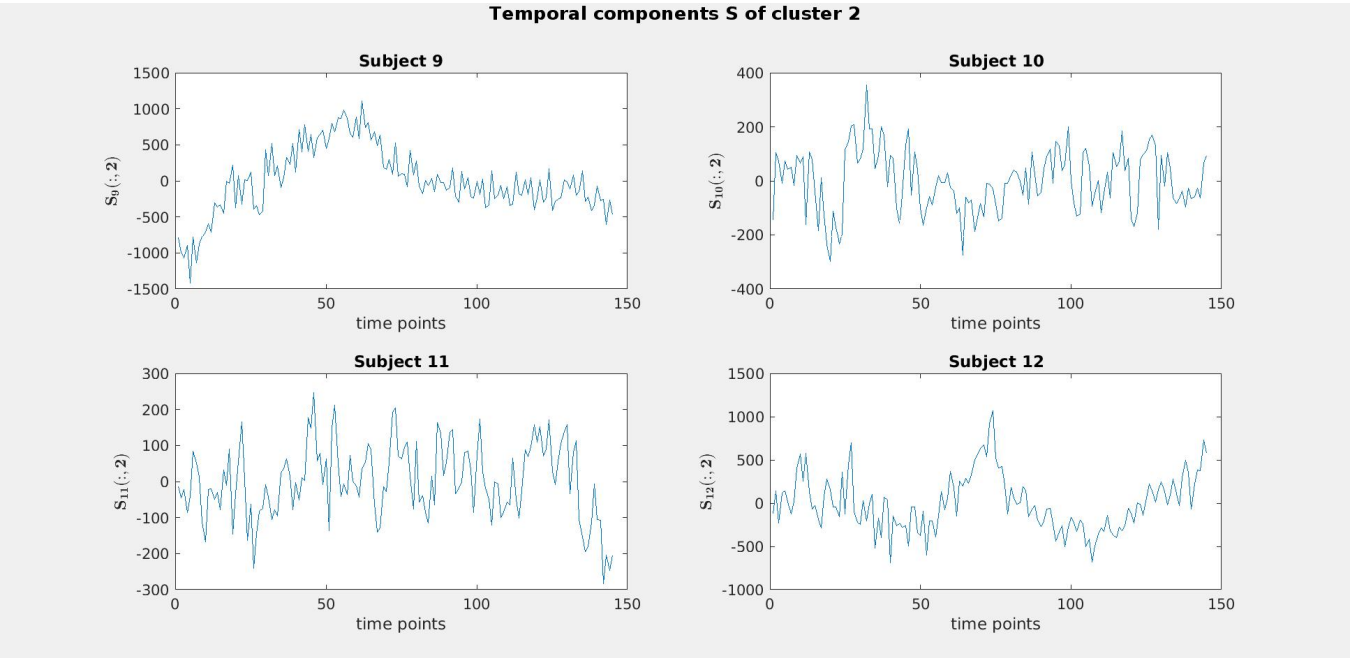


Figure B.88: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 9,\ldots,12$ and R=27.

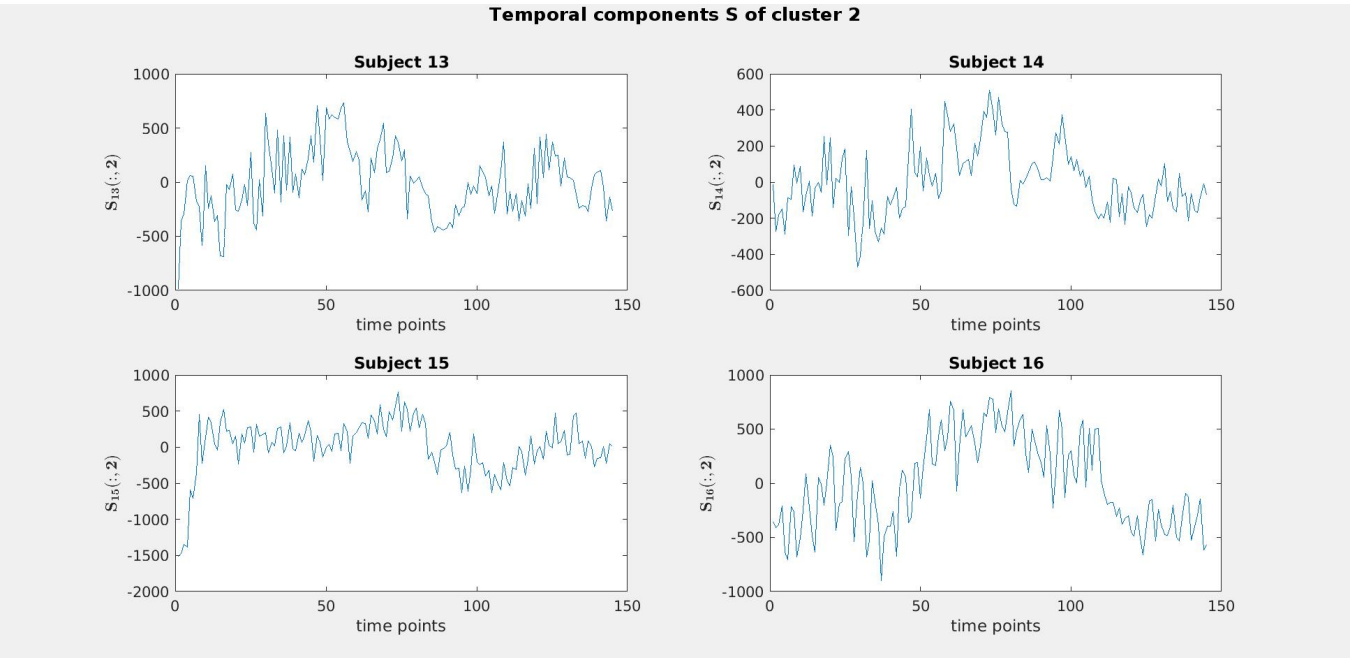


Figure B.89: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 13,\ldots,16$ and R=27.

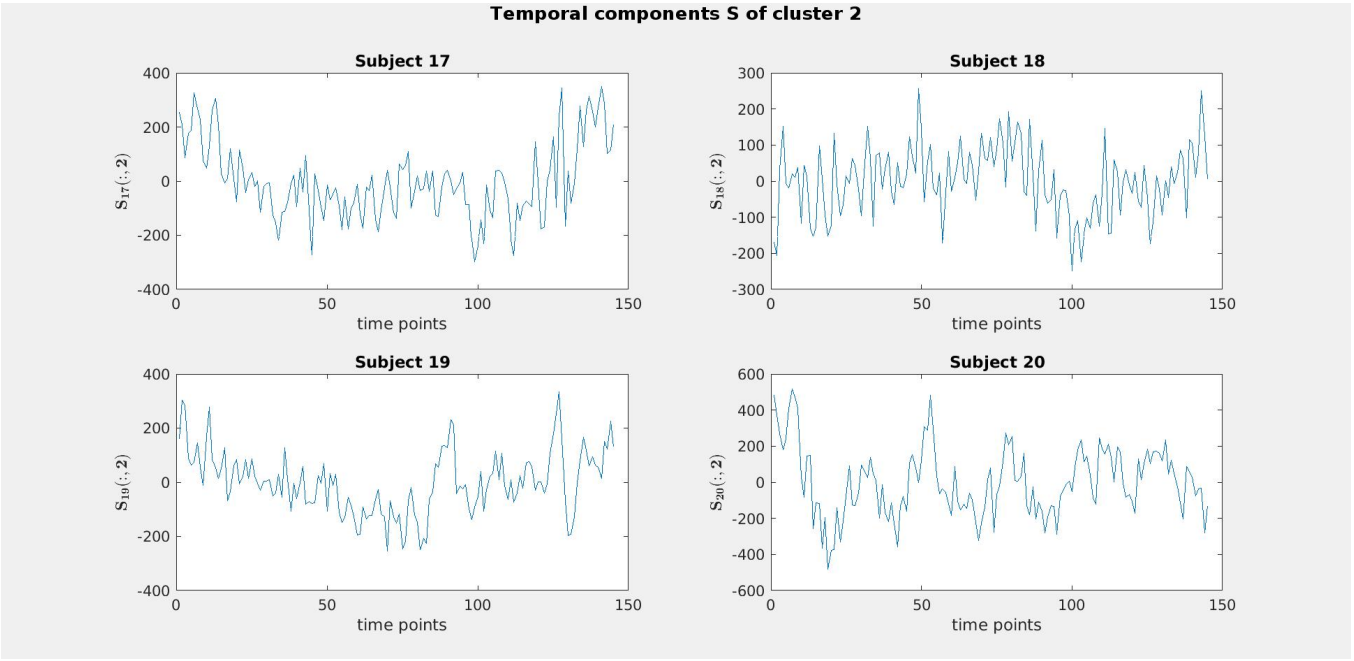


Figure B.90: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 17,\ldots,20$ and R=27.

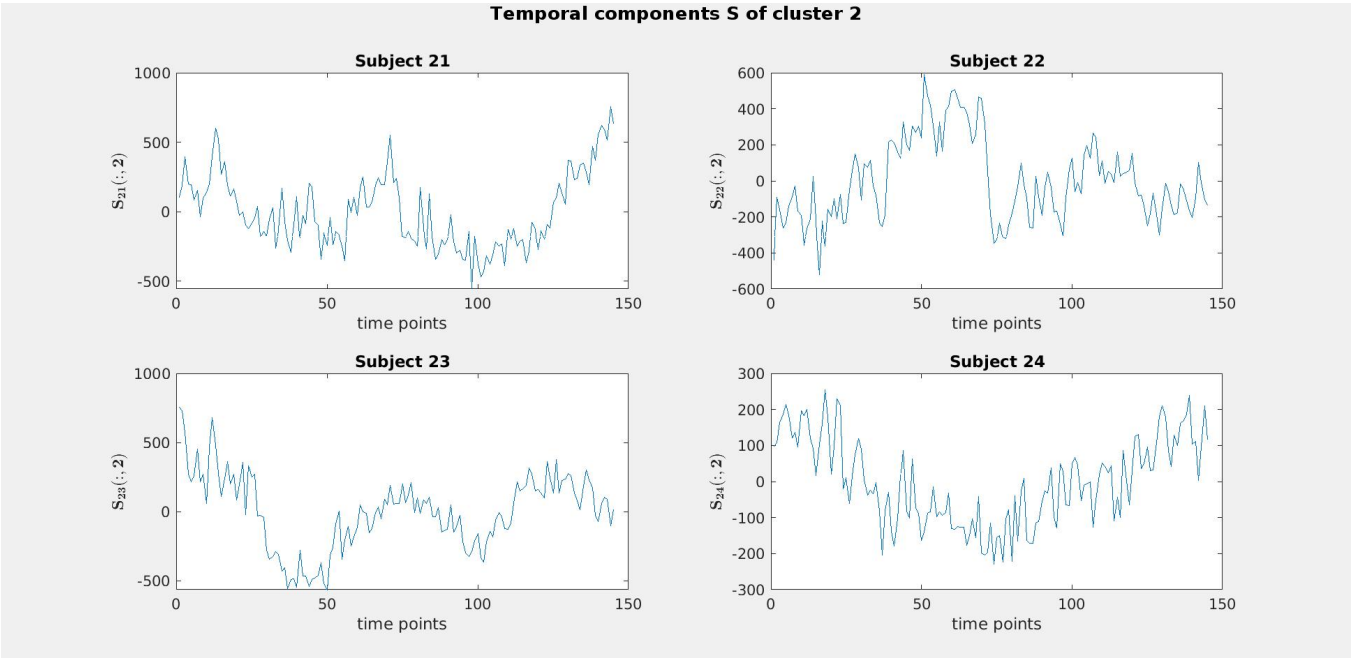


Figure B.91: Waveform of temporal components $\mathbf{S}_k(:,12)$ for $k = 21,\ldots,24$ and R=27.

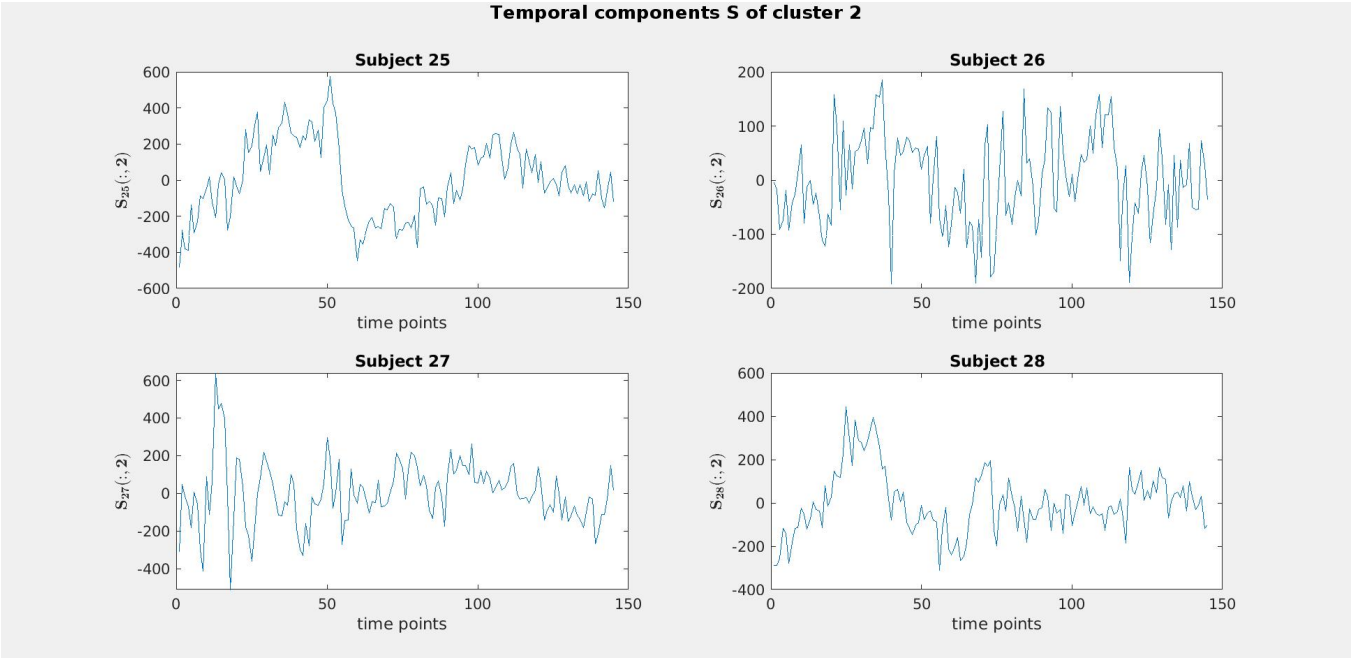


Figure B.92: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 25,\ldots,28$ and R=27.

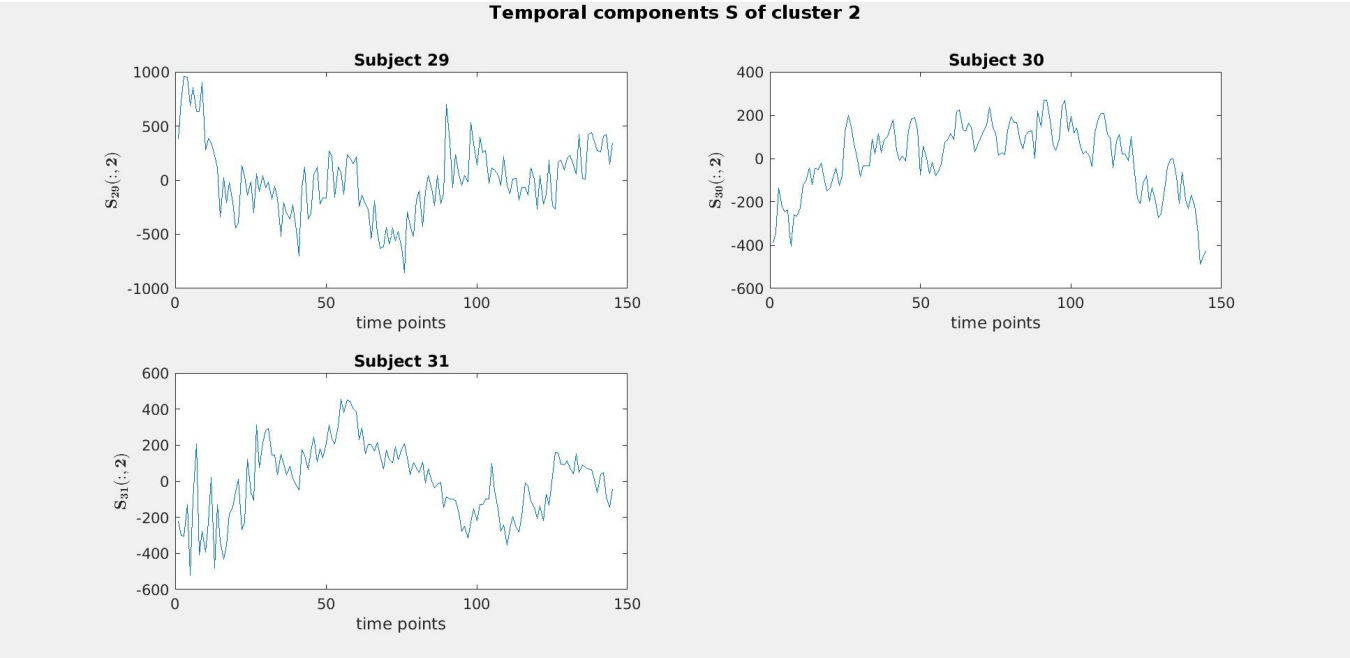


Figure B.93: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 29, \ldots, 31$ and R=27.

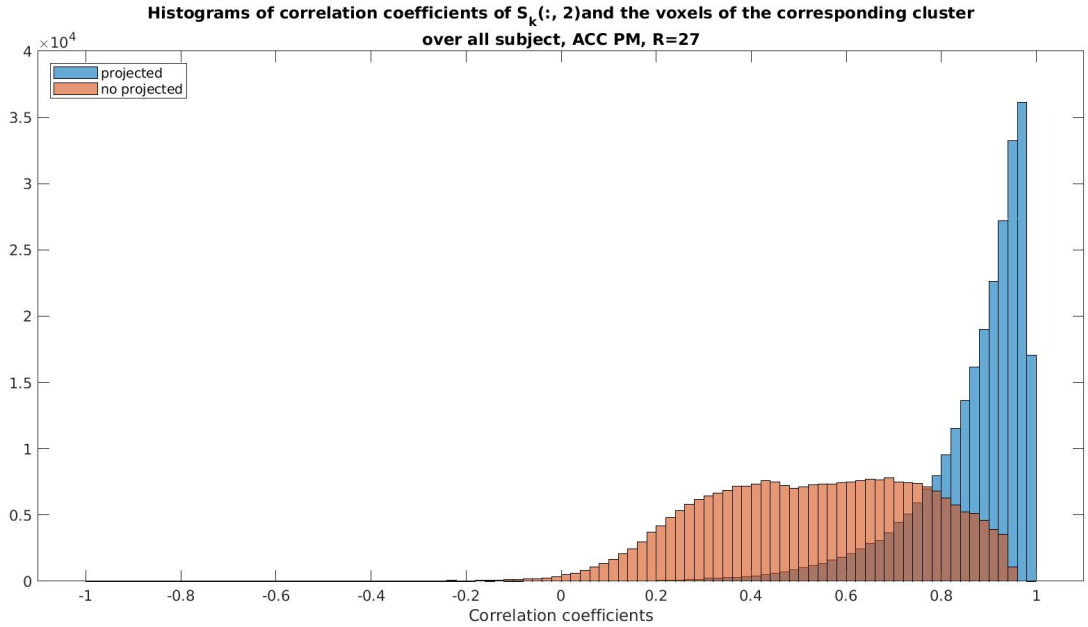


Figure B.94: For the column 2 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,2)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

8. **Cluster 9**
   Cluster 9 contains 10302 voxels, which correspond to 4.559456% of the whole brain (background, soft tissue, bone are excluded).

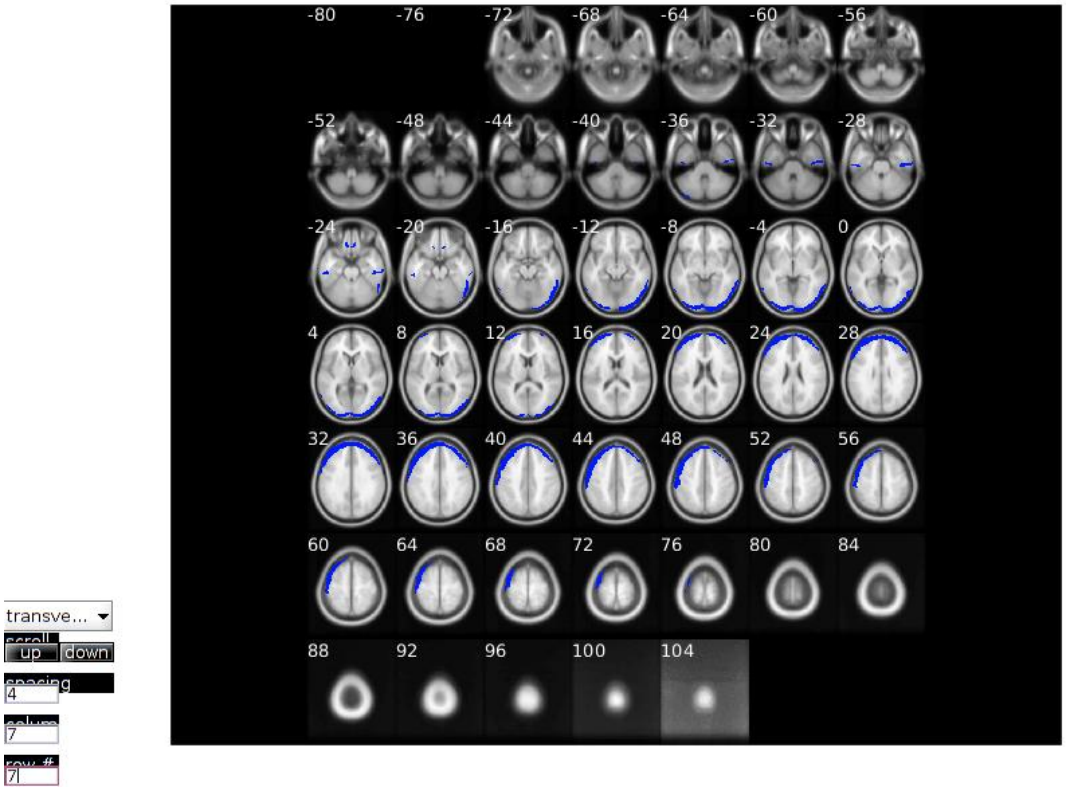


Figure B.95: Cluster 9. Blue areas indicate the voxels which belong to cluster 9.

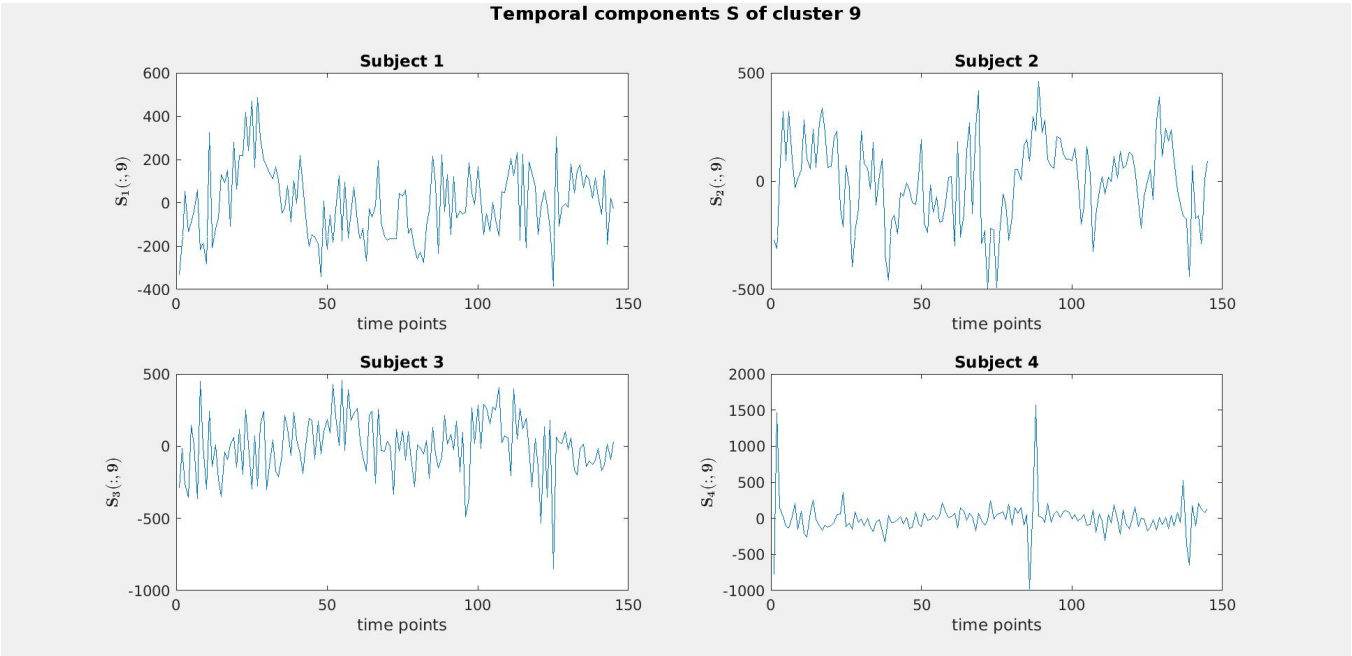


Figure B.96: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 1, \ldots, 4$ and R=27.

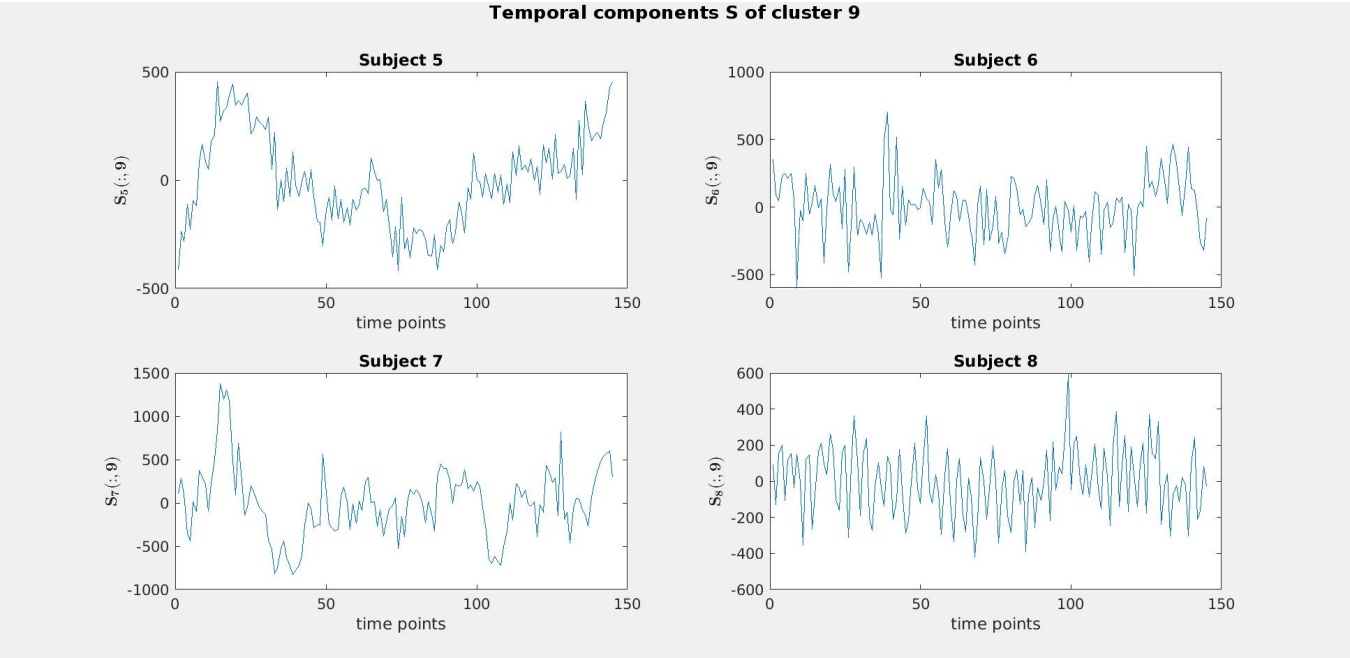


Figure B.97: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 5, \ldots, 8$ and R=27.

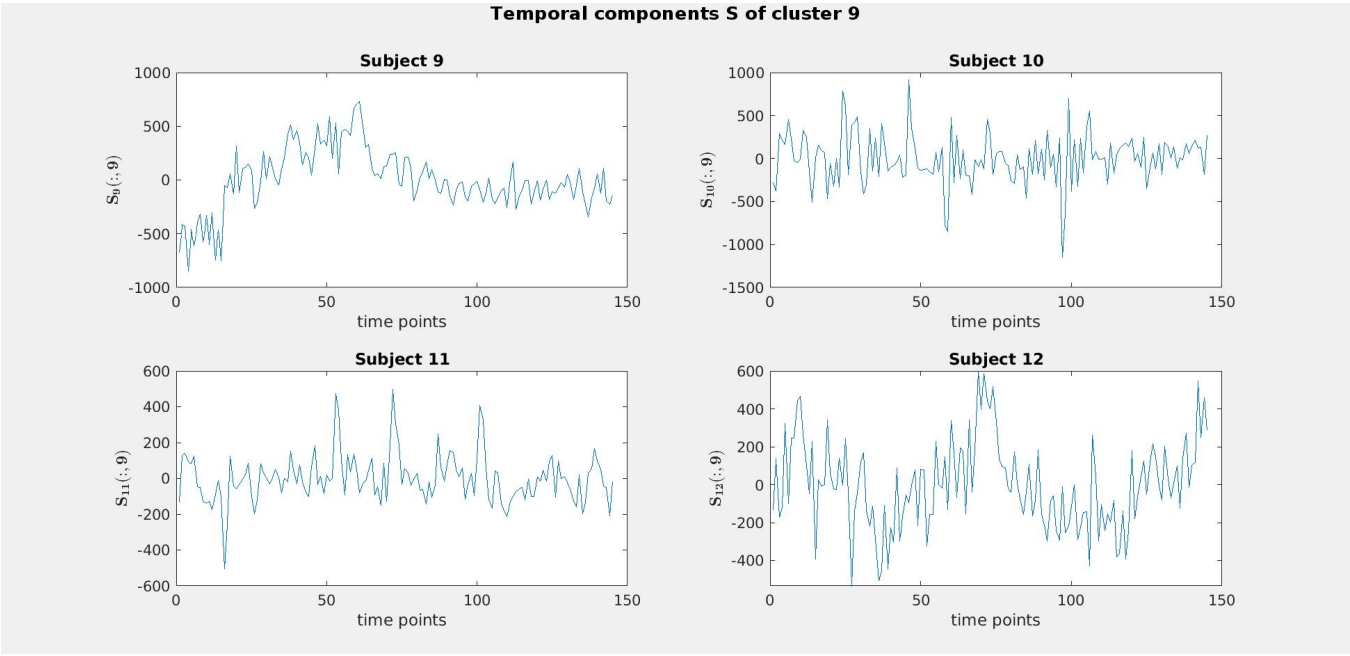


Figure B.98: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 9, \ldots, 12$ and R=27.

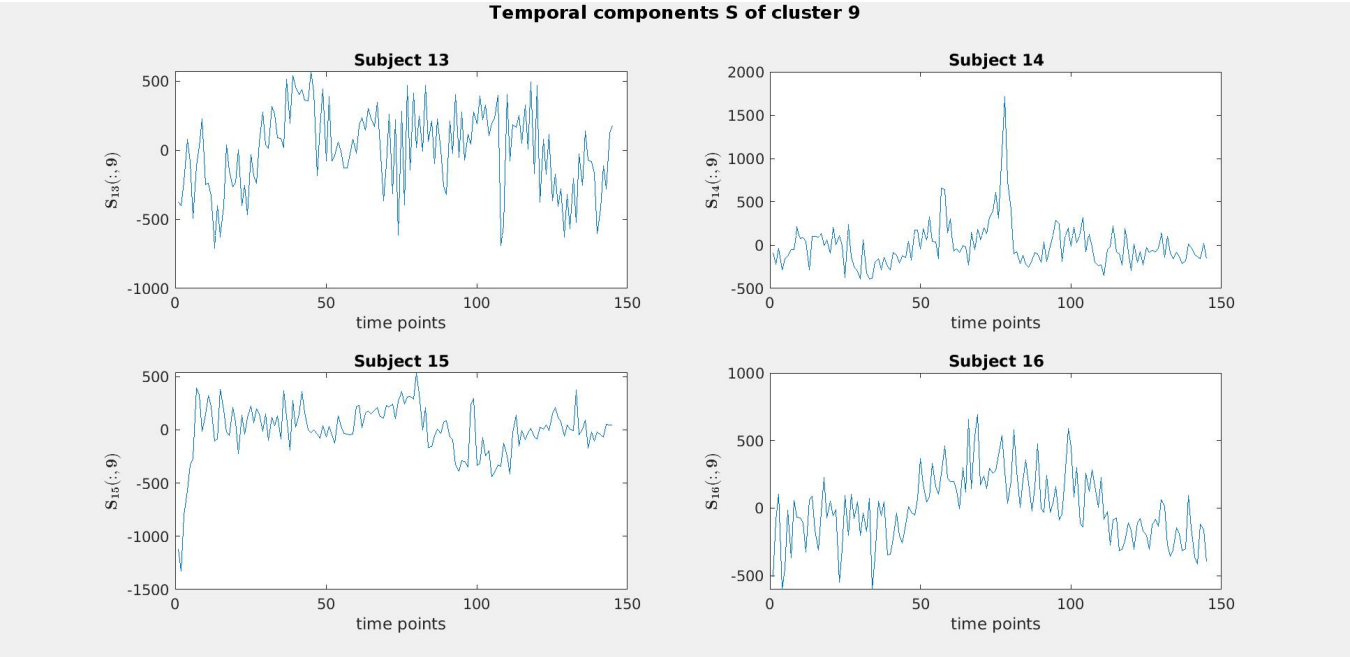


Figure B.99: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 13, \ldots, 16$ and R=27.

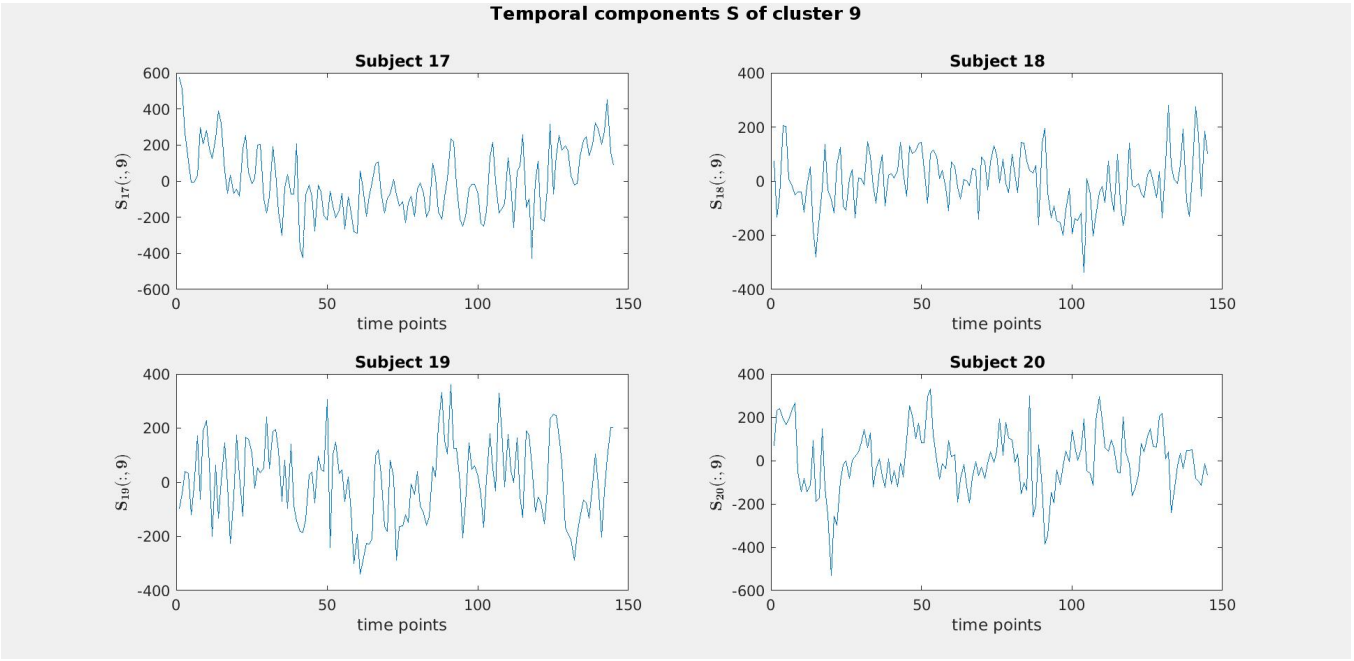


Figure B.100: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 17, \ldots, 20$ and R=27.

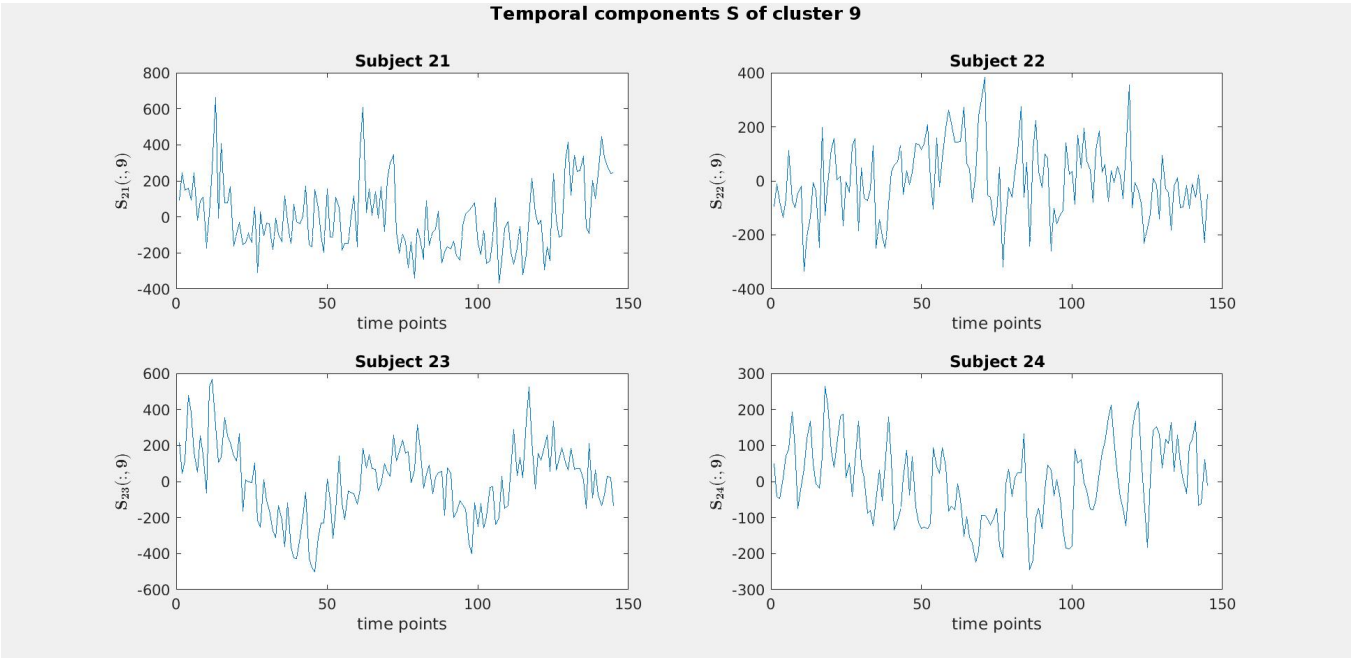


Figure B.101: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 21, \ldots, 24$ and R=27.

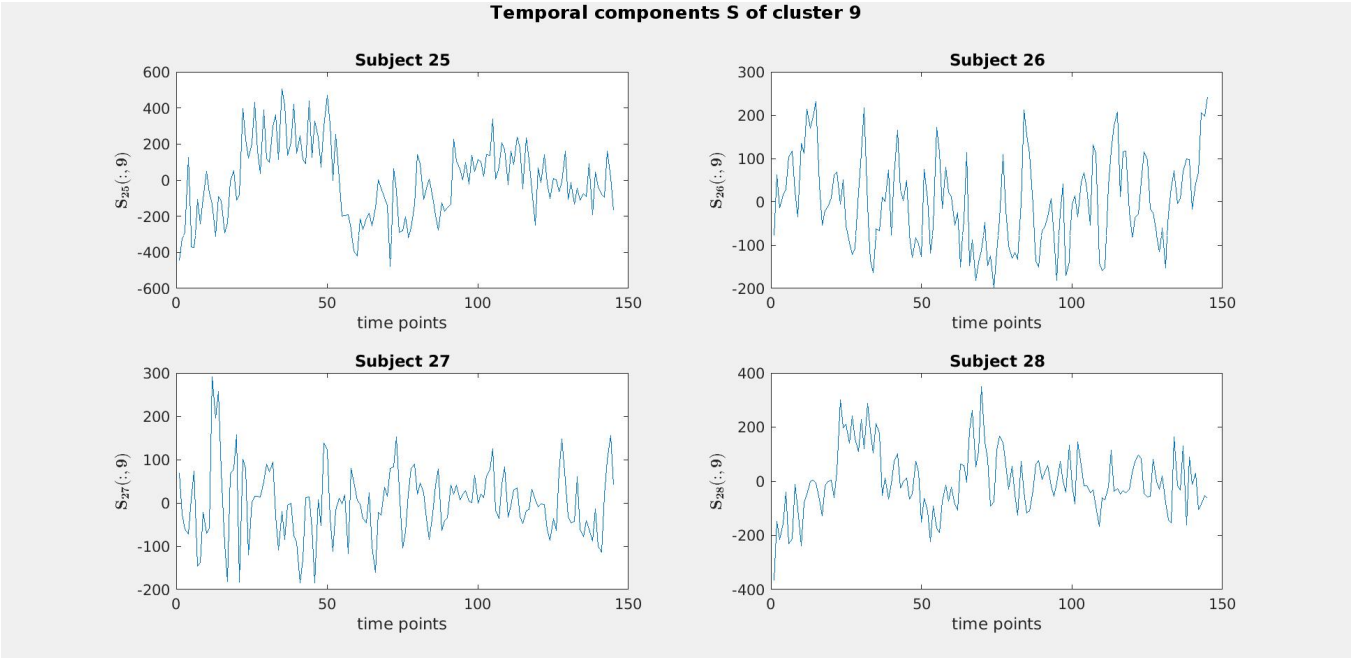


Figure B.102: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 25, \ldots, 28$ and R=27.

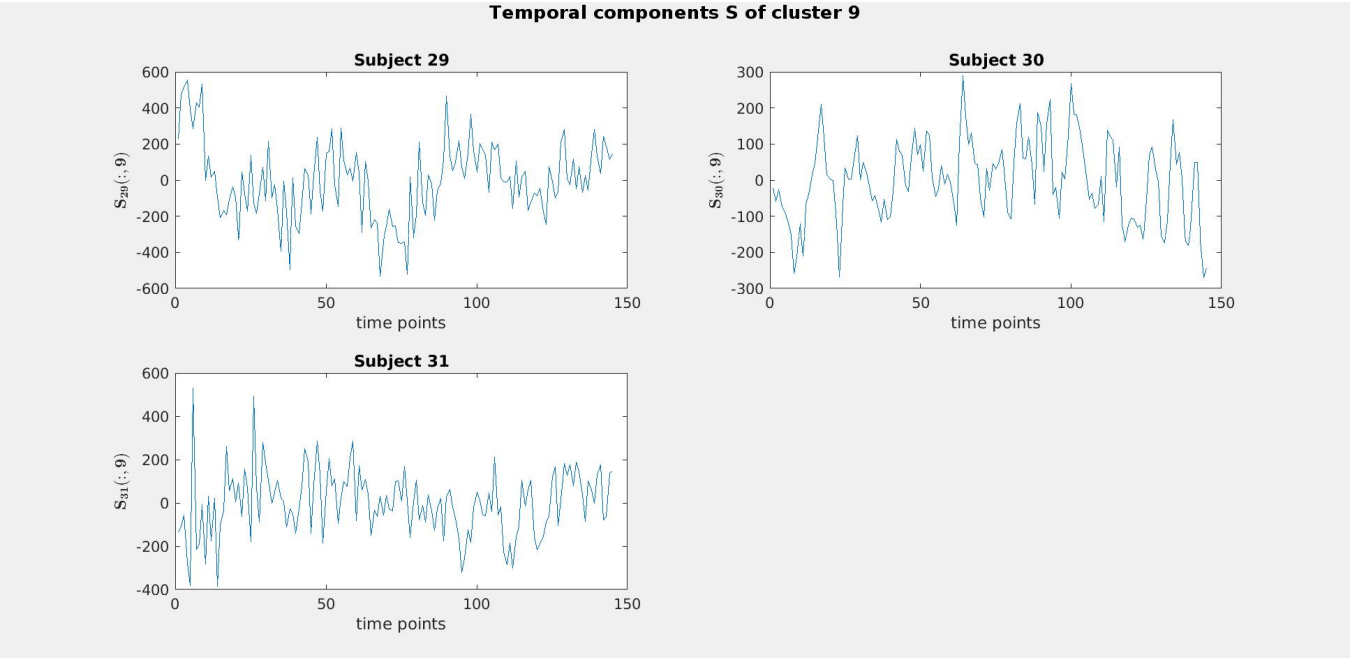


Figure B.103: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 29, \ldots, 31$ and R=27.

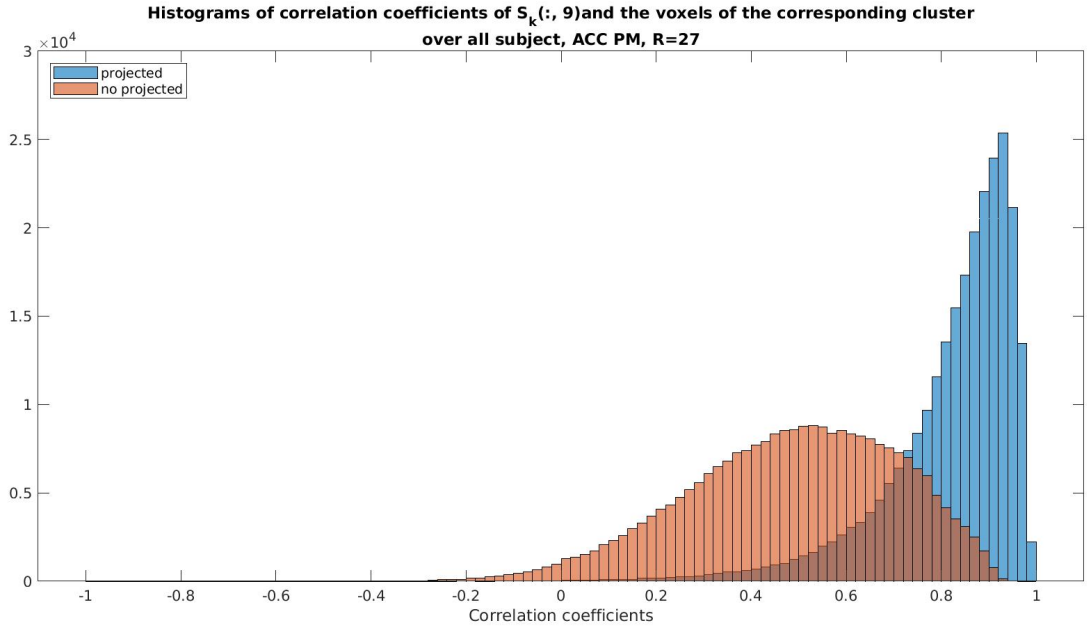


Figure B.104: For the column 9 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 9)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

9. **Cluster 14**
   Cluster 14 contains 10264 voxels, which correspond to 4.542638% of the whole brain (background, soft tissue, bone are excluded).

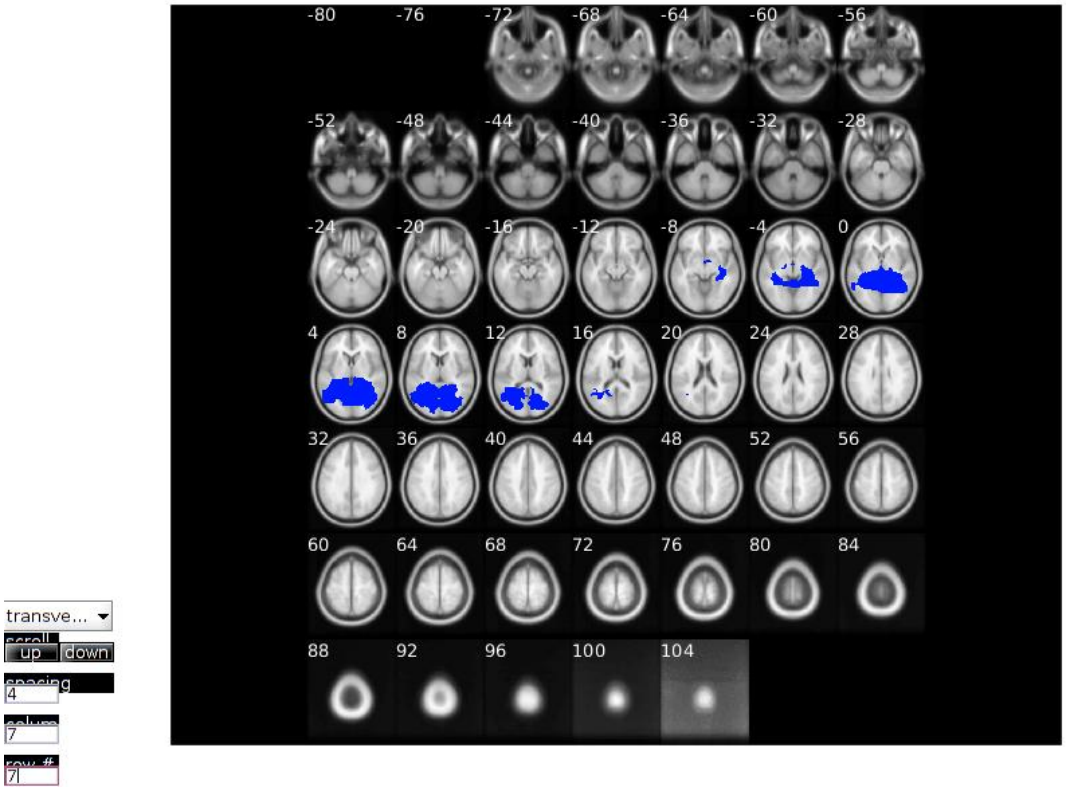


Figure B.105: Cluster 14. Blue areas indicate the voxels which belong to cluster 14.

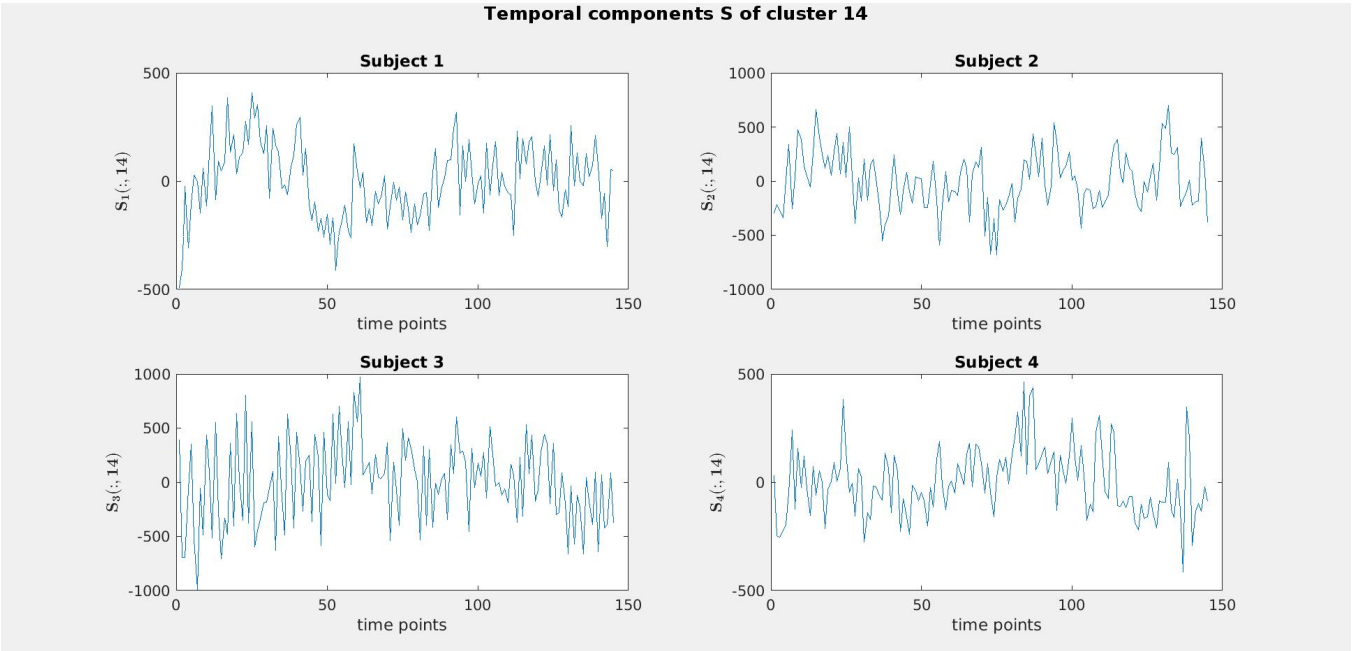


Figure B.106: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 1, \ldots, 4$ and R=27.

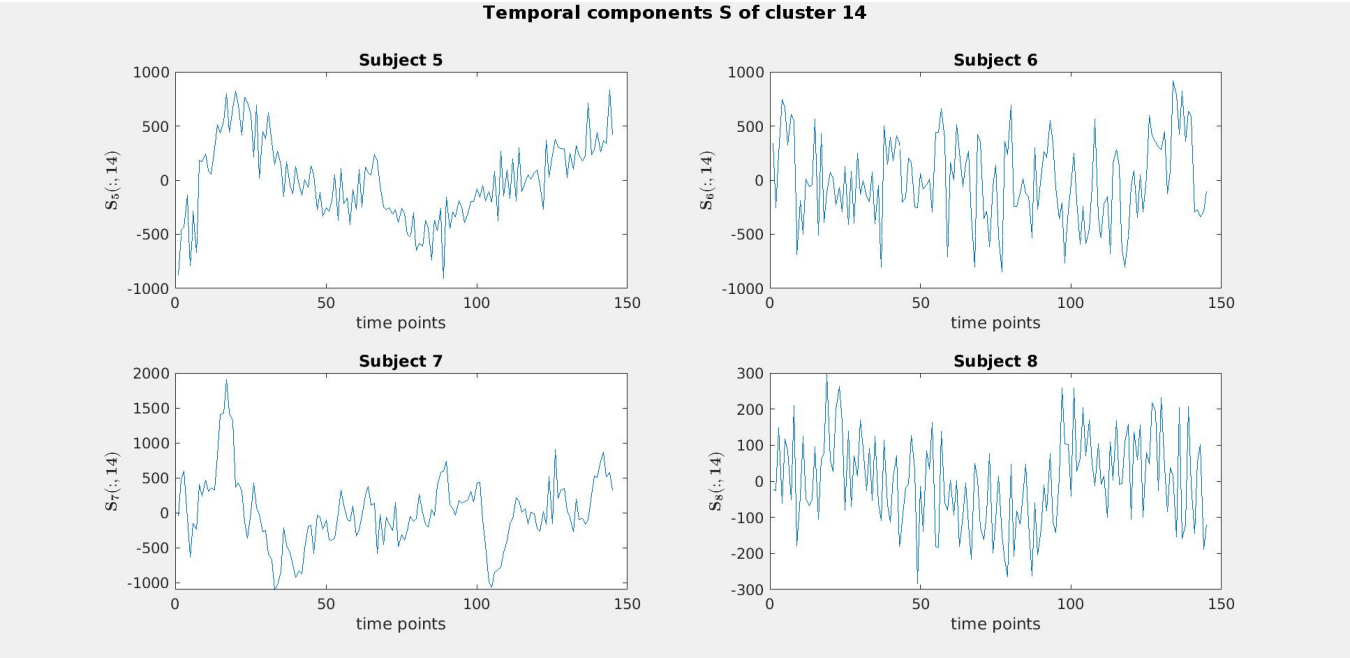


Figure B.107: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 5, \ldots, 8$ and R=27.

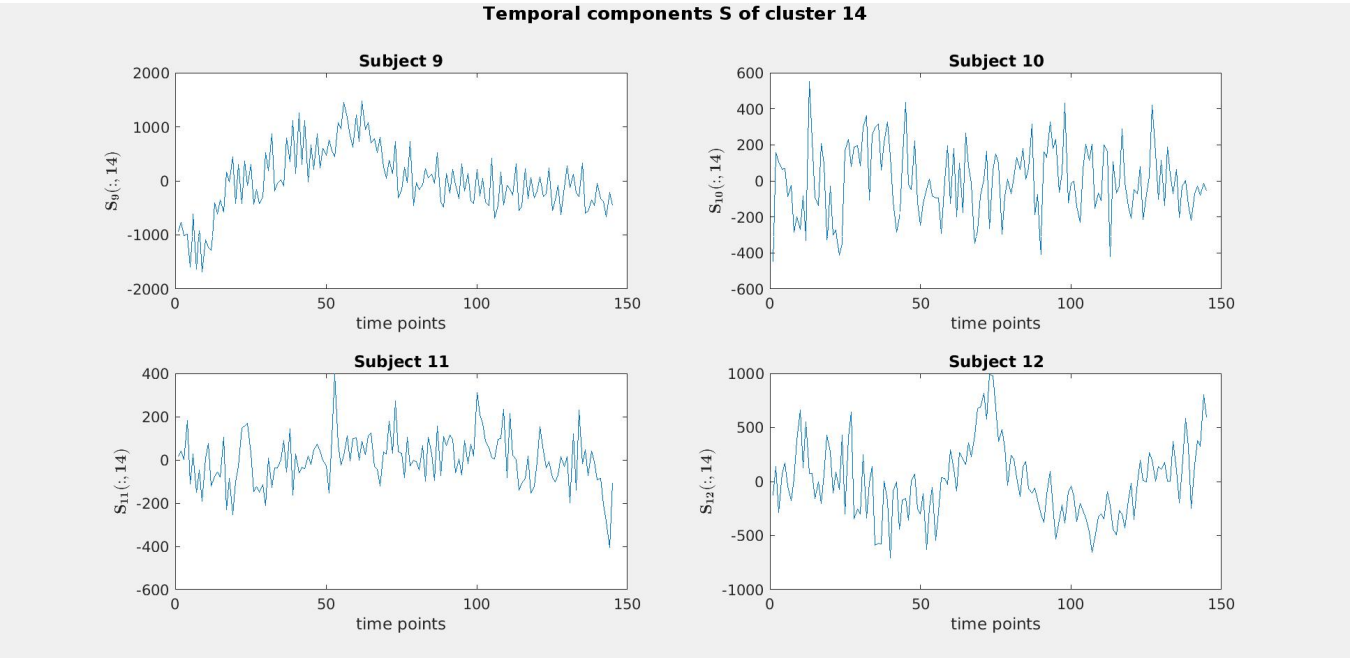


Figure B.108: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 9, \ldots, 12$ and R=27.

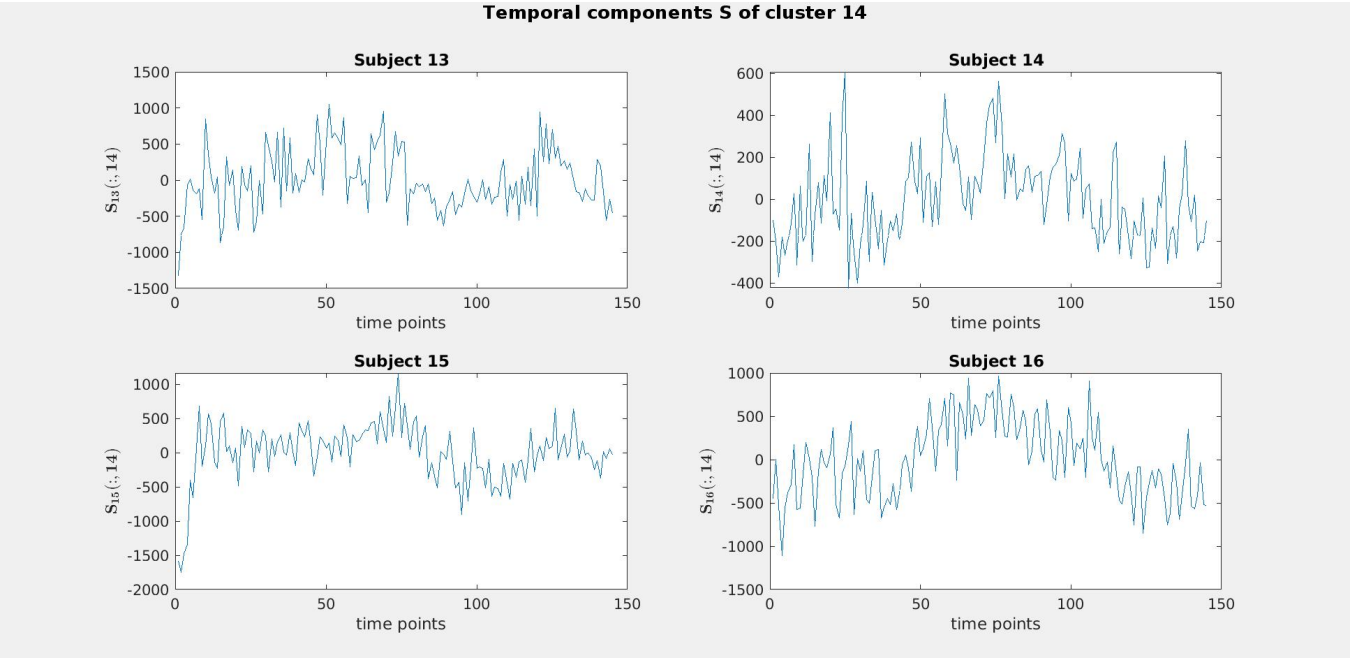


Figure B.109: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 13, \ldots, 16$ and R=27.

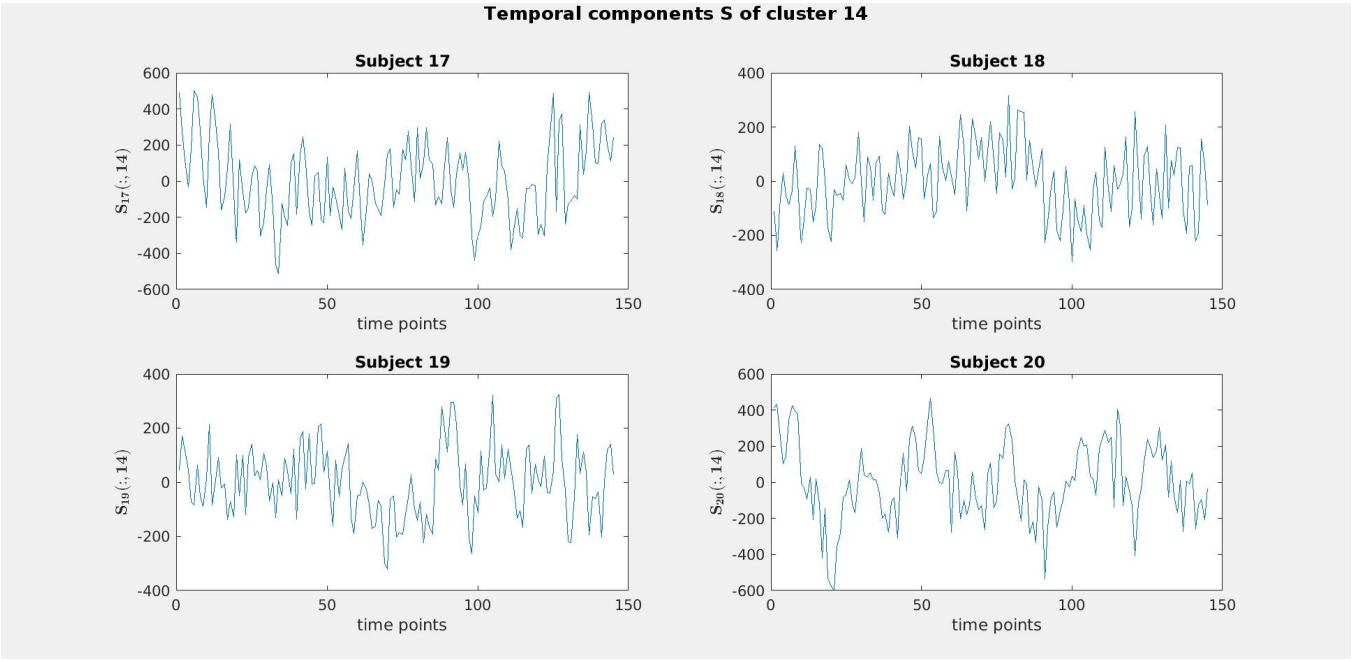


Figure B.110: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 17, \ldots, 20$ and R=27.

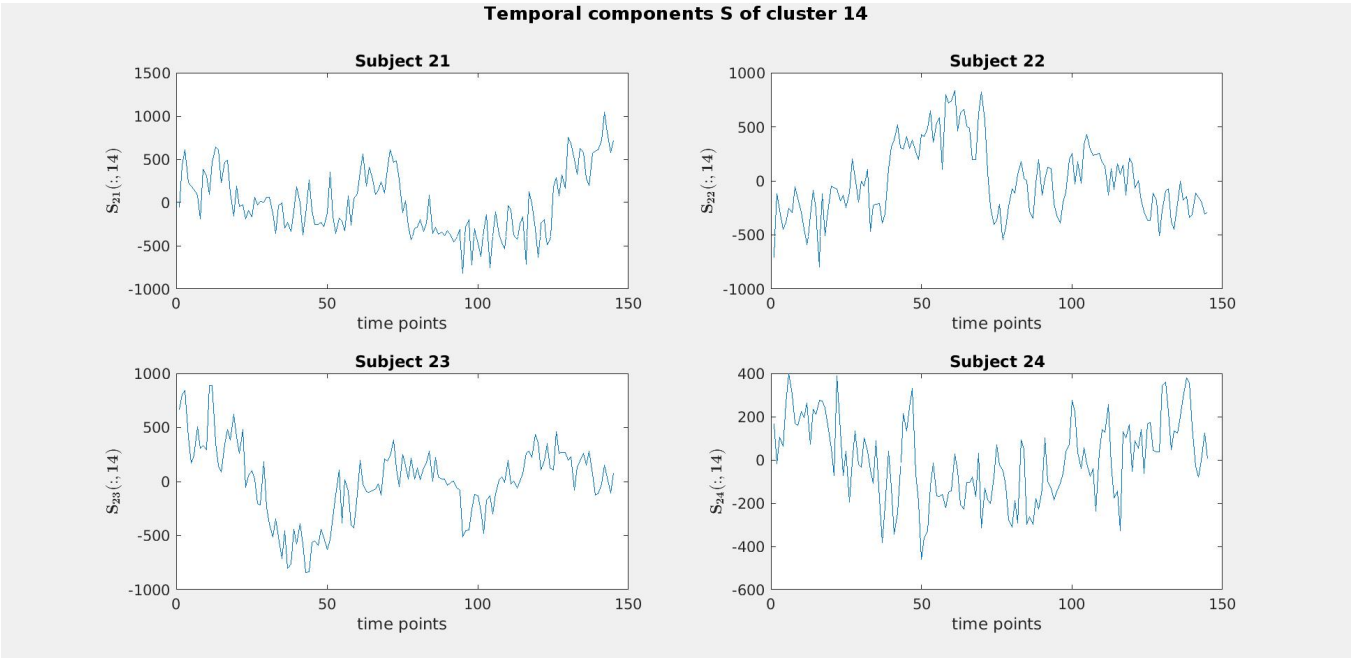


Figure B.111: Waveform of temporal components $\mathbf{S}_k(:, 114)$ for $k = 21, \ldots, 24$ and R=27.

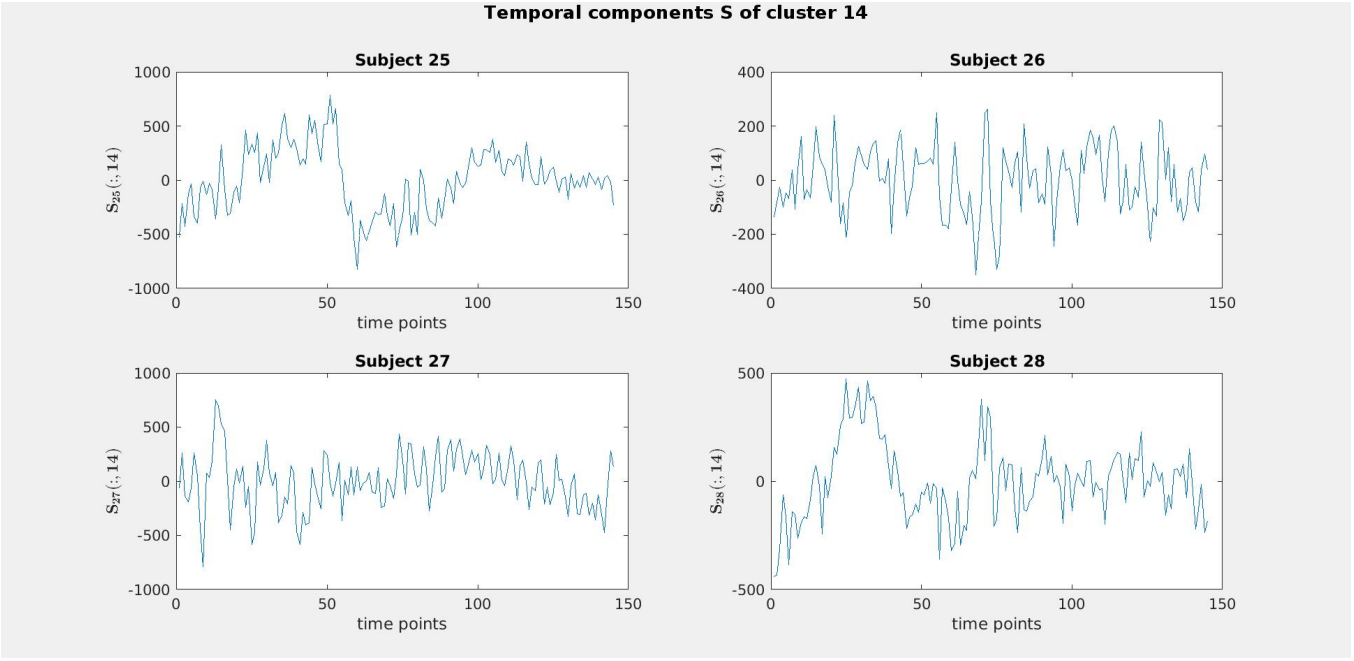


Figure B.112: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 25, \ldots, 28$ and R=27.

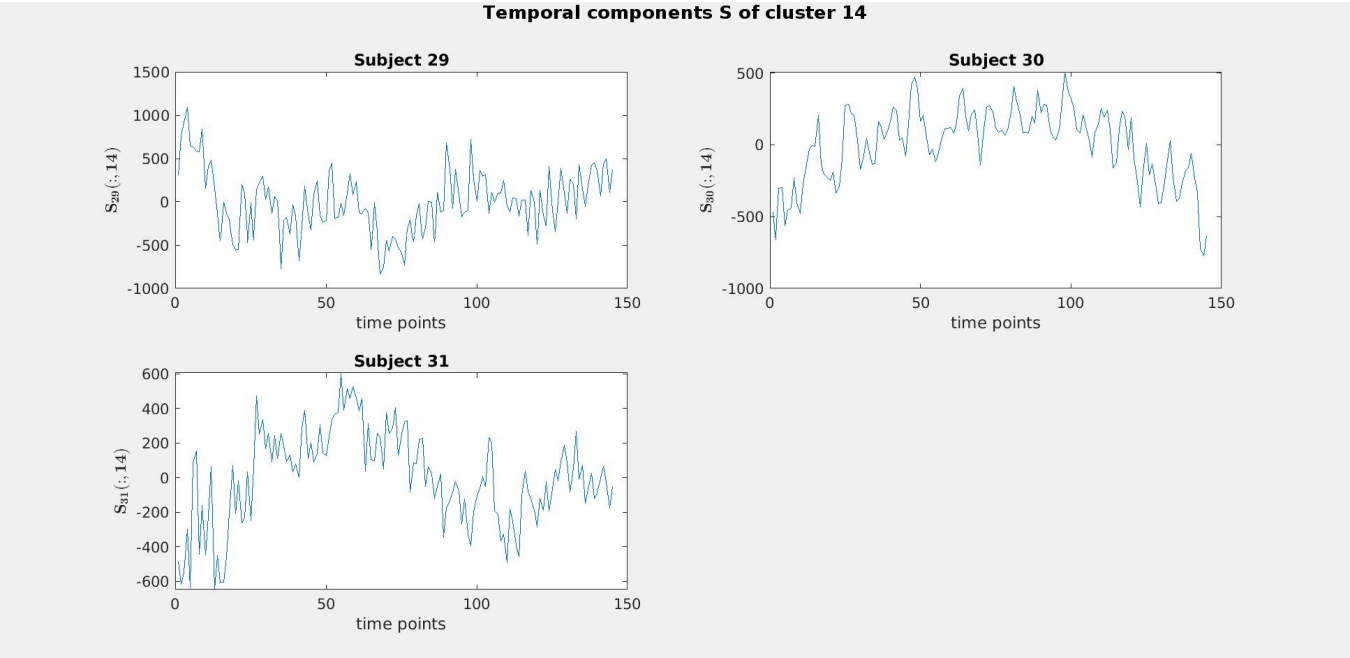


Figure B.113: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 29, \ldots, 31$ and R=27.

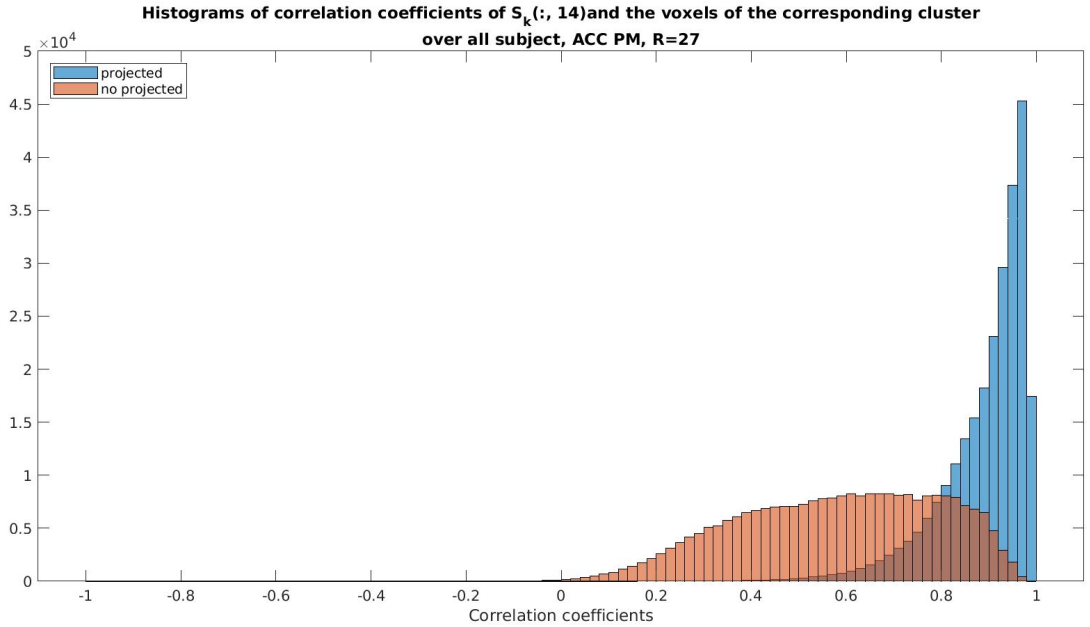


Figure B.114: For the column 14 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 14)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

10. **Cluster 24**

    Cluster 24 contains 10183 voxels, which correspond to 4.506789% of the whole brain (background, soft tissue, bone are excluded).

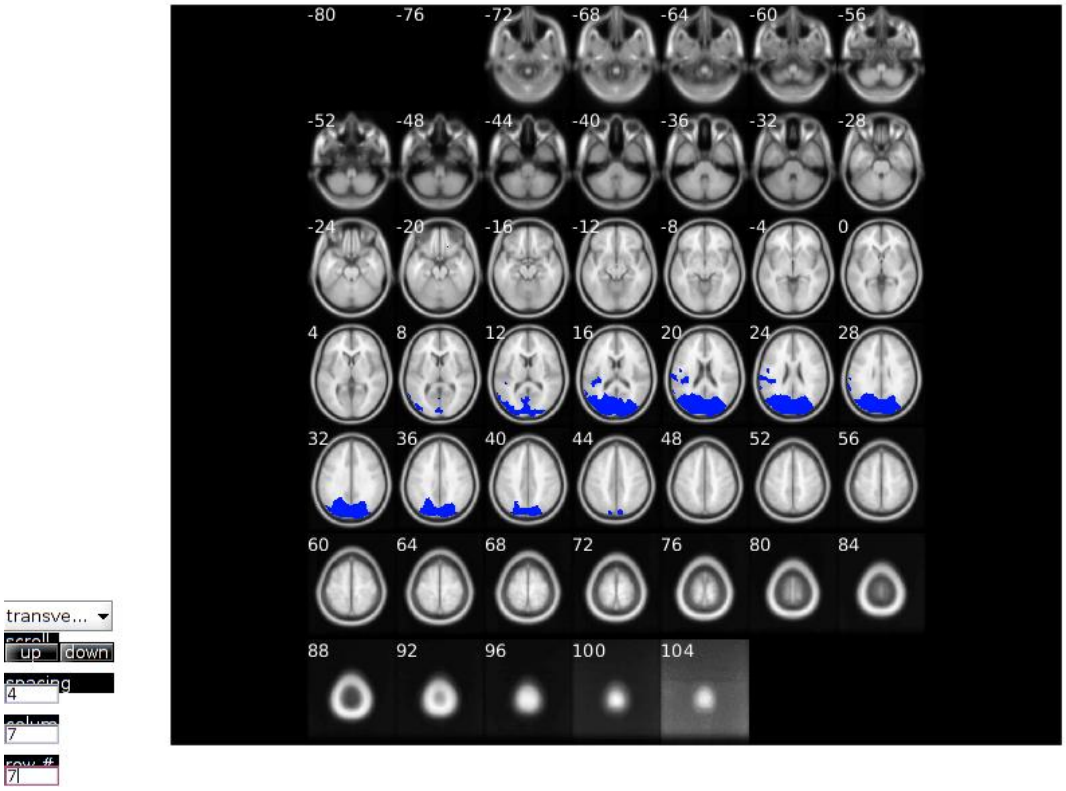


Figure B.115: Cluster 24. Blue areas indicate the voxels which belong to cluster 24.

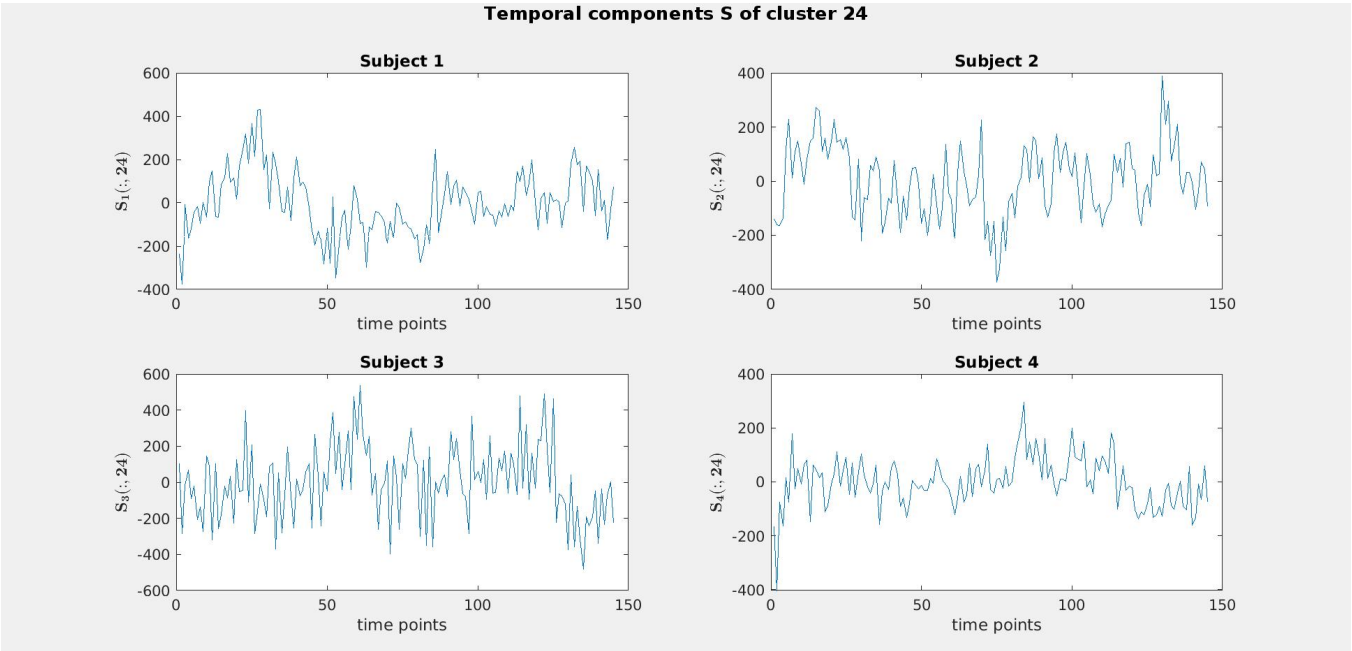


Figure B.116: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 1, \ldots, 4$ and R=27.

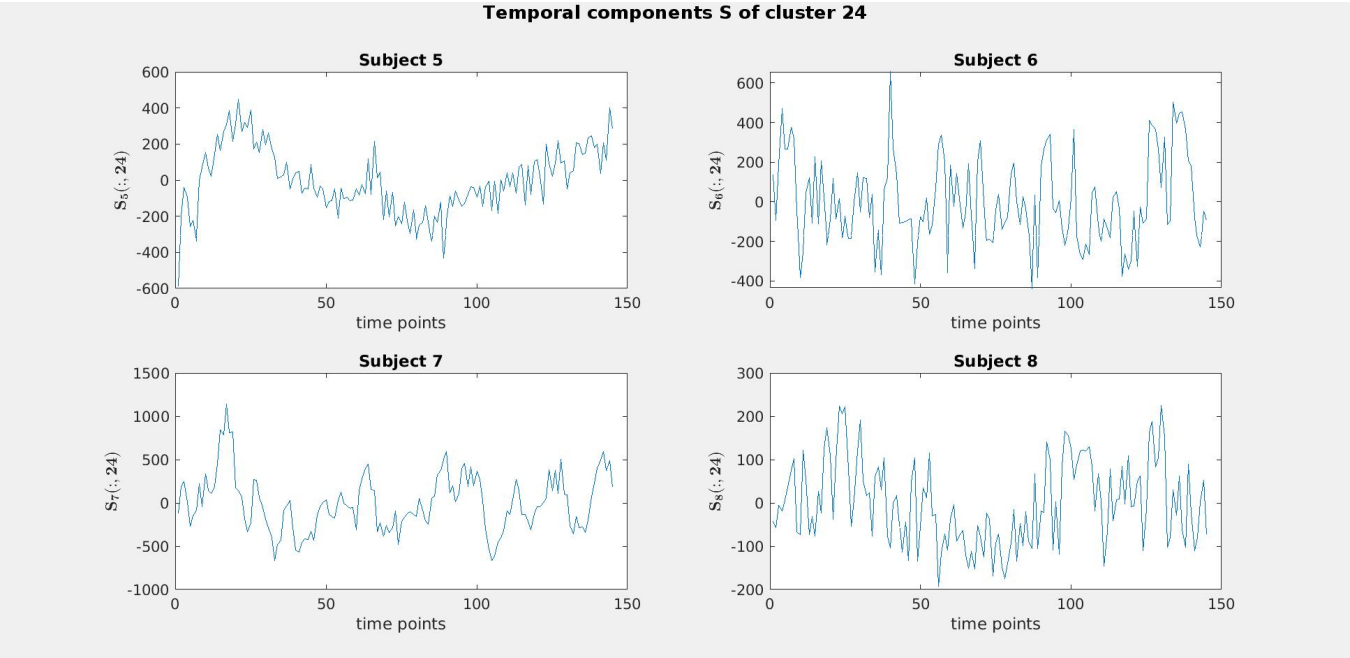


Figure B.117: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 5, \ldots, 8$ and R=27.

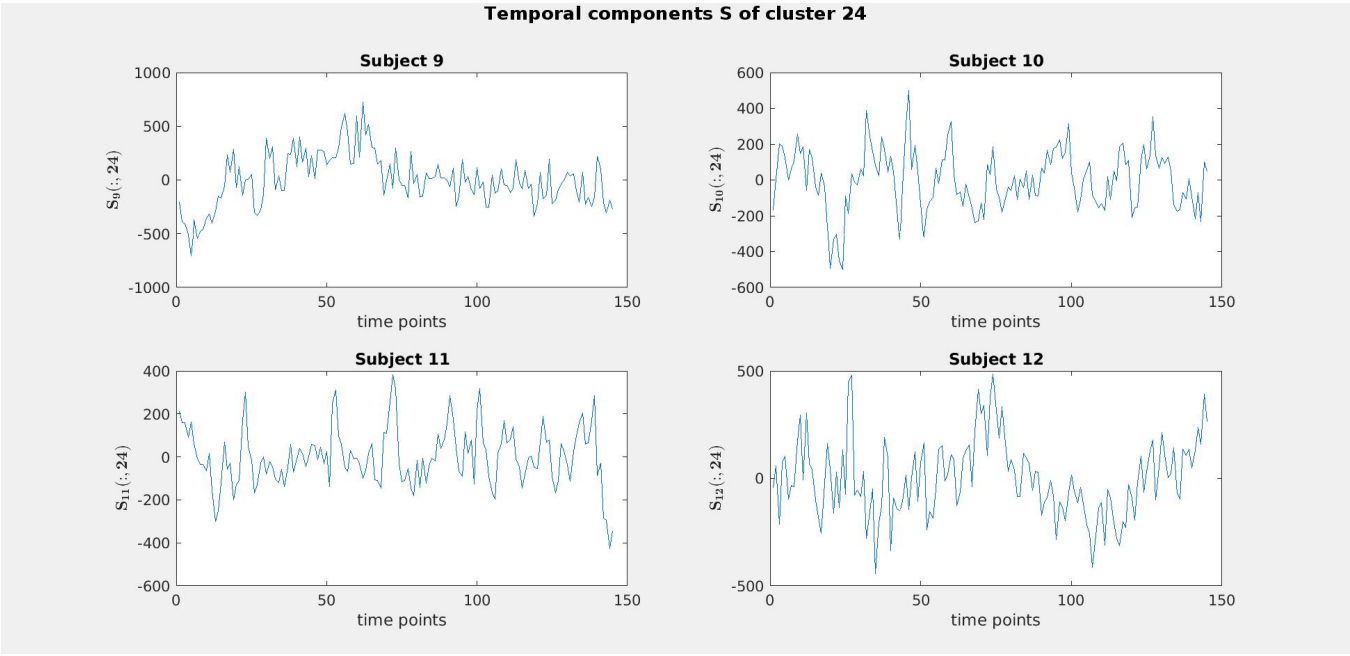


Figure B.118: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 9, \ldots, 12$ and R=27.

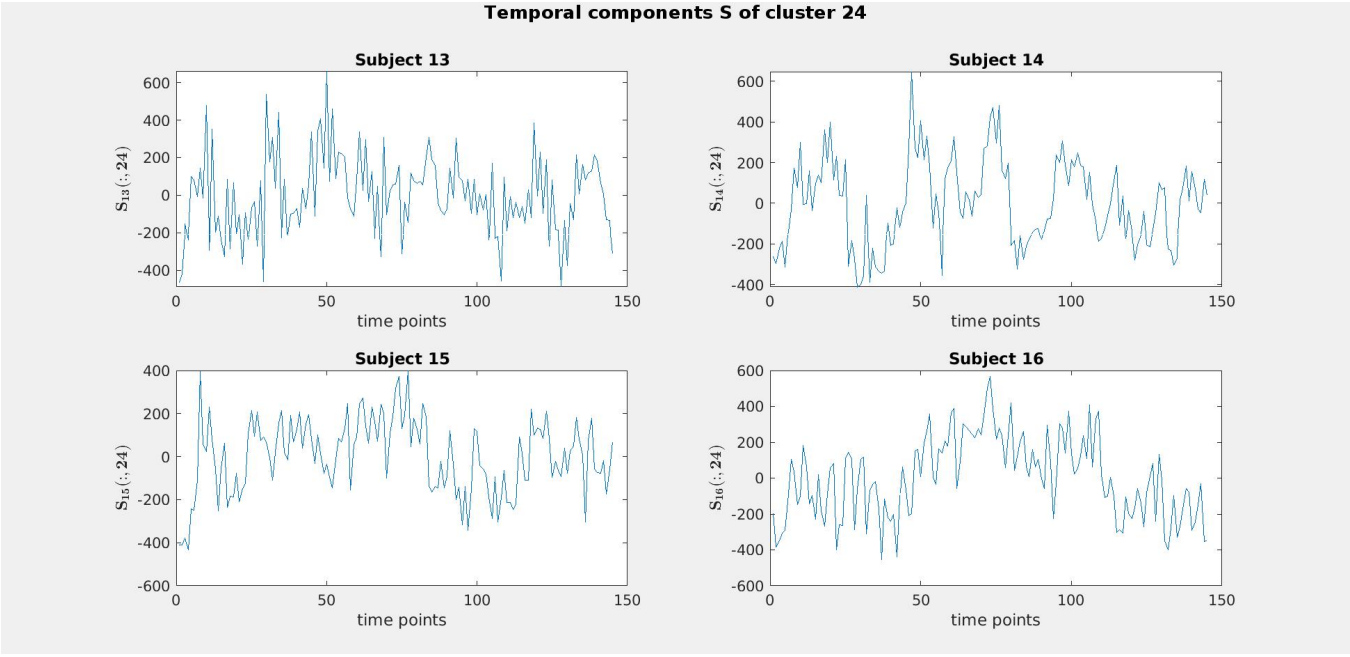


Figure B.119: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 13, \ldots, 16$ and R=27.

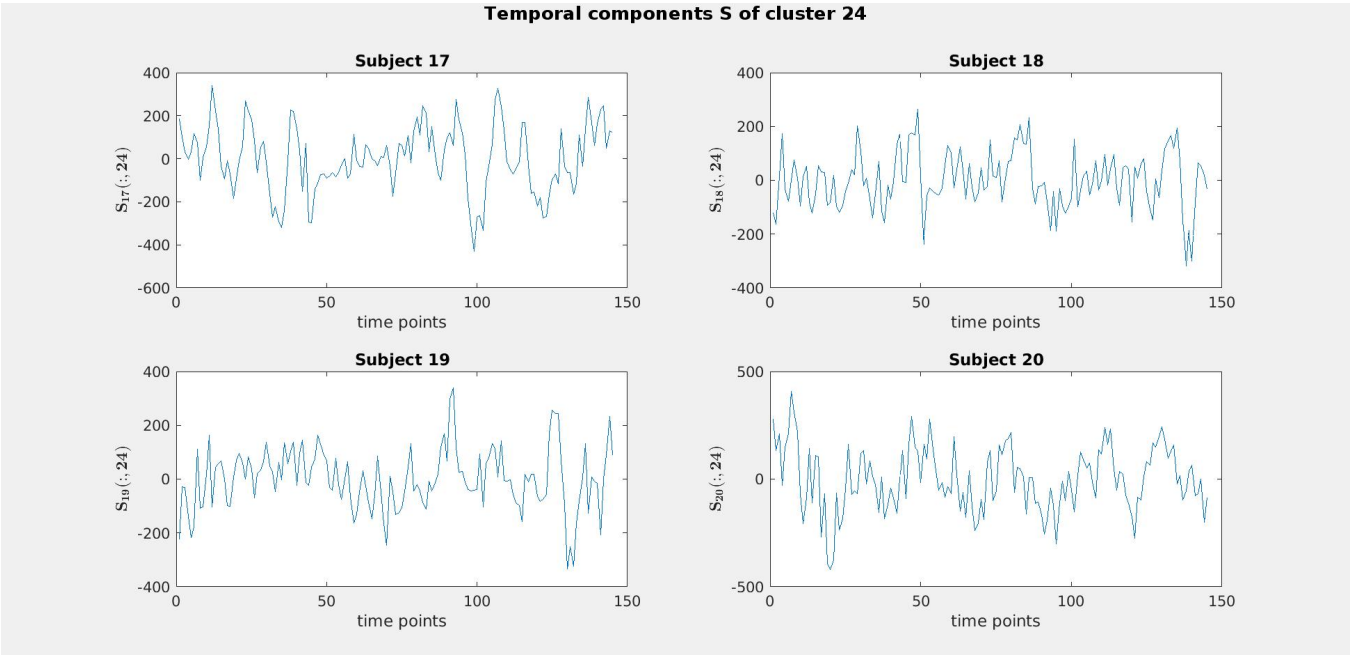


Figure B.120: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 17, \ldots, 20$ and R=27.

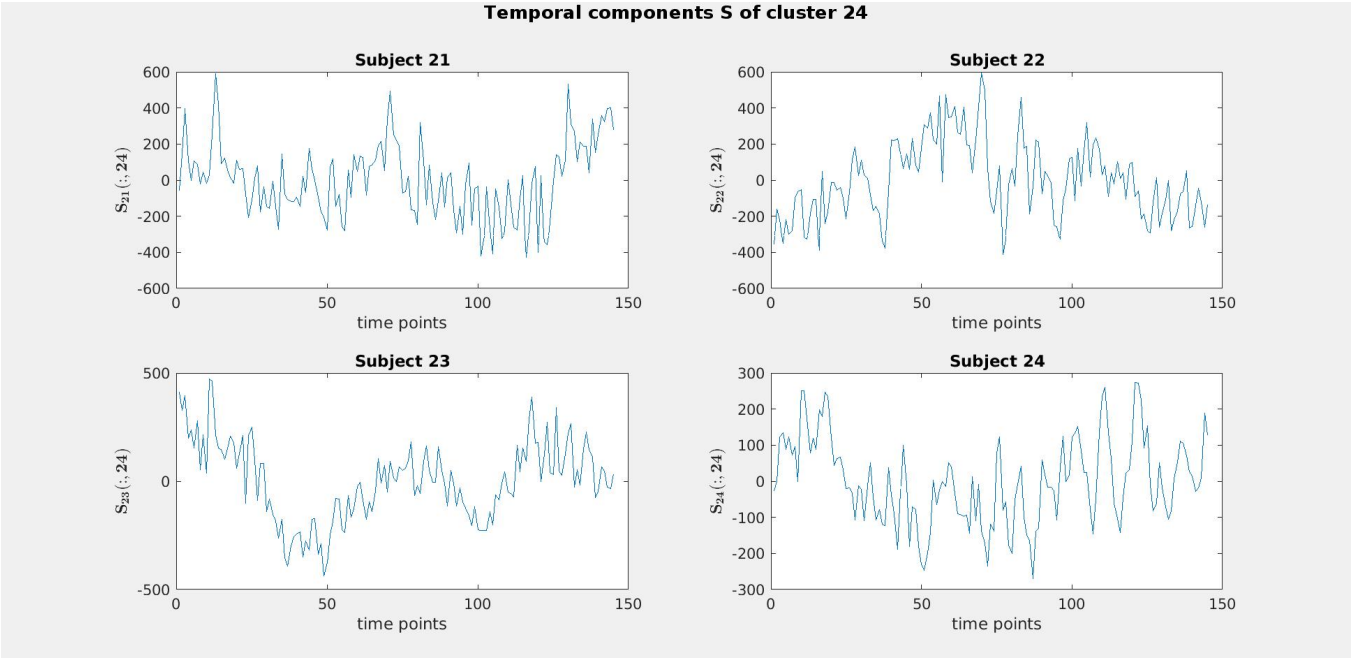


Figure B.121: Waveform of temporal components $\mathbf{S}_k(:, 124)$ for $k = 21, \ldots, 24$ and R=27.

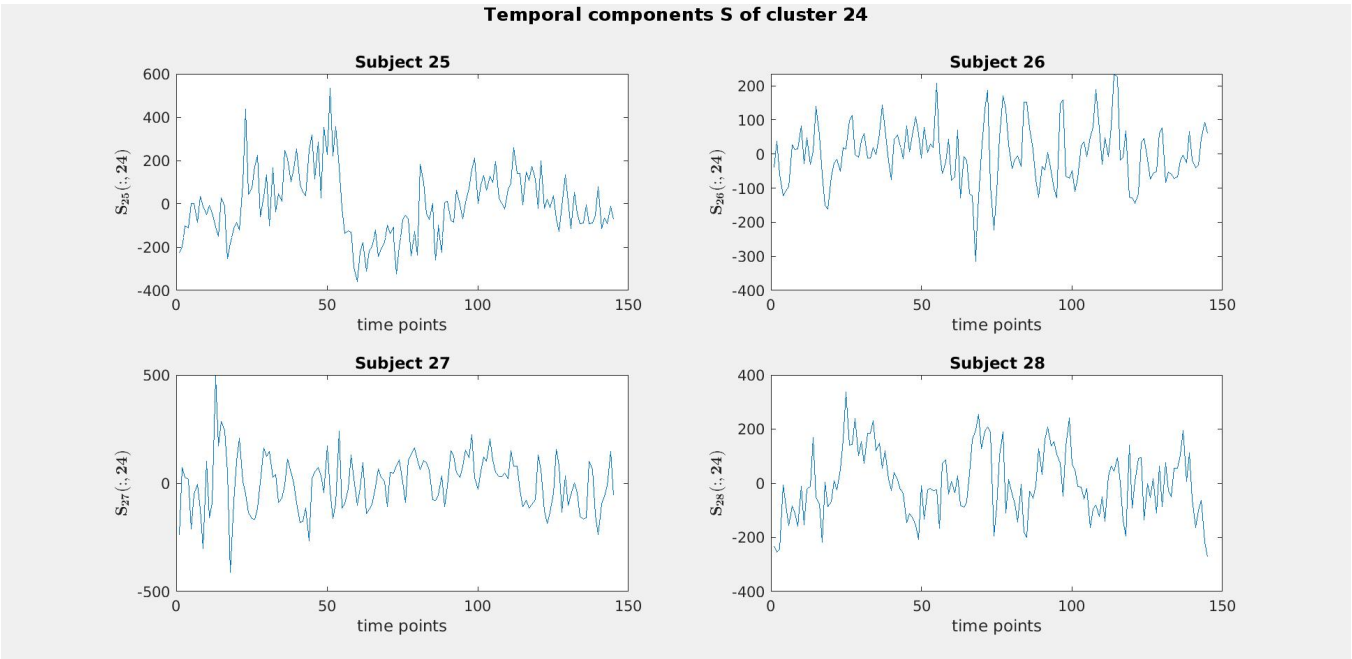


Figure B.122: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 25, \ldots, 28$ and R=27.

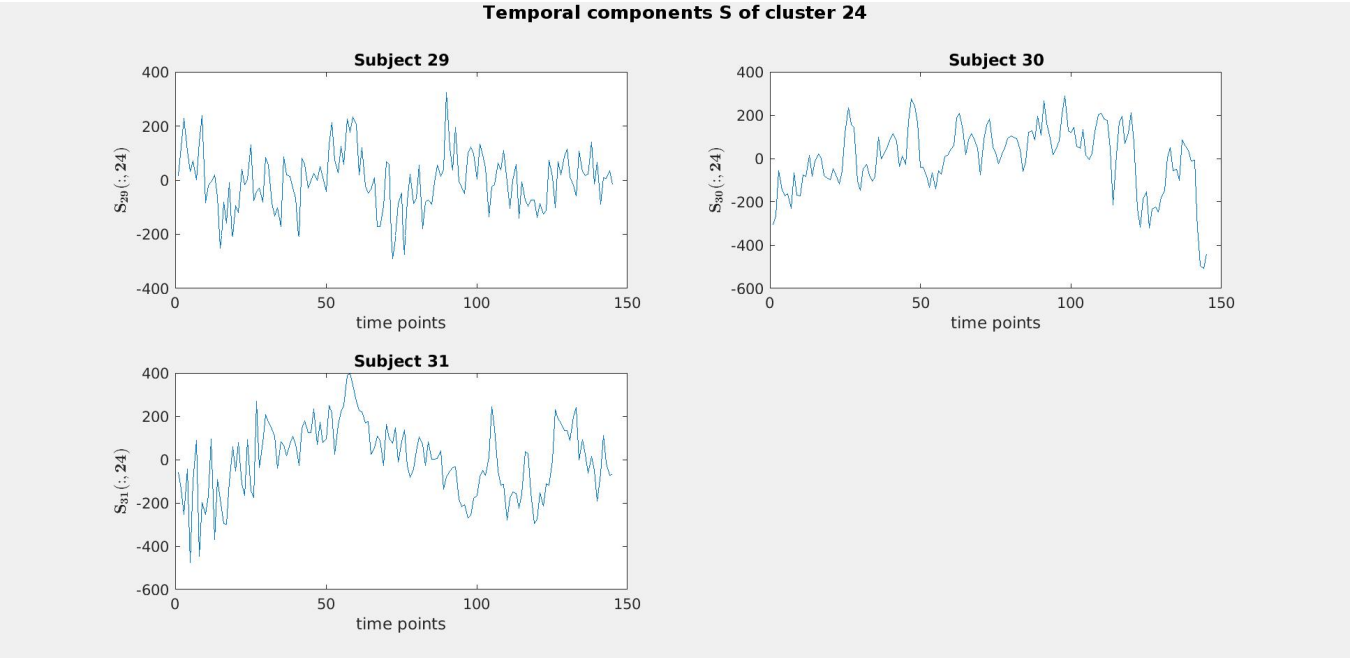


Figure B.123: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 29, \ldots, 31$ and R=27.

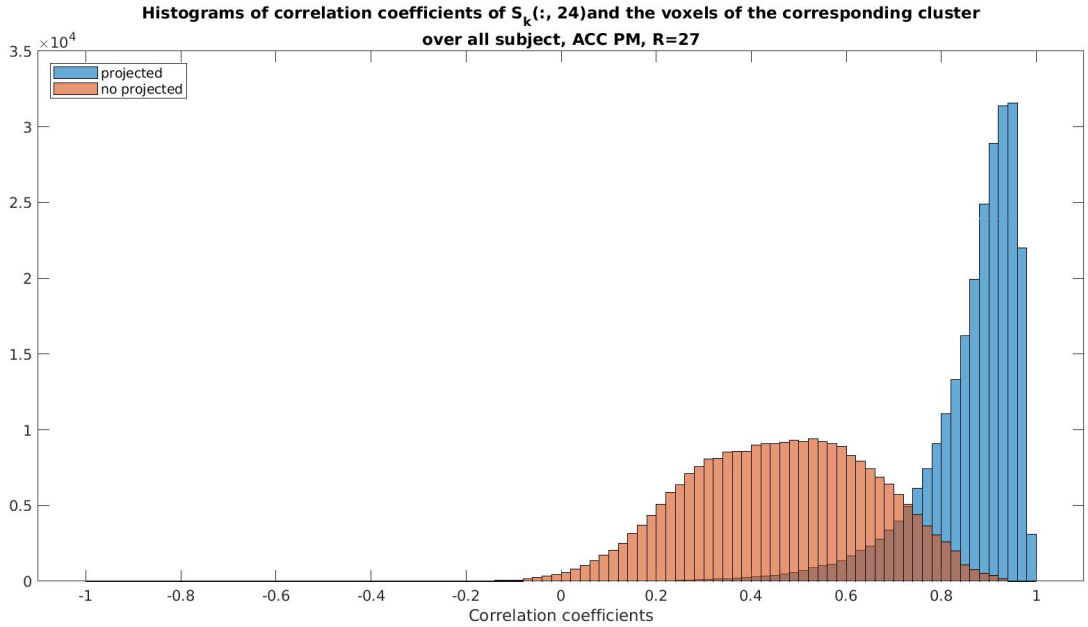


Figure B.124: For the column 24 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 24)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

11. **Cluster 6**

    Cluster 6 contains 9826 voxels, which correspond to 4.348788% of the whole brain (background, soft tissue, bone are excluded).

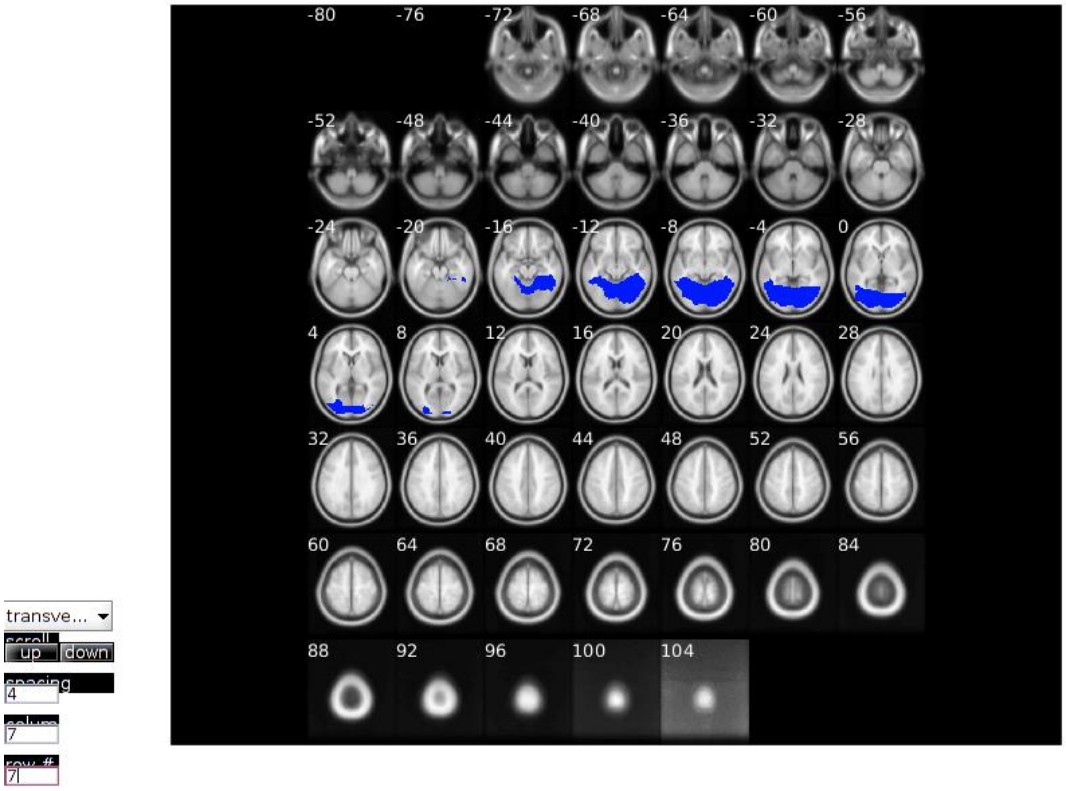


Figure B.125: Cluster 6. Blue areas indicate the voxels which belong to cluster 6.

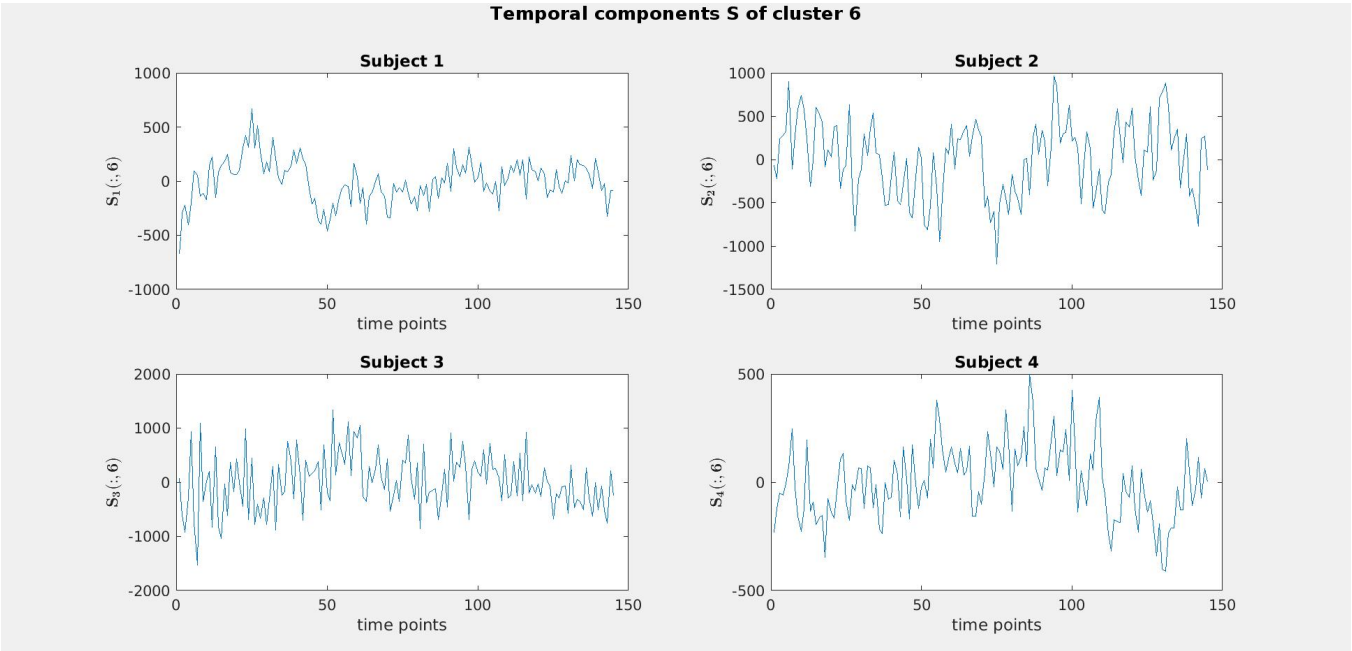


Figure B.126: Waveform of temporal components $\mathbf{S}_k(:, 6)$ for $k = 1, \ldots, 4$ and R=27.

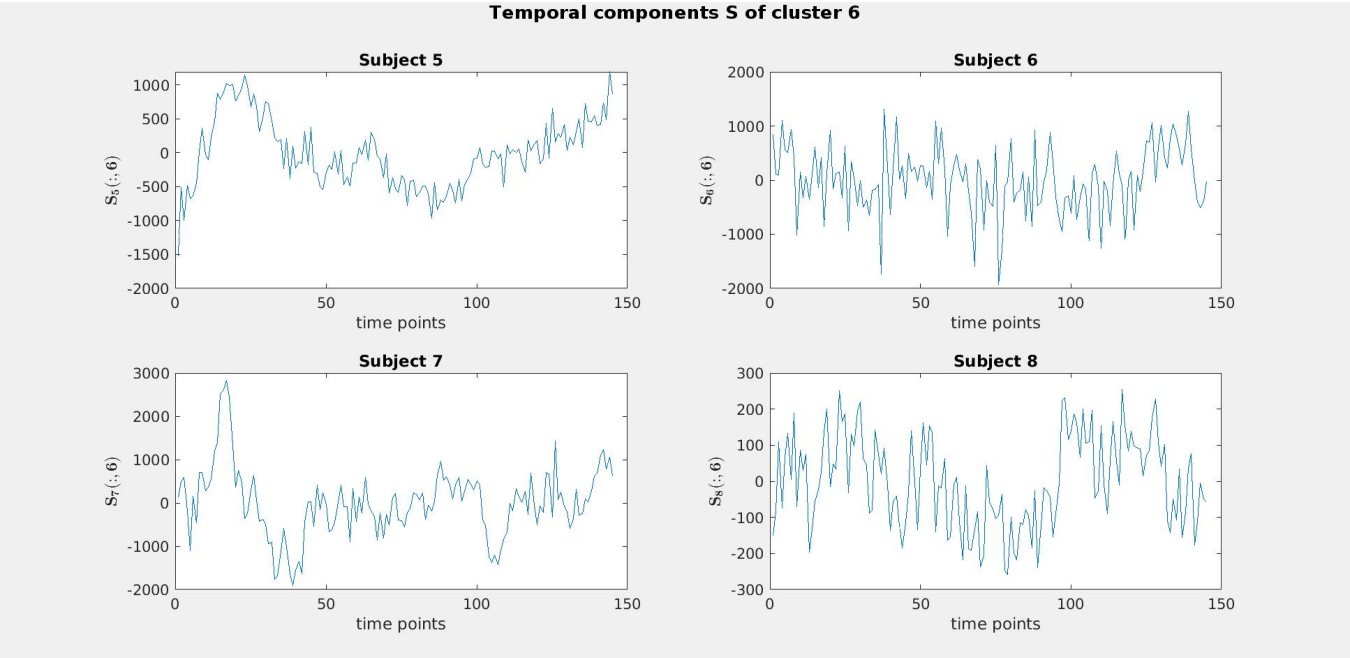


Figure B.127: Waveform of temporal components $\mathbf{S}_k(:,6)$ for $k = 5,\ldots,8$ and R=27.

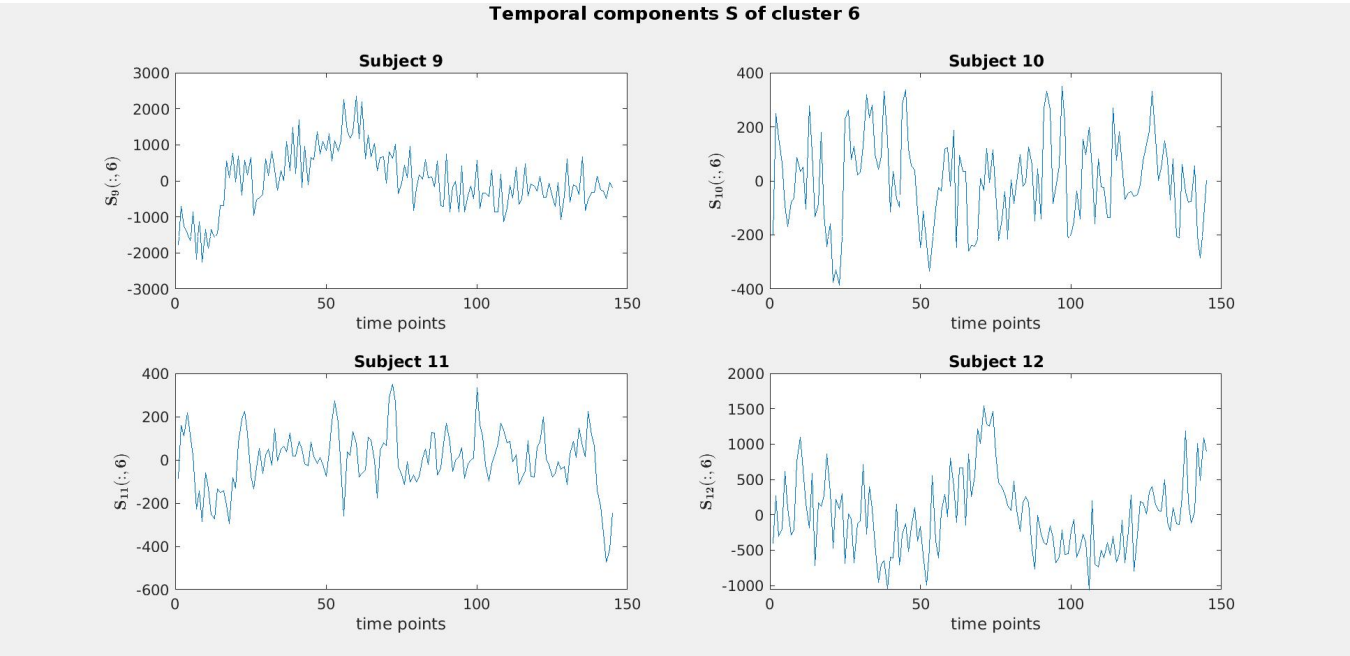


Figure B.128: Waveform of temporal components $\mathbf{S}_k(:,6)$ for $k = 9,\ldots,12$ and R=27.

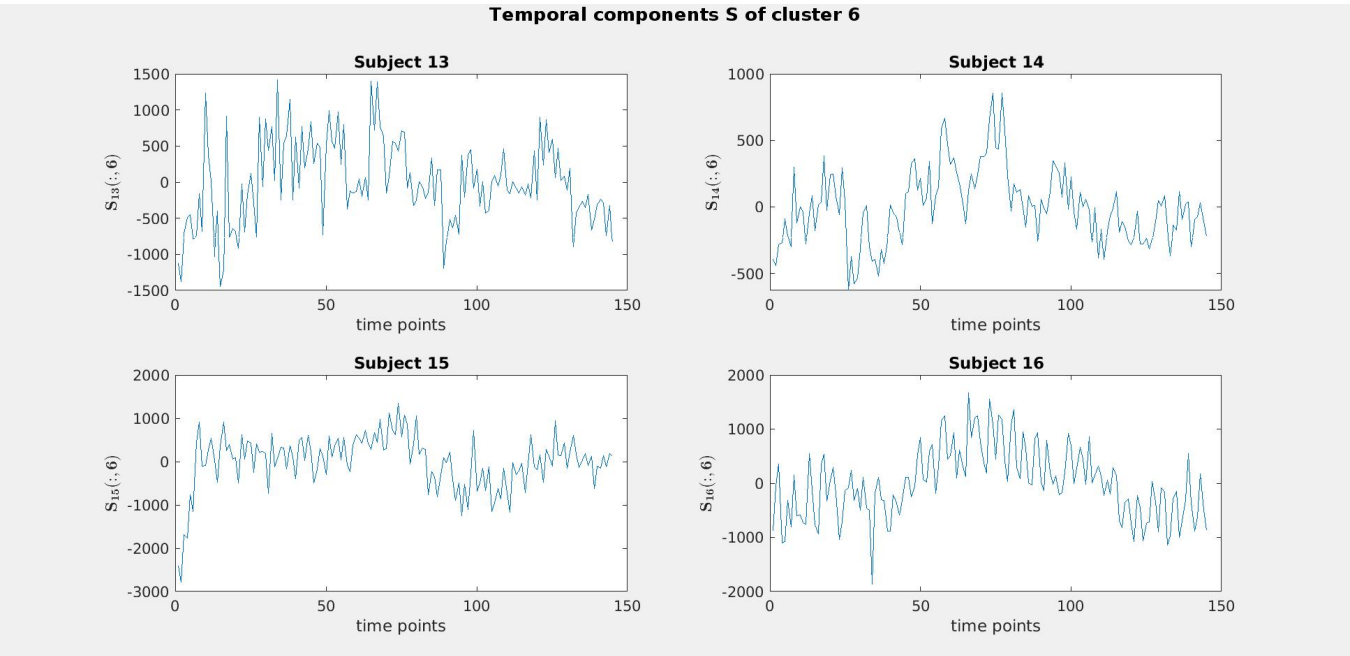


Figure B.129: Waveform of temporal components $\mathbf{S}_k(:,6)$ for $k = 13,\ldots,16$ and R=27.

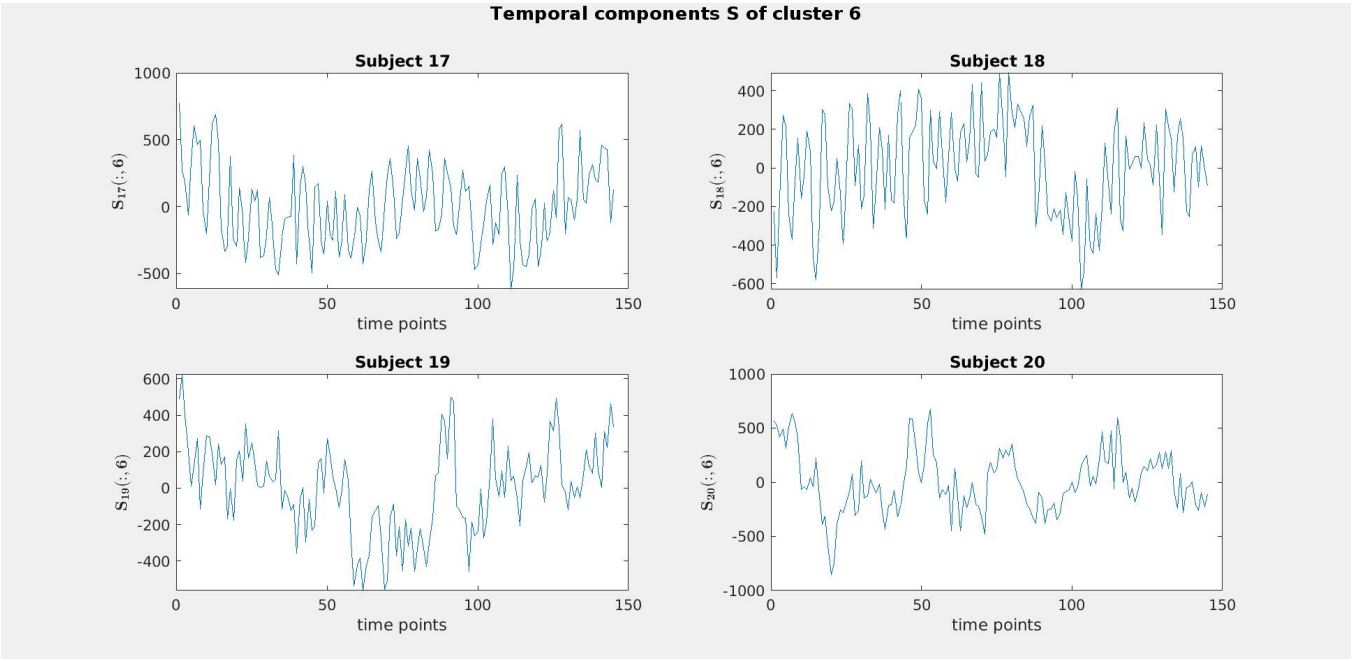


Figure B.130: Waveform of temporal components $\mathbf{S}_k(:, 6)$ for $k = 17, \ldots, 20$ and R=27.

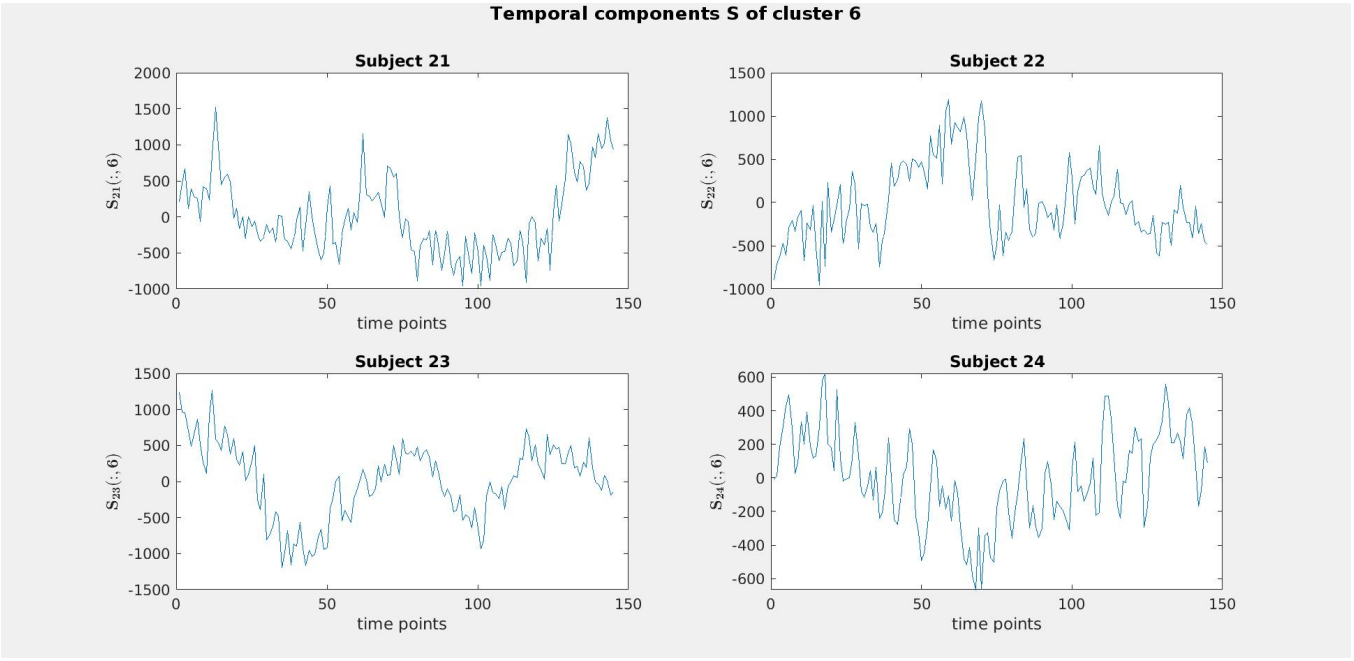


Figure B.131: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 21, \ldots, 24$ and R=27.

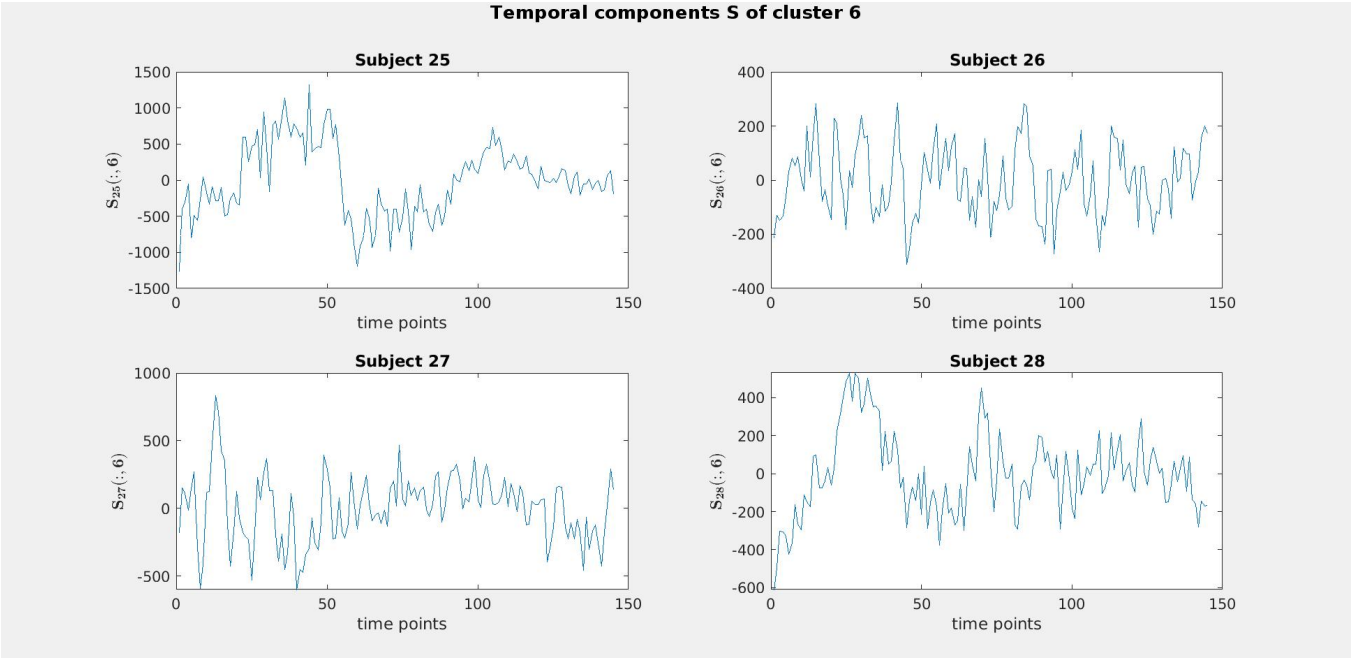


Figure B.132: Waveform of temporal components $\mathbf{S}_k(:, 6)$ for $k = 25, \ldots, 28$ and R=27.

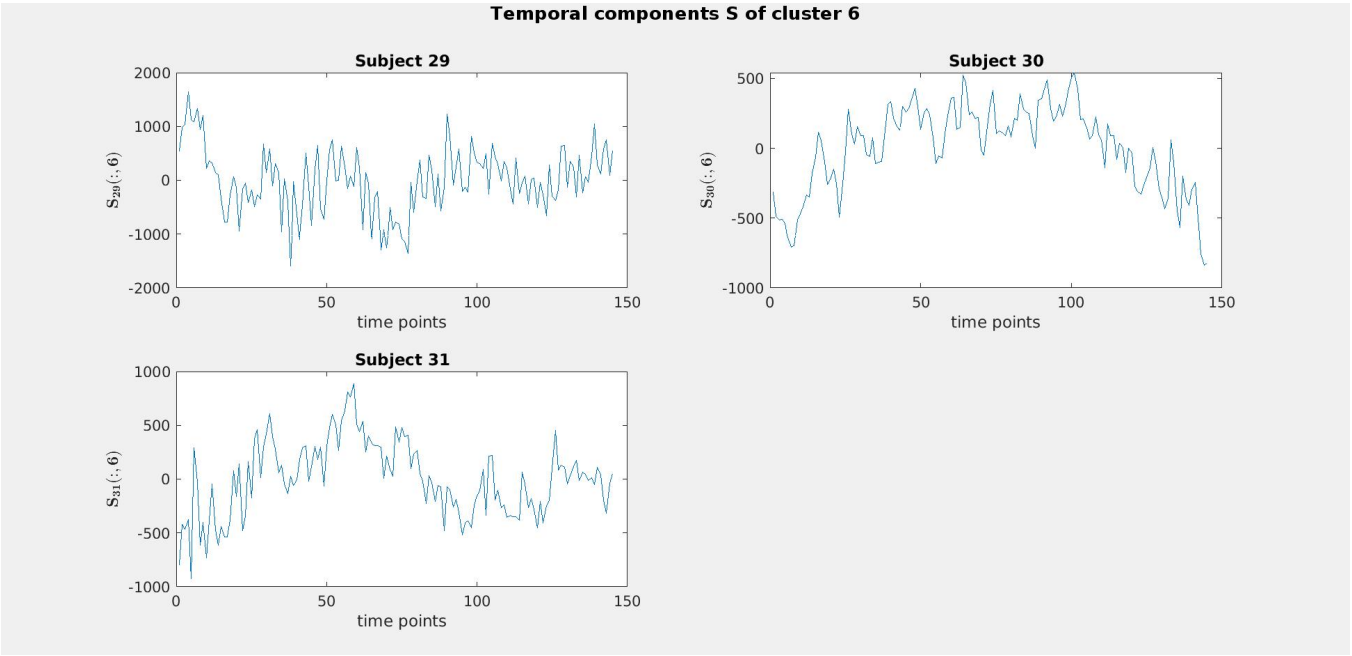


Figure B.133: Waveform of temporal components $\mathbf{S}_k(:, 6)$ for $k = 29, \ldots, 31$ and R=27.

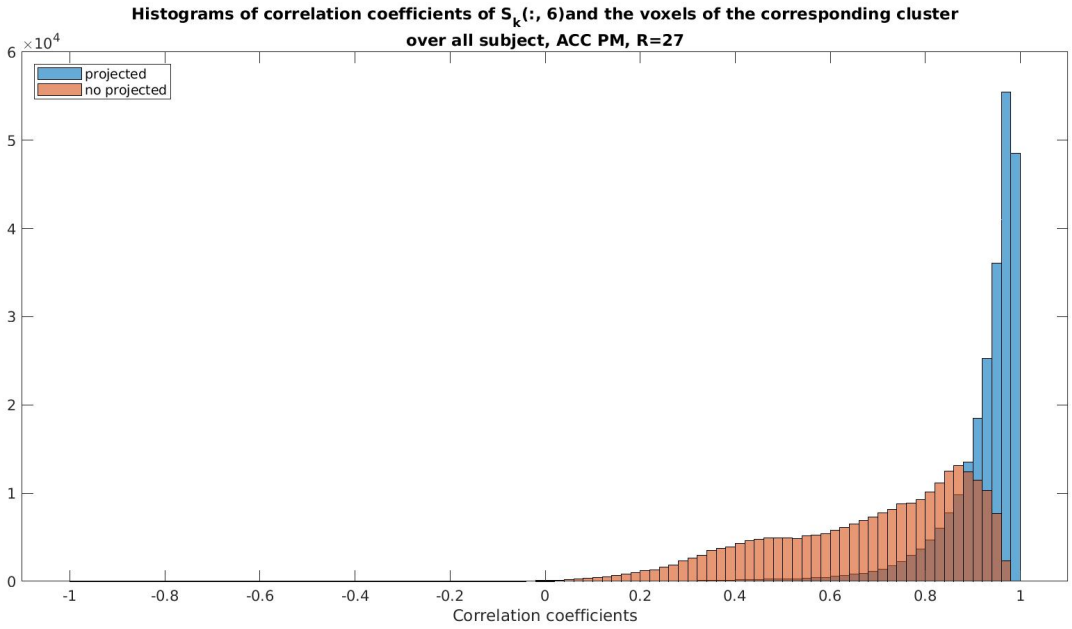


Figure B.134: For the column 6 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 6)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

12. **Cluster 23**

    Cluster 23 contains 7343 voxels, which correspond to 3.249863% of the whole brain (background, soft tissue, bone are excluded).

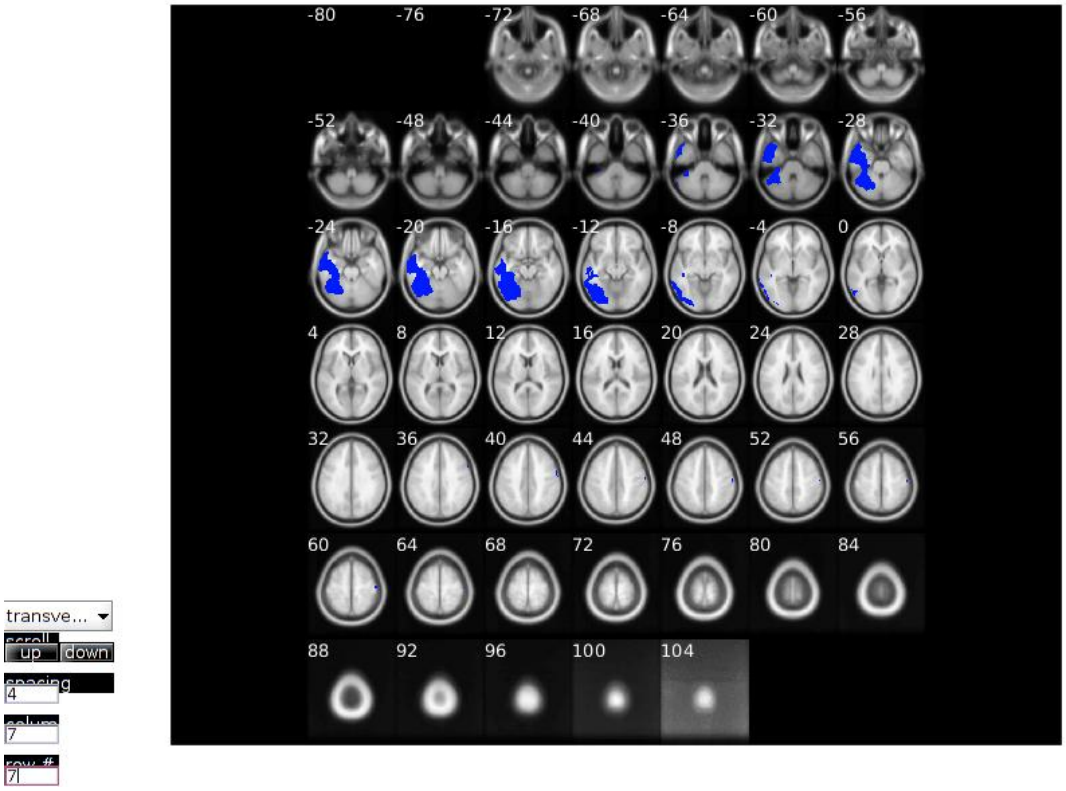


Figure B.135: Cluster 23. Blue areas indicate the voxels which belong to cluster 23.

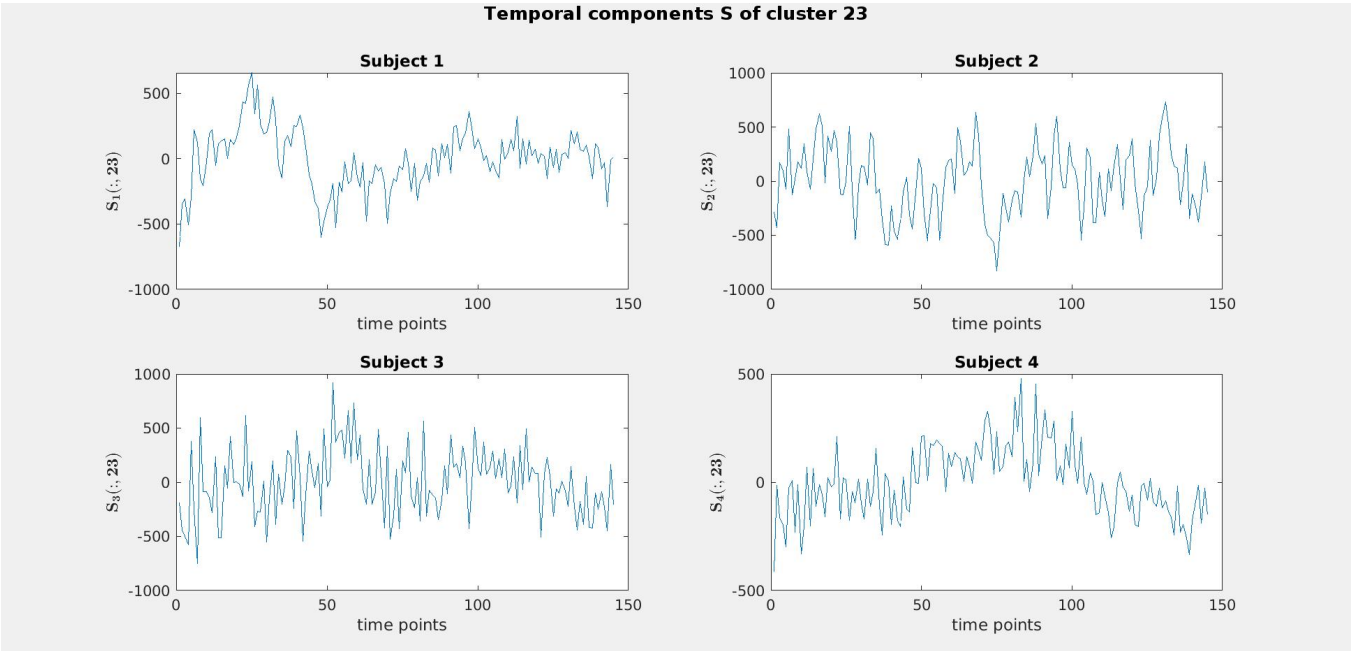


Figure B.136: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 1, \ldots, 4$ and R=27.

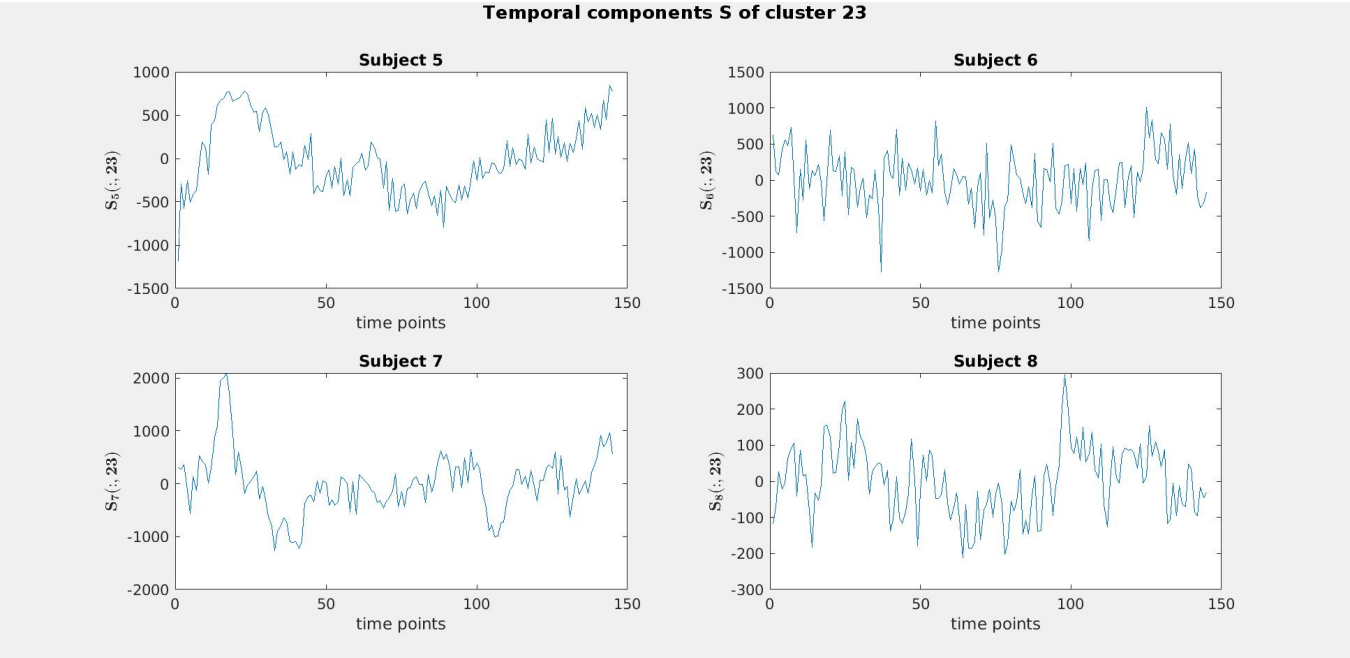


Figure B.137: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 5, \ldots, 8$ and R=27.

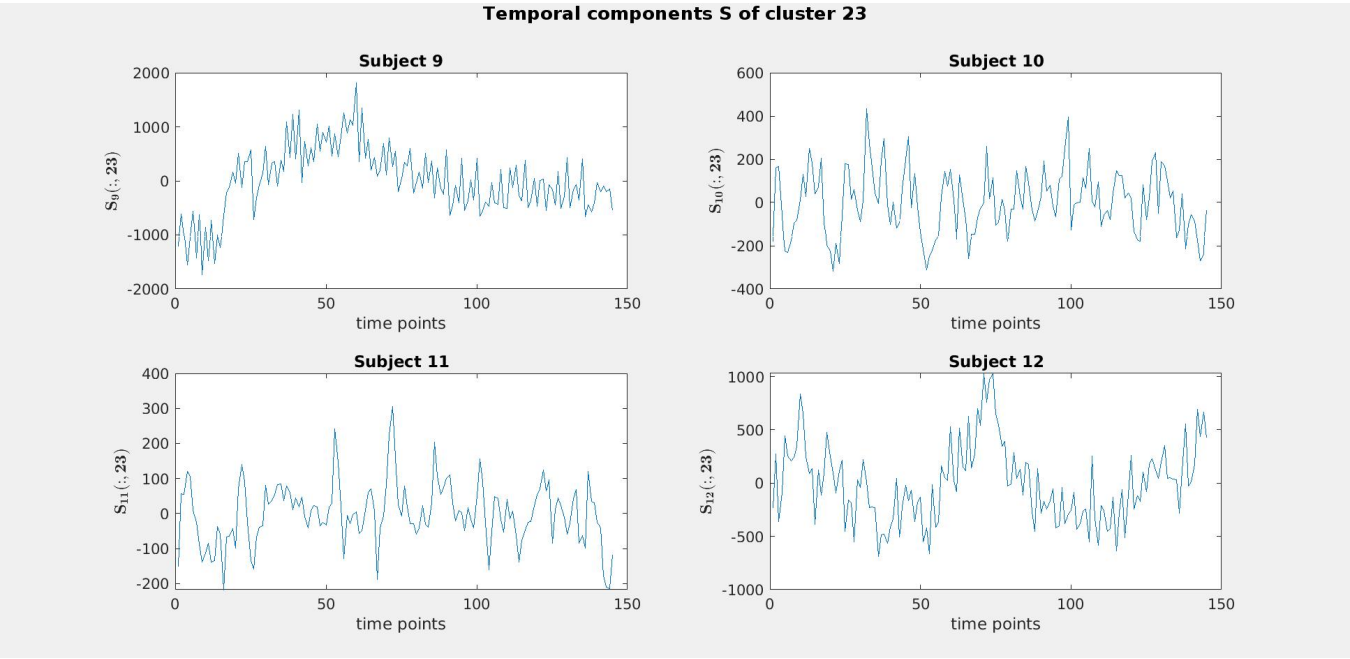


Figure B.138: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 9, \ldots, 12$ and R=27.

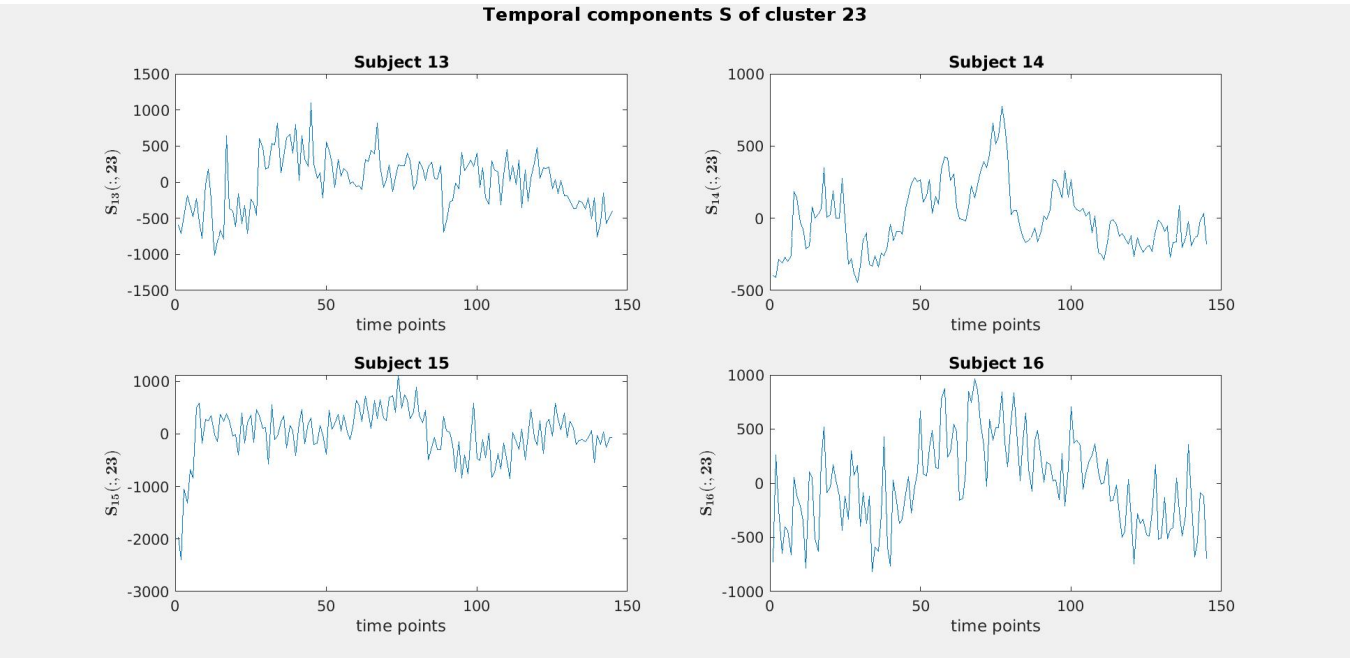


Figure B.139: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 13, \ldots, 16$ and R=27.

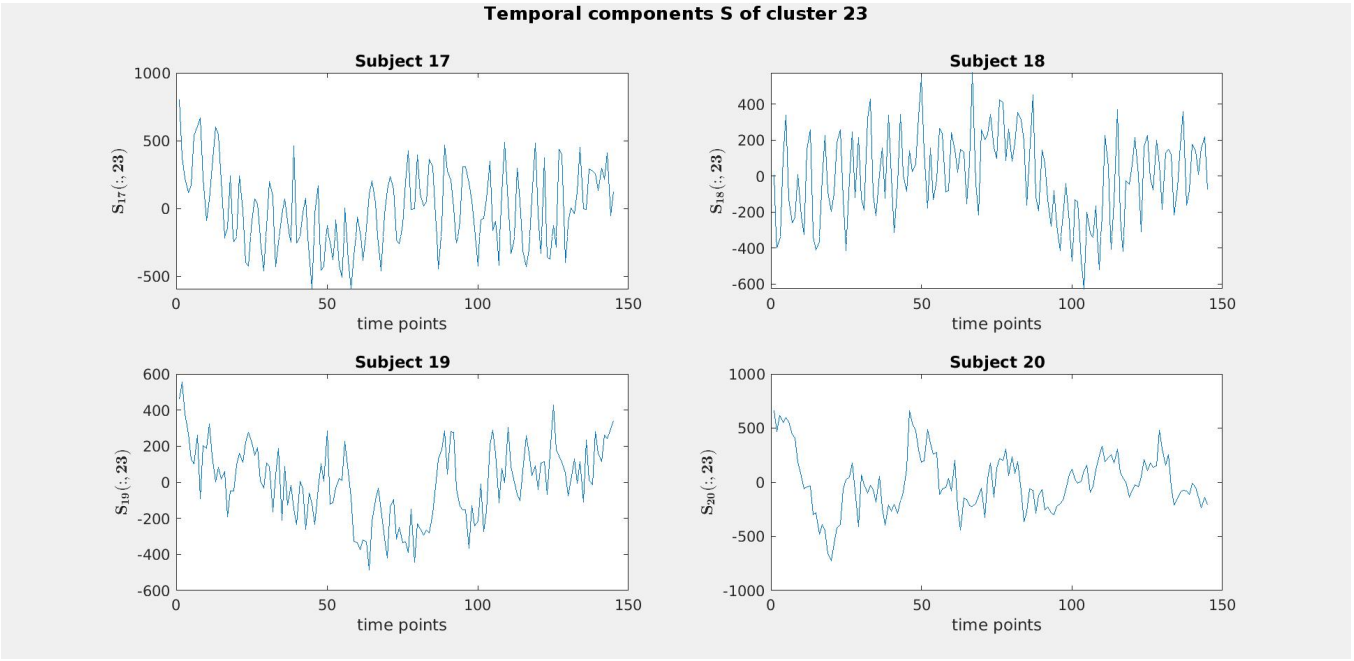


Figure B.140: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 17, \ldots, 20$ and R=27.

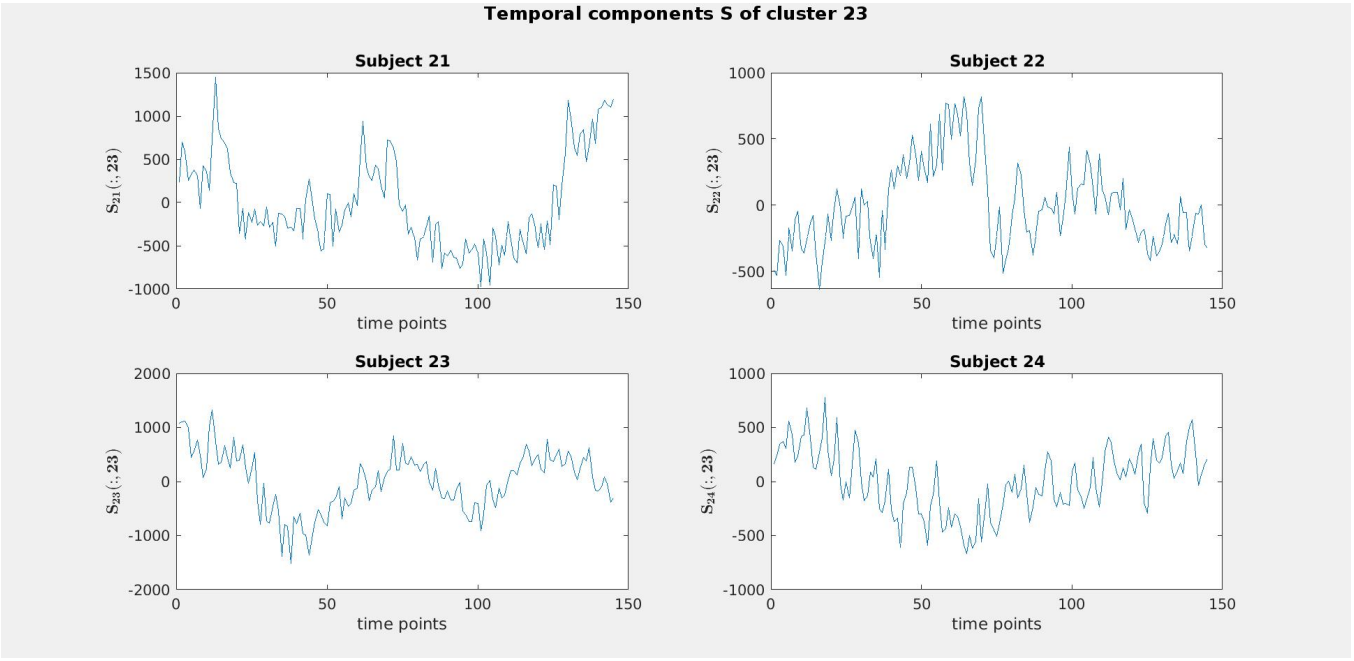


Figure B.141: Waveform of temporal components $\mathbf{S}_k(:, 123)$ for $k = 21, \ldots, 24$ and R=27.

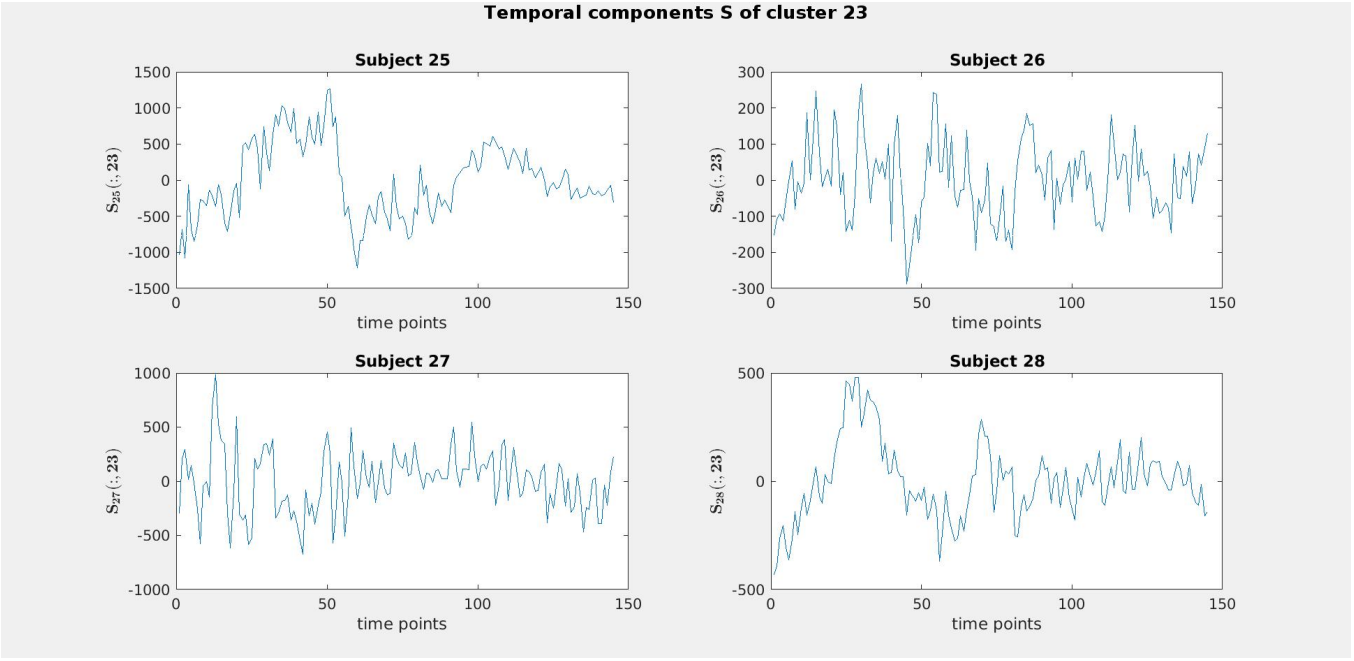


Figure B.142: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 25, \ldots, 28$ and R=27.

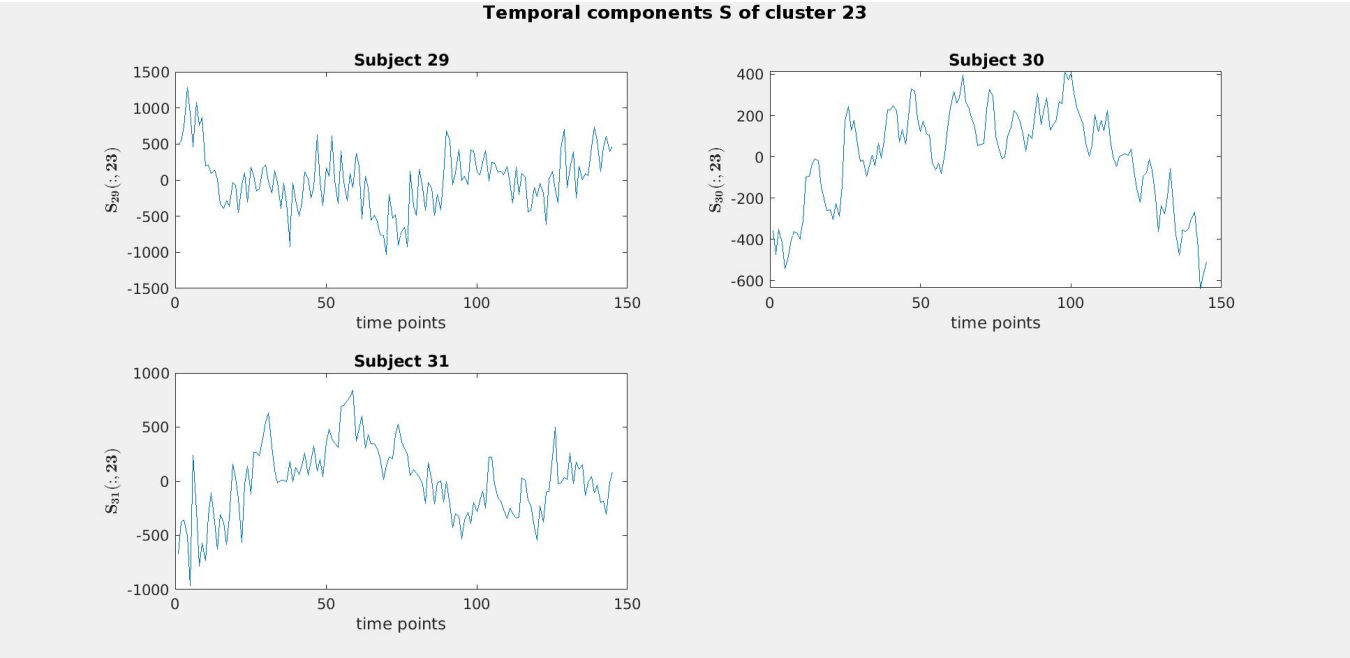


Figure B.143: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 29, \ldots, 31$ and R=27.

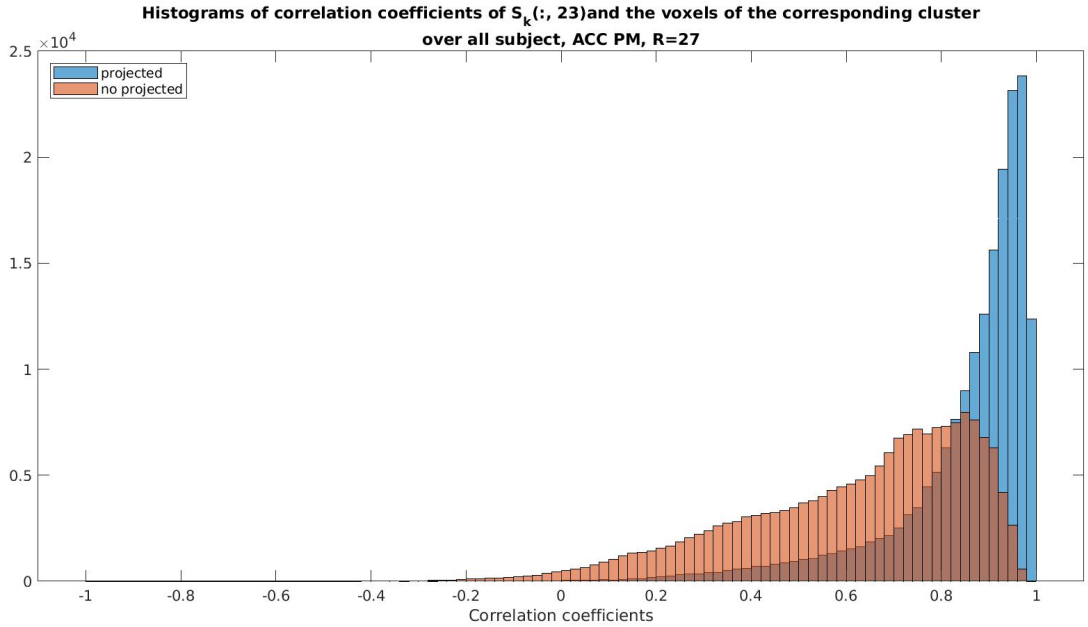


Figure B.144: For the column 23 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 23)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

13. **Cluster 5**

    Cluster 5 contains 7308 voxels, which correspond to 3.234373% of the whole brain (background, soft tissue, bone are excluded).

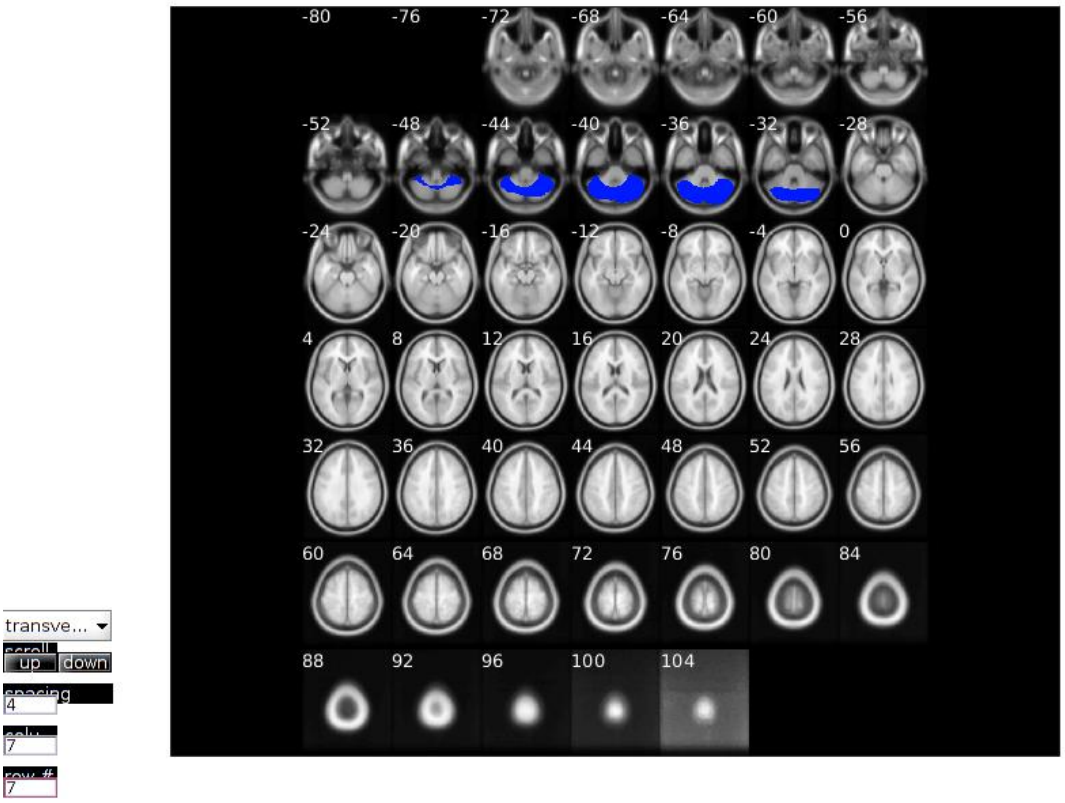


Figure B.145: Cluster 5. Blue areas indicate the voxels which belong to cluster 5.

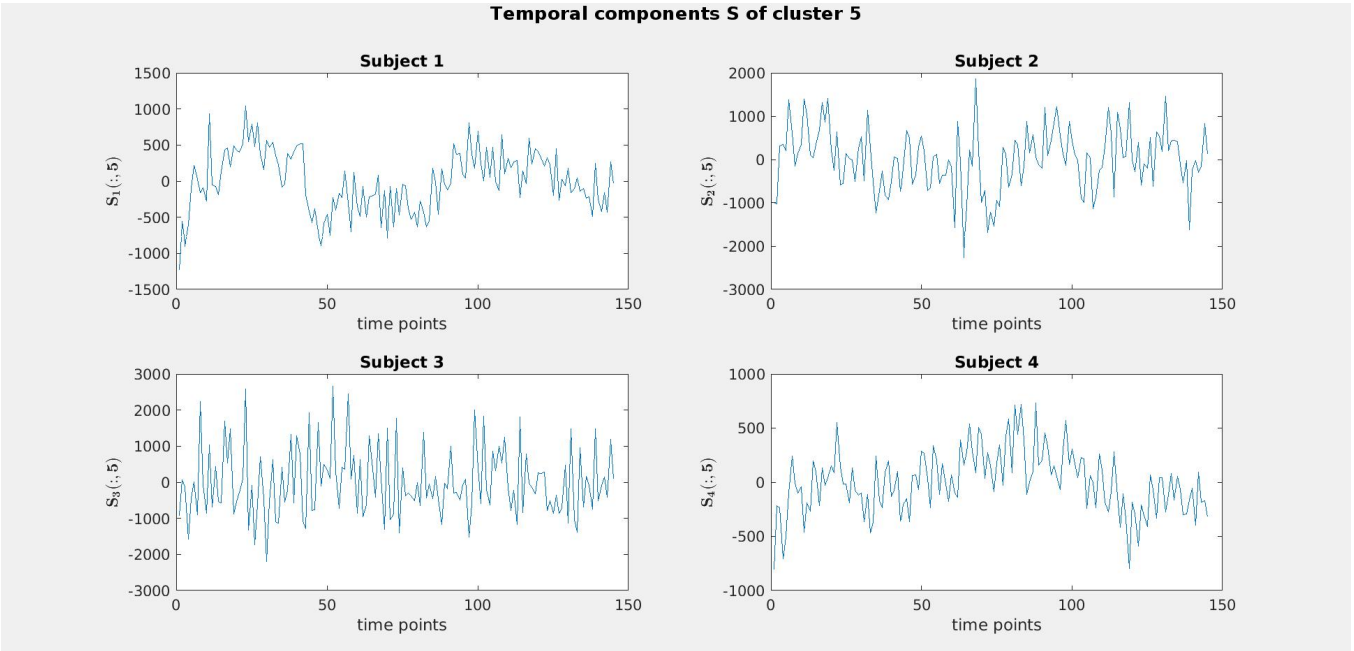


Figure B.146: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 1,\ldots,4$ and R=27.

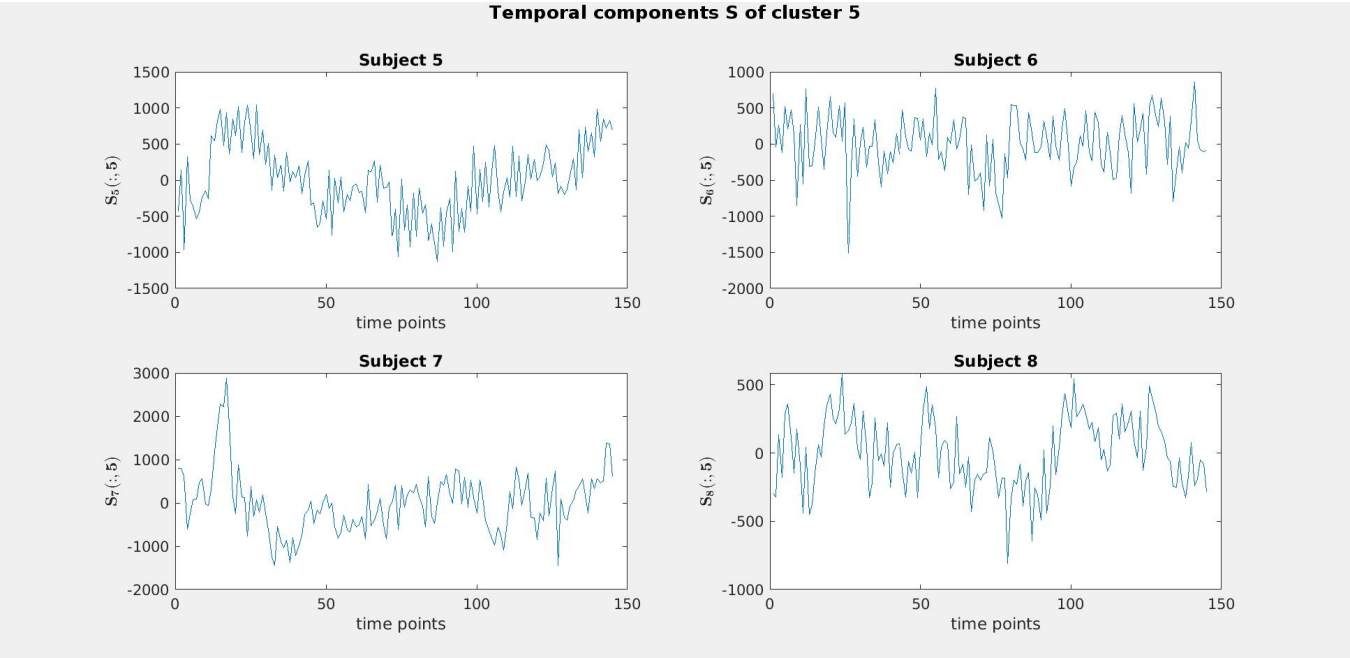


Figure B.147: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 5,\ldots,8$ and R=27.

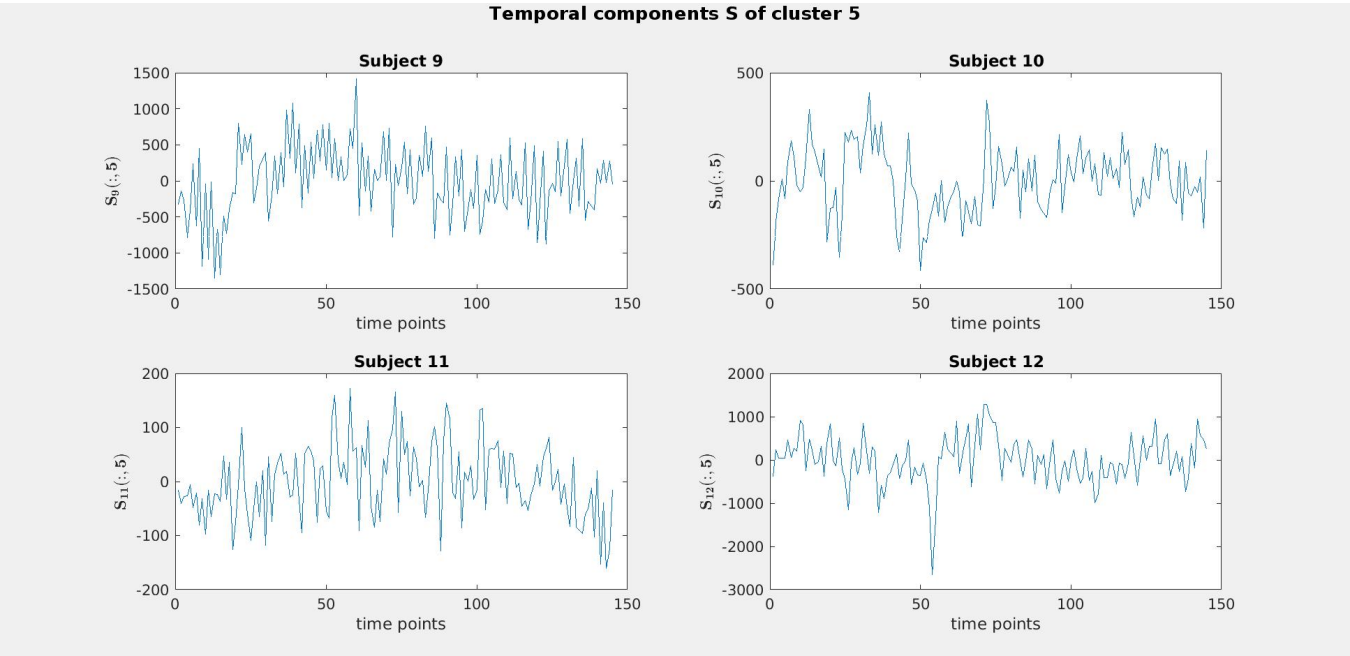


Figure B.148: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 9,\ldots,12$ and R=27.

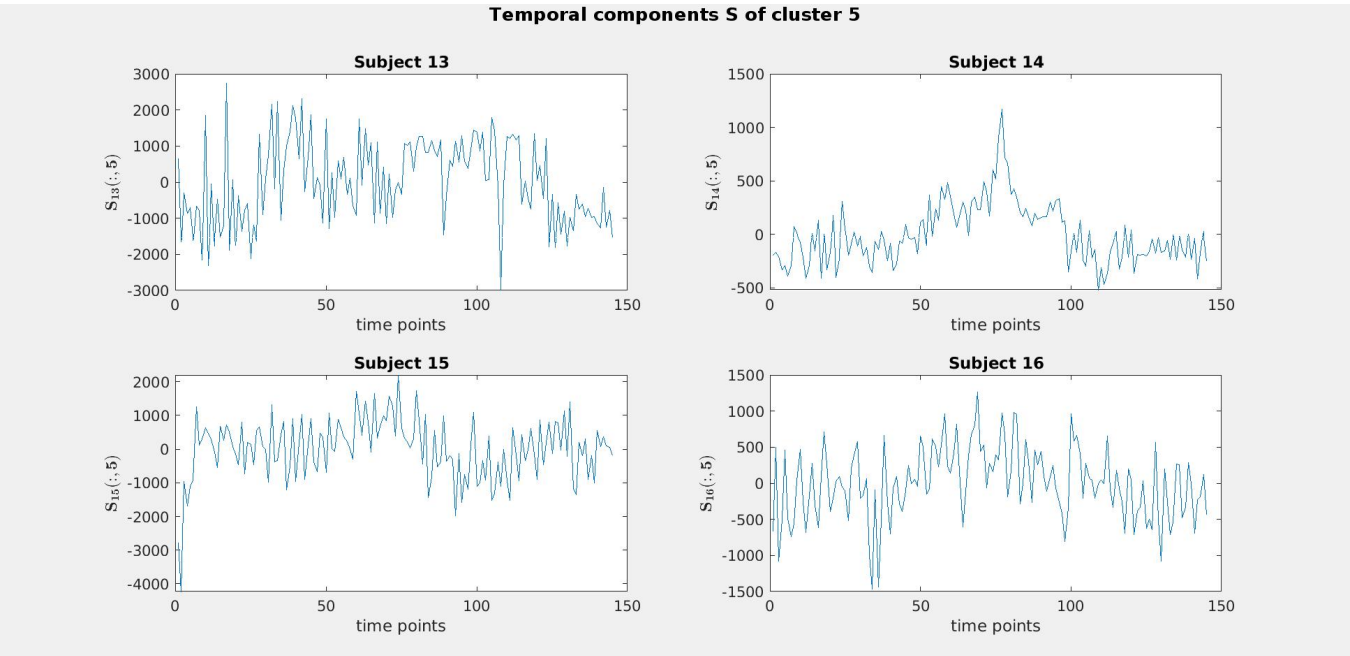


Figure B.149: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 13,\ldots,16$ and R=27.

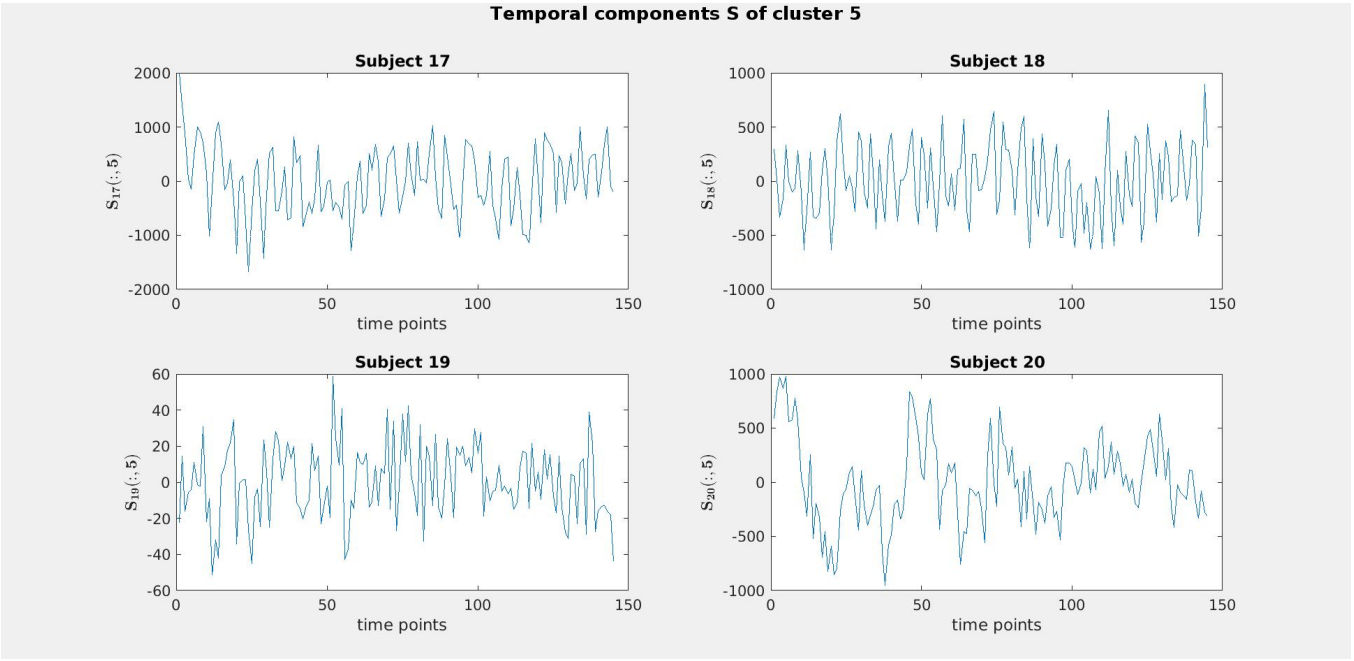


Figure B.150: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 17,\ldots,20$ and R=27.

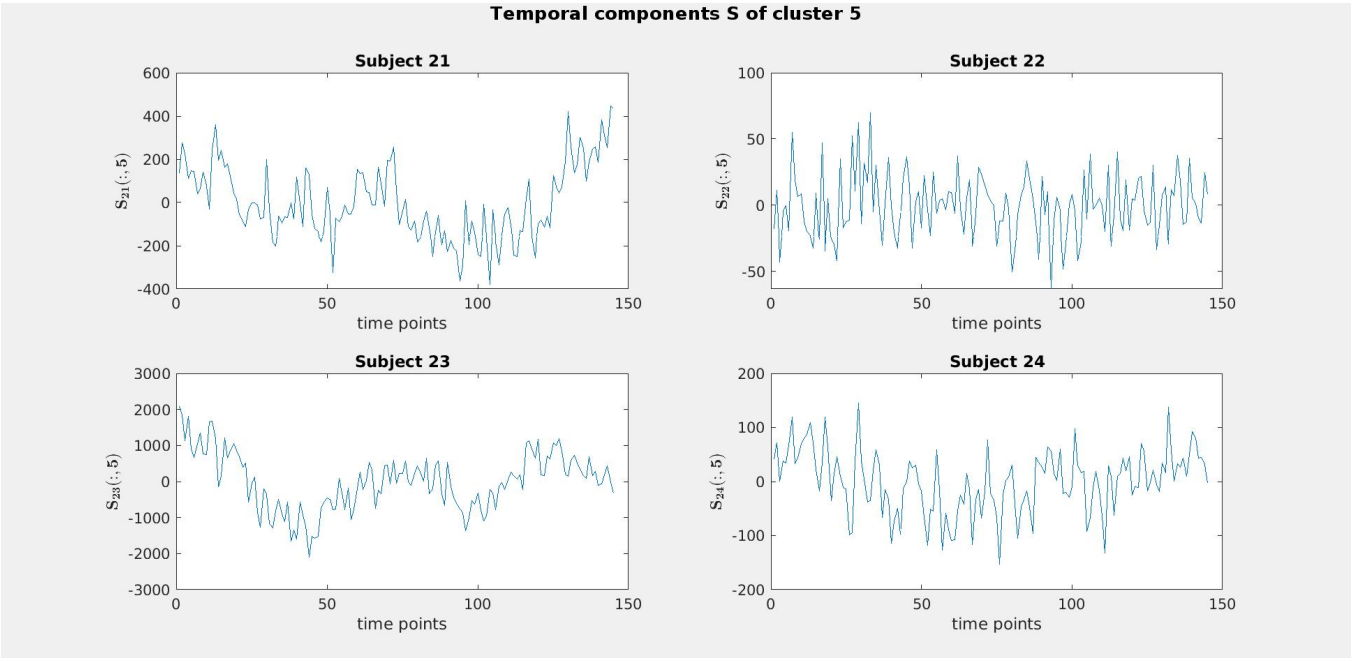


Figure B.151: Waveform of temporal components $\mathbf{S}_k(:,15)$ for $k = 21,\ldots,24$ and R=27.

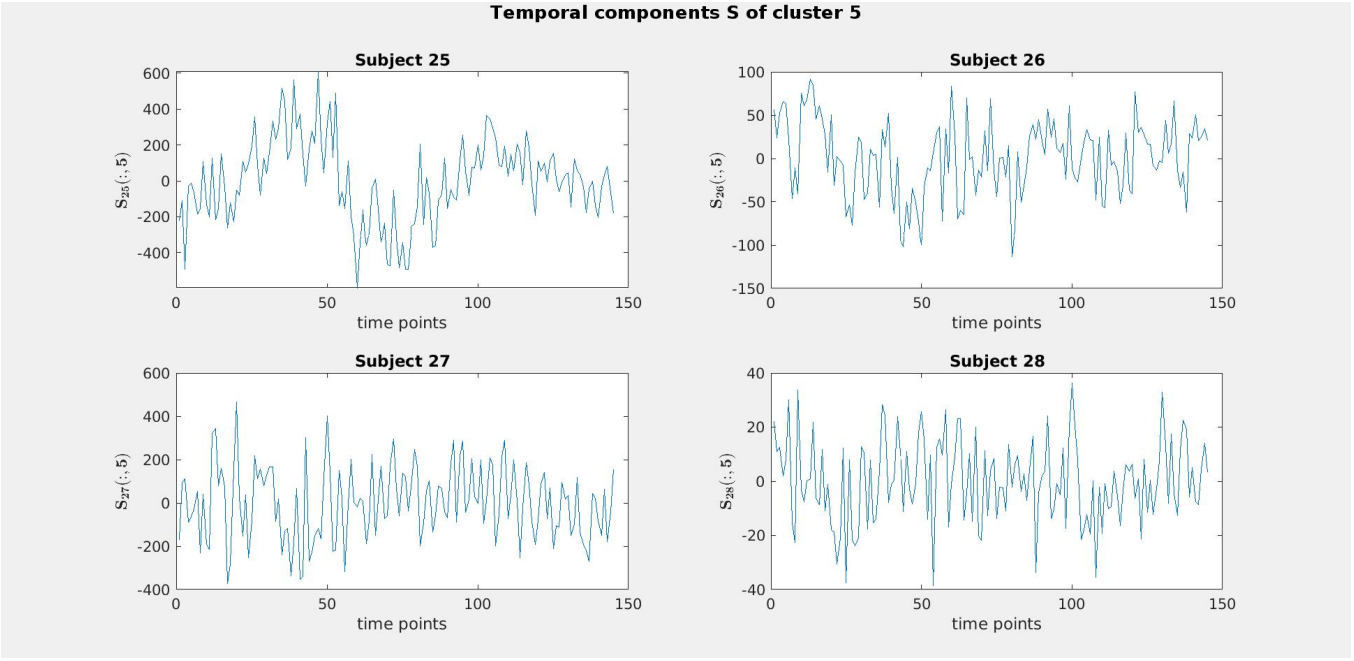


Figure B.152: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 25,\ldots,28$ and R=27.

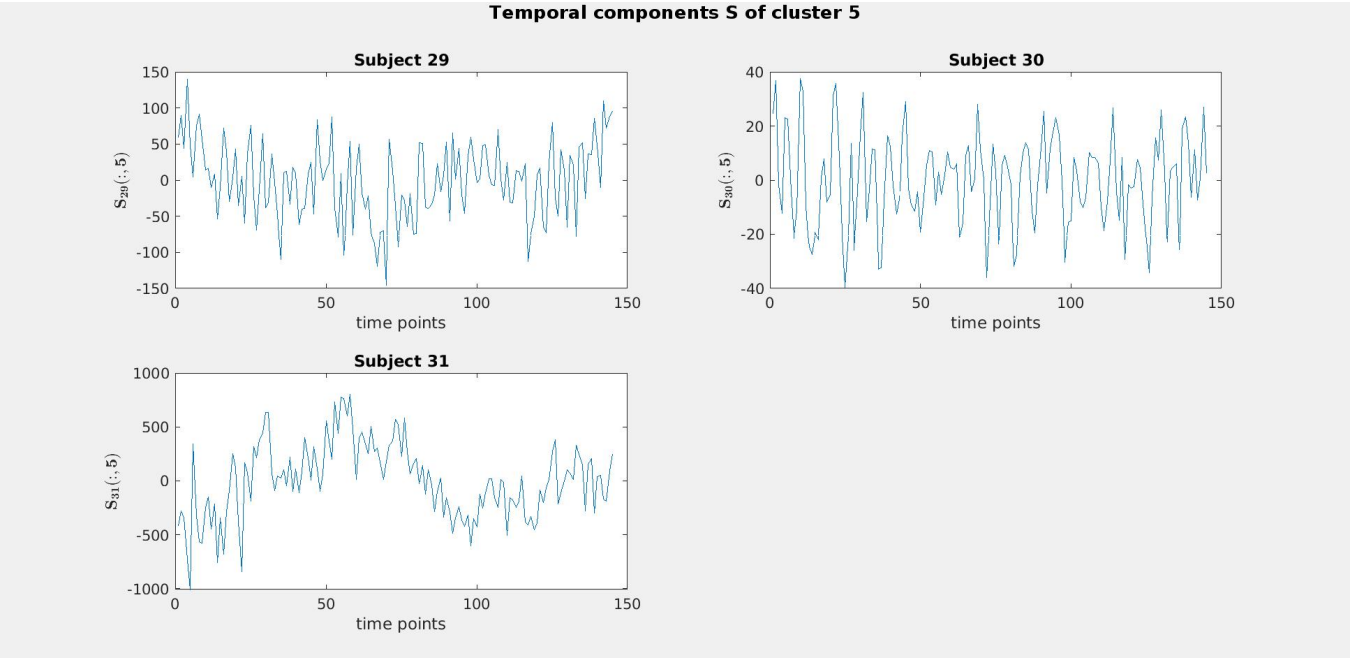


Figure B.153: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 29, \ldots, 31$ and R=27.

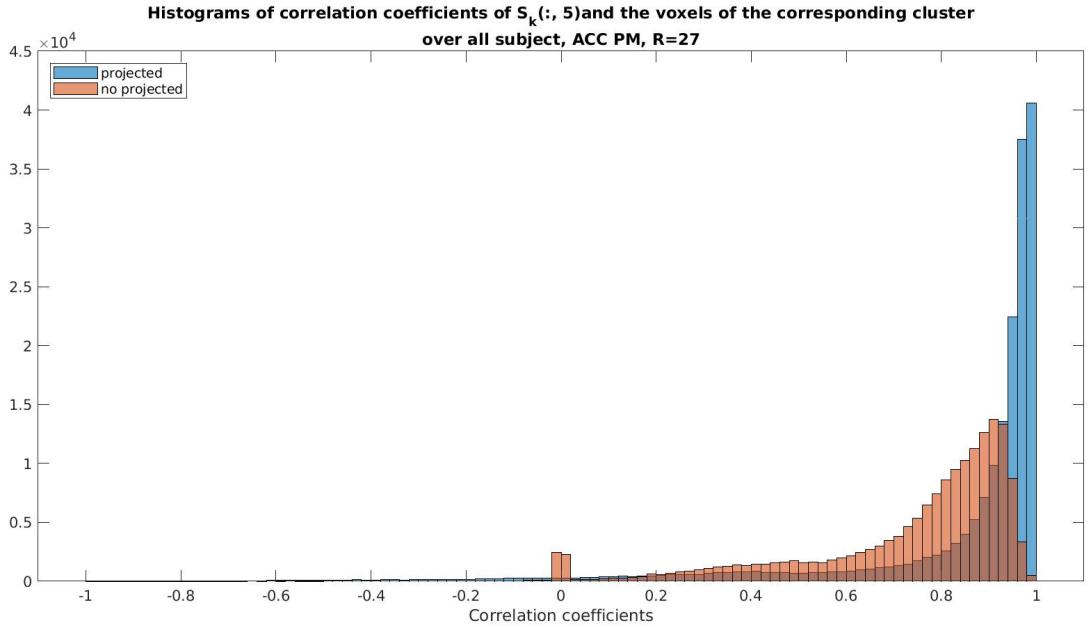


Figure B.154: For the column 5 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,5)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

14. **Cluster 20**

    Cluster 20 contains 6614 voxels, which correspond to 2.927222% of the whole brain (background, soft tissue, bone are excluded).

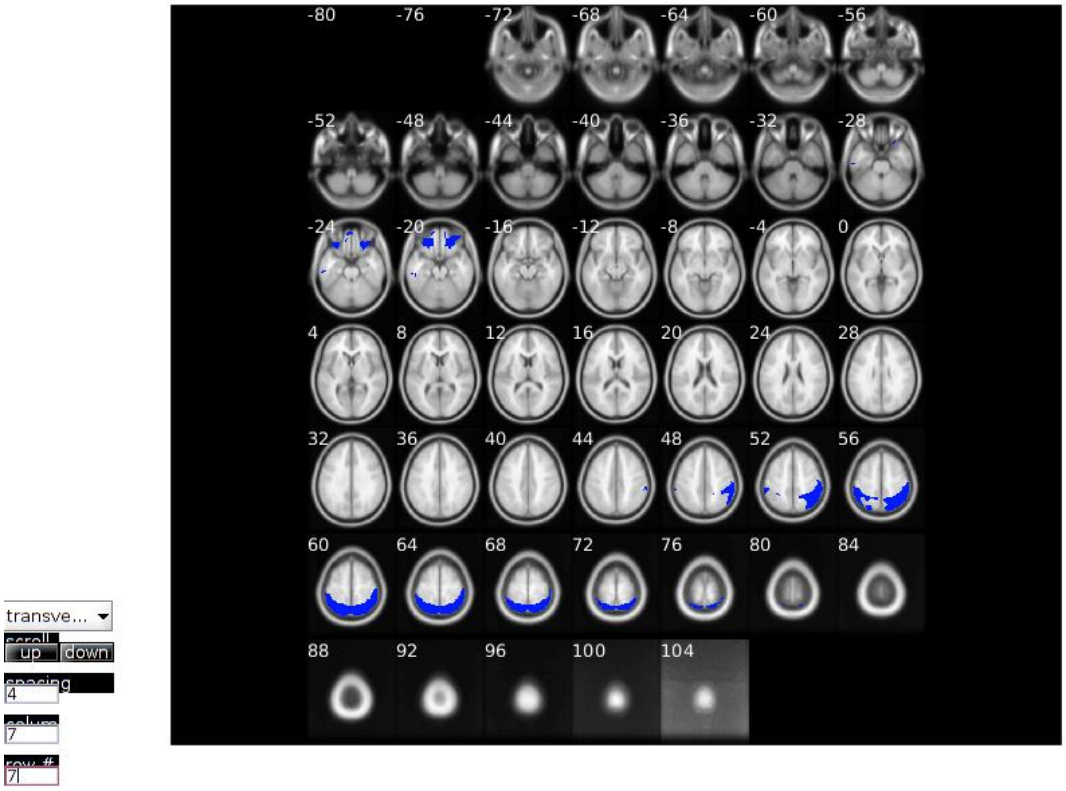


Figure B.155: Cluster 20. Blue areas indicate the voxels which belong to cluster 20.

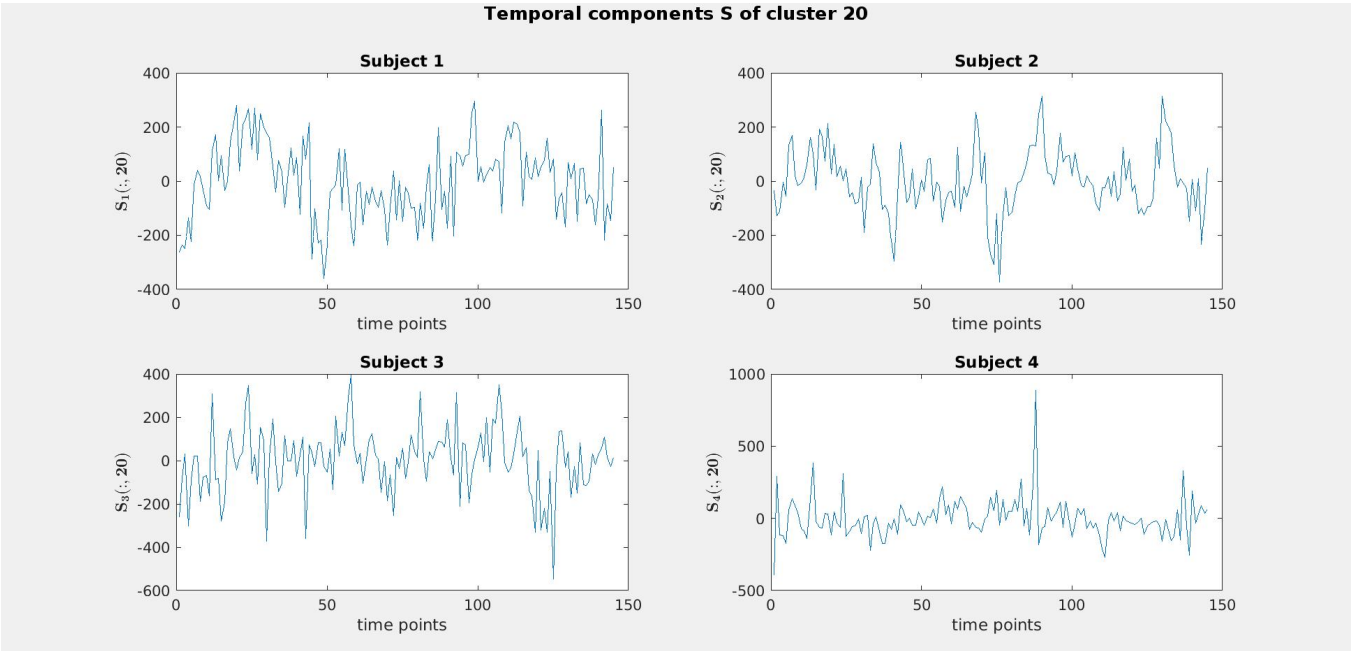


Figure B.156: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 1, \ldots, 4$ and R=27.

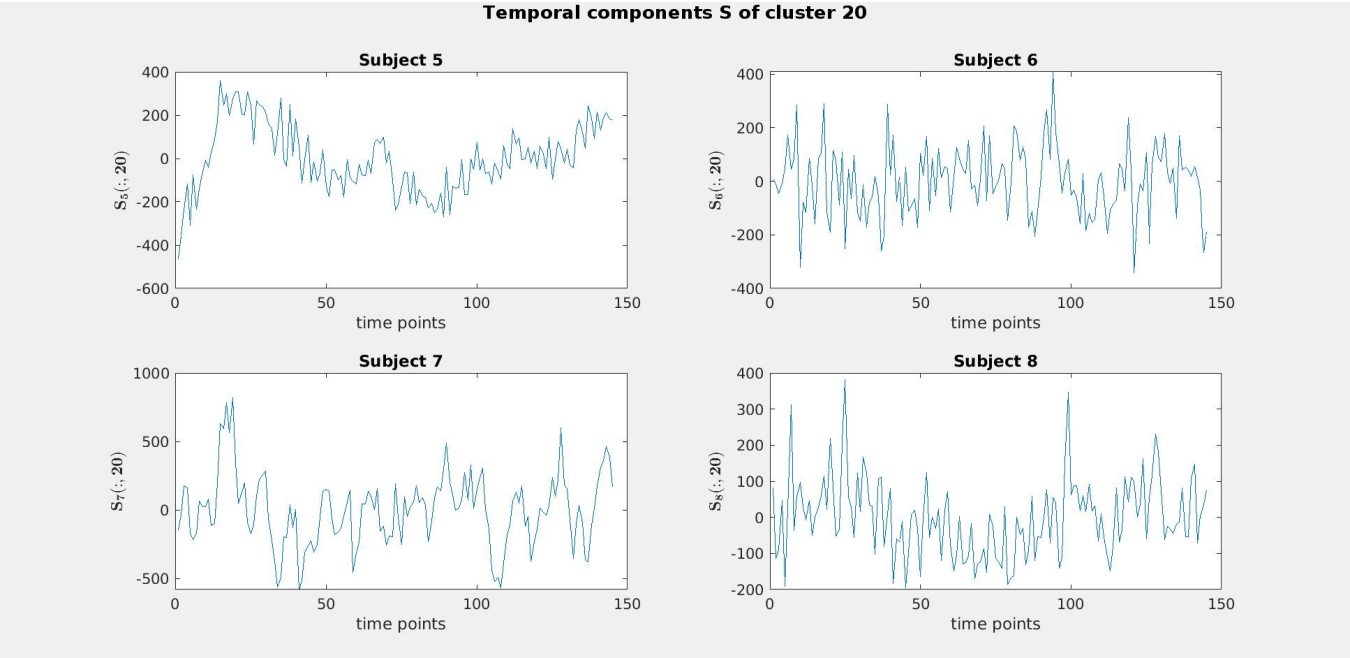


Figure B.157: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 5, \ldots, 8$ and R=27.

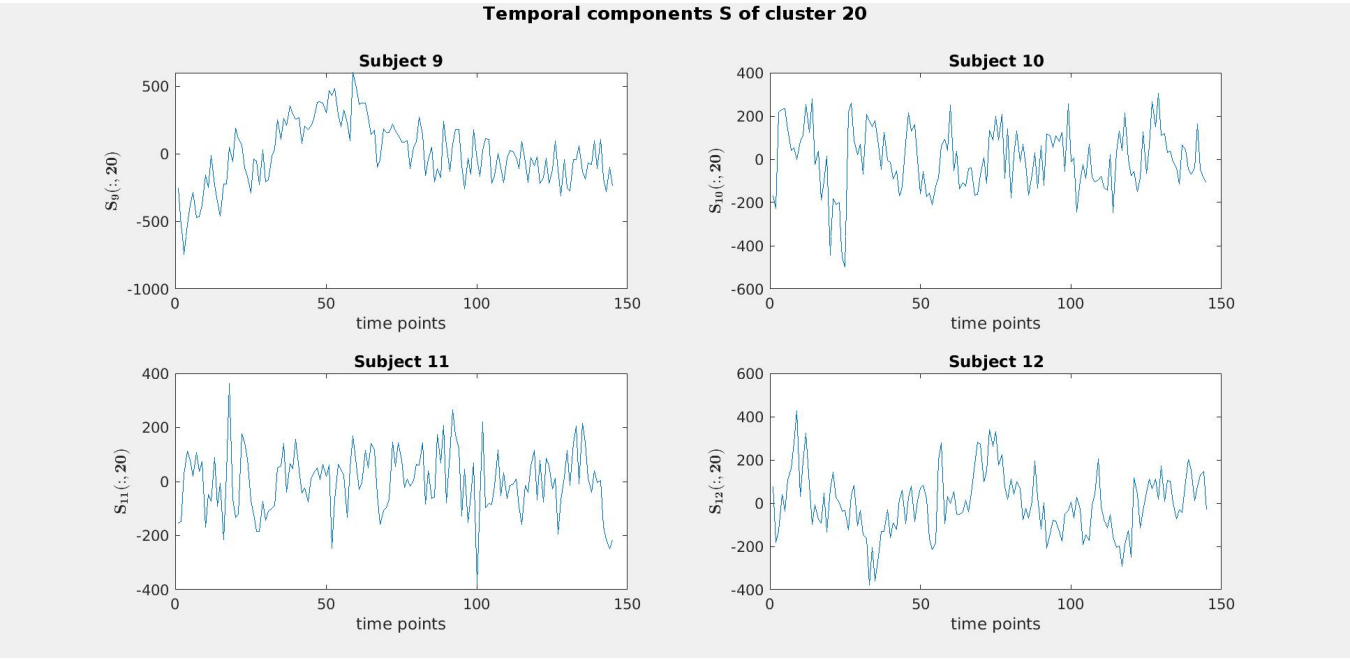


Figure B.158: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 9, \ldots, 12$ and R=27.

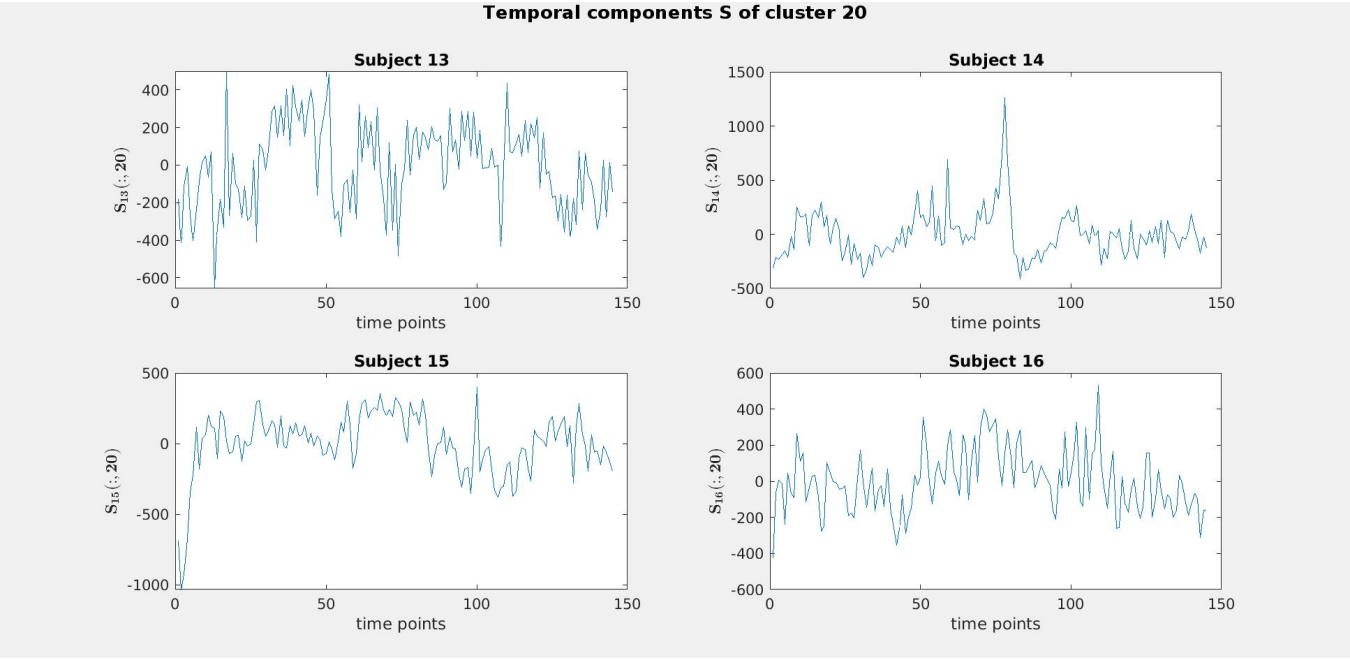


Figure B.159: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 13, \ldots, 16$ and R=27.

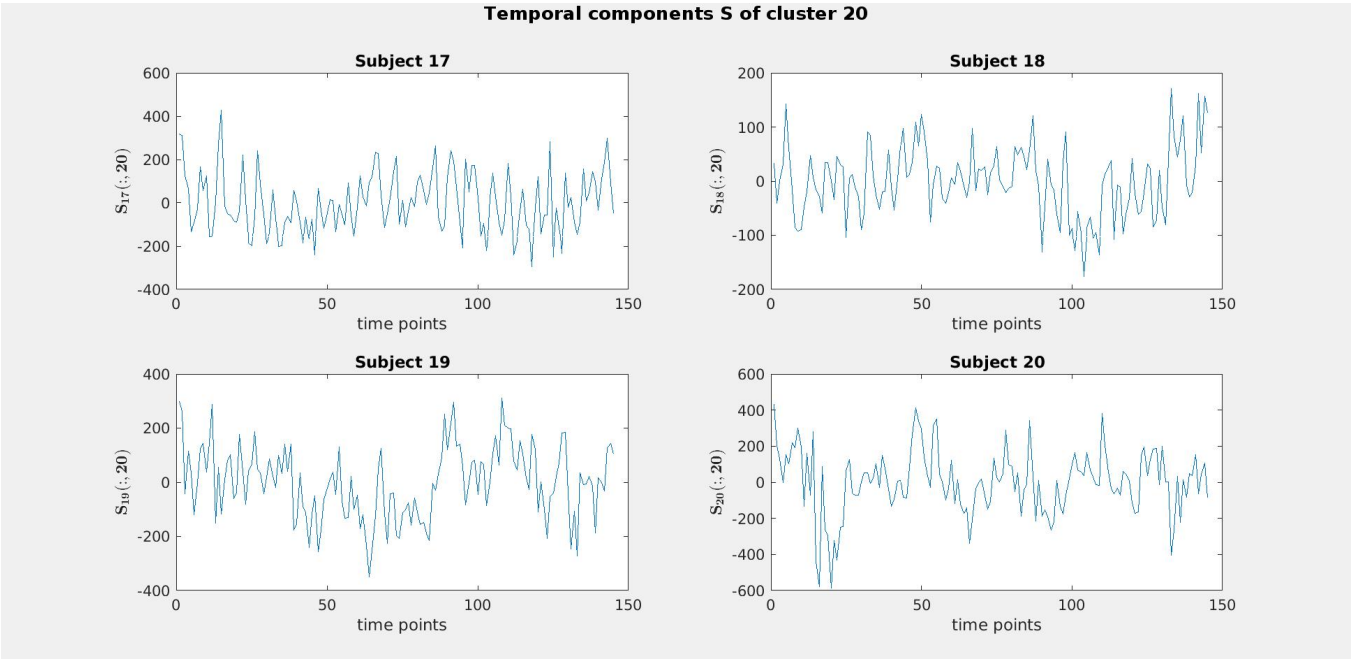


Figure B.160: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 17, \ldots, 20$ and R=27.

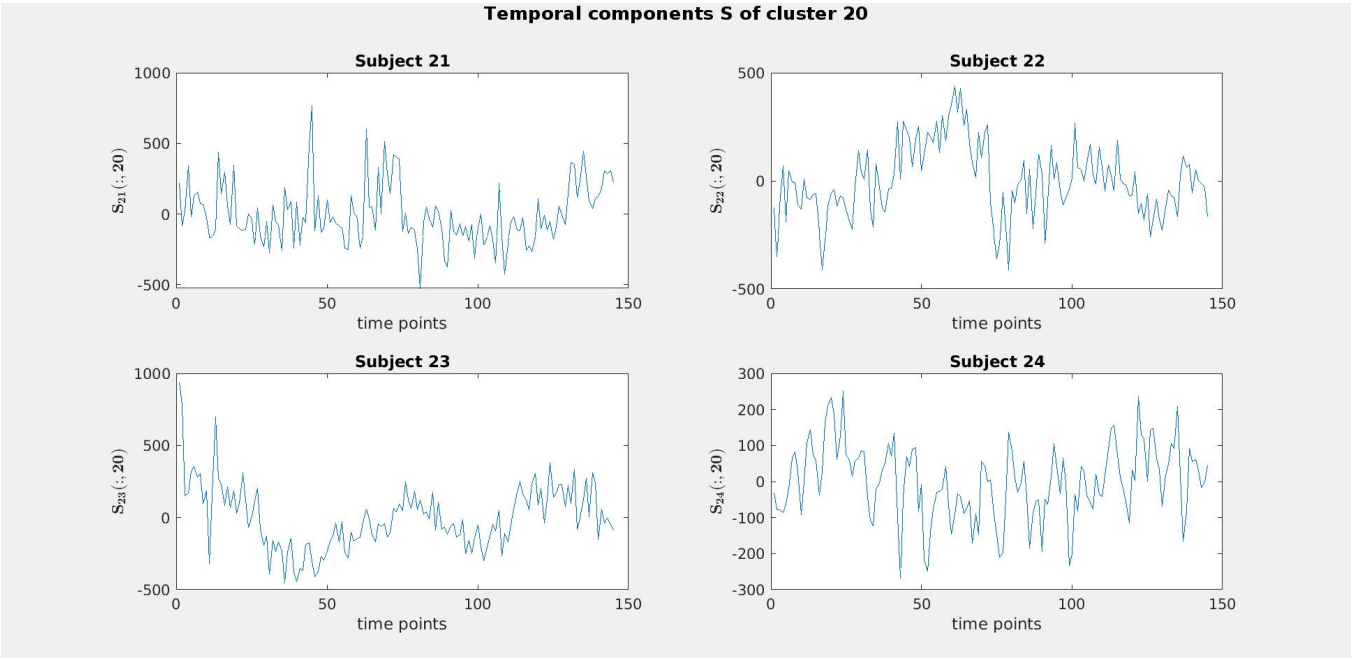


Figure B.161: Waveform of temporal components $\mathbf{S}_k(:, 120)$ for $k = 21, \ldots, 24$ and R=27.

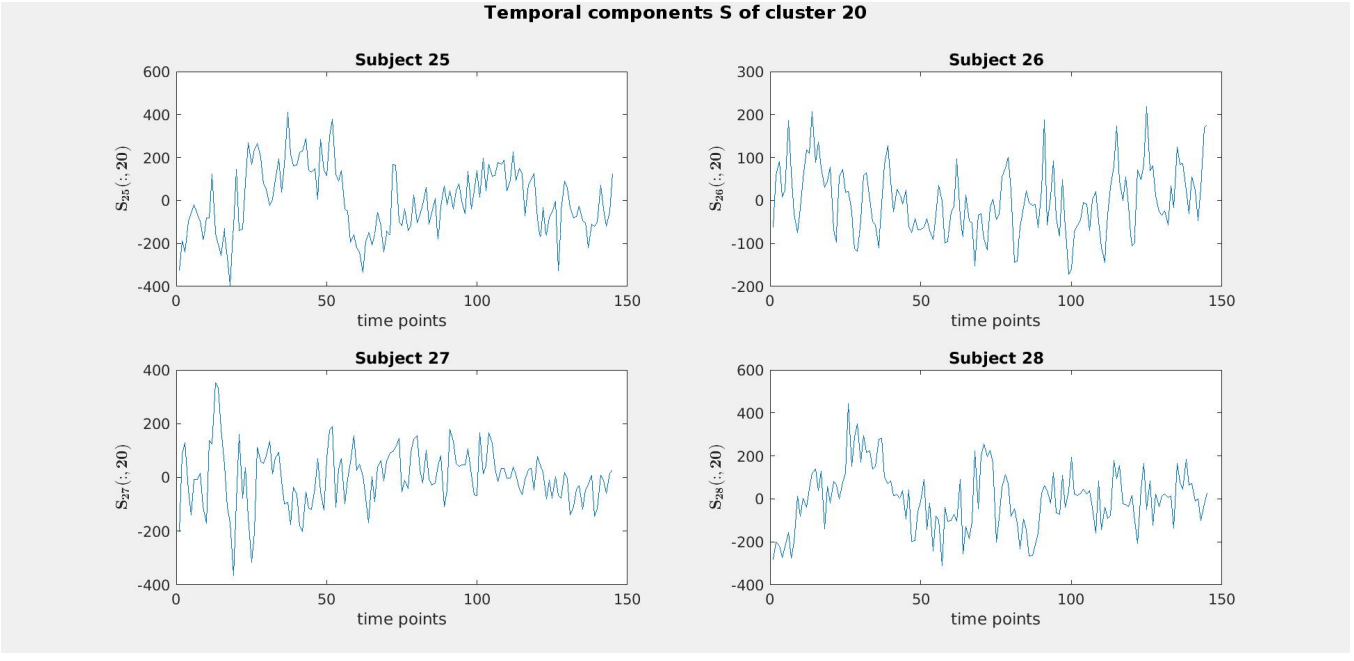


Figure B.162: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 25, \ldots, 28$ and R=27.

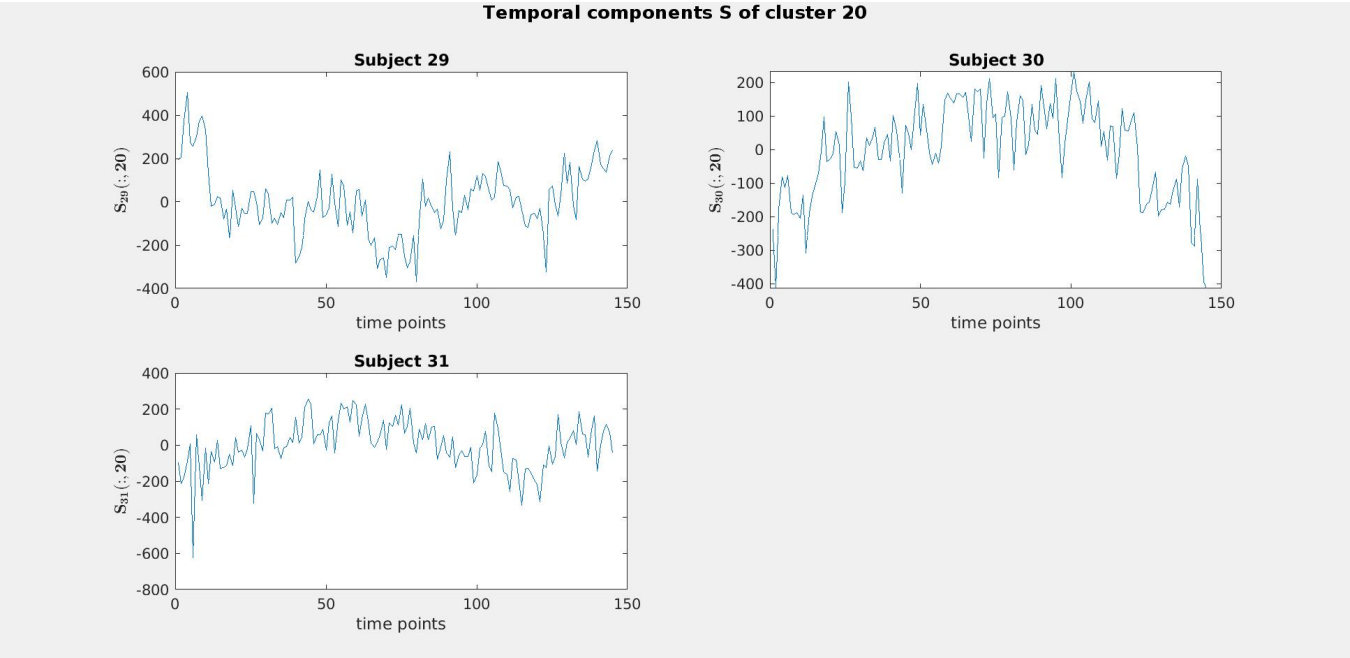


Figure B.163: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 29, \ldots, 31$ and R=27.

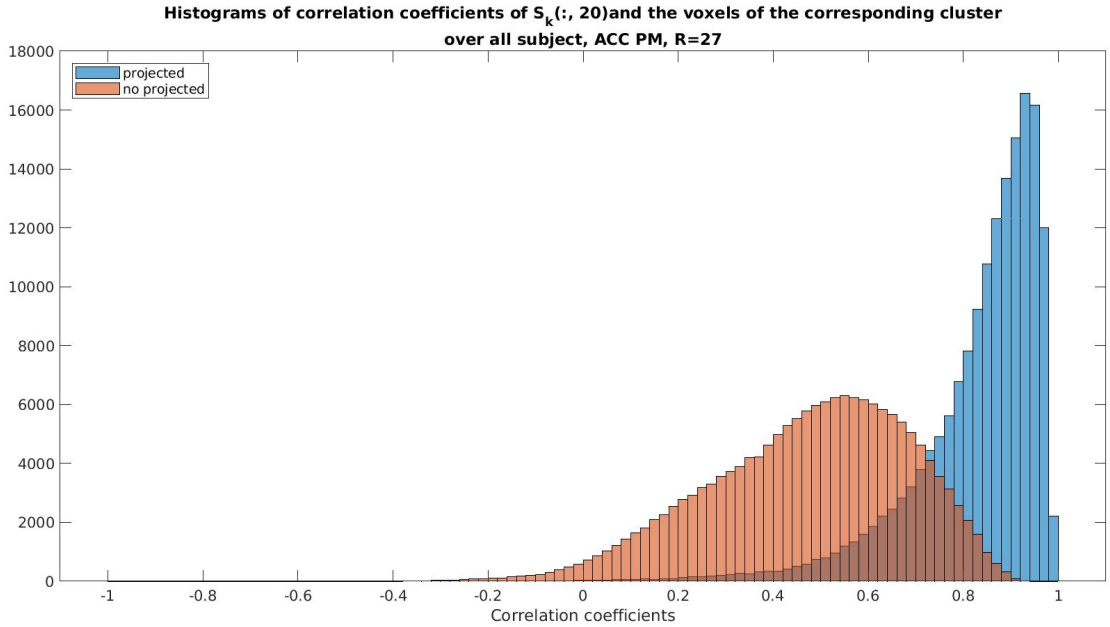


Figure B.164: For the column 20 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 20)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

15. **Cluster 11**

    Cluster 11 contains 6101 voxels, which correspond to 2.700179% of the whole brain (background, soft tissue, bone are excluded).

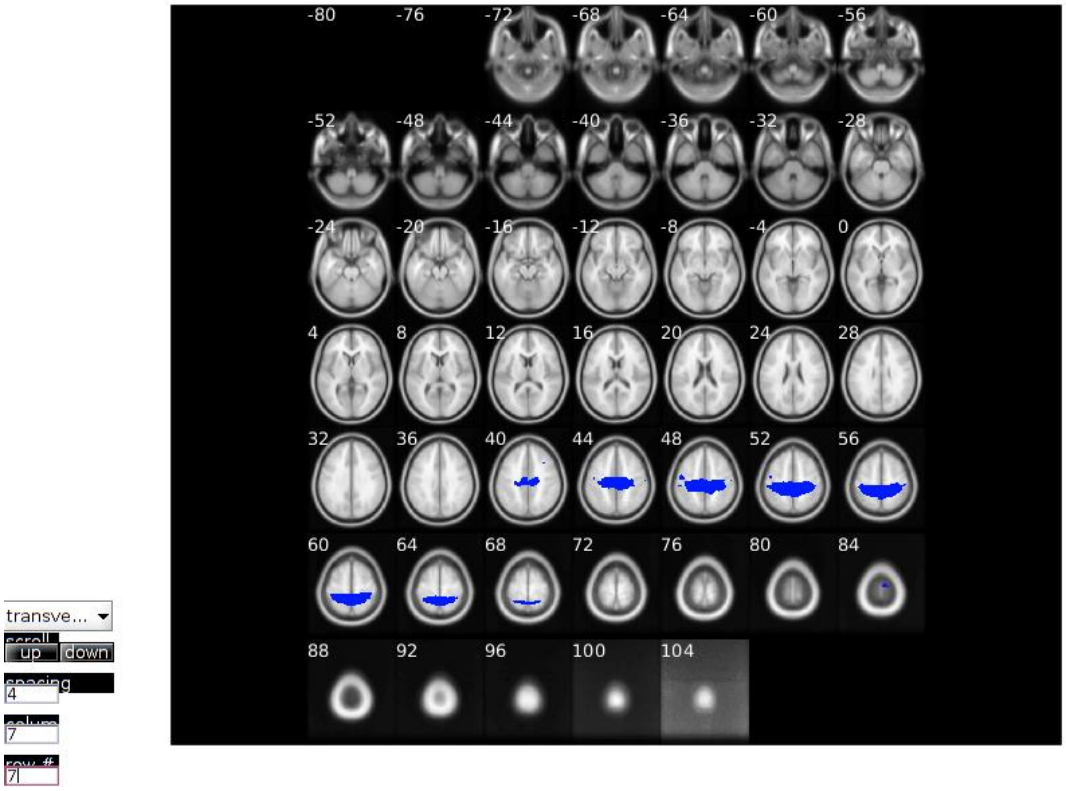


Figure B.165: Cluster 11. Blue areas indicate the voxels which belong to cluster 11.

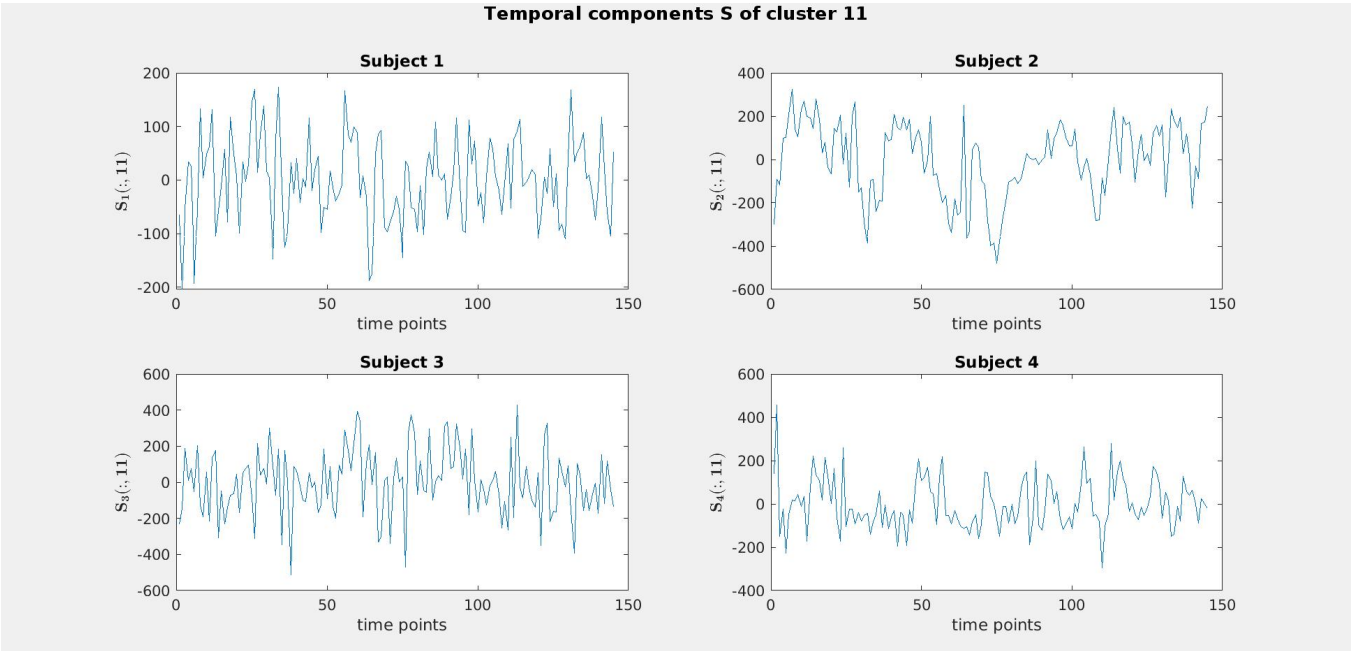


Figure B.166: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 1, \ldots, 4$ and R=27.

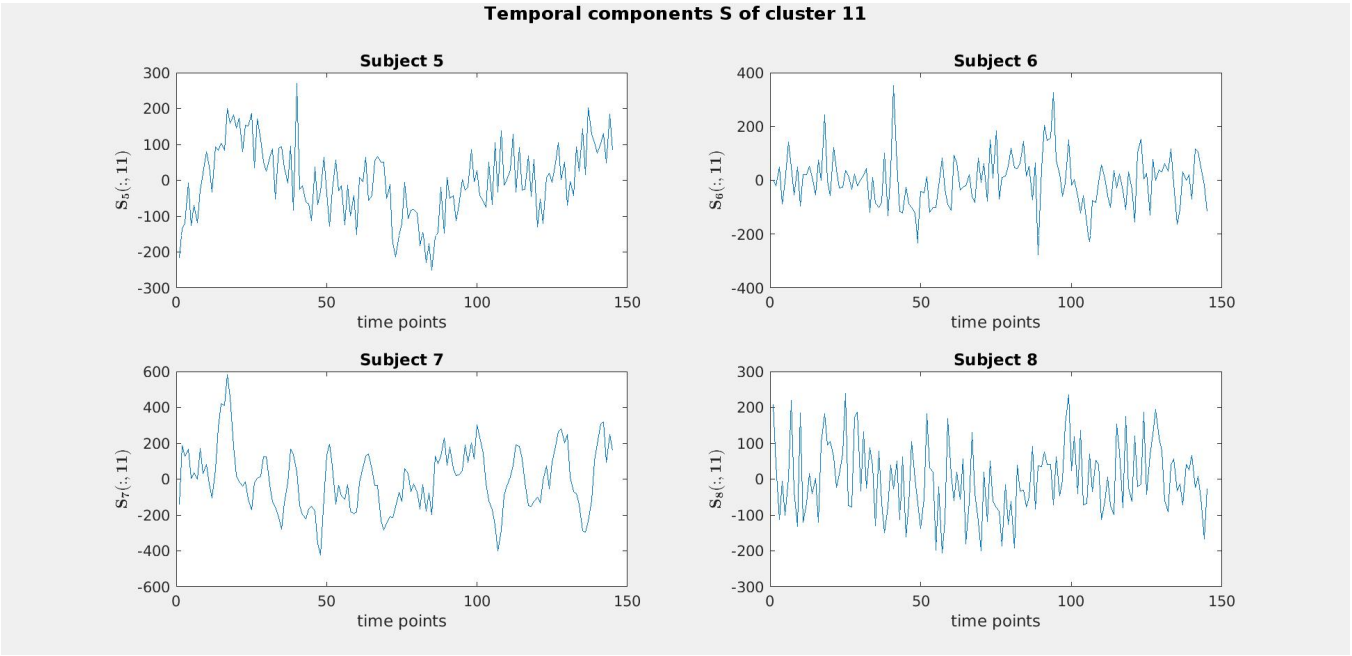


Figure B.167: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 5, \ldots, 8$ and R=27.

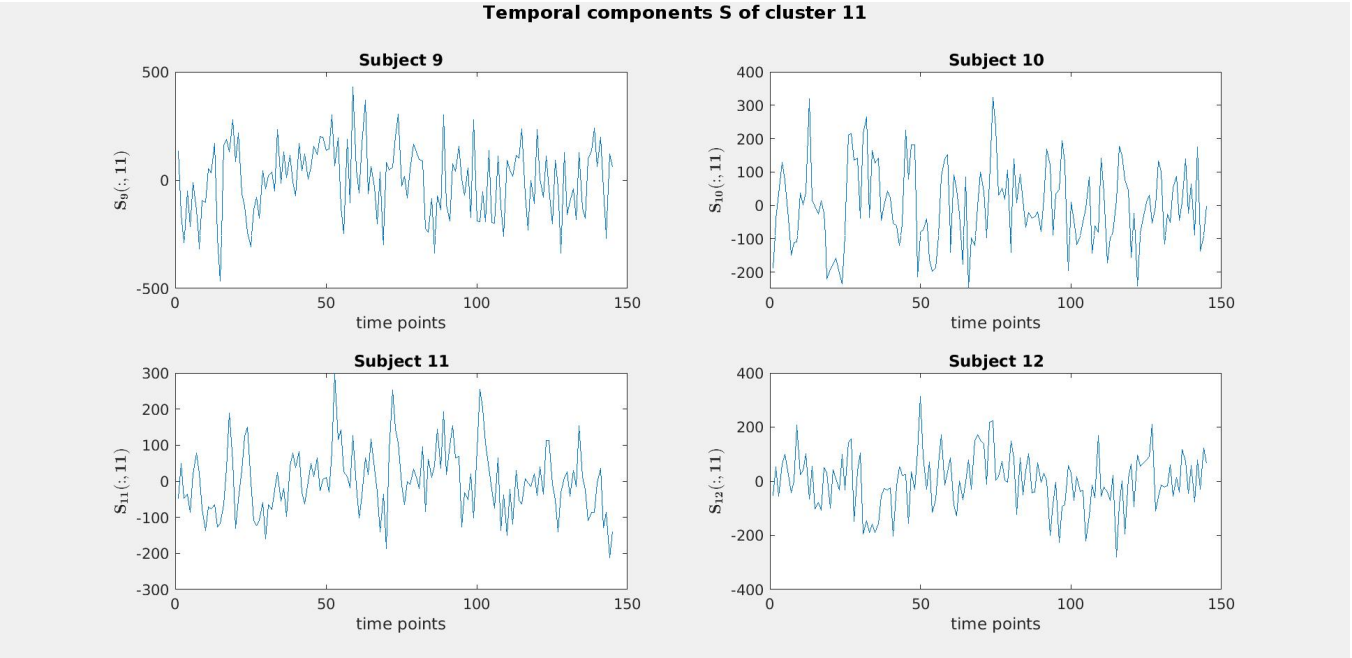


Figure B.168: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 9, \ldots, 12$ and R=27.

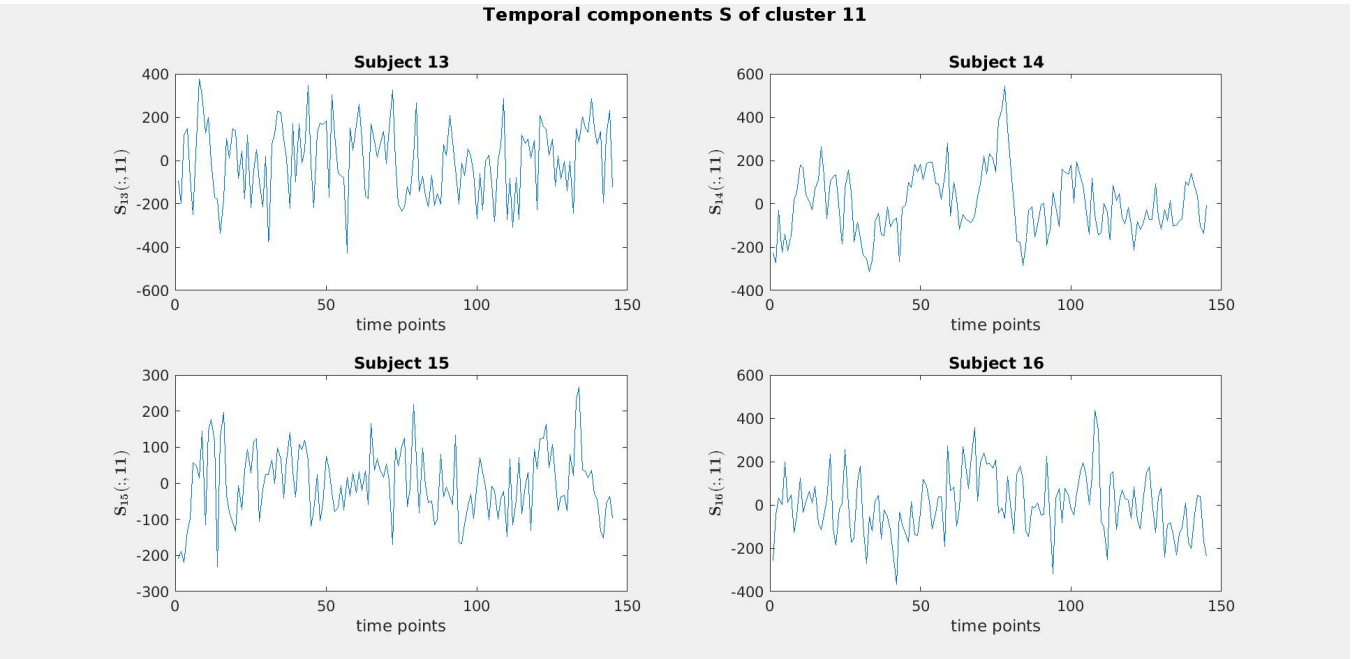


Figure B.169: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 13, \ldots, 16$ and R=27.

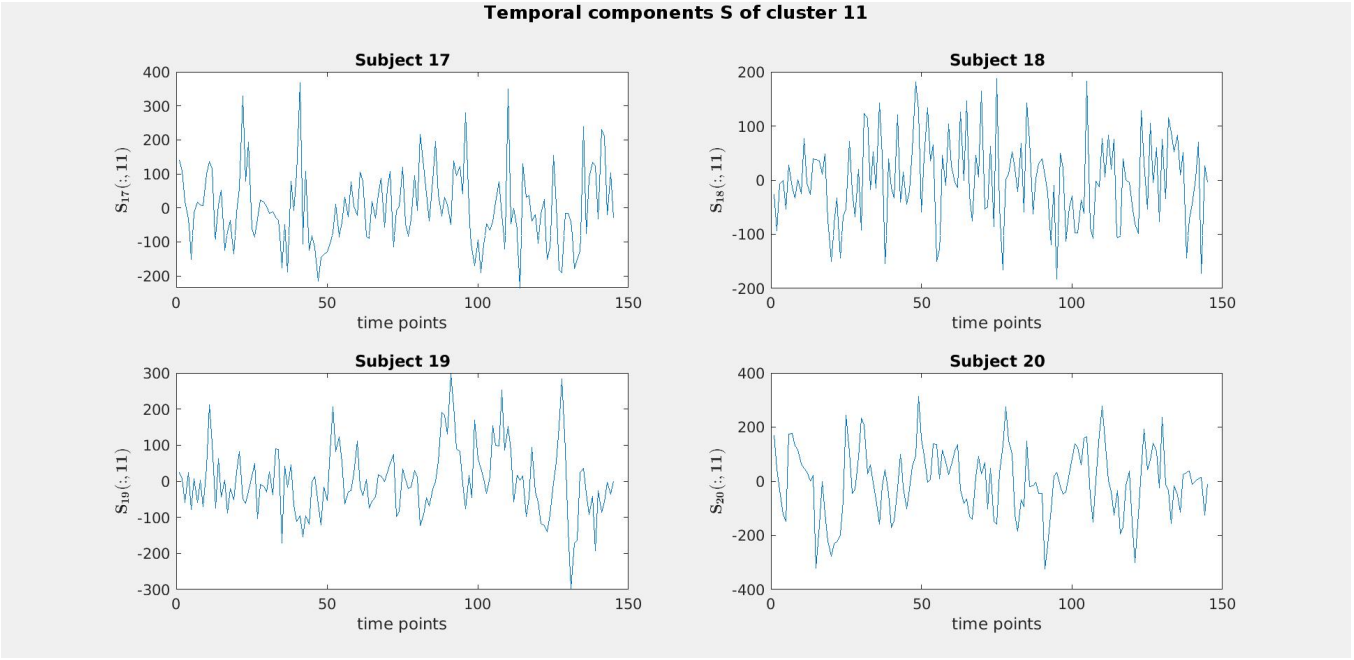


Figure B.170: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 17, \ldots, 20$ and R=27.

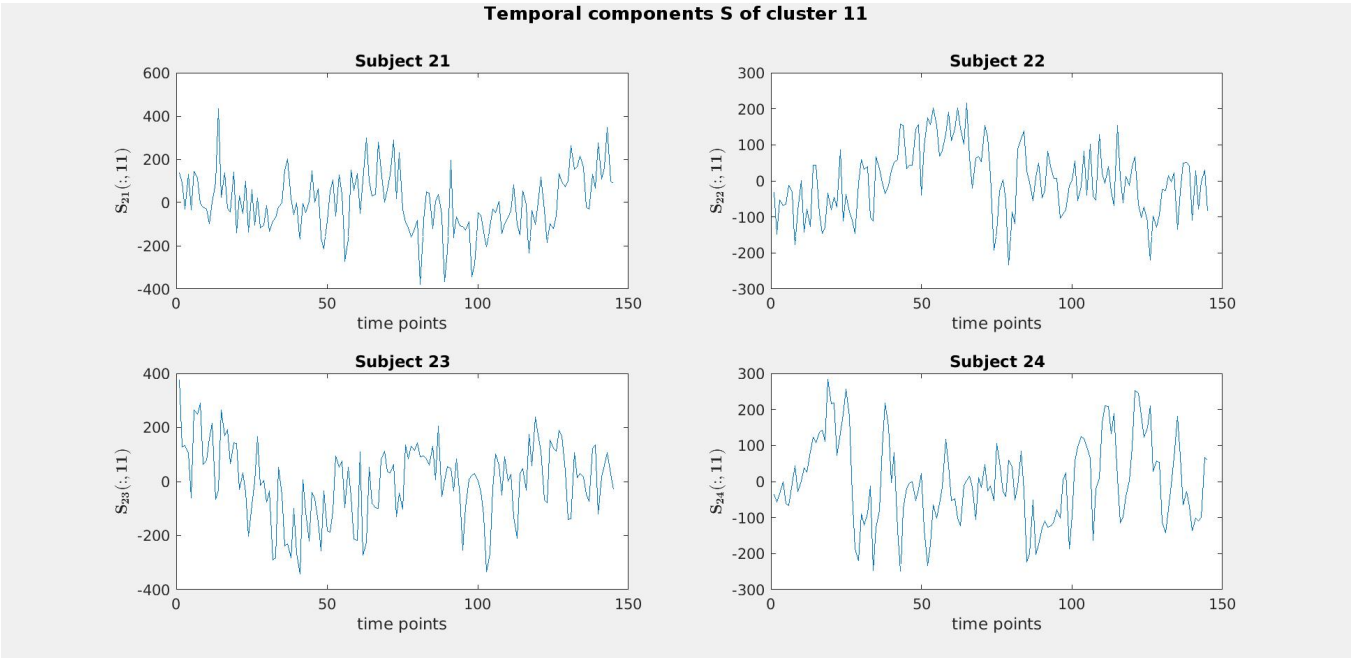


Figure B.171: Waveform of temporal components $\mathbf{S}_k(:, 111)$ for $k = 21, \ldots, 24$ and R=27.

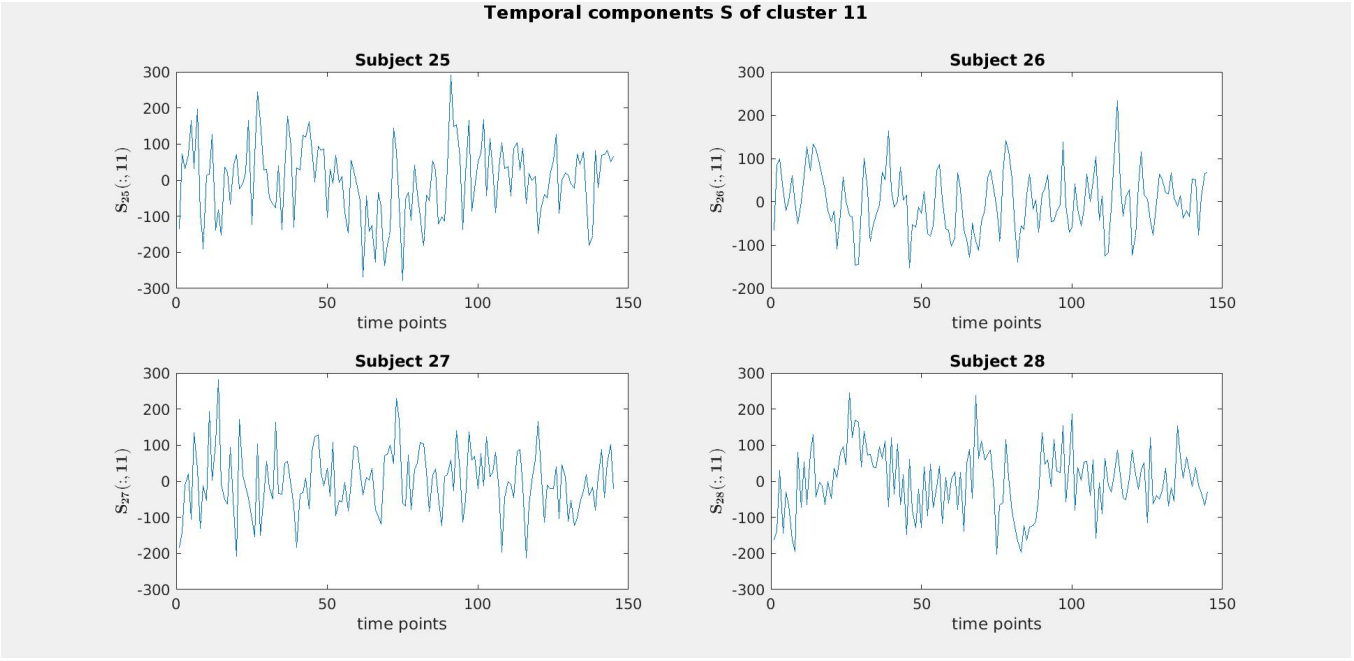


Figure B.172: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 25, \ldots, 28$ and R=27.

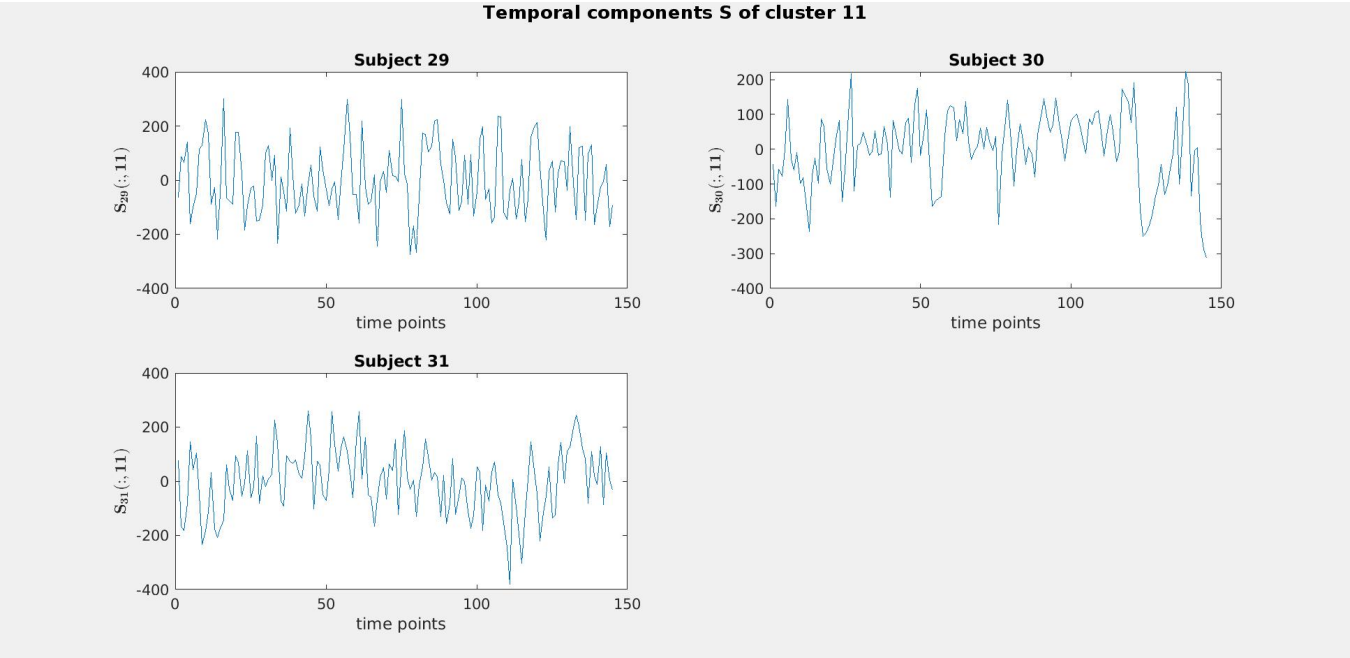


Figure B.173: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 29, \ldots, 31$ and R=27.

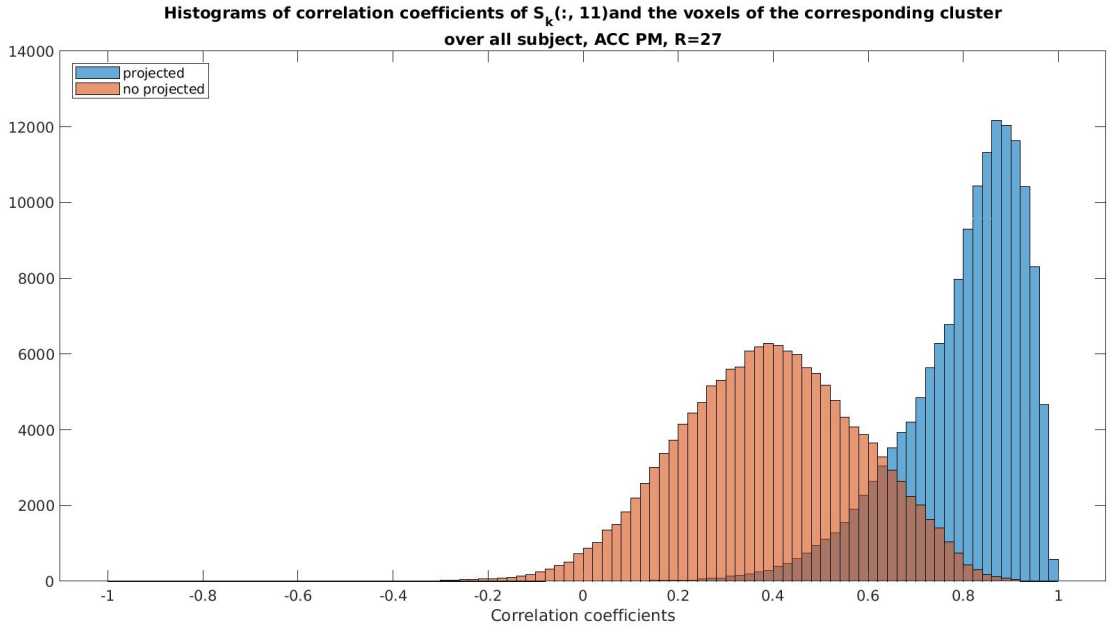


Figure B.174: For the column 11 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 11)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

16. **Cluster 17**

    Cluster 17 contains 5987 voxels, which correspond to 2.649725% of the whole brain (background, soft tissue, bone are excluded).

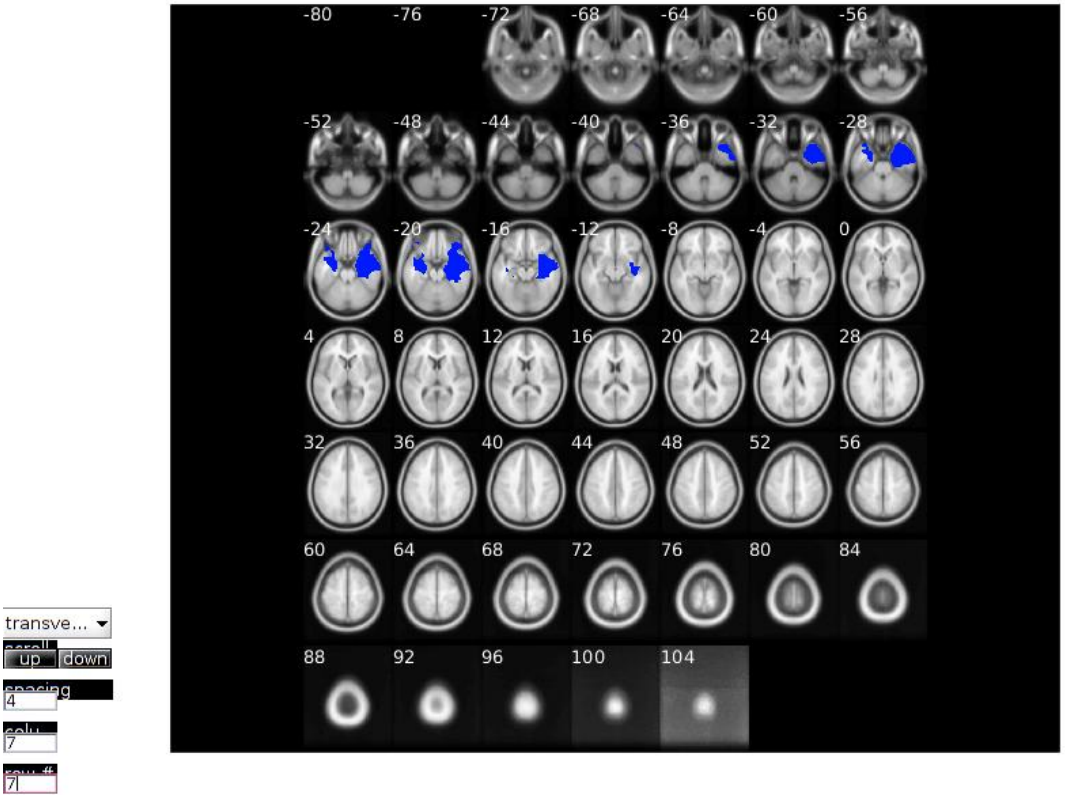


Figure B.175: Cluster 17. Blue areas indicate the voxels which belong to cluster 17.

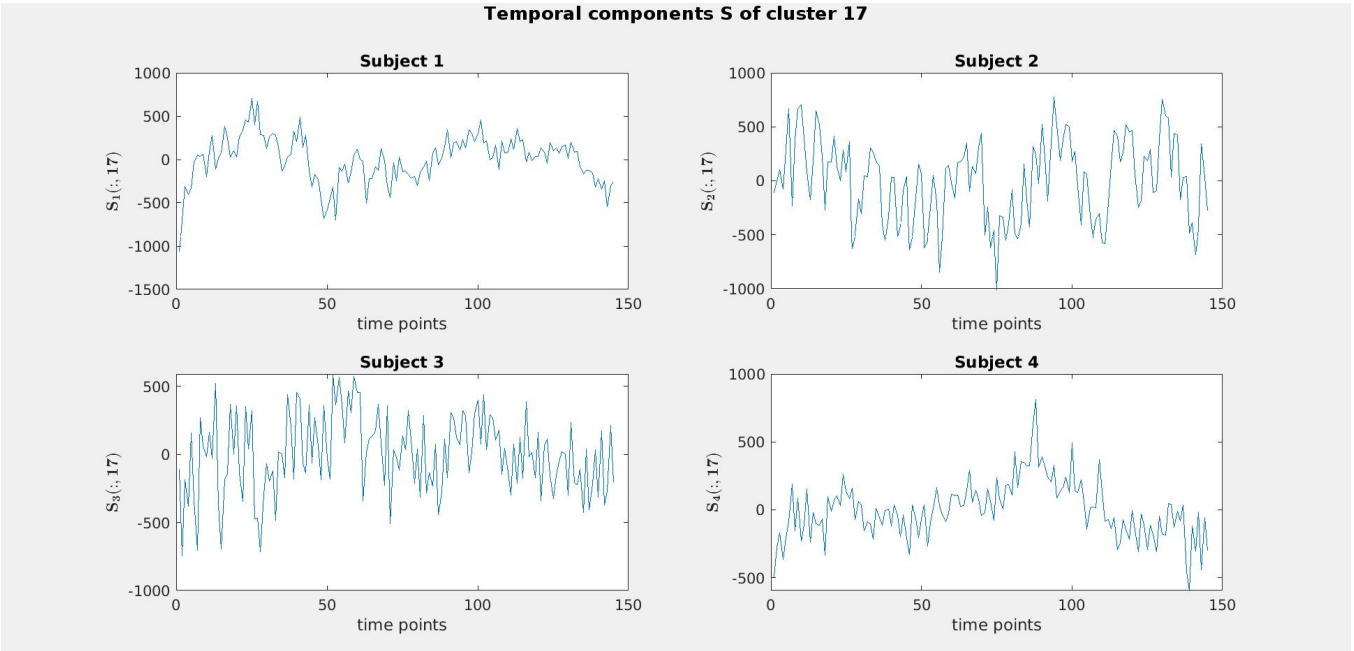


Figure B.176: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 1, \ldots, 4$ and R=27.

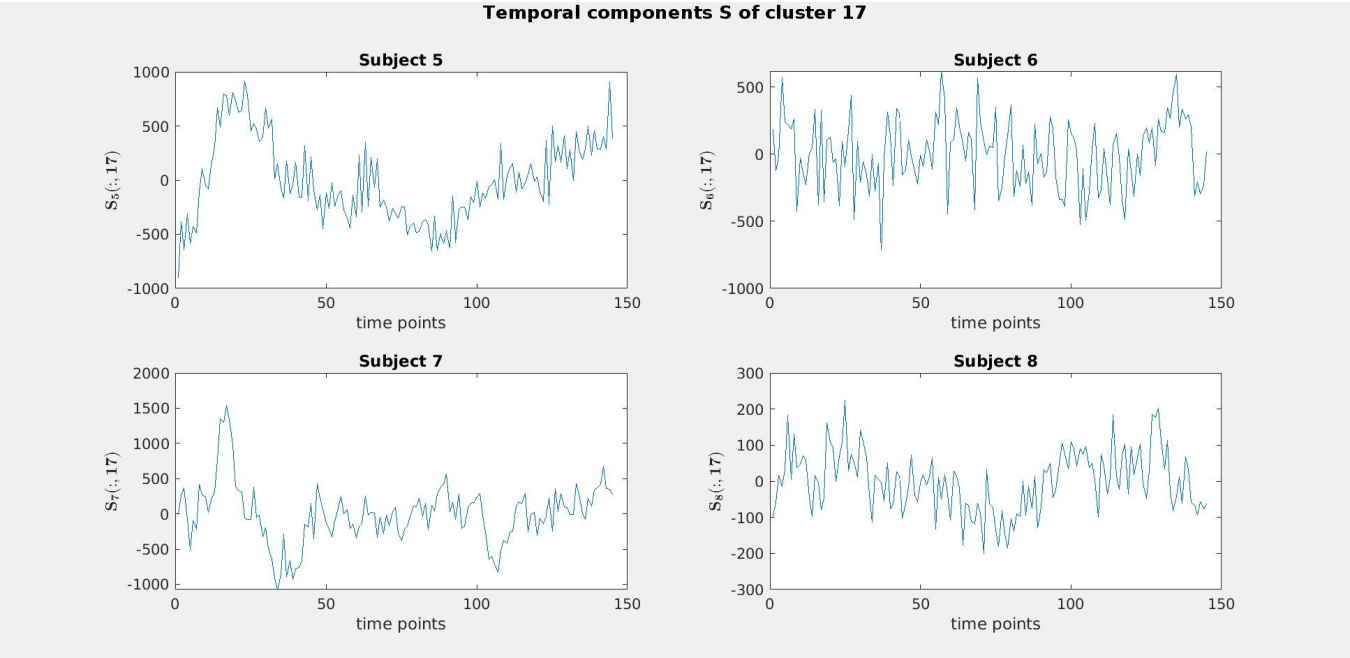


Figure B.177: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 5, \ldots, 8$ and R=27.

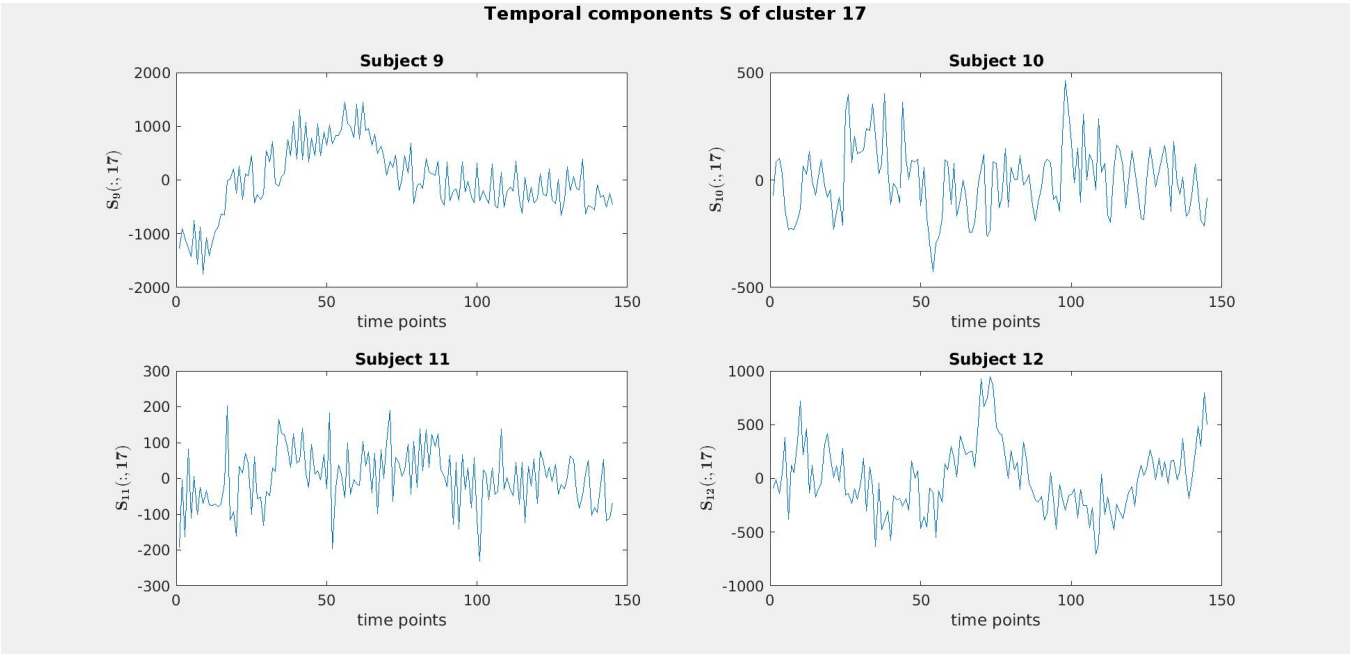


Figure B.178: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 9, \ldots, 12$ and R=27.

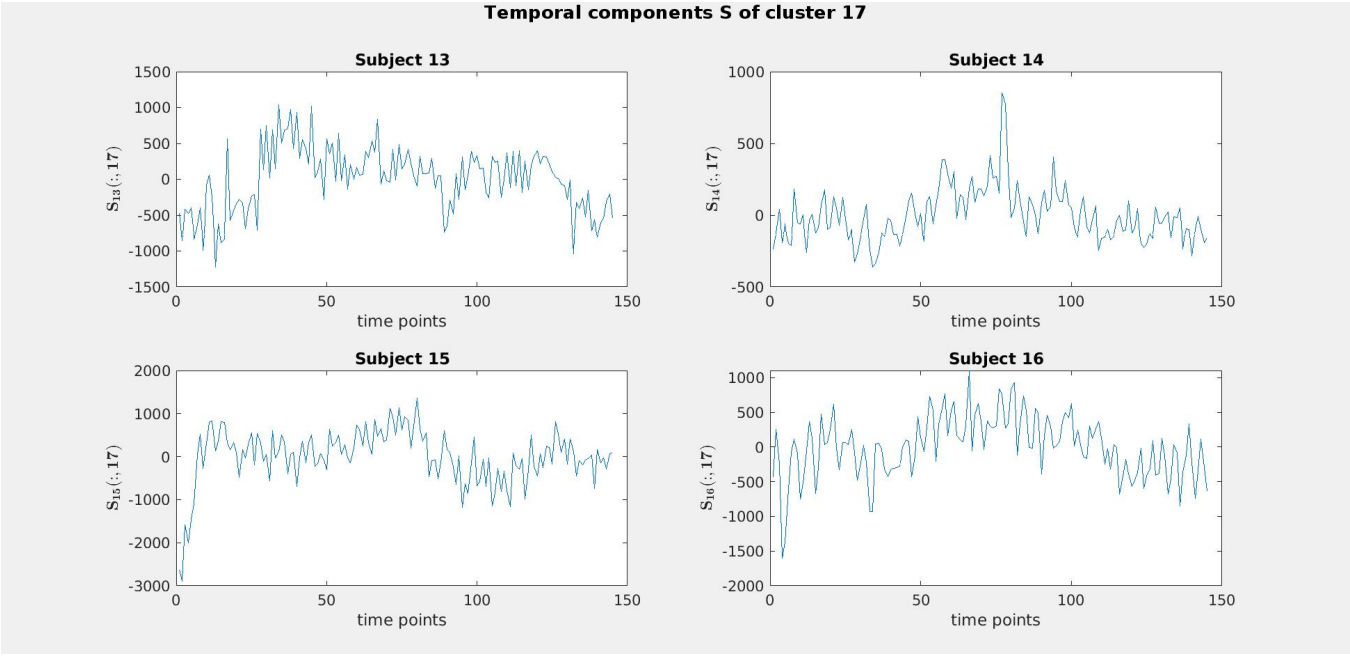


Figure B.179: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 13, \ldots, 16$ and R=27.

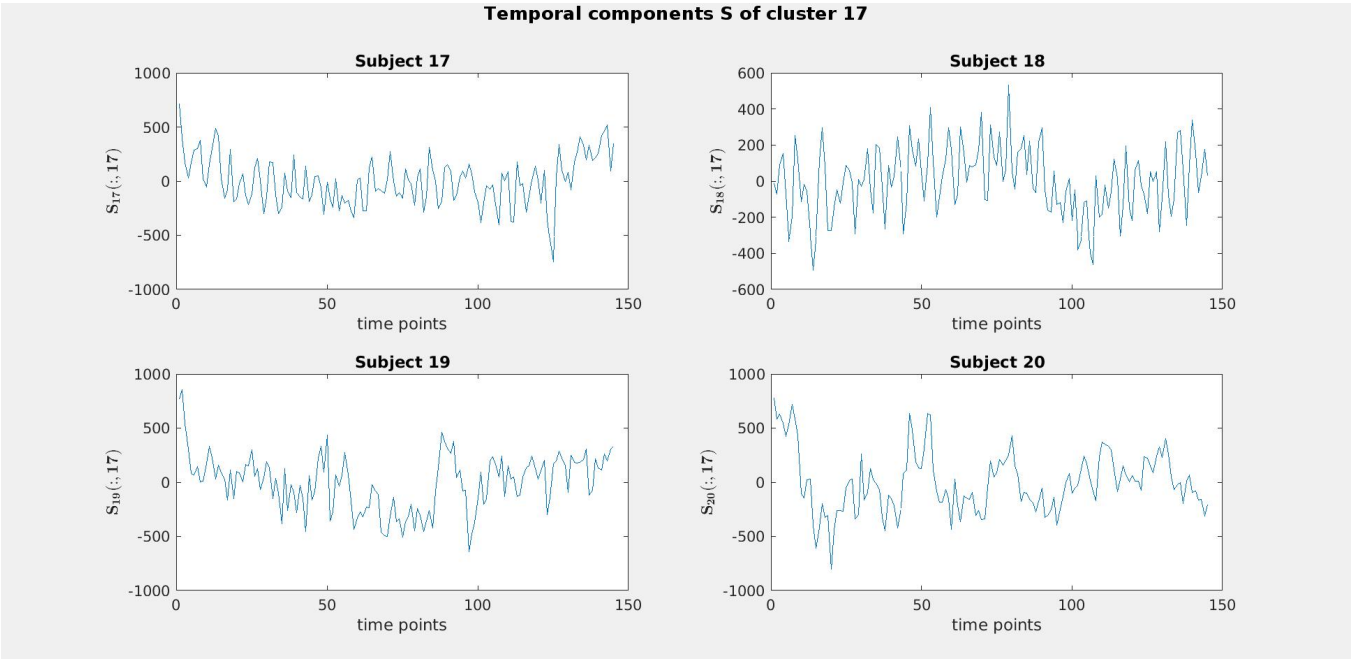


Figure B.180: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 17, \ldots, 20$ and R=27.

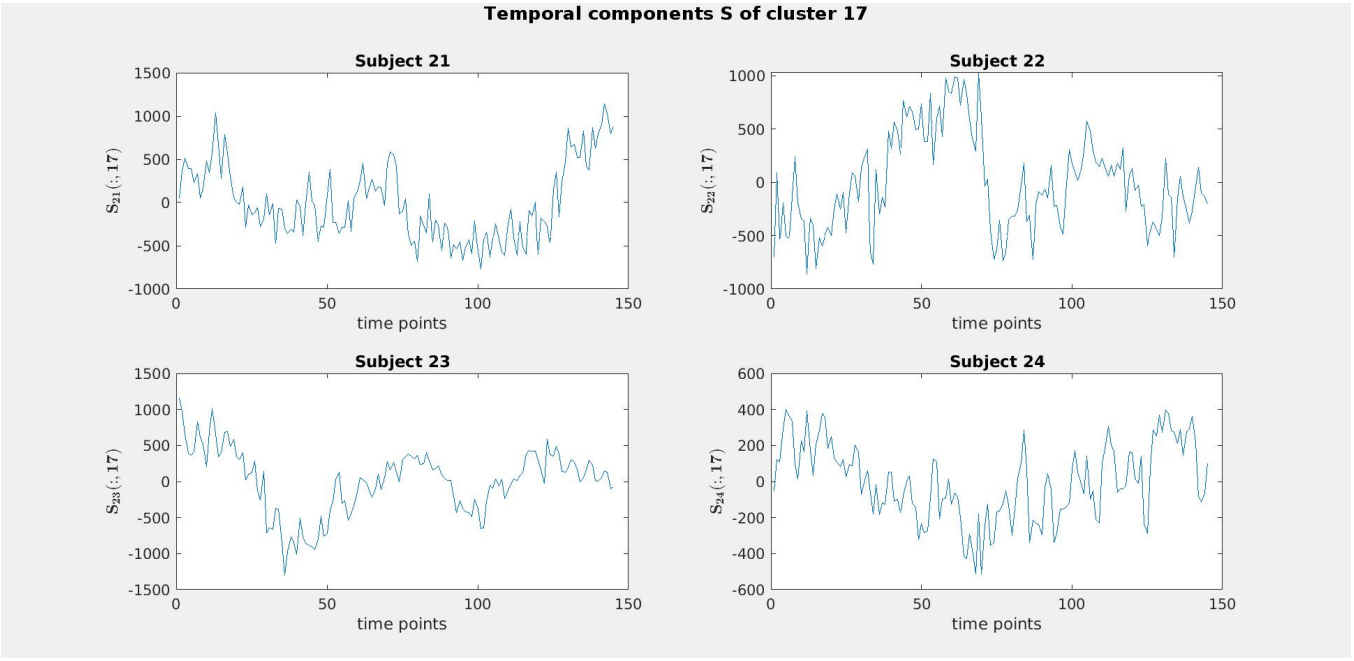


Figure B.181: Waveform of temporal components $\mathbf{S}_k(:, 117)$ for $k = 21, \ldots, 24$ and R=27.

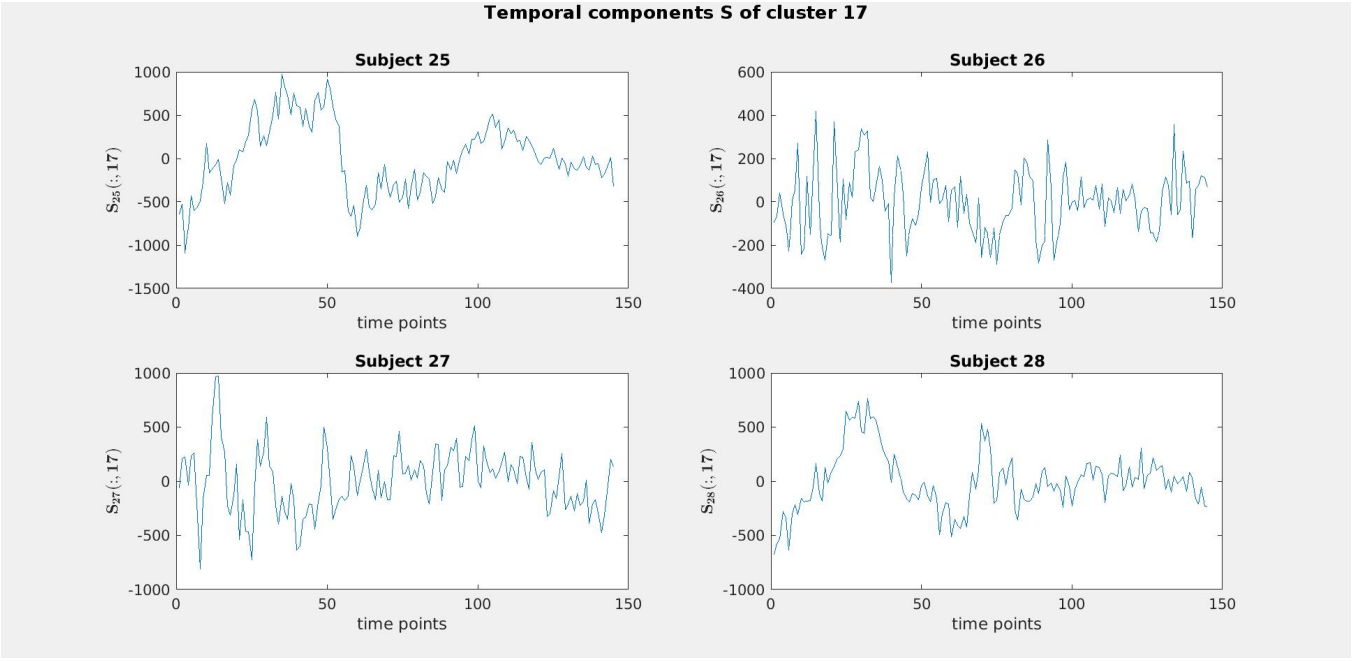


Figure B.182: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 25, \ldots, 28$ and R=27.

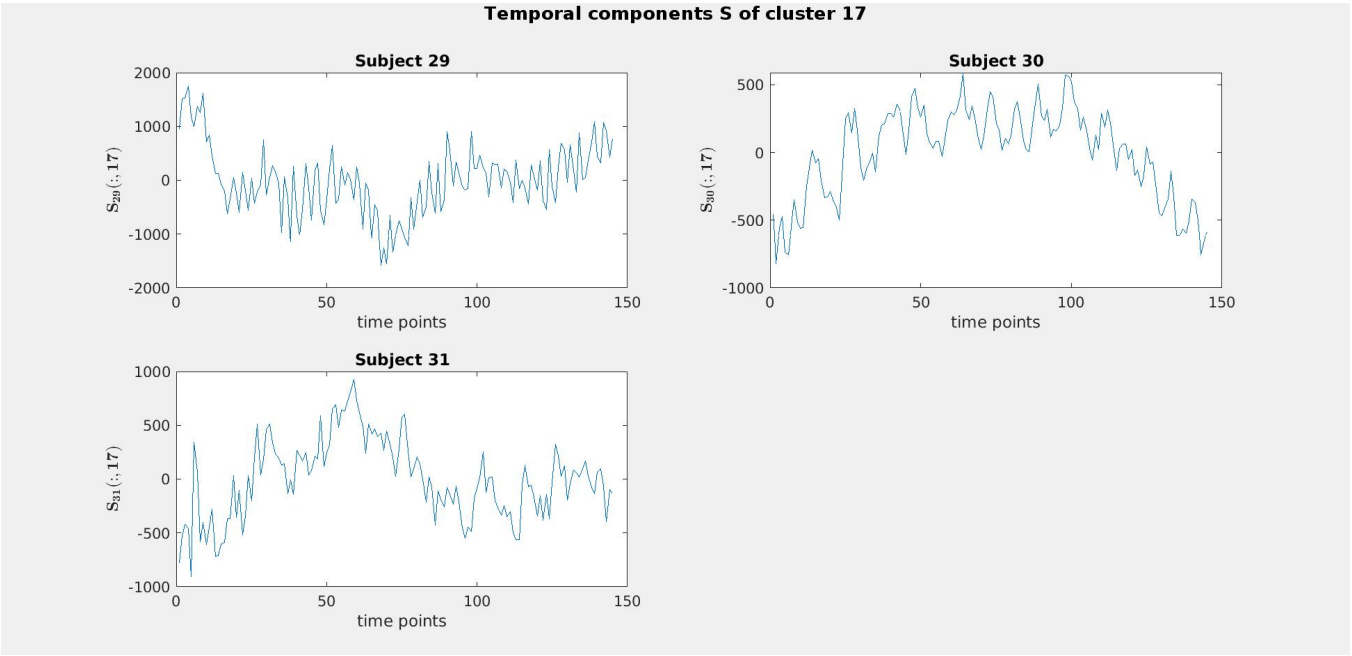


Figure B.183: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 29, \ldots, 31$ and R=27.

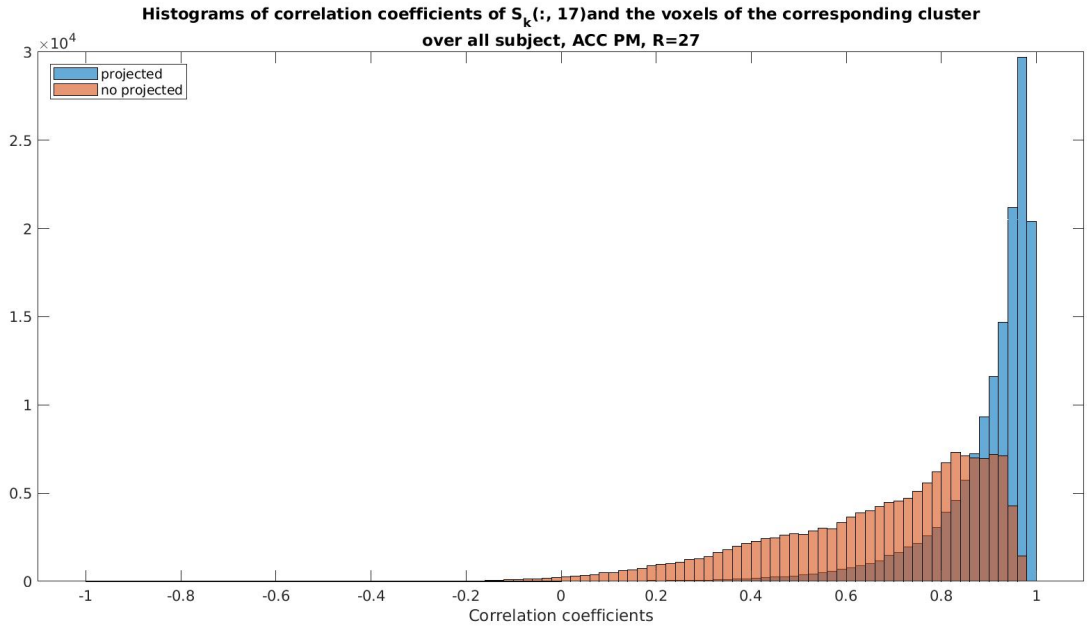


Figure B.184: For the column 17 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 17)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

17. **Cluster 21**

Cluster 21 contains 5530 voxels, which correspond to 2.447466% of the whole brain (background, soft tissue, bone are excluded).

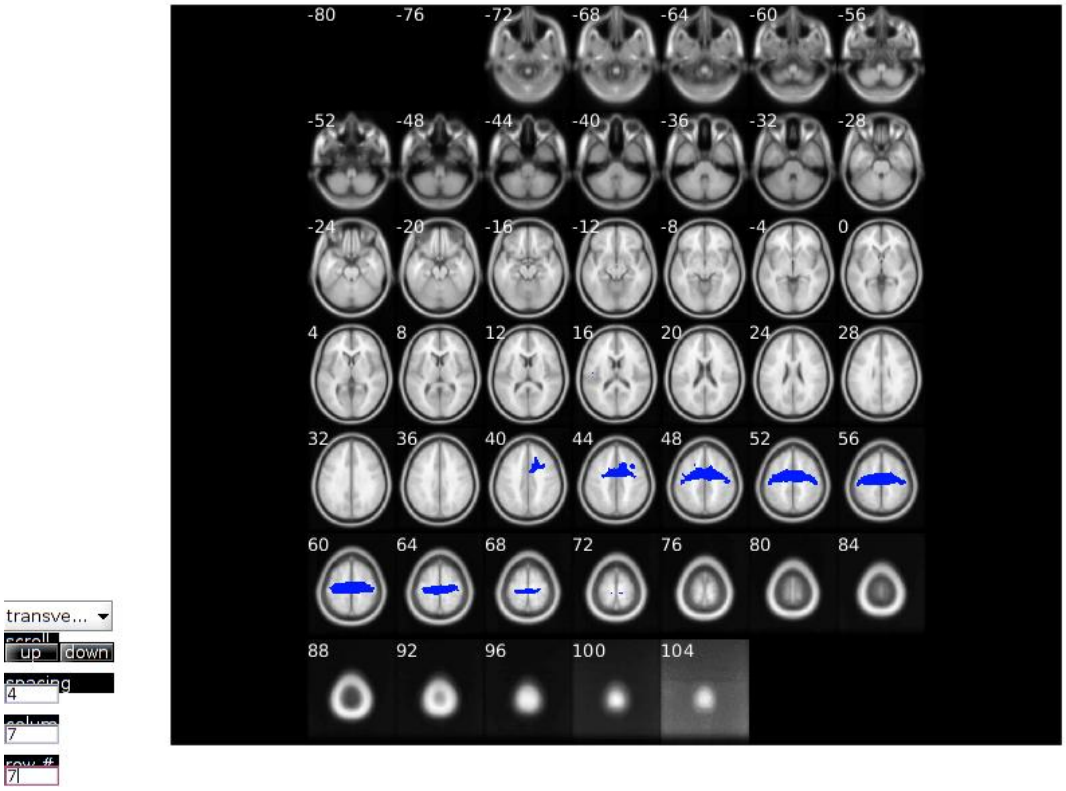


Figure B.185: Cluster 21. Blue areas indicate the voxels which belong to cluster 21.

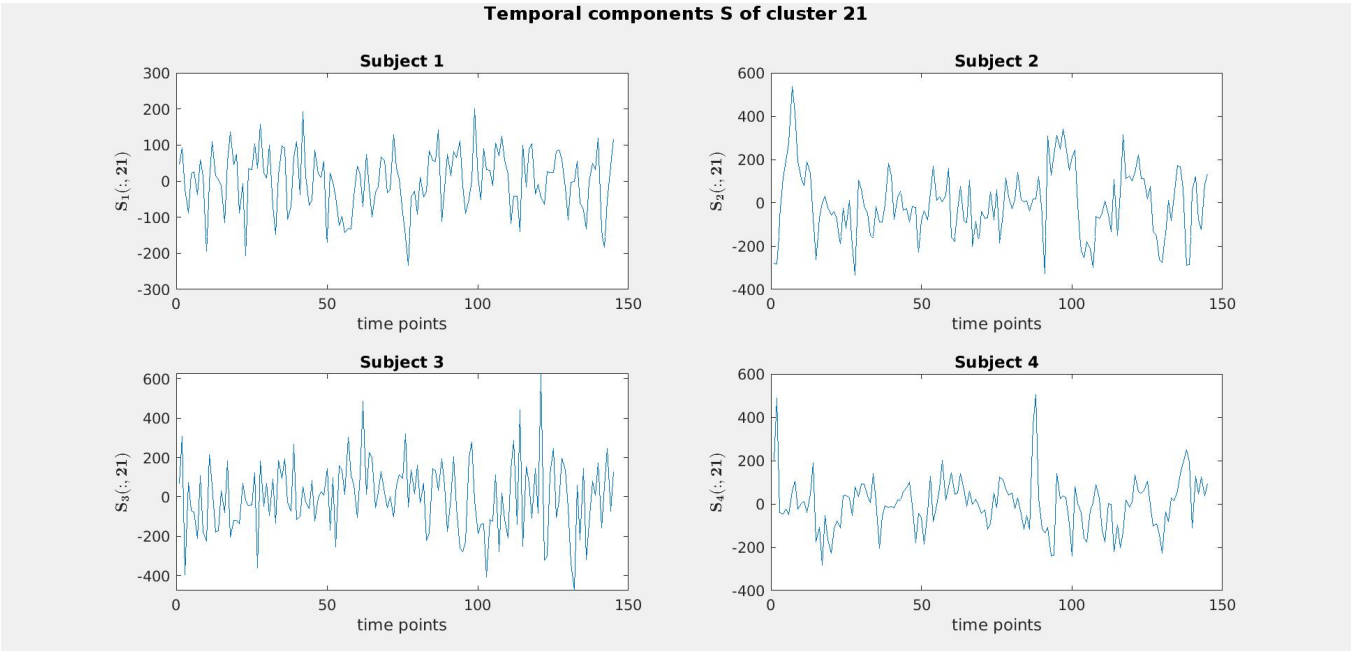


Figure B.186: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 1, \ldots, 4$ and R=27.

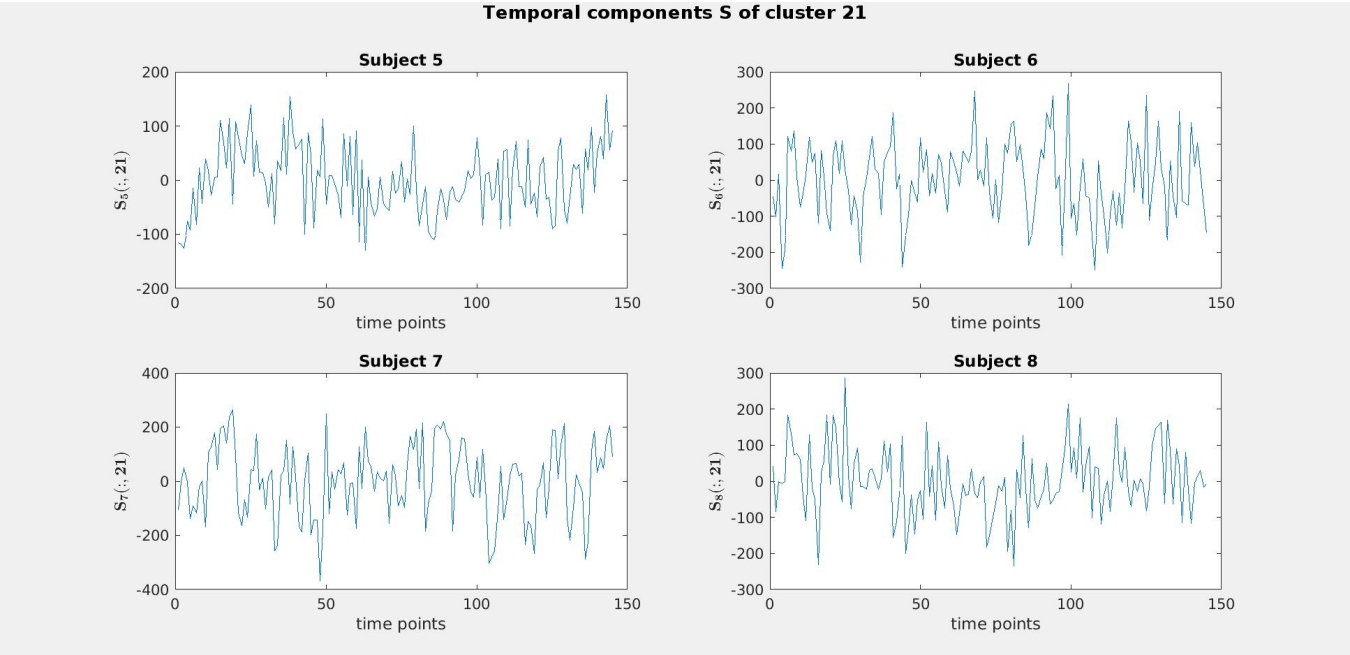


Figure B.187: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 5, \ldots, 8$ and R=27.

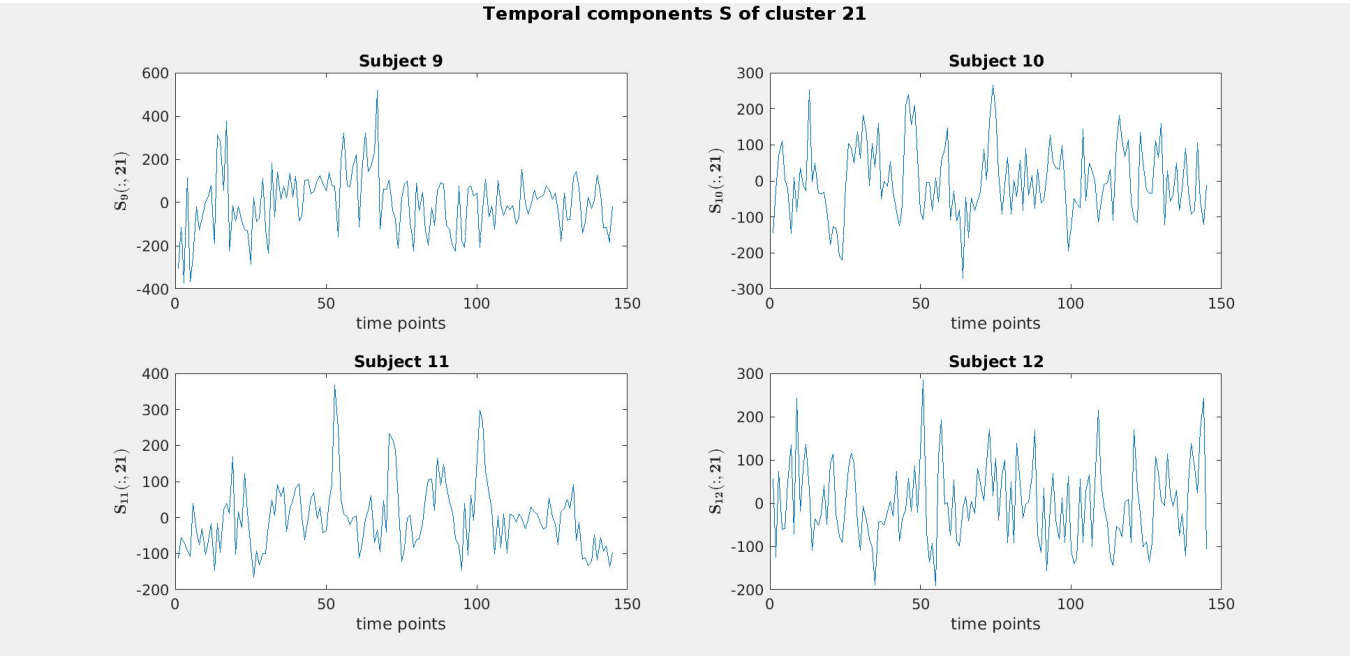


Figure B.188: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 9, \ldots, 12$ and R=27.

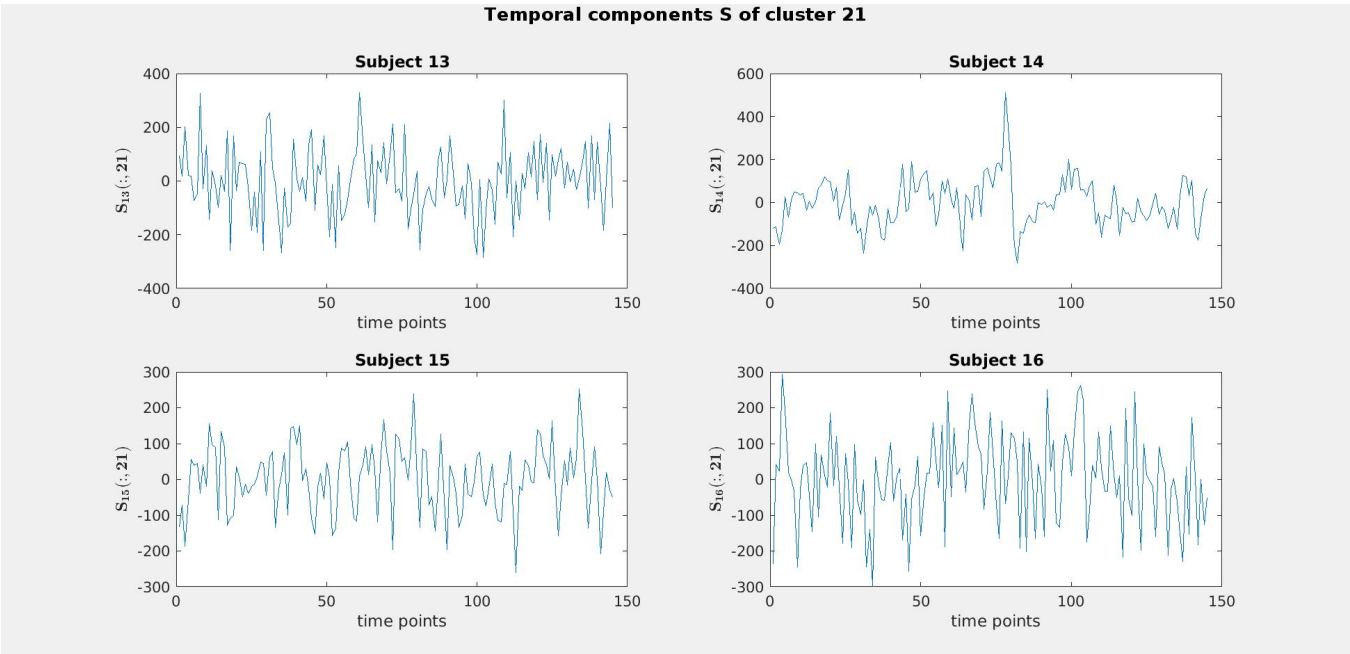


Figure B.189: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 13, \ldots, 16$ and R=27.

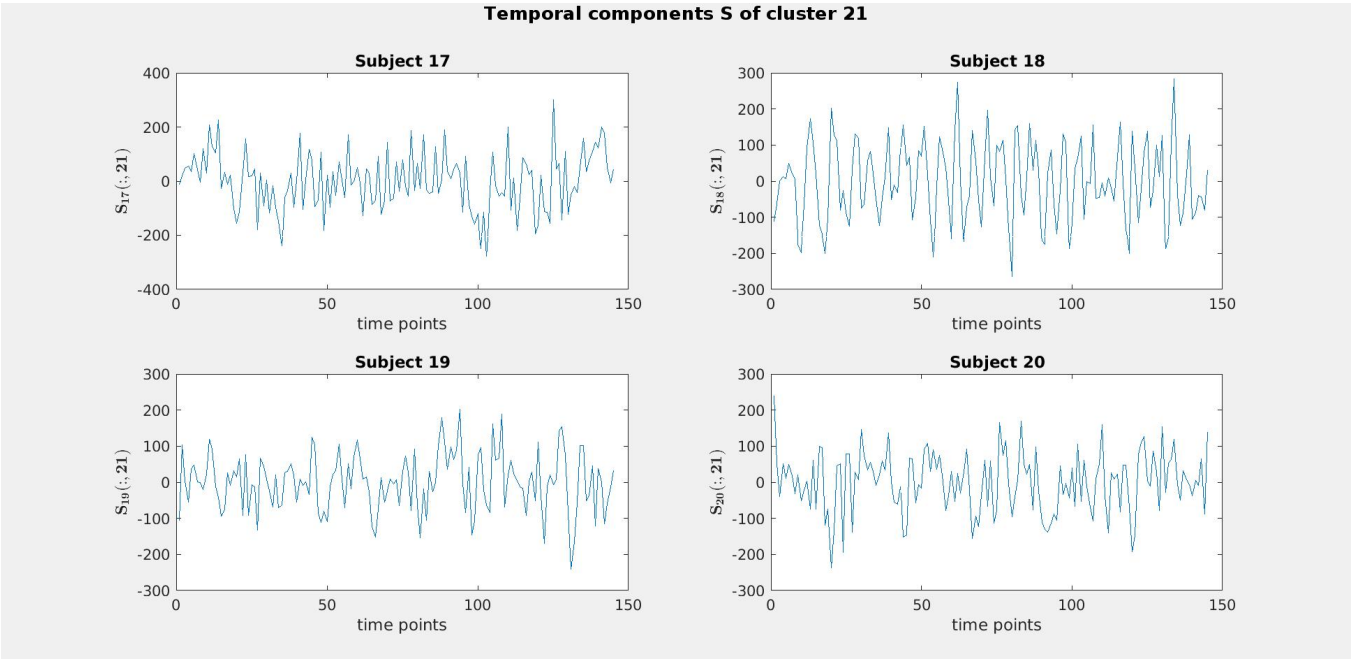


Figure B.190: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 17, \ldots, 20$ and R=27.

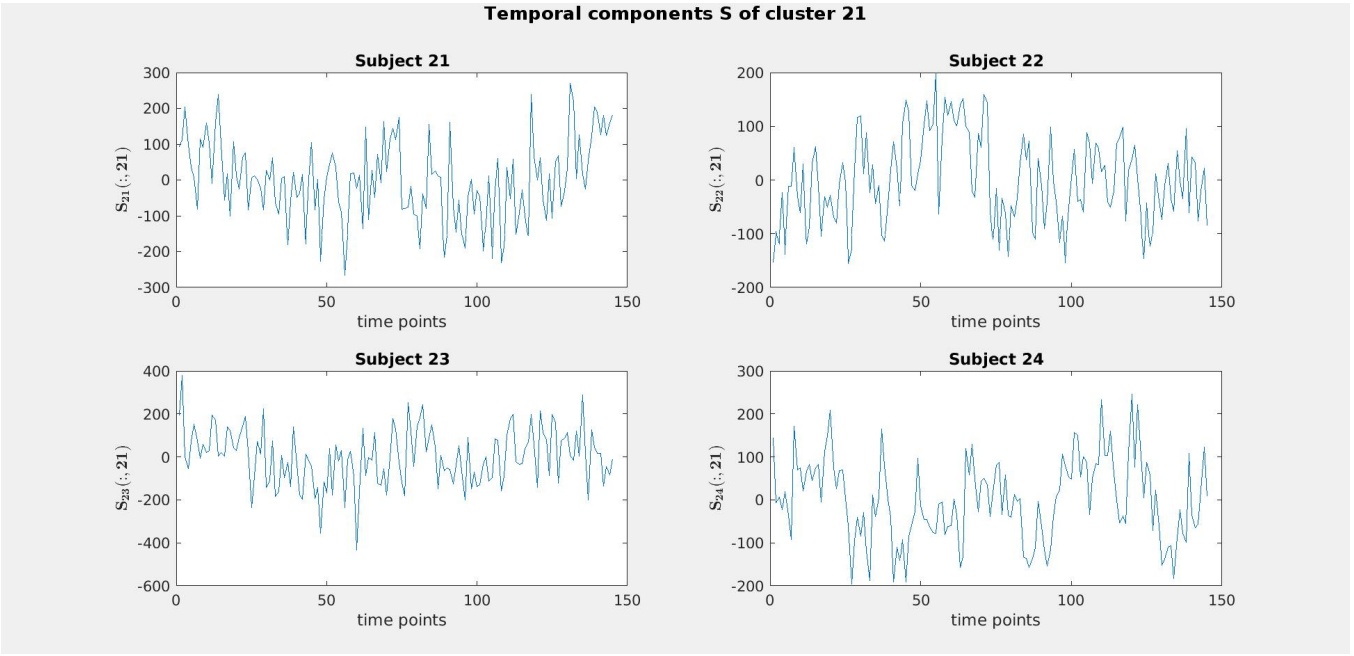


Figure B.191: Waveform of temporal components $\mathbf{S}_k(:, 121)$ for $k = 21, \ldots, 24$ and R=27.

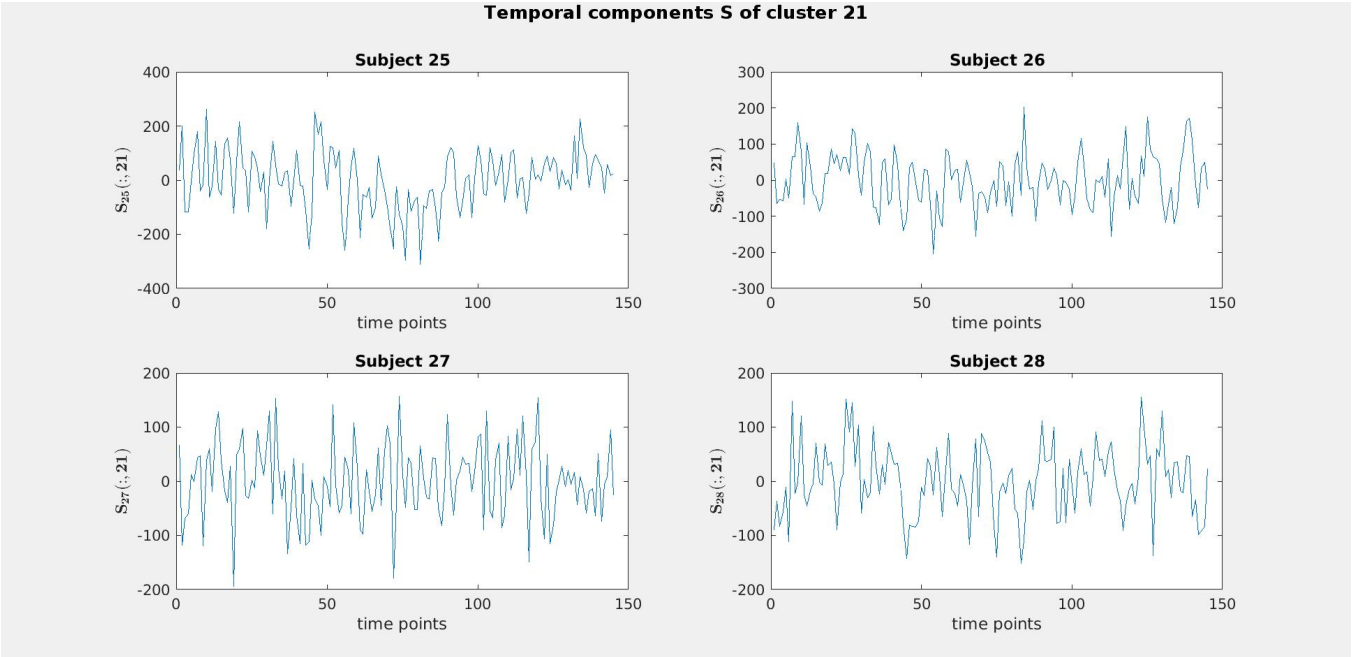


Figure B.192: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 25, \ldots, 28$ and R=27.

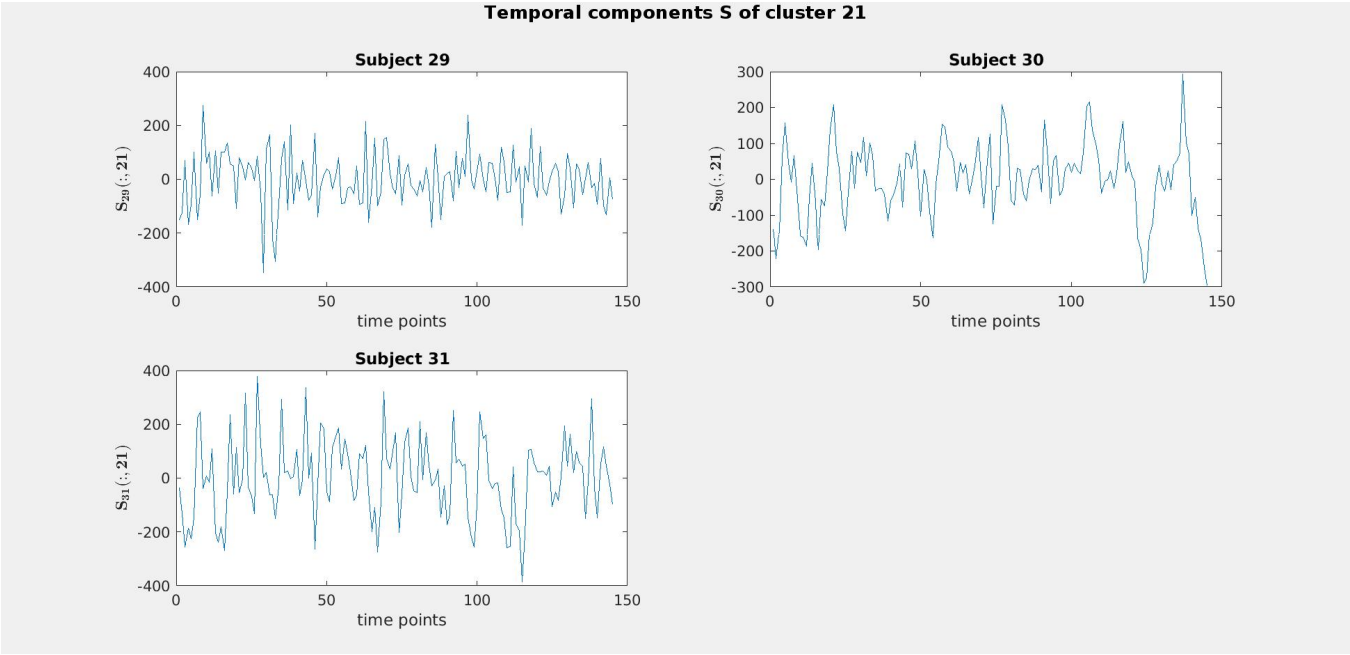


Figure B.193: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 29, \ldots, 31$ and R=27.

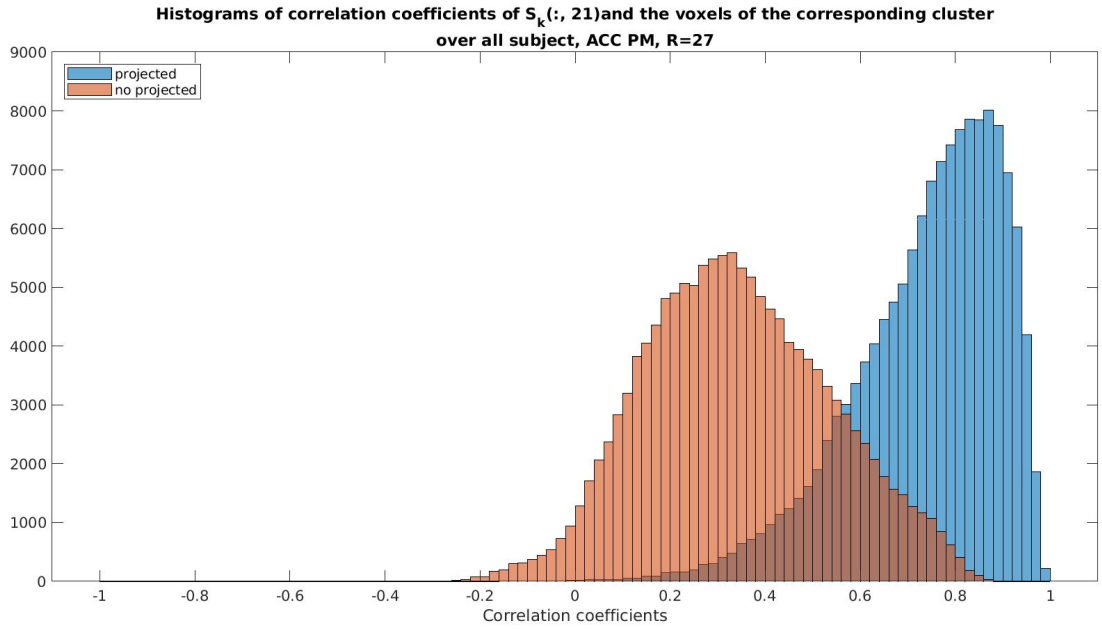


Figure B.194: For the column 21 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 21)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

18. **Cluster 4**
    Cluster 4 contains 5282 voxels, which correspond to 2.337706% of the whole brain (background, soft tissue, bone are excluded).

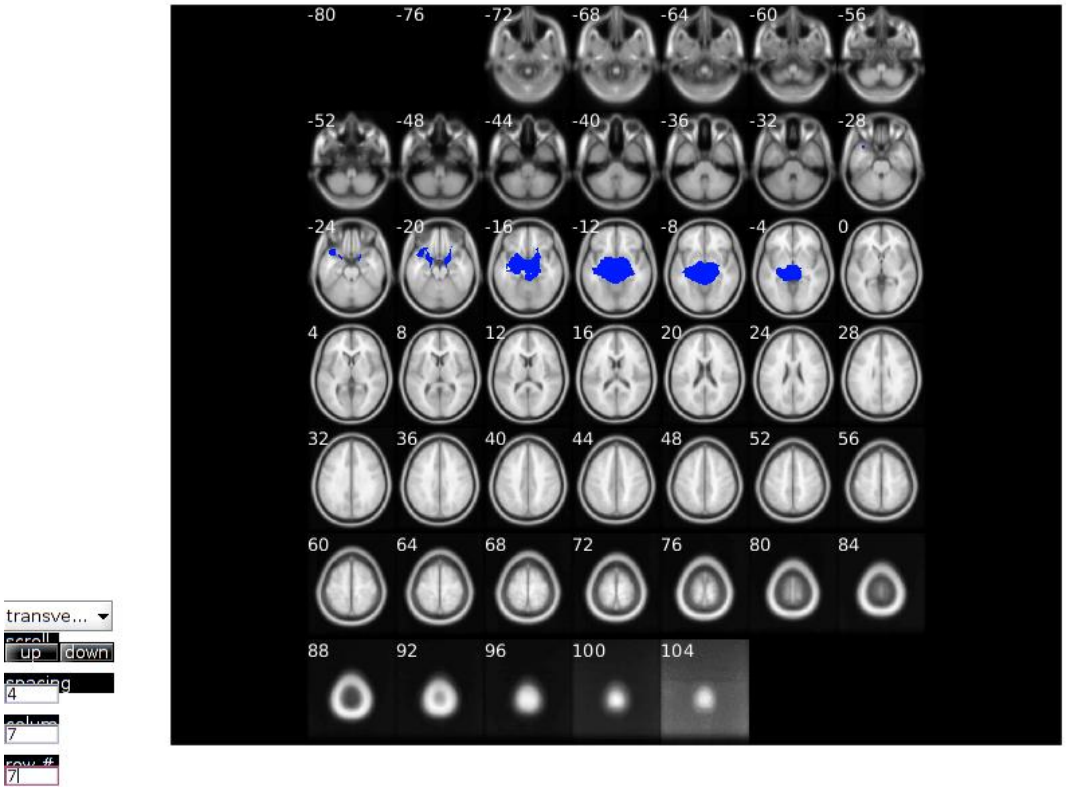


Figure B.195: Cluster 4. Blue areas indicate the voxels which belong to cluster 4.

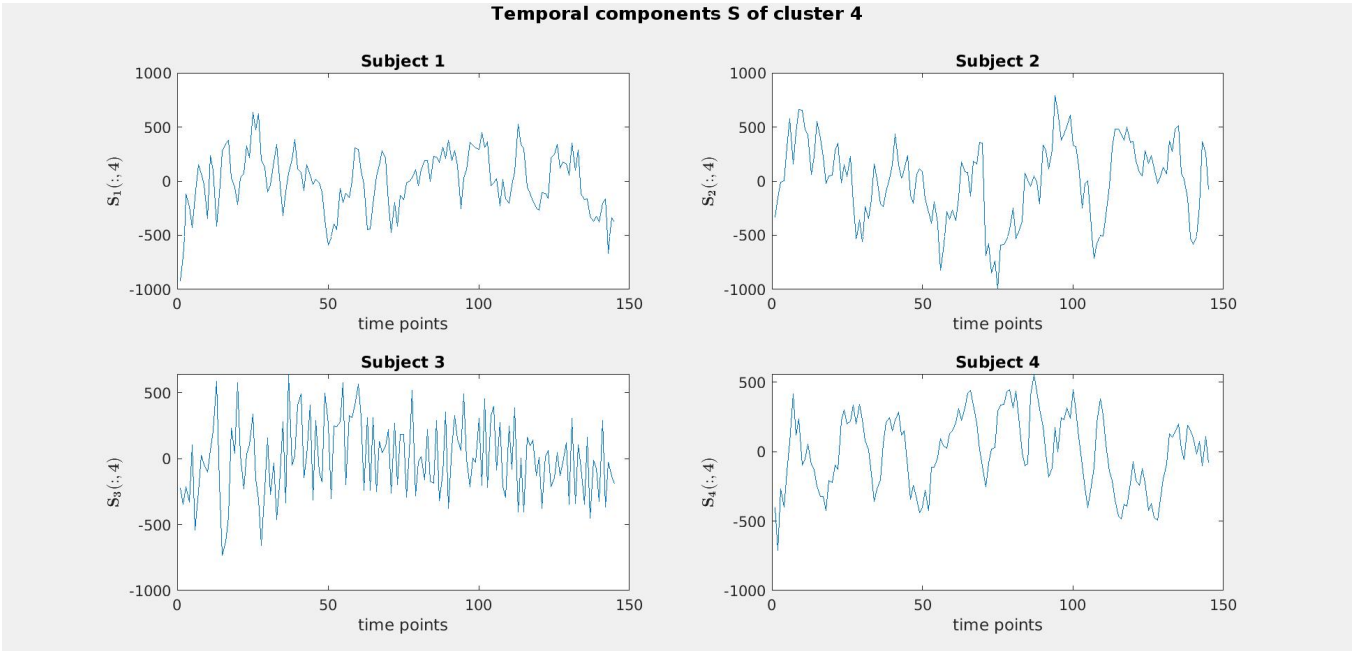


Figure B.196: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 1,\ldots,4$ and R=27.

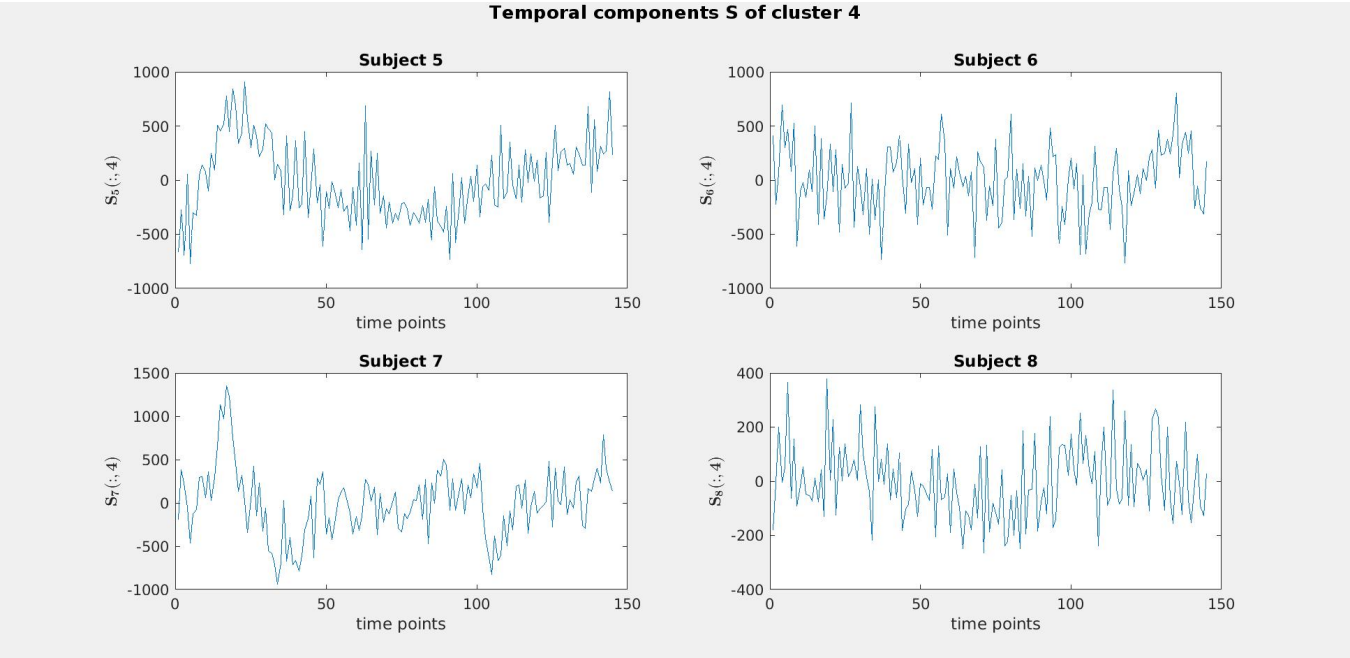


Figure B.197: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 5,\ldots,8$ and R=27.

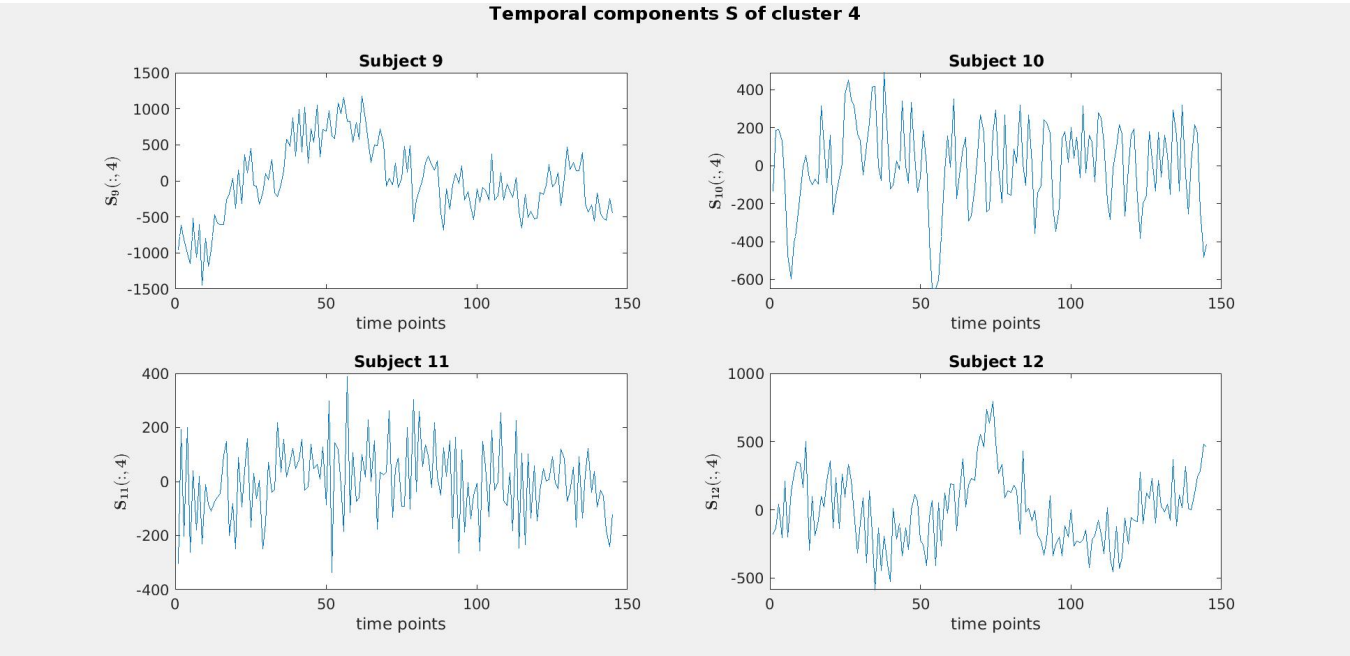


Figure B.198: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 9,\ldots,12$ and R=27.

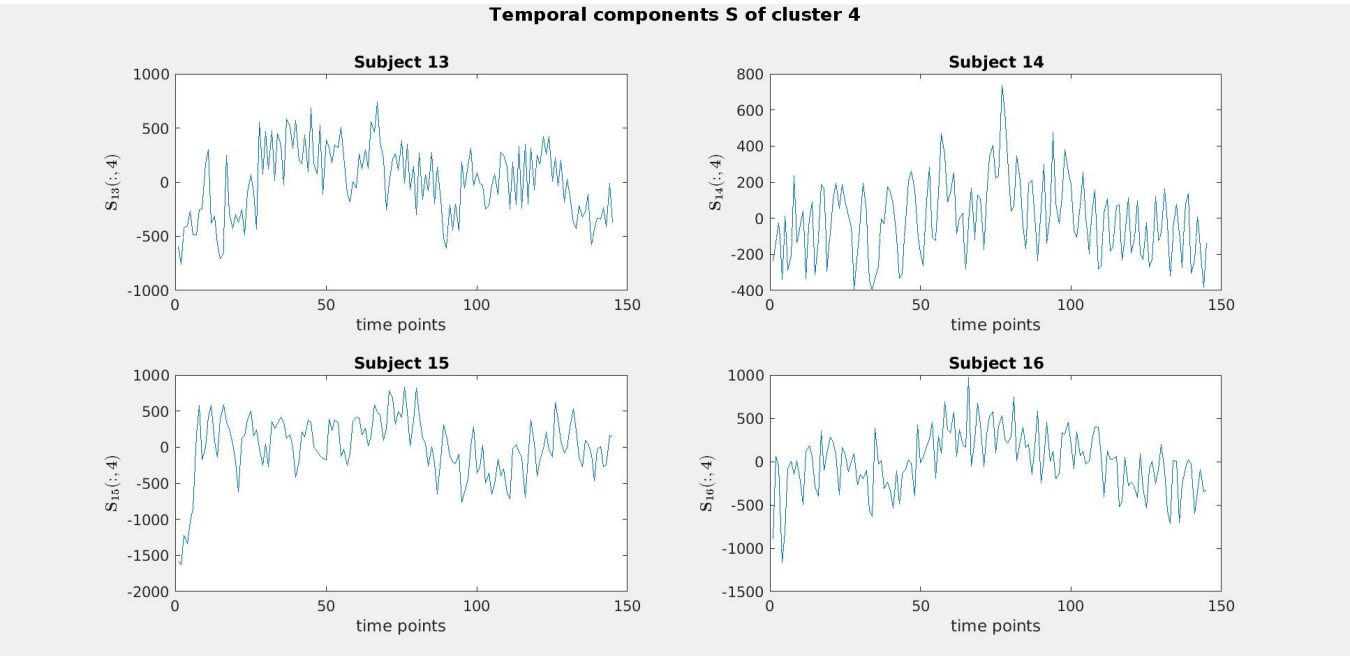


Figure B.199: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 13,\ldots,16$ and R=27.

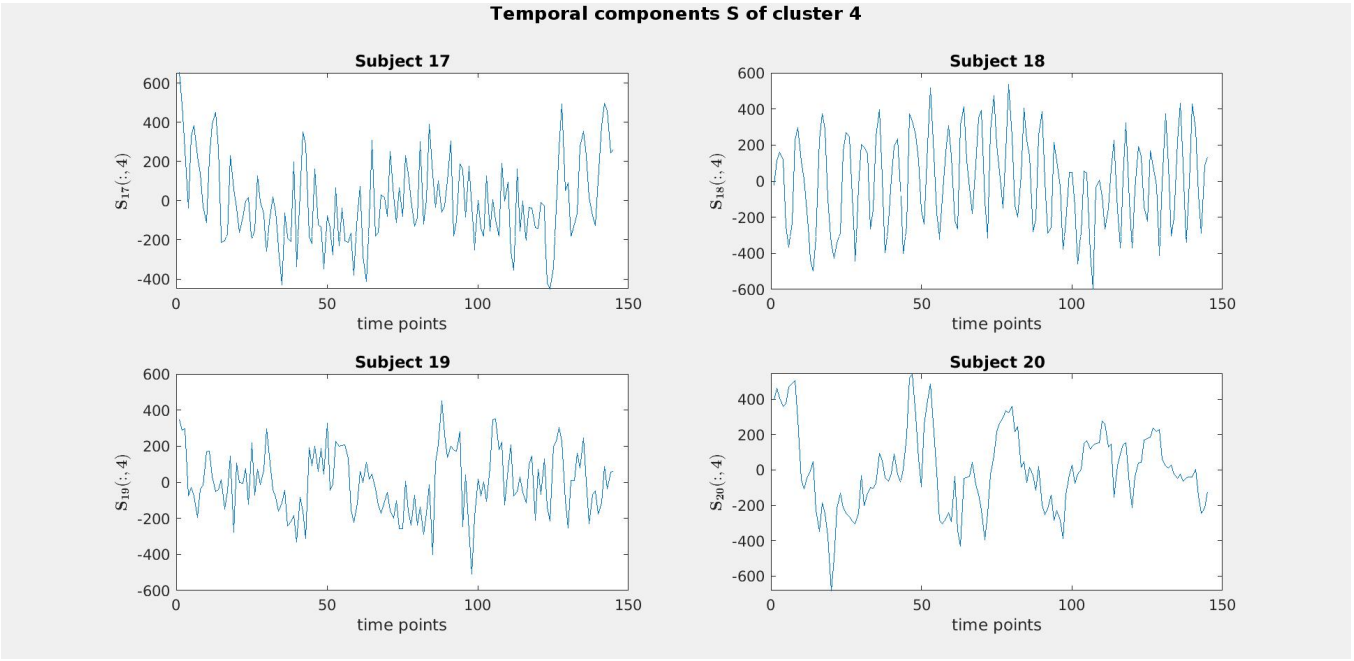


Figure B.200: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 17,\ldots,20$ and R=27.

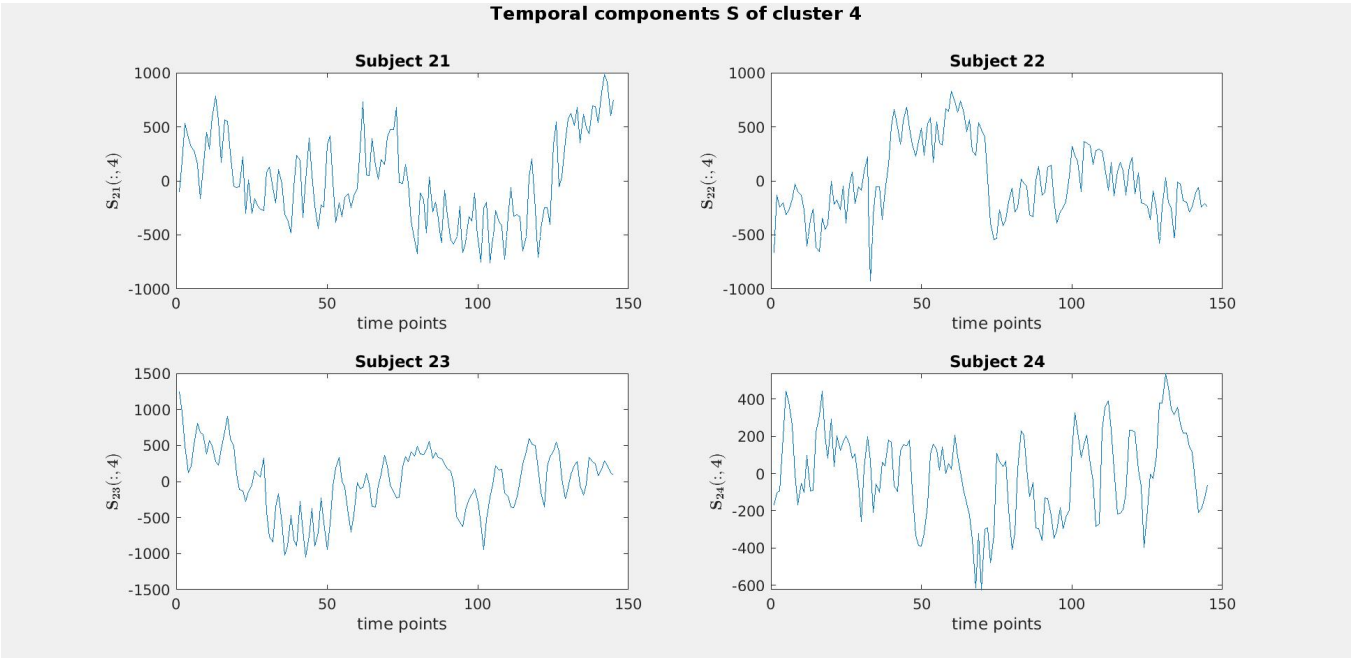


Figure B.201: Waveform of temporal components $\mathbf{S}_k(:,14)$ for $k = 21,\ldots,24$ and R=27.

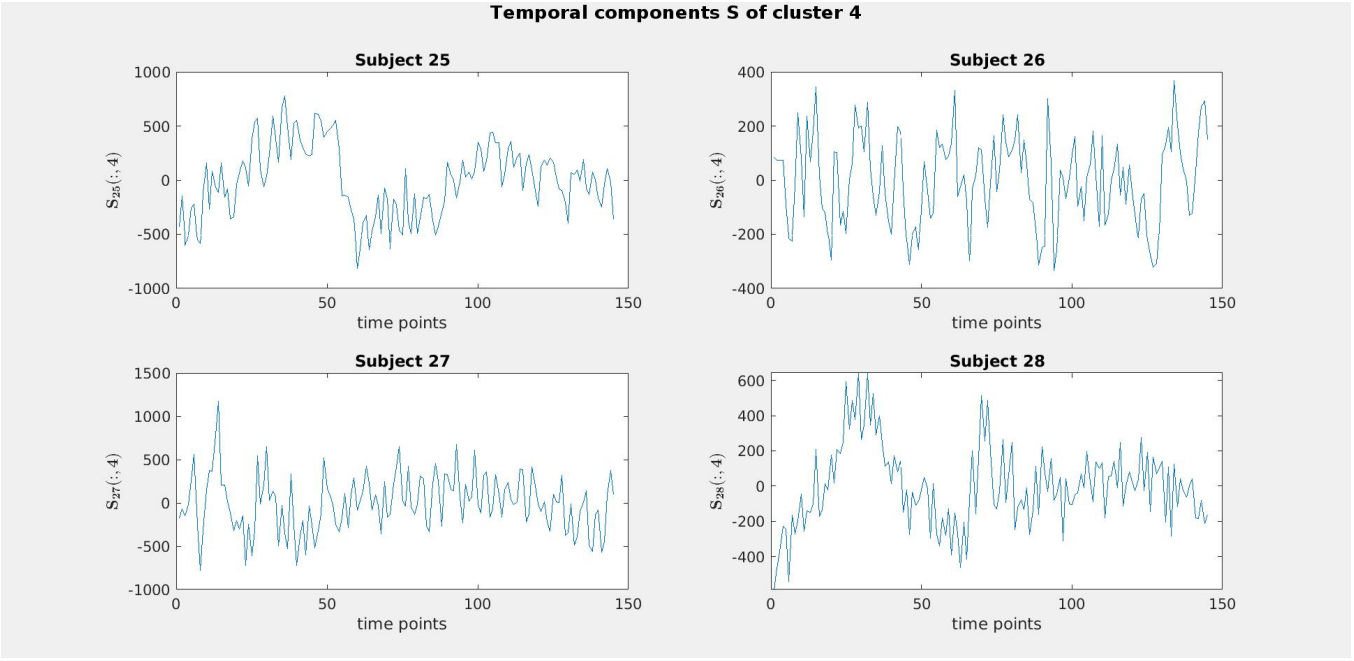


Figure B.202: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 25,\ldots,28$ and R=27.

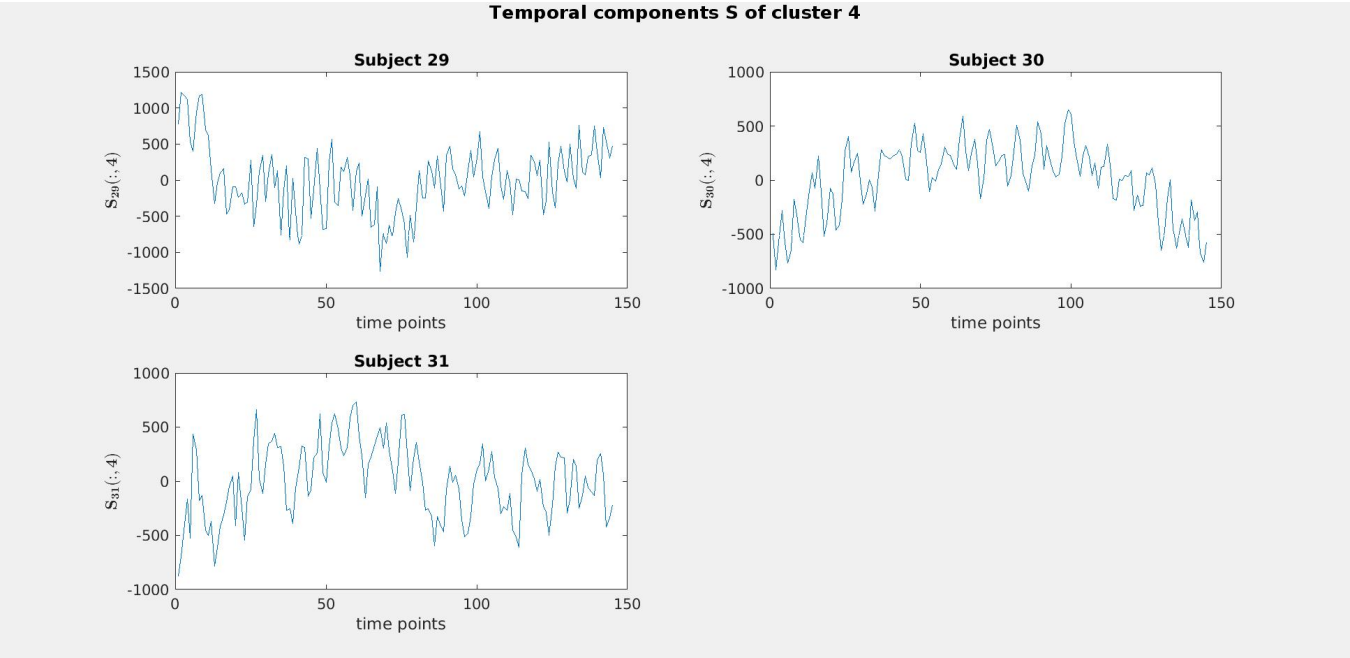


Figure B.203: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 29, \ldots, 31$ and R=27.

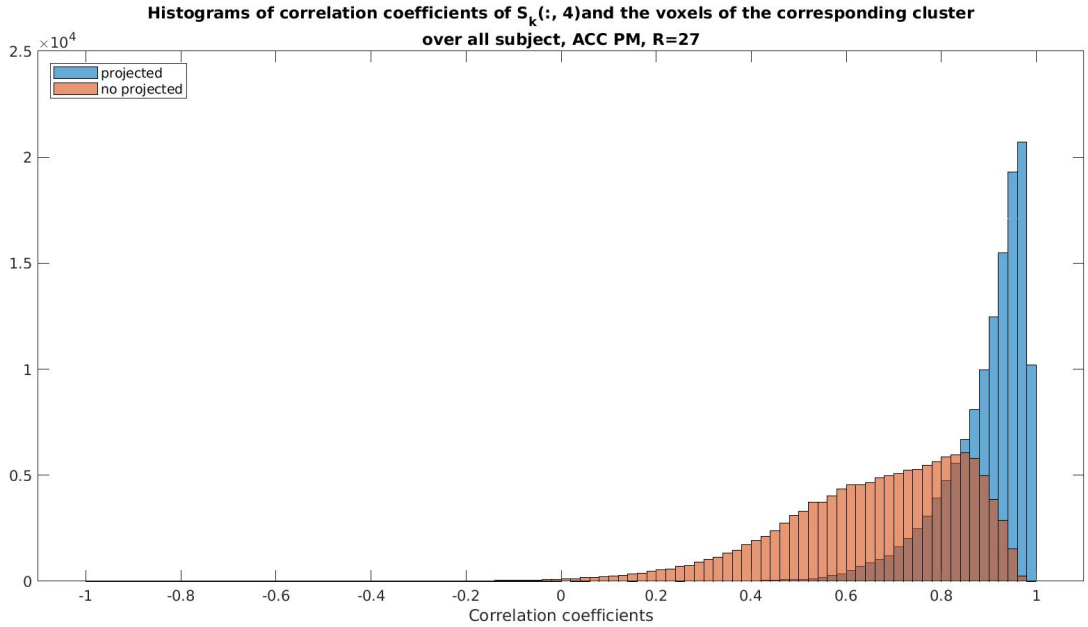


Figure B.204: For the column 4 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,4)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

19. **Cluster 16**

Cluster 16 contains 4412 voxels, which correspond to 1.952662% of the whole brain (background, soft tissue, bone are excluded).

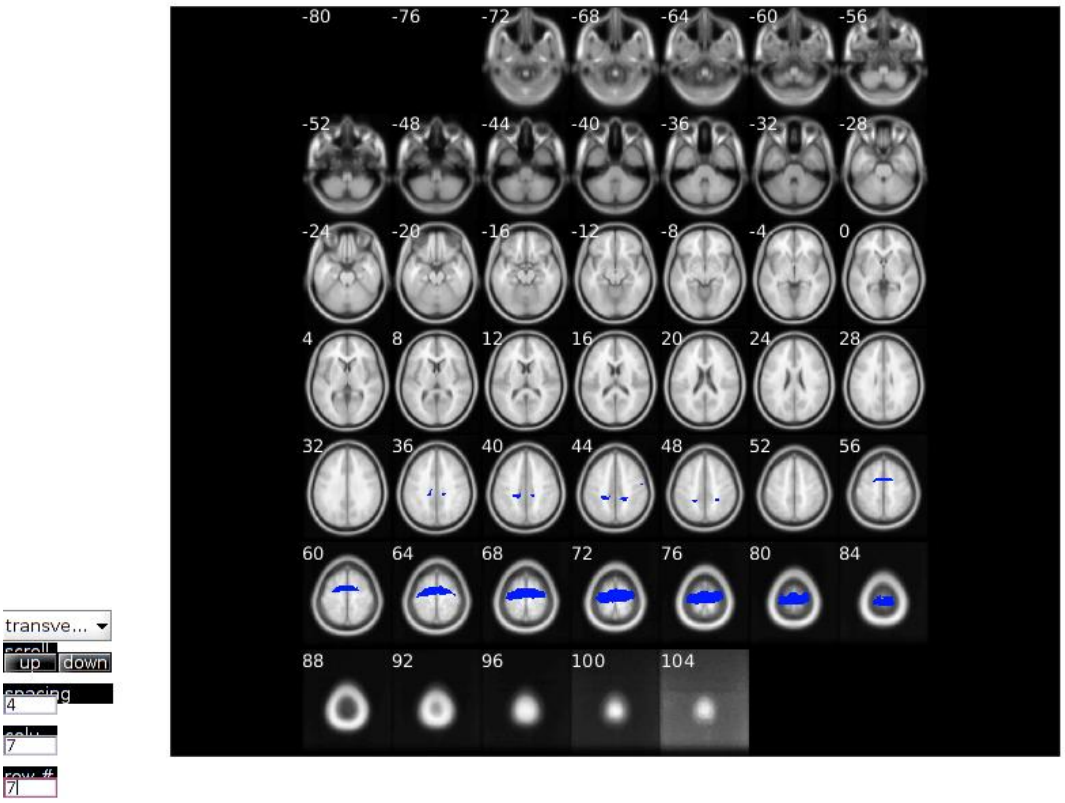


Figure B.205: Cluster 16. Blue areas indicate the voxels which belong to cluster 16.

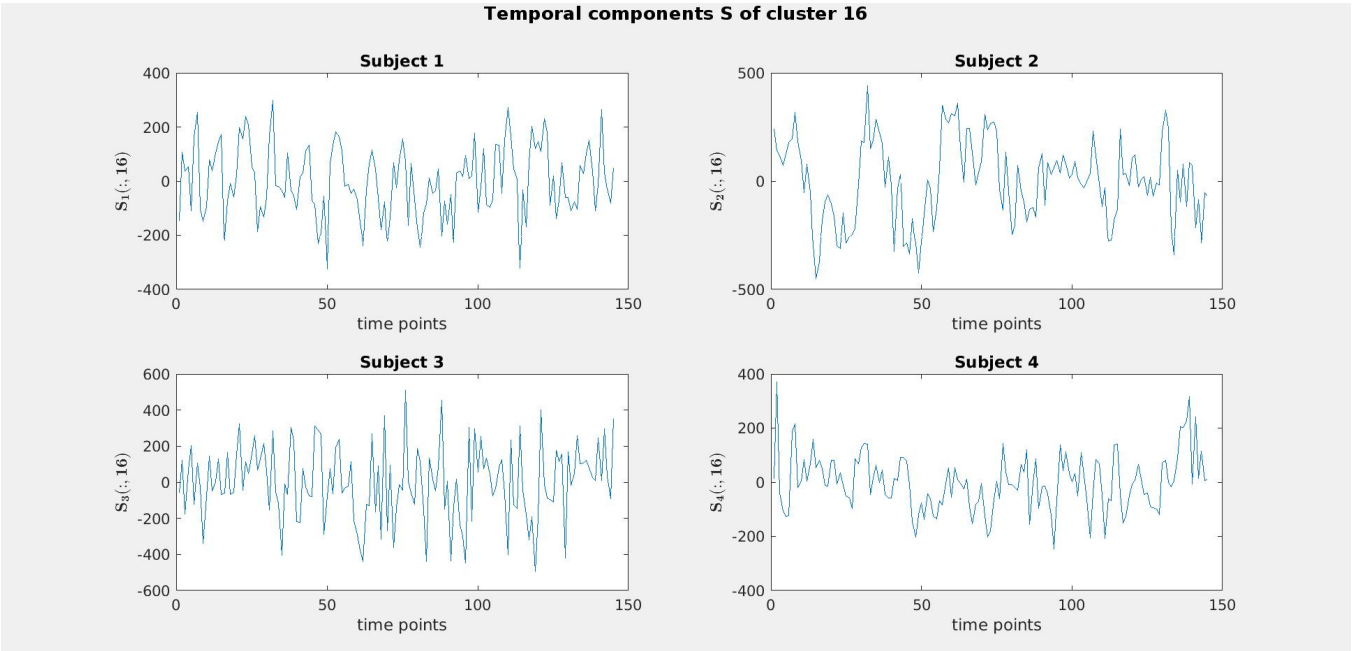


Figure B.206: Waveform of temporal components $\mathbf{S}_k(:,16)$ for $k = 1,\ldots,4$ and R=27.

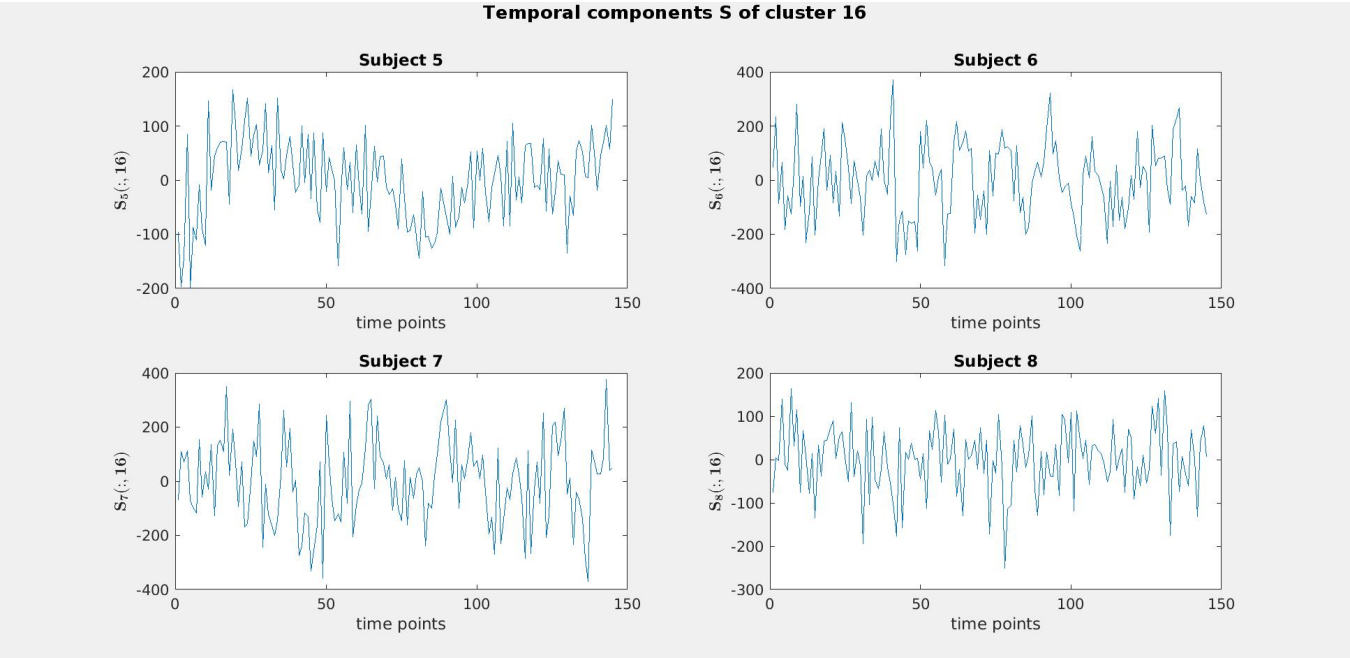


Figure B.207: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 5, \ldots, 8$ and R=27.

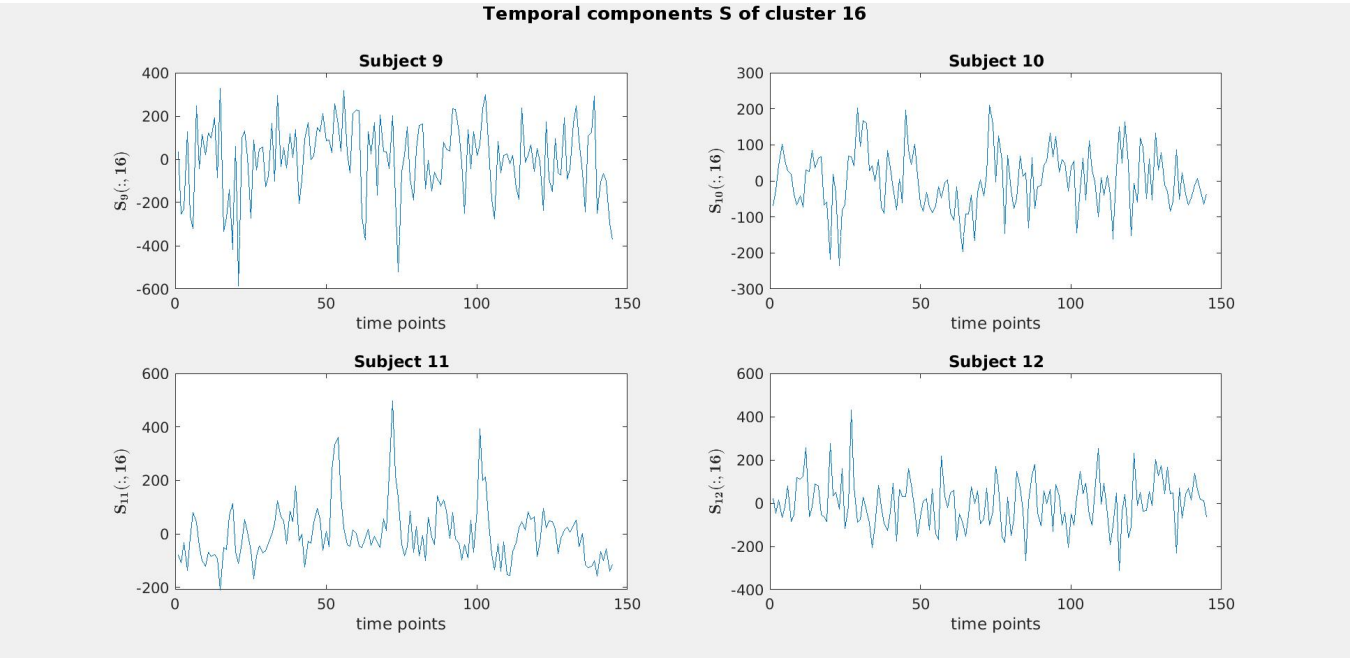


Figure B.208: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 9, \ldots, 12$ and R=27.

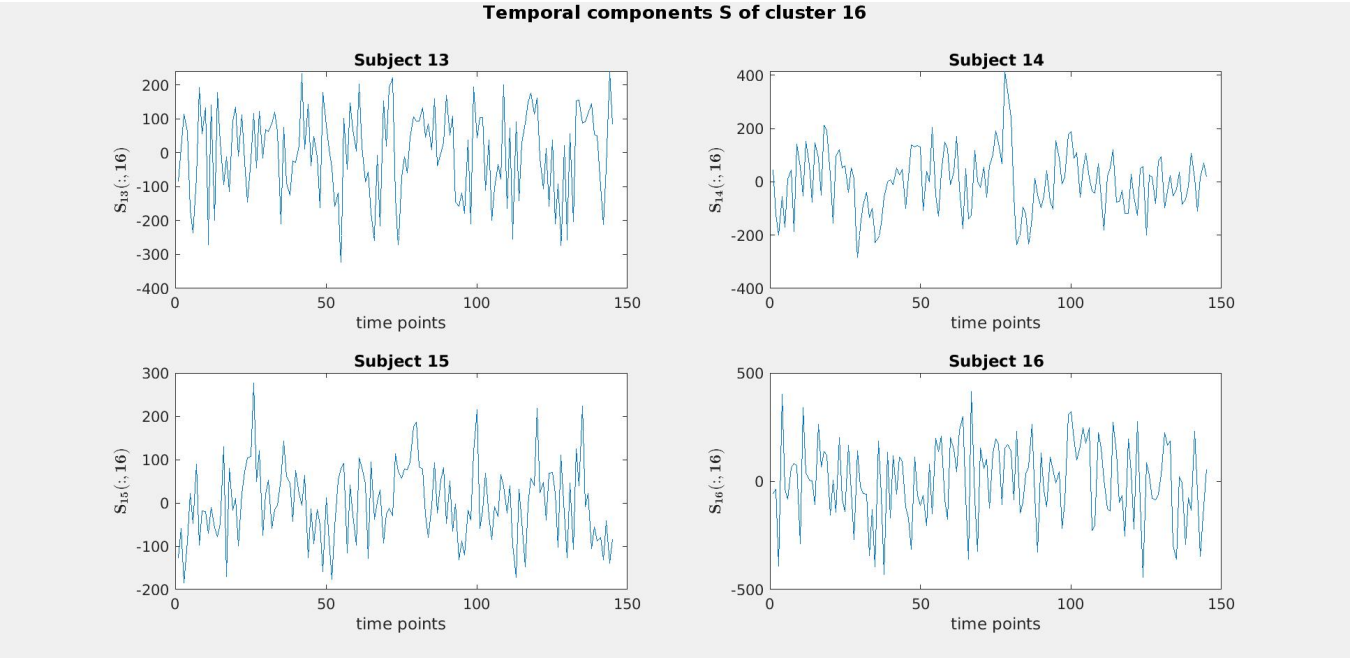


Figure B.209: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 13, \ldots, 16$ and R=27.

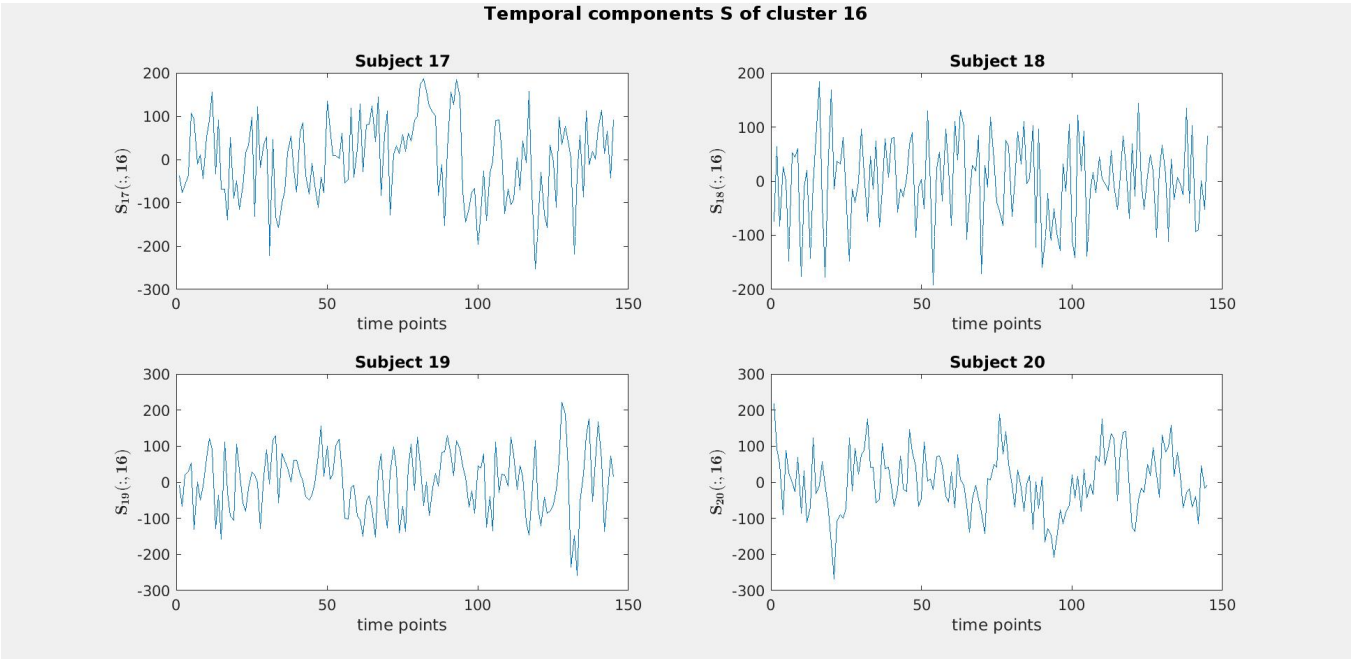


Figure B.210: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 17, \ldots, 20$ and R=27.

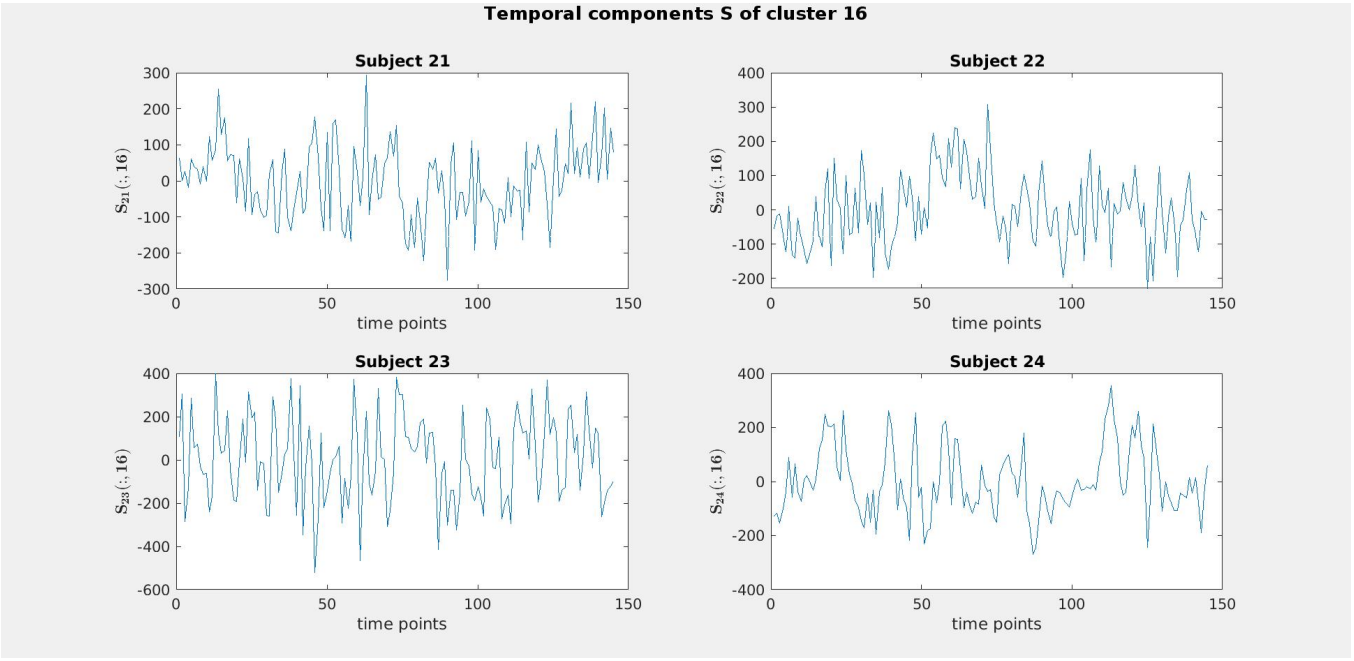


Figure B.211: Waveform of temporal components $\mathbf{S}_k(:, 116)$ for $k = 21, \ldots, 24$ and R=27.

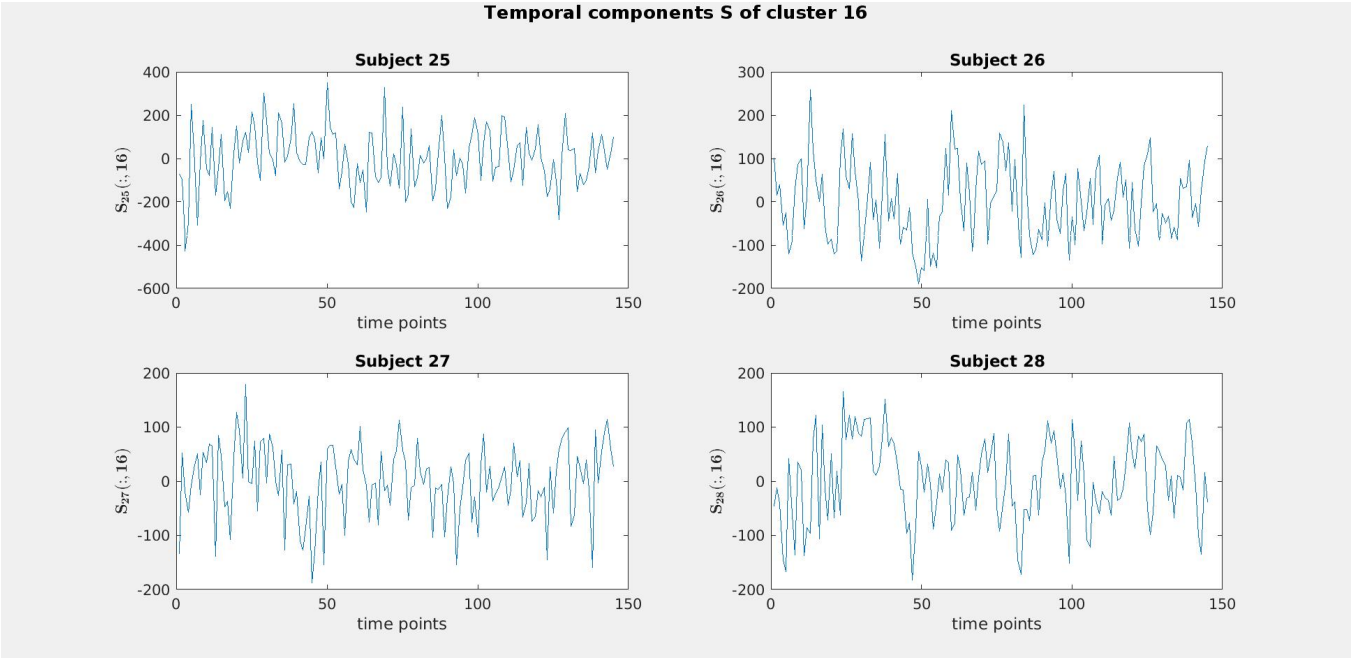


Figure B.212: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 25, \ldots, 28$ and R=27.

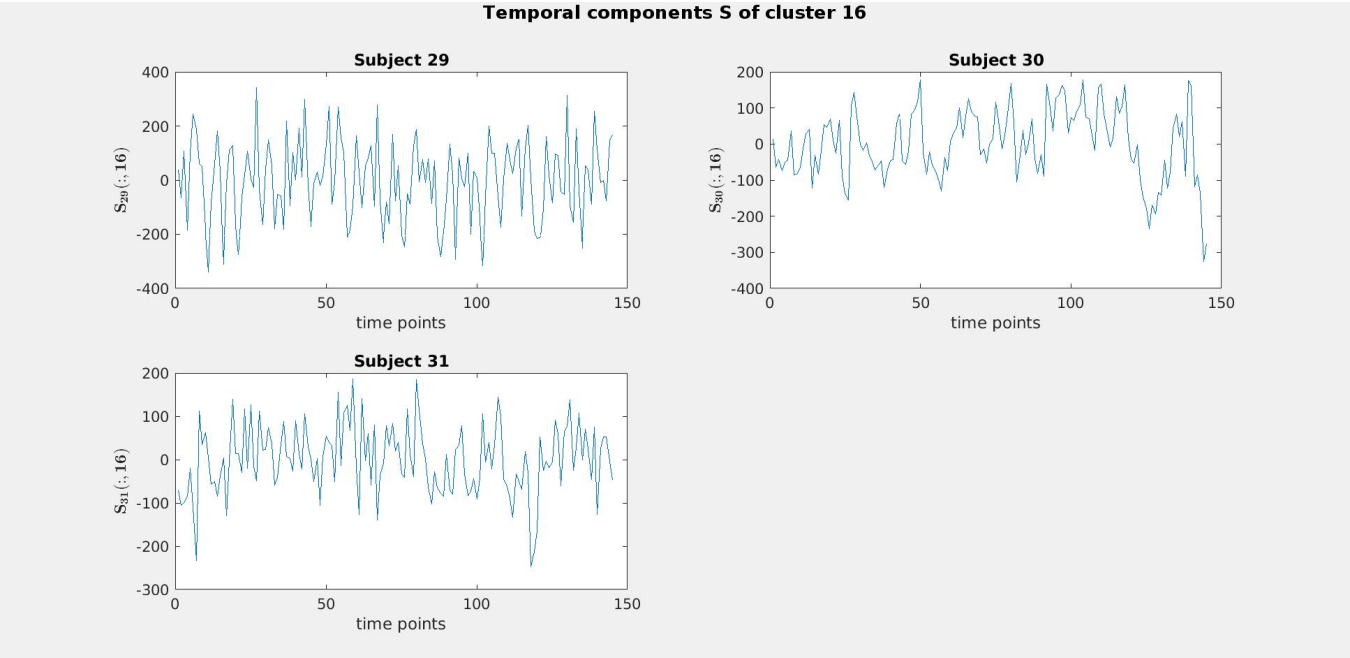


Figure B.213: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 29, \ldots, 31$ and R=27.

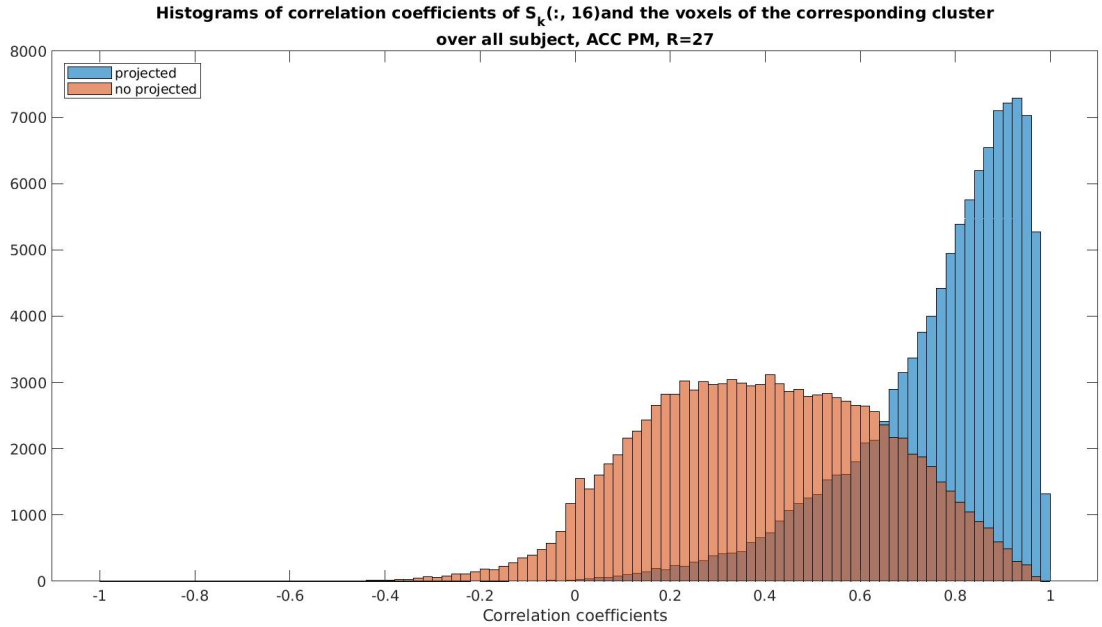


Figure B.214: For the column 16 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 16)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

20. **Cluster 3**

    Cluster 3 contains 4230 voxels, which correspond to 1.872112% of the whole brain (background, soft tissue, bone are excluded).

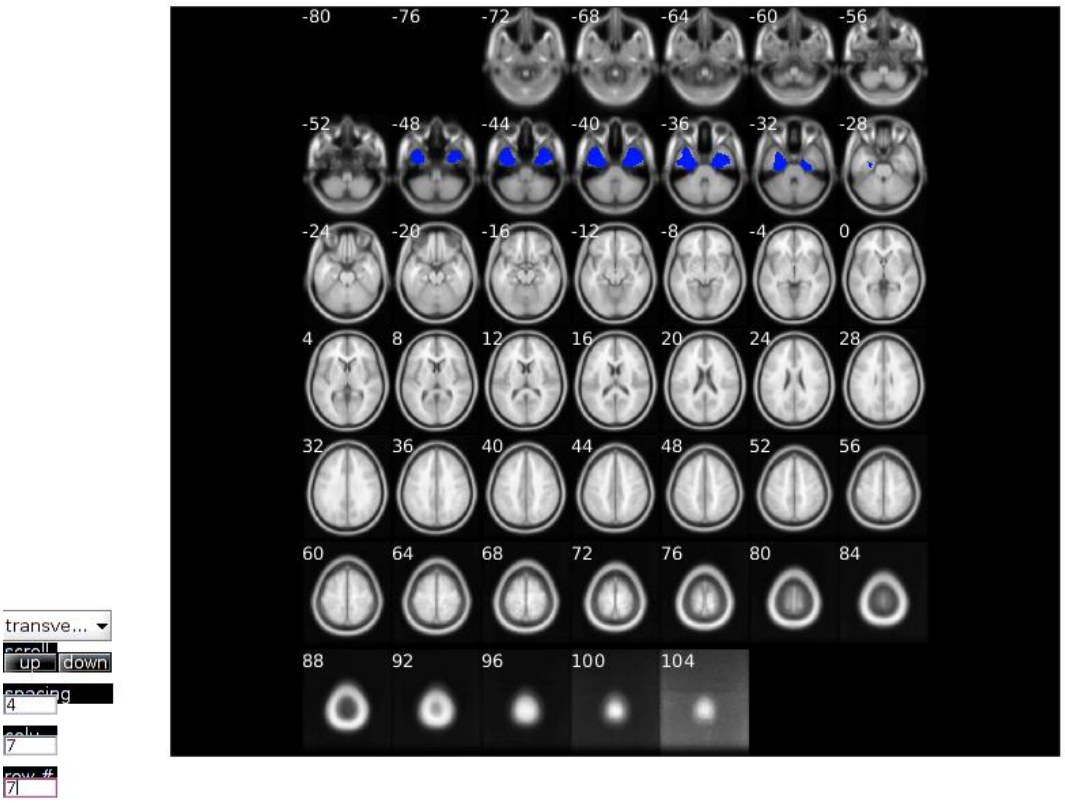


Figure B.215: Cluster 3. Blue areas indicate the voxels which belong to cluster 3.

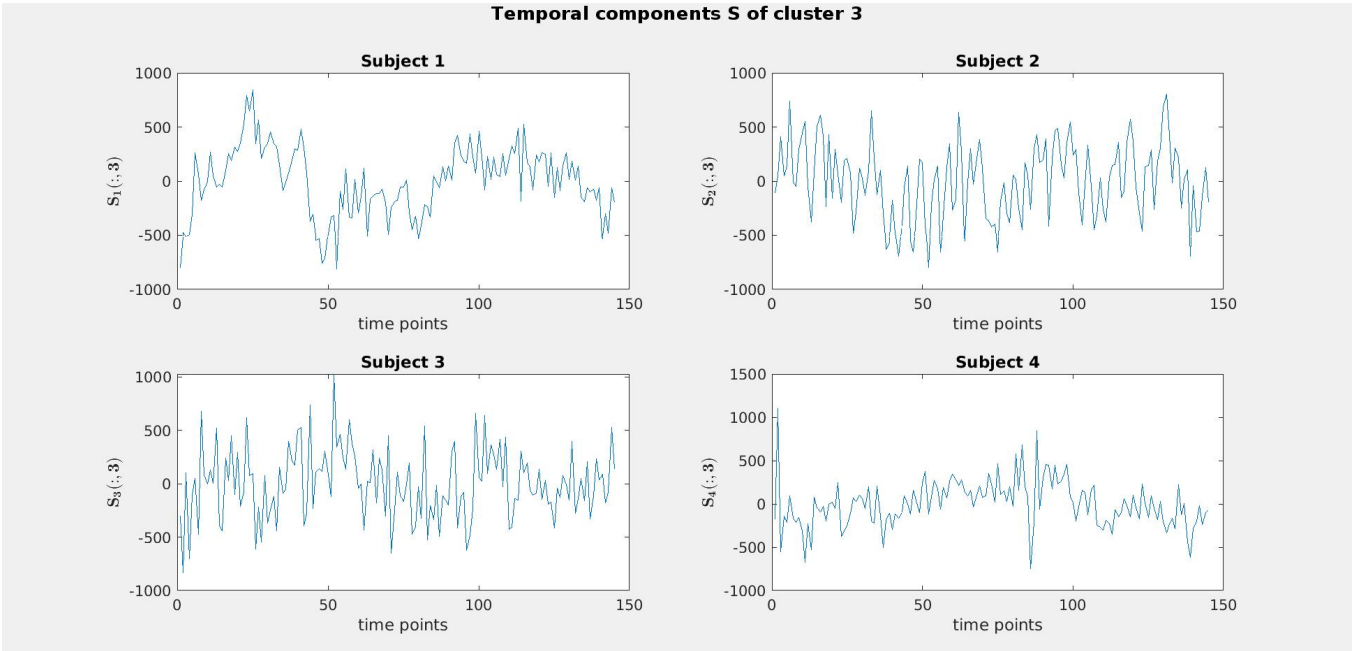


Figure B.216: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 1,\ldots,4$ and R=27.

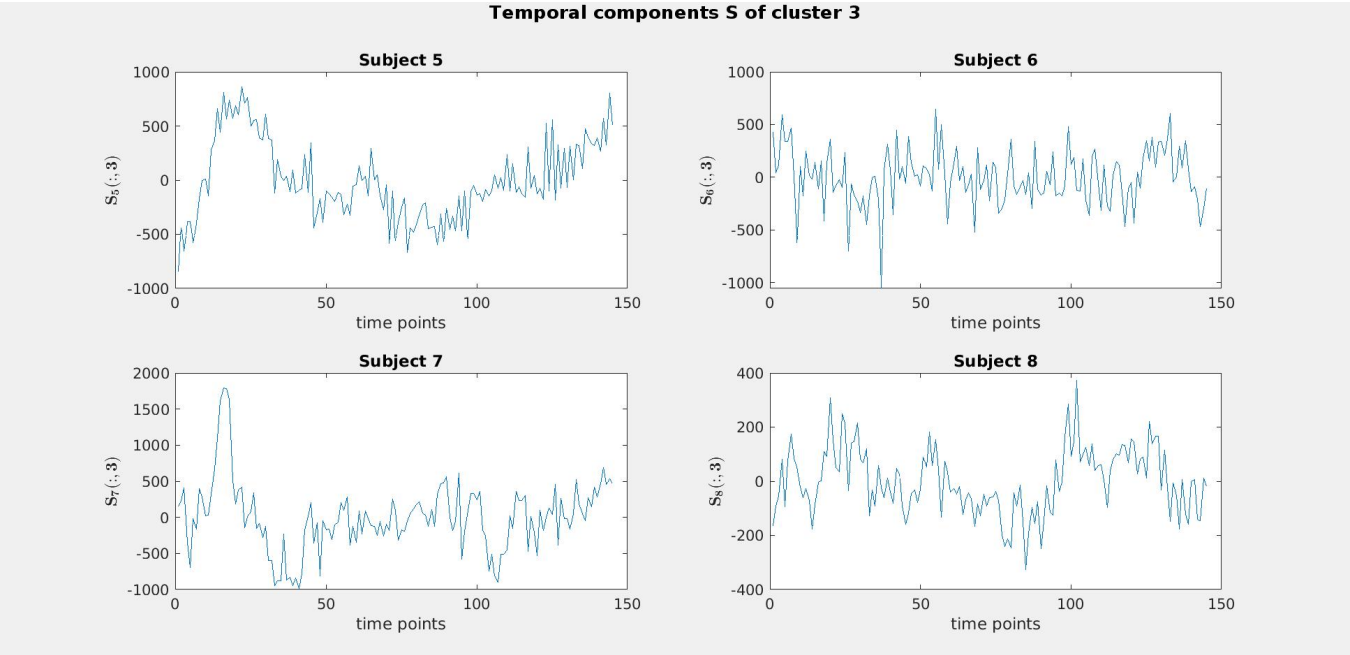


Figure B.217: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 5,\ldots,8$ and R=27.

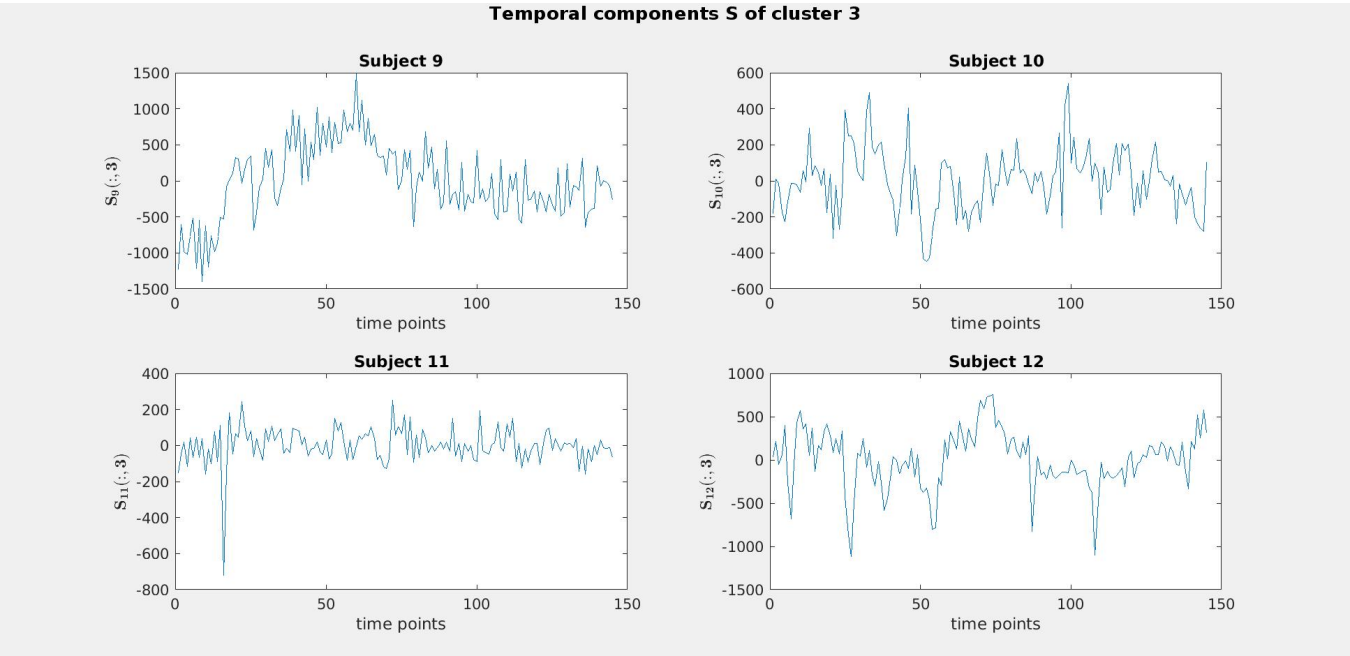


Figure B.218: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 9,\ldots,12$ and R=27.

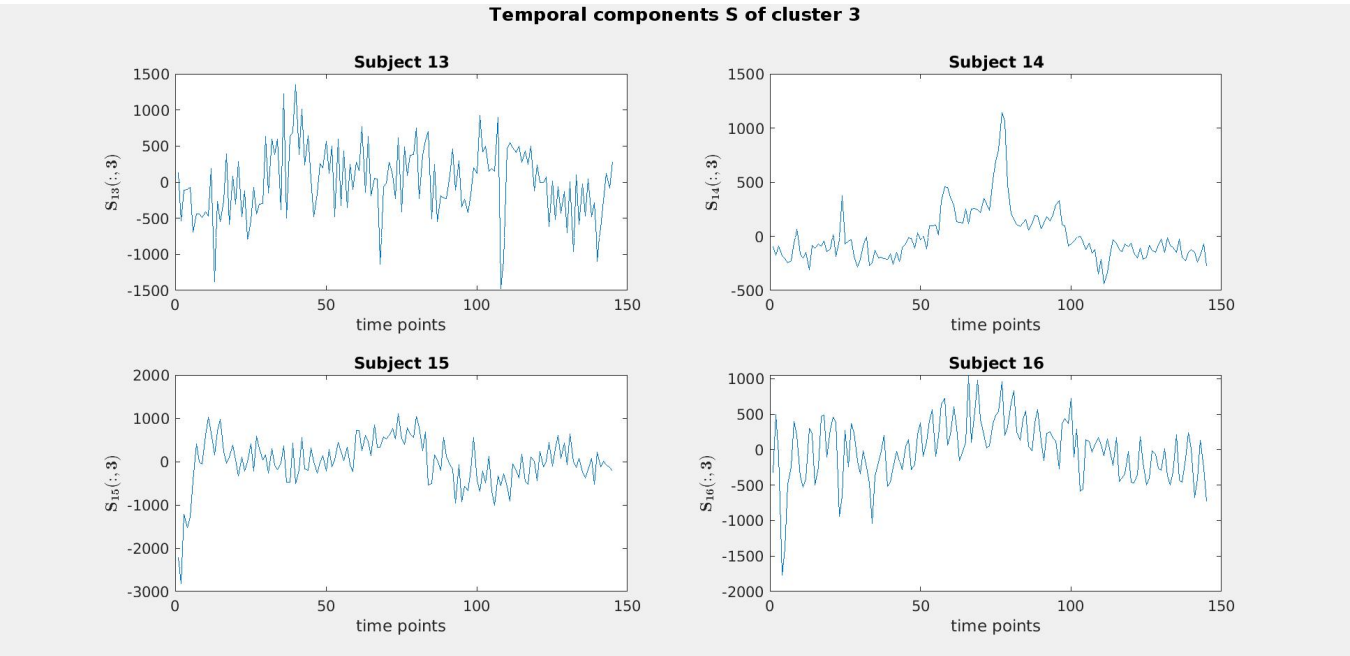


Figure B.219: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 13,\ldots,16$ and R=27.

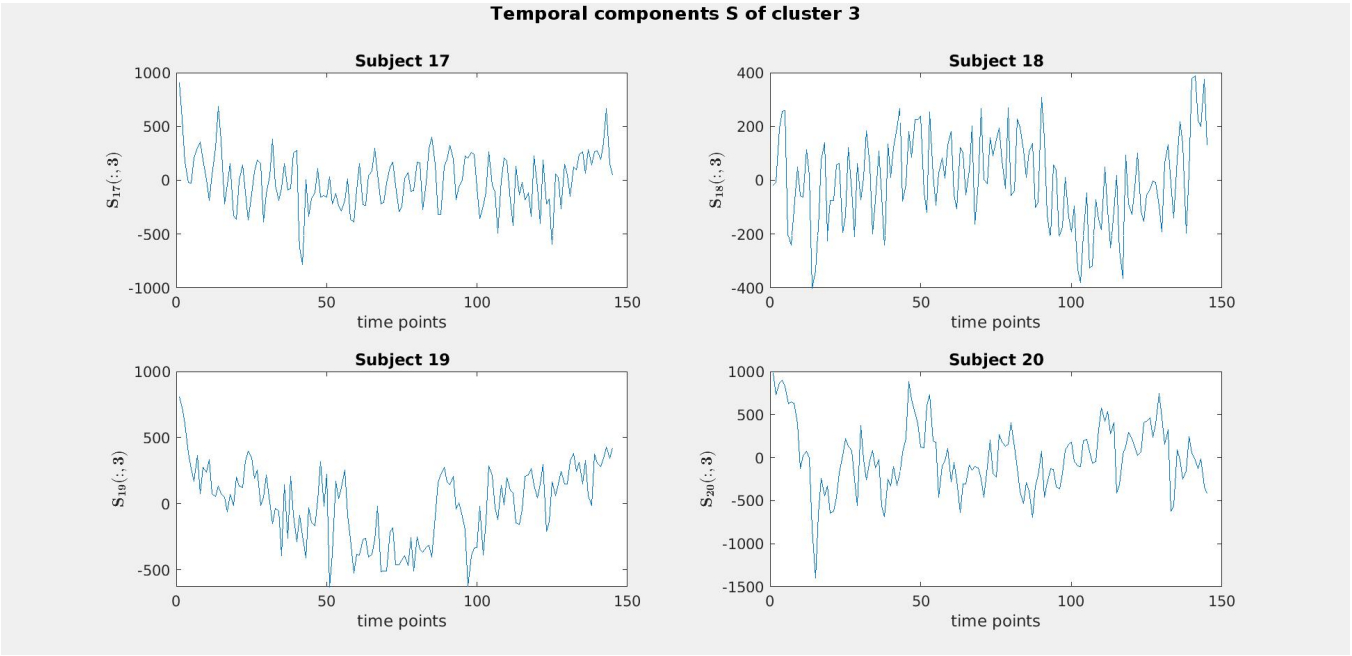


Figure B.220: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 17,\ldots,20$ and R=27.

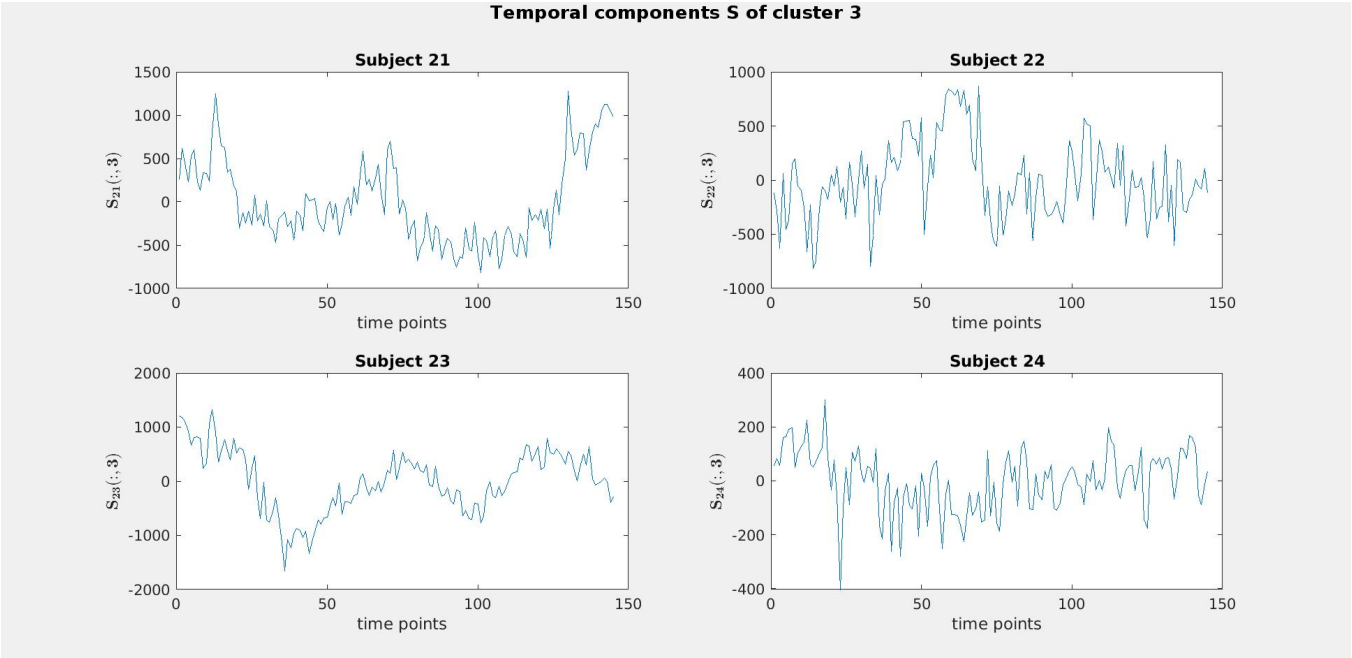


Figure B.221: Waveform of temporal components $\mathbf{S}_k(:,13)$ for $k = 21,\ldots,24$ and R=27.

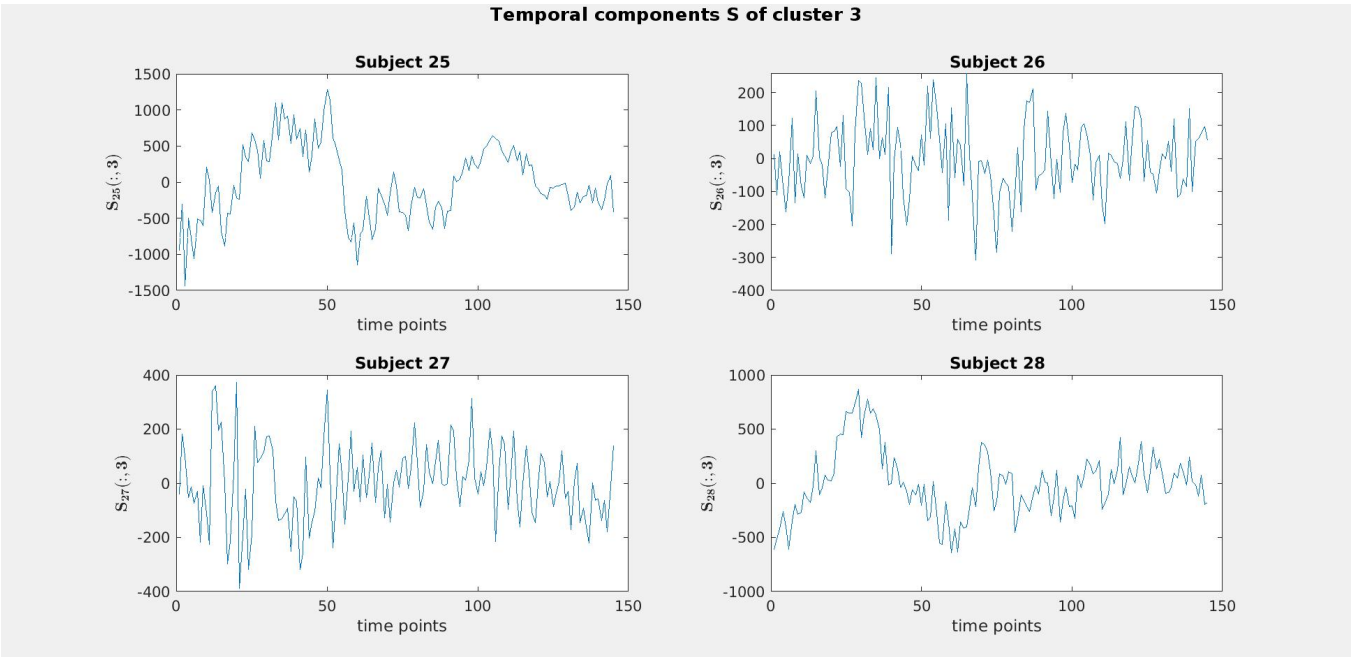


Figure B.222: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 25,\ldots,28$ and R=27.

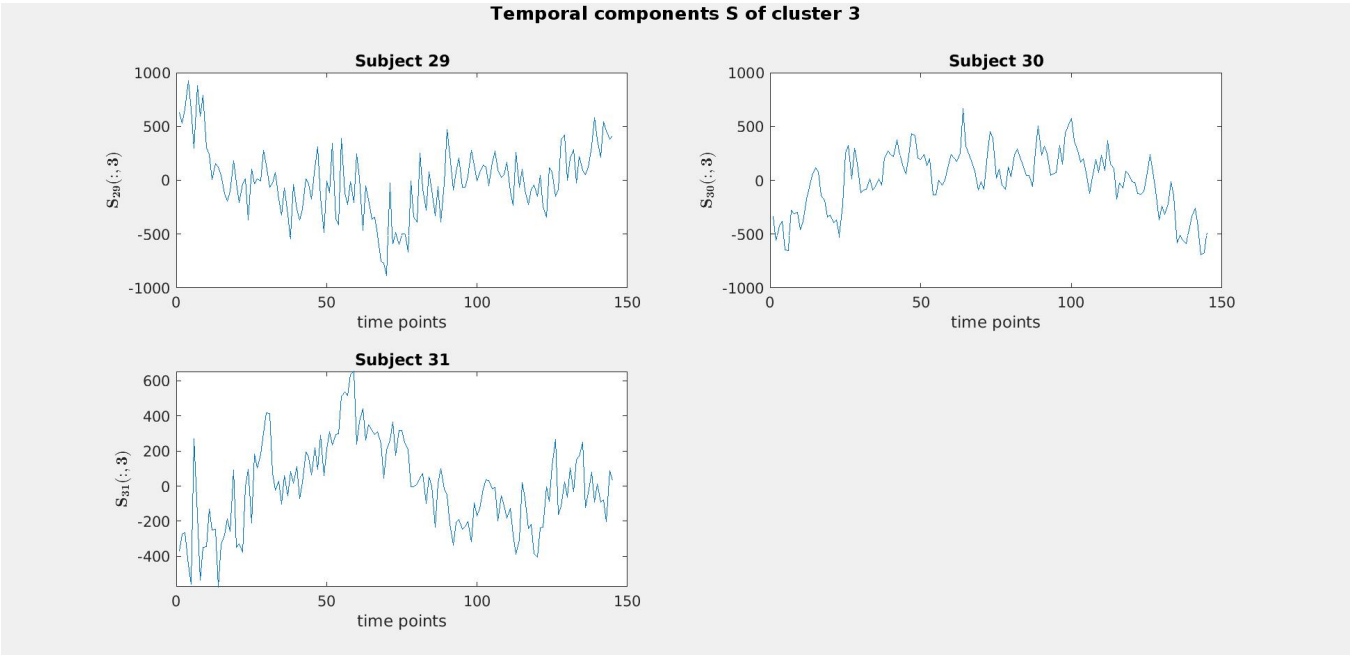


Figure B.223: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 29, \ldots, 31$ and R=27.

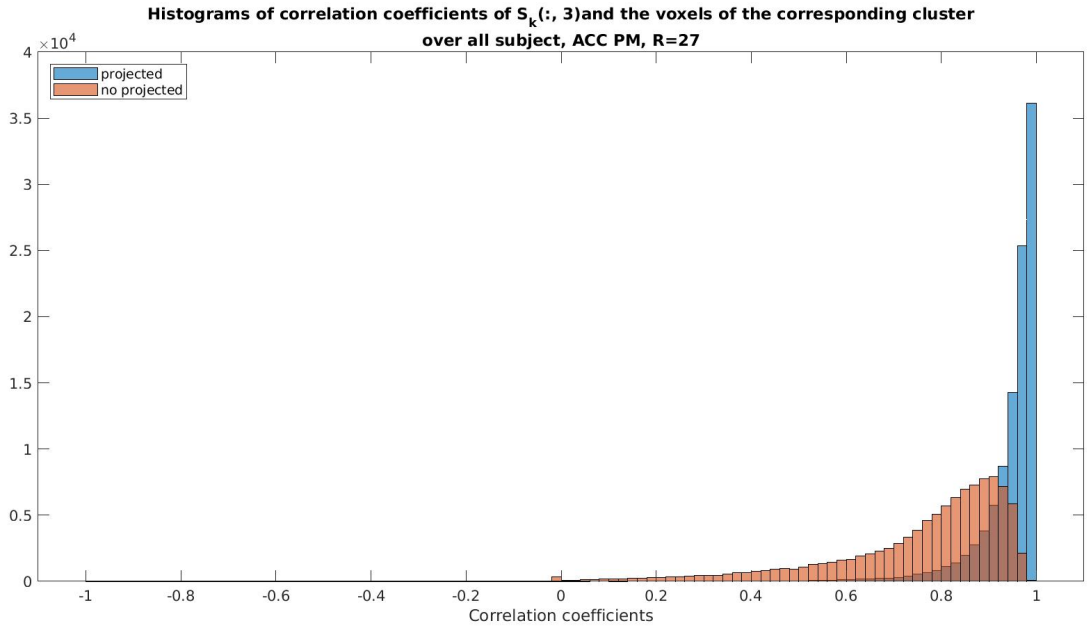


Figure B.224: For the column 3 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,3)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

21. **Cluster 25**

Cluster 25 contains 3971 voxels, which correspond to 1.757484% of the whole brain (background, soft tissue, bone are excluded).

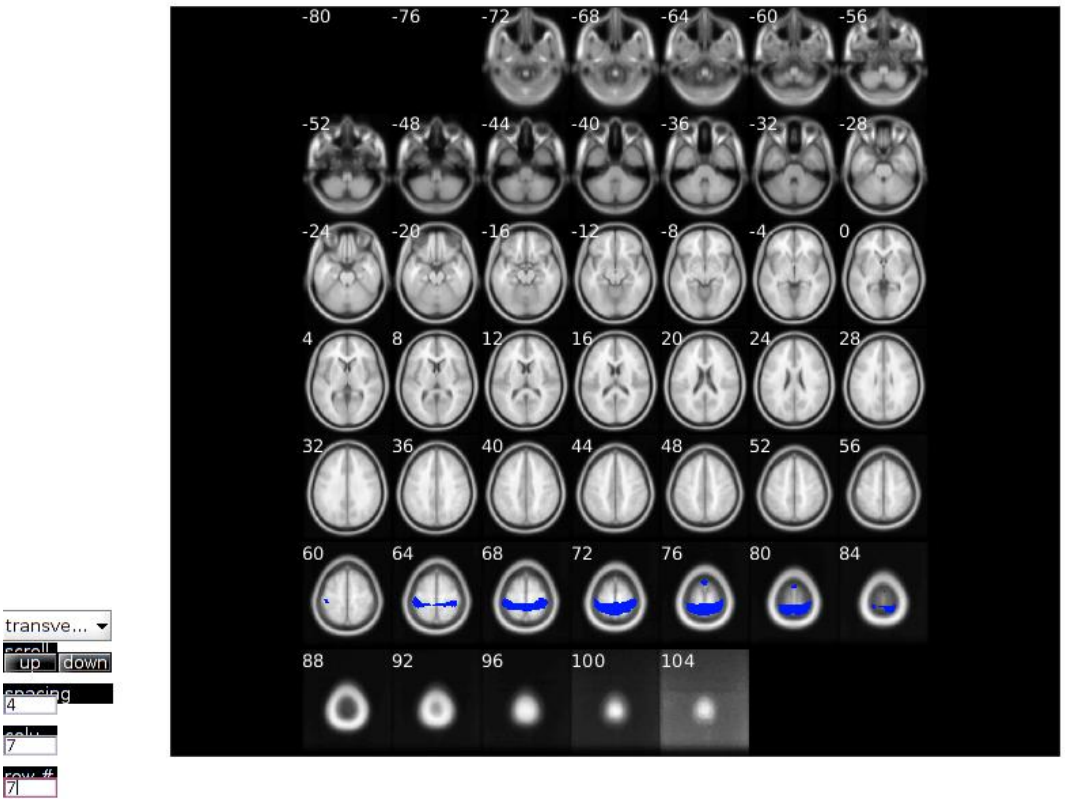


Figure B.225: Cluster 25. Blue areas indicate the voxels which belong to cluster 25.

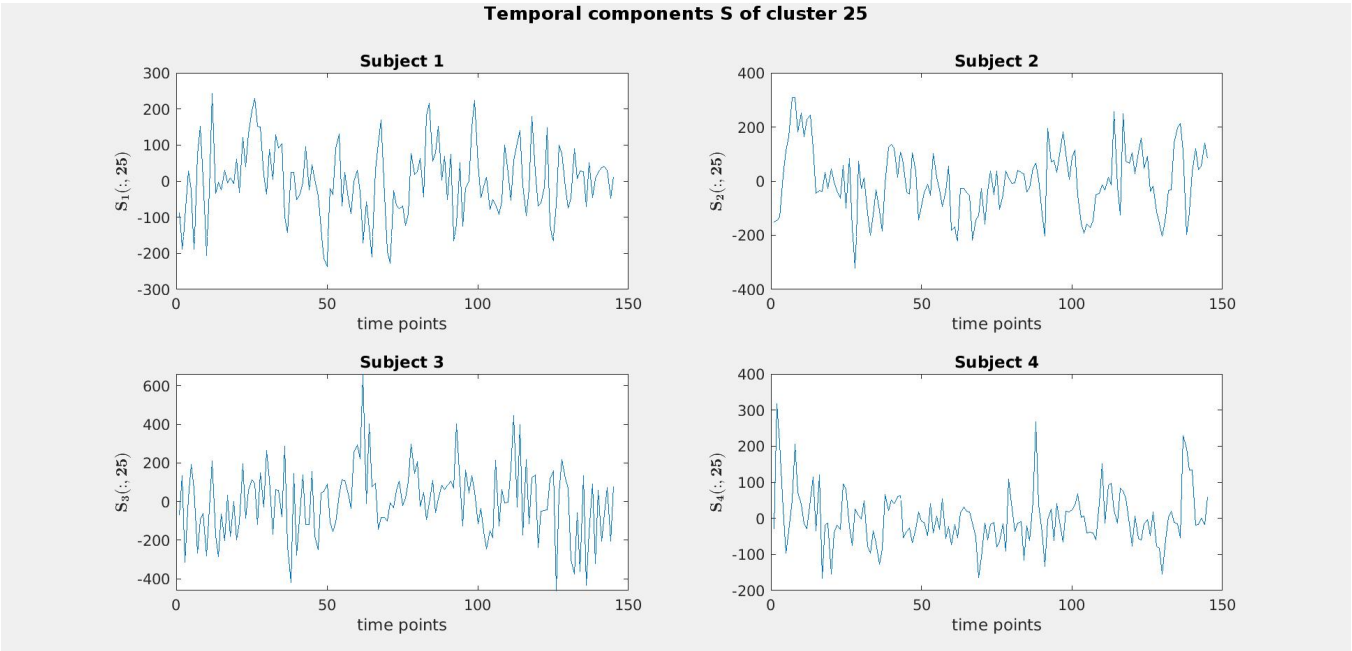


Figure B.226: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 1, \ldots, 4$ and R=27.

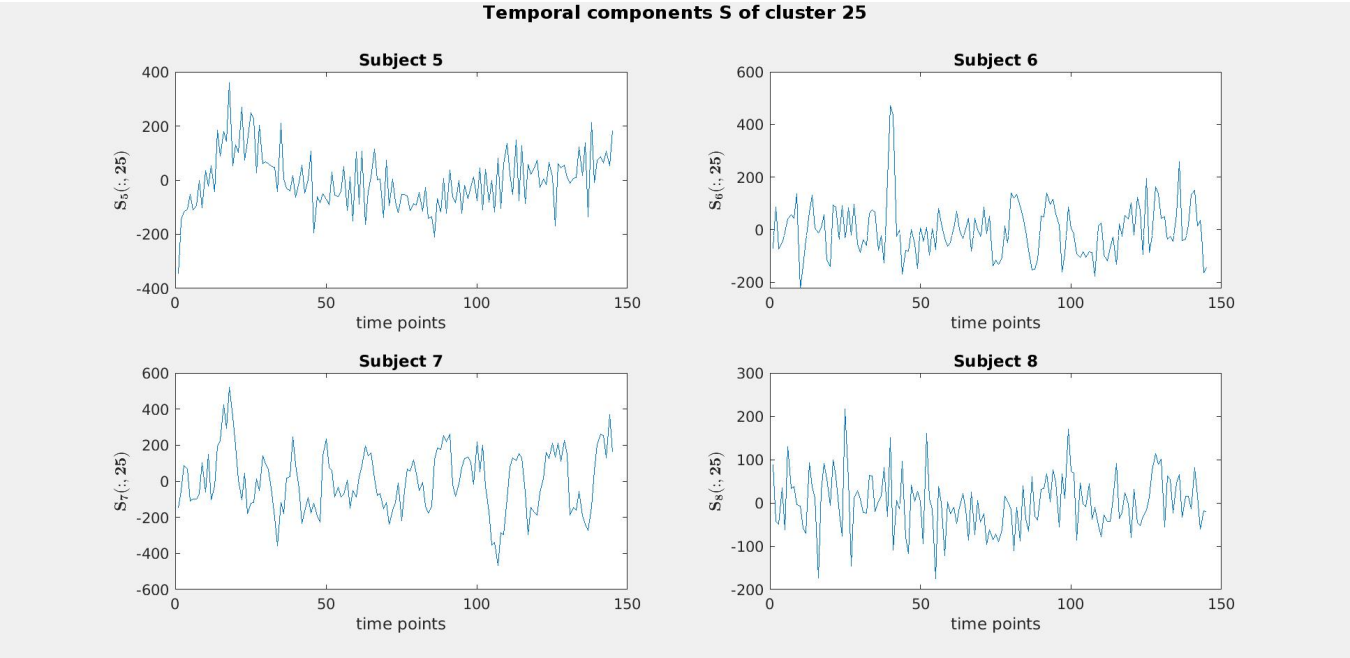


Figure B.227: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 5, \ldots, 8$ and R=27.

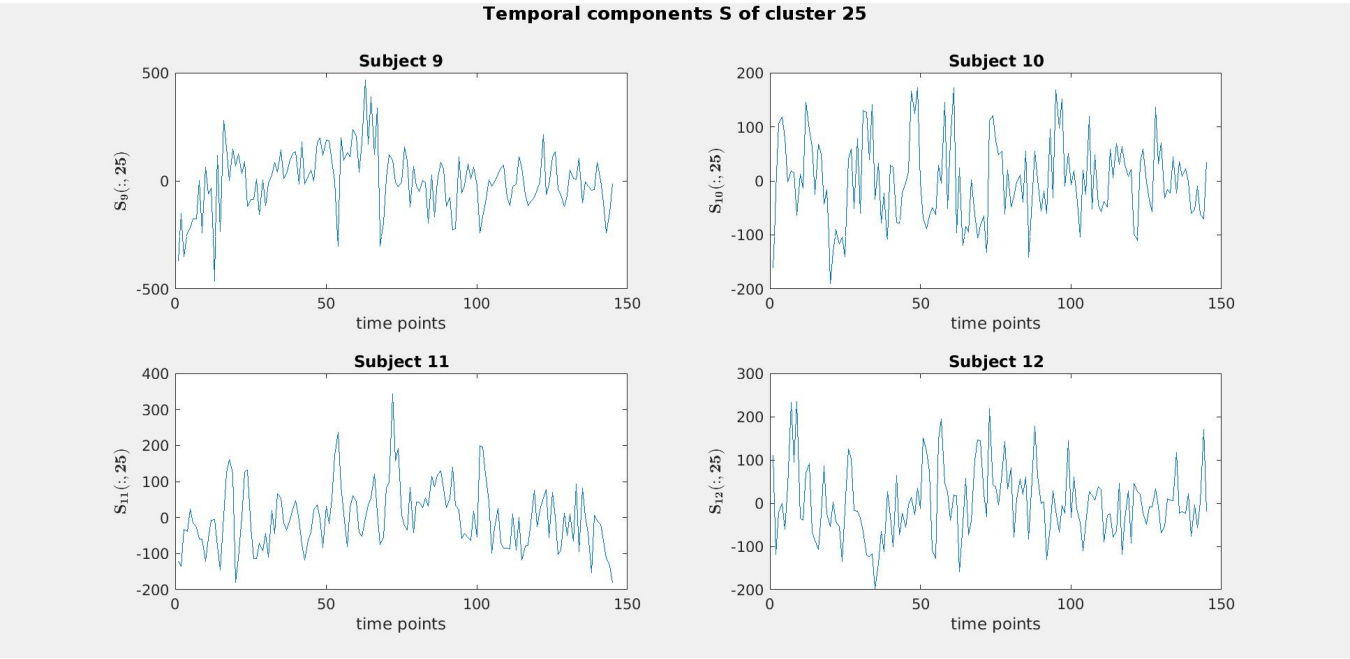


Figure B.228: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 9, \ldots, 12$ and R=27.

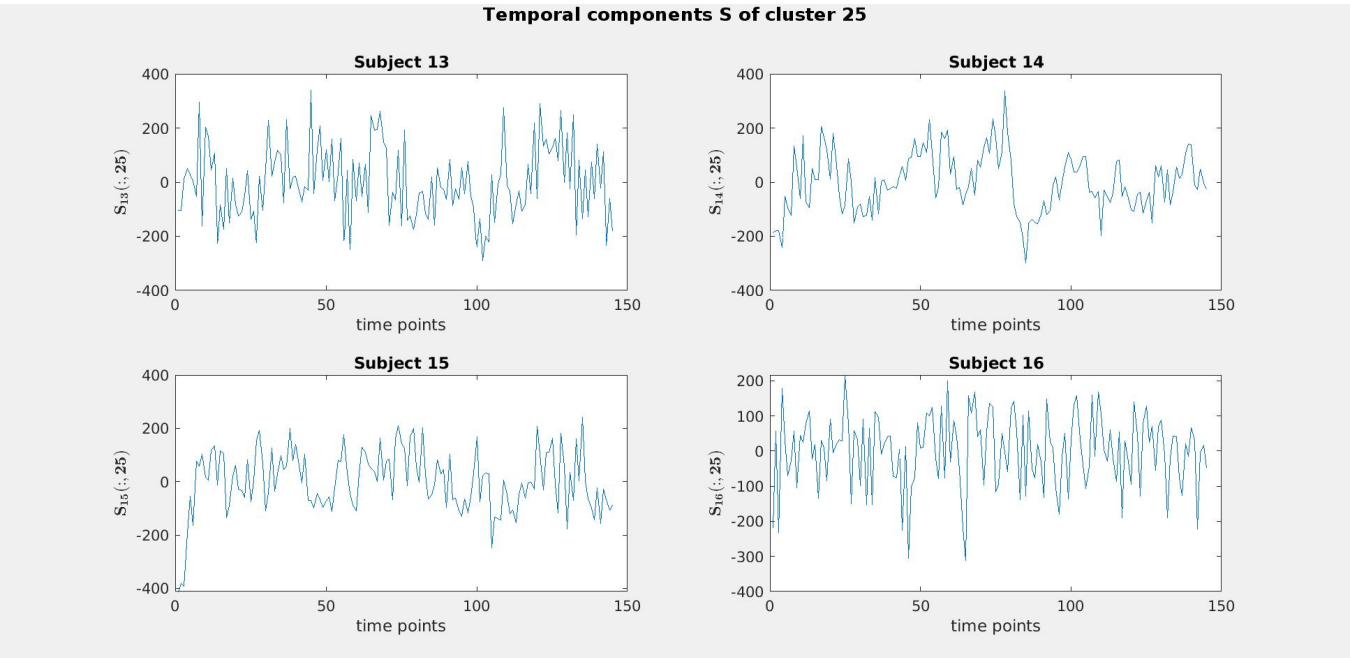


Figure B.229: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 13, \ldots, 16$ and R=27.

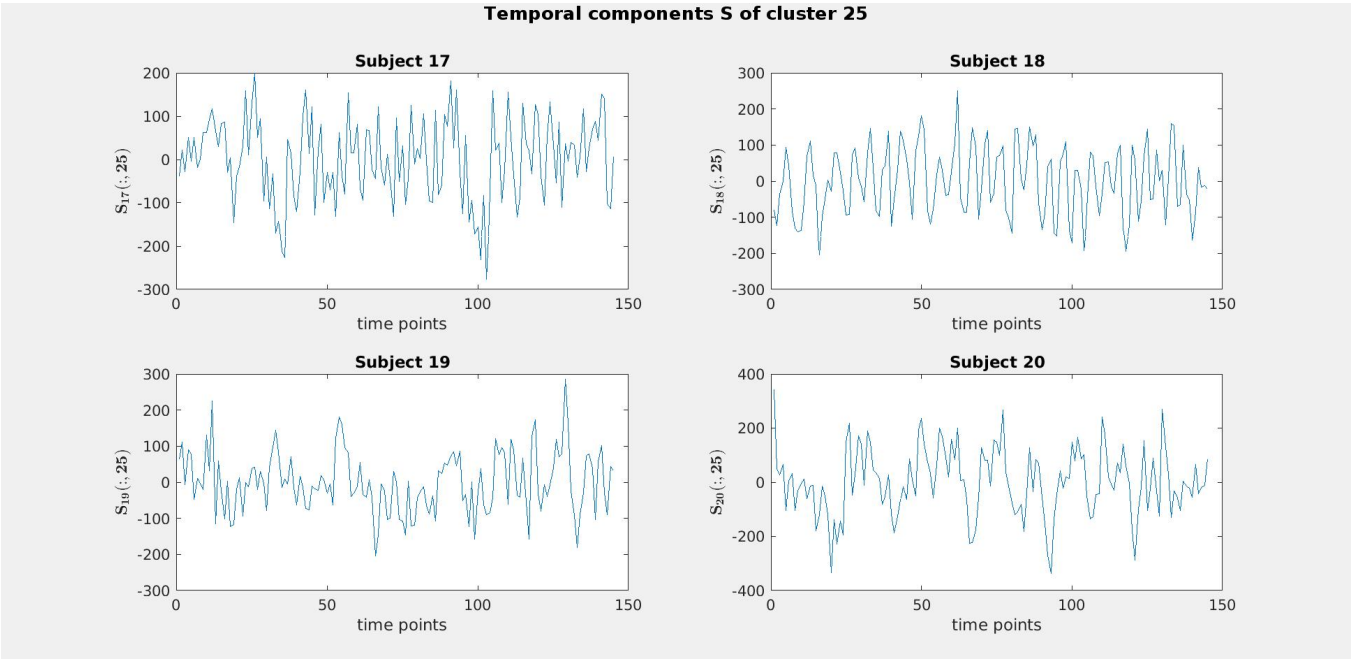


Figure B.230: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 17, \ldots, 20$ and R=27.

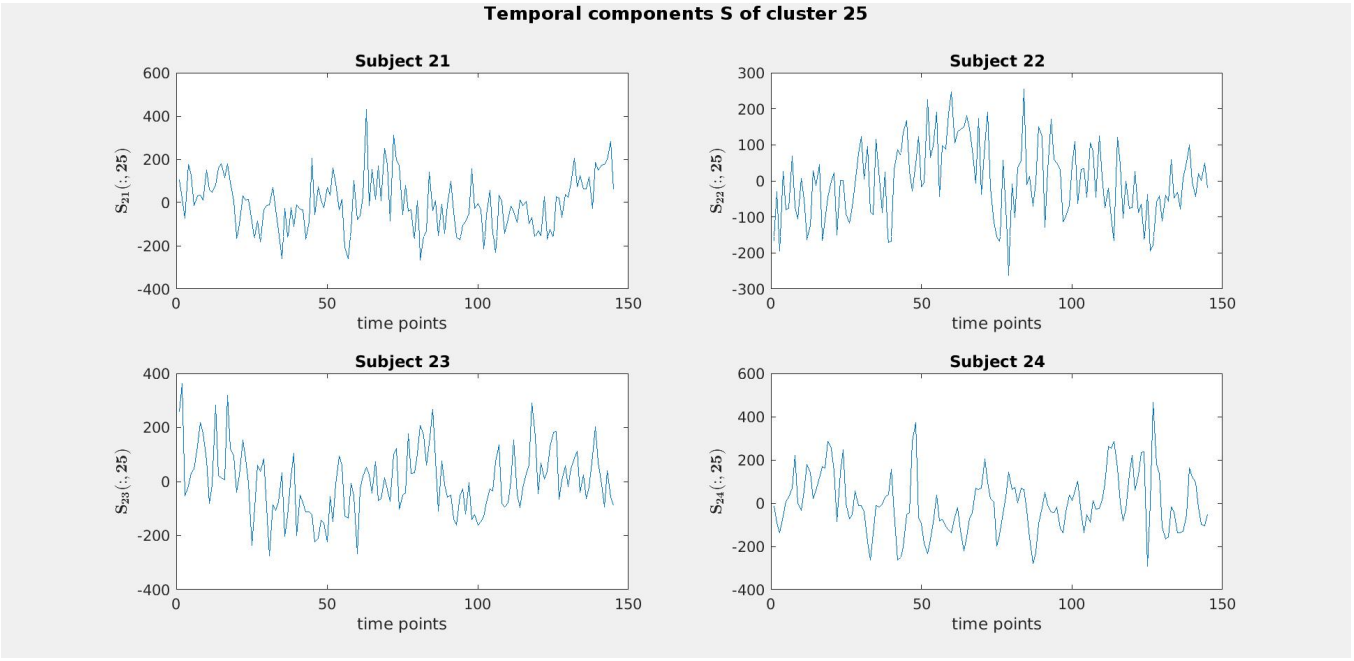


Figure B.231: Waveform of temporal components $\mathbf{S}_k(:, 125)$ for $k = 21, \ldots, 24$ and R=27.

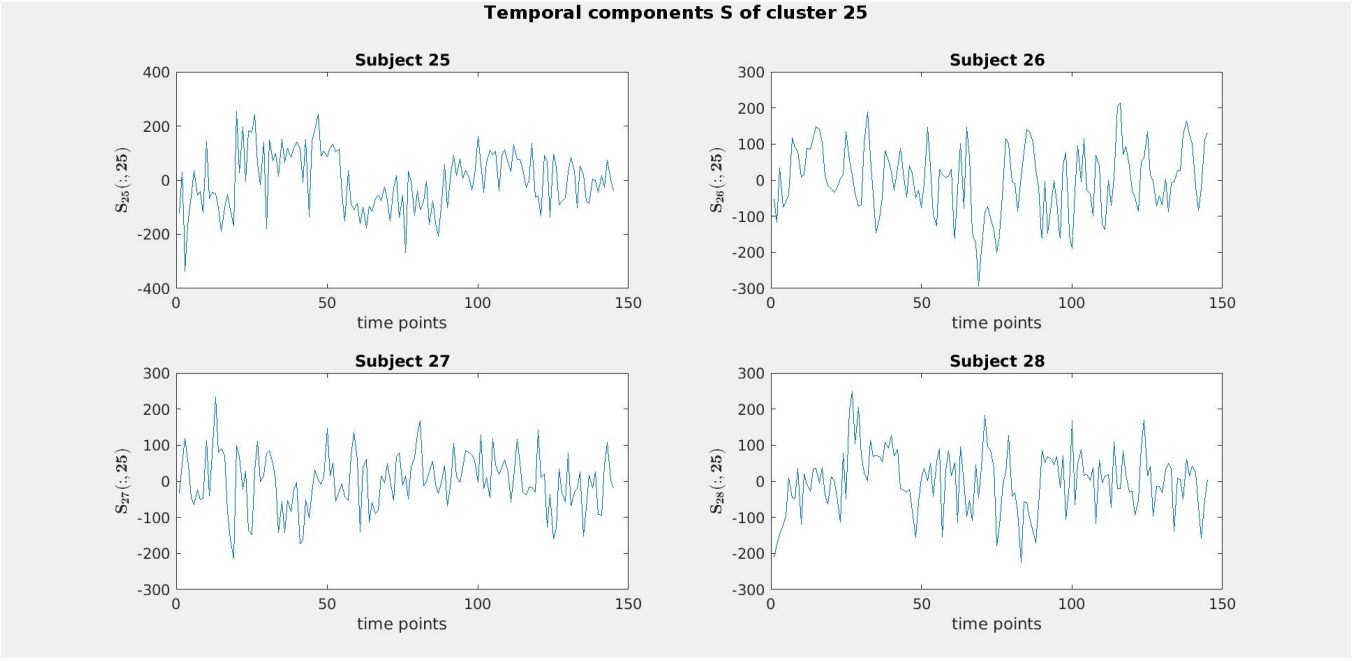


Figure B.232: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 25, \ldots, 28$ and R=27.

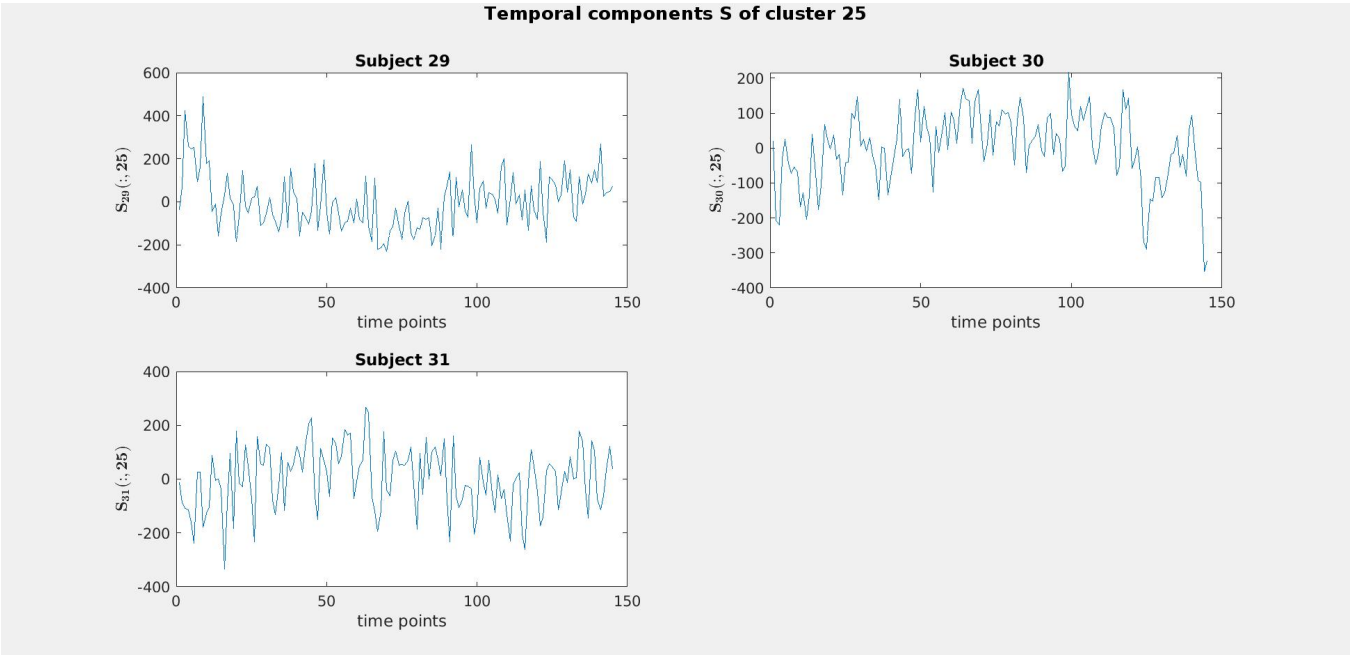


Figure B.233: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 29, \ldots, 31$ and R=27.

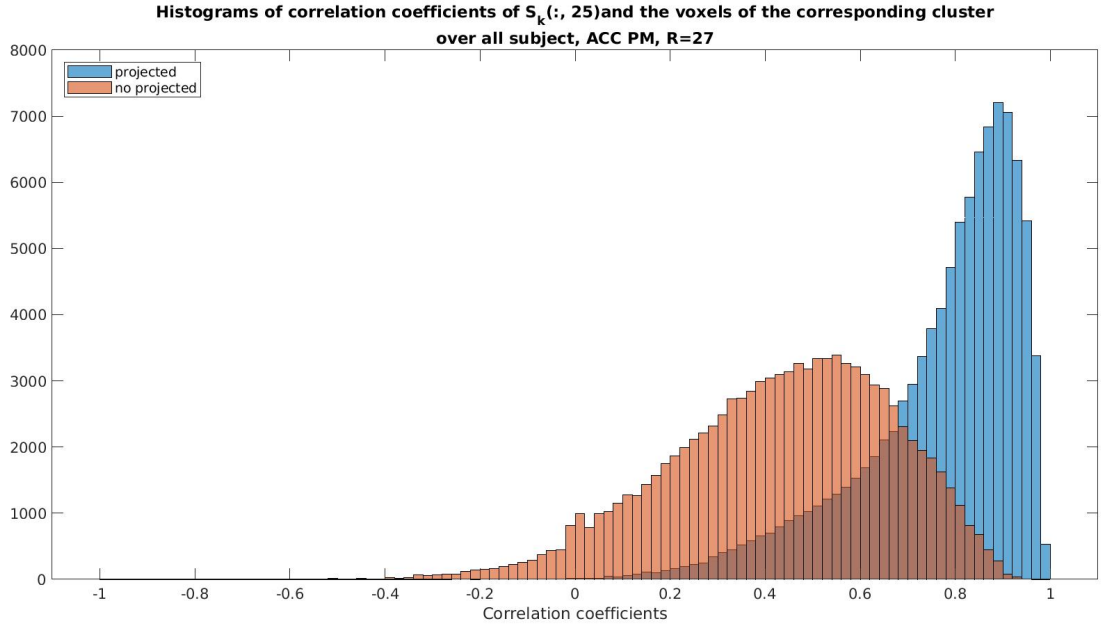


Figure B.234: For the column 25 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 25)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

22. **Cluster 1**

Cluster 1 contains 3927 voxels, which correspond to 1.738011% of the whole brain (background, soft tissue, bone are excluded).

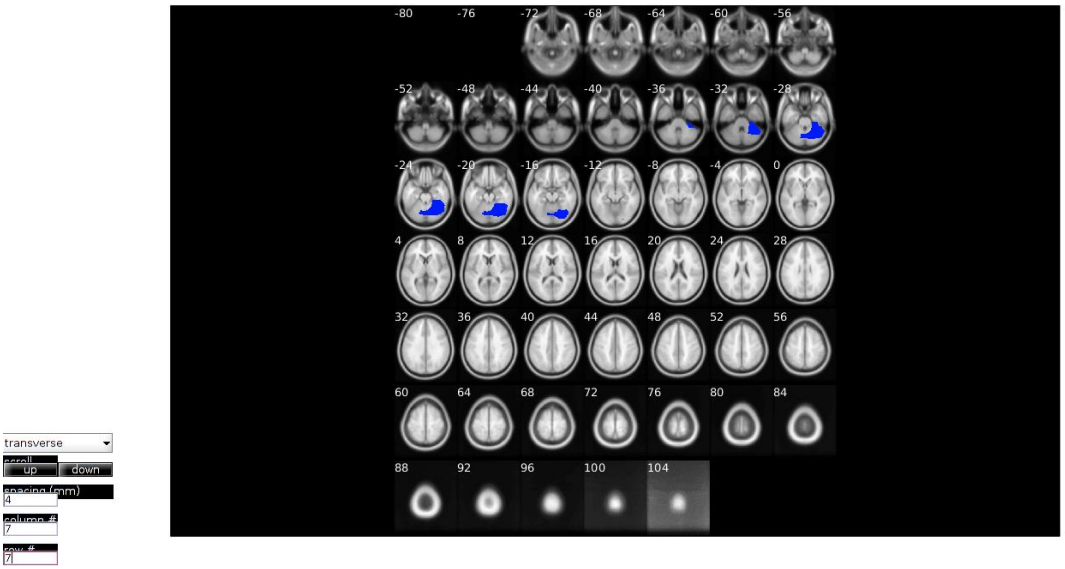


Figure B.235: Cluster 1. Blue areas indicate the voxels which belong to cluster 1.

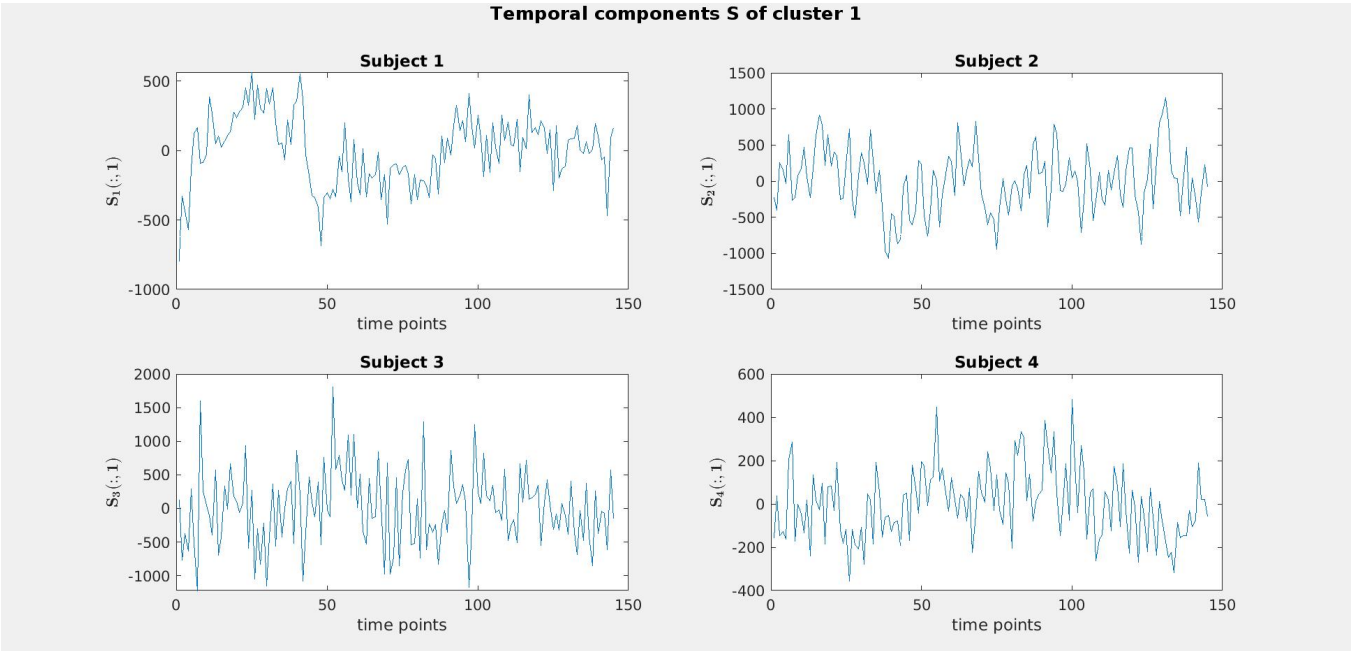


Figure B.236: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 1, \ldots, 4$ and R=27.

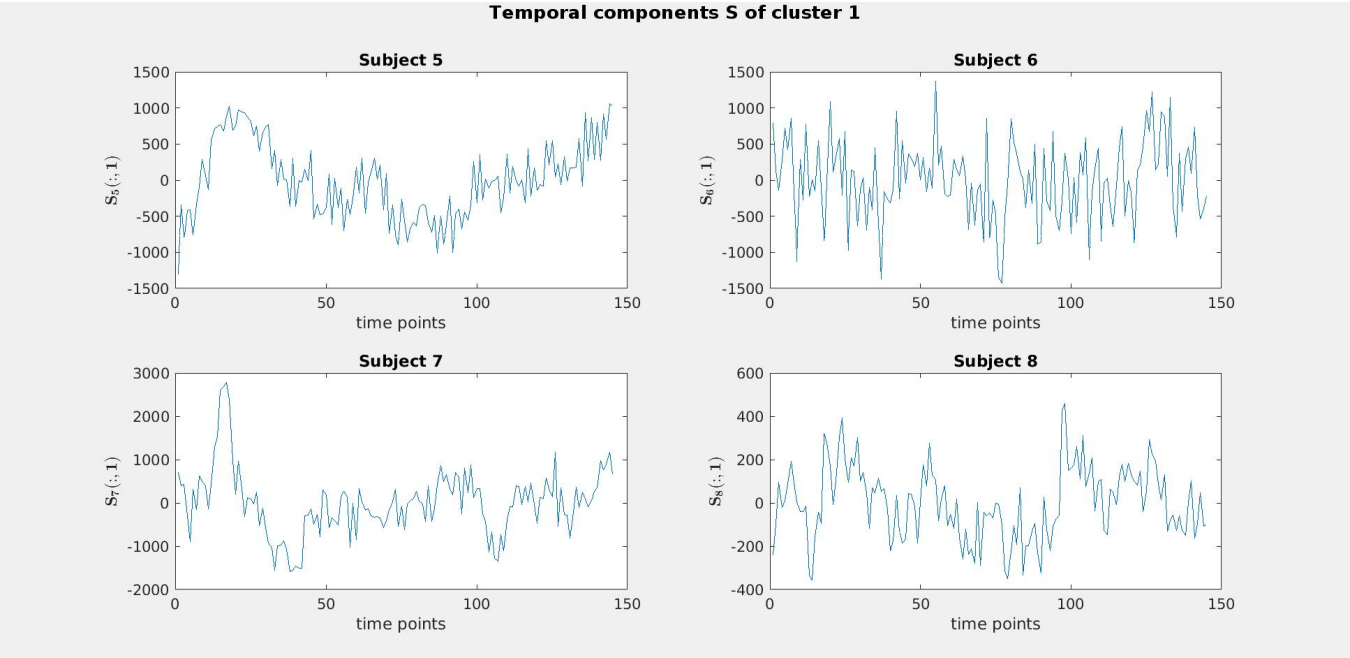


Figure B.237: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 5, \ldots, 8$ and R=27.

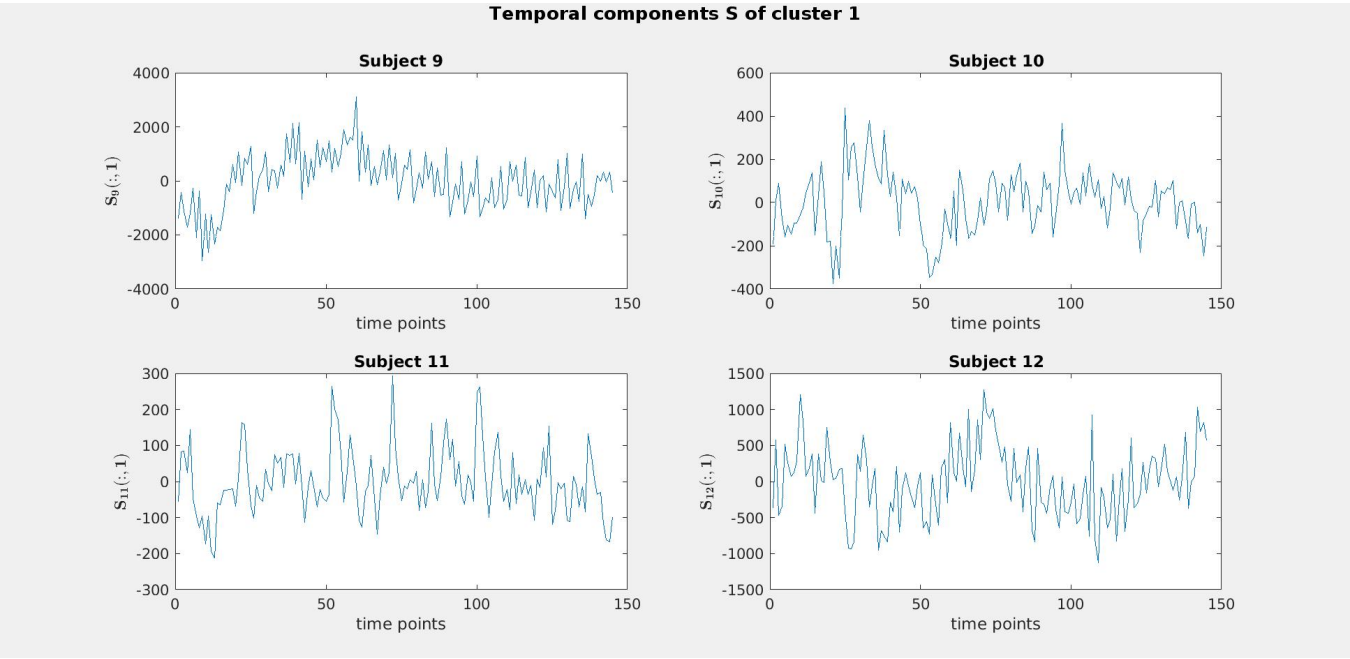


Figure B.238: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 9, \ldots, 12$ and R=27.

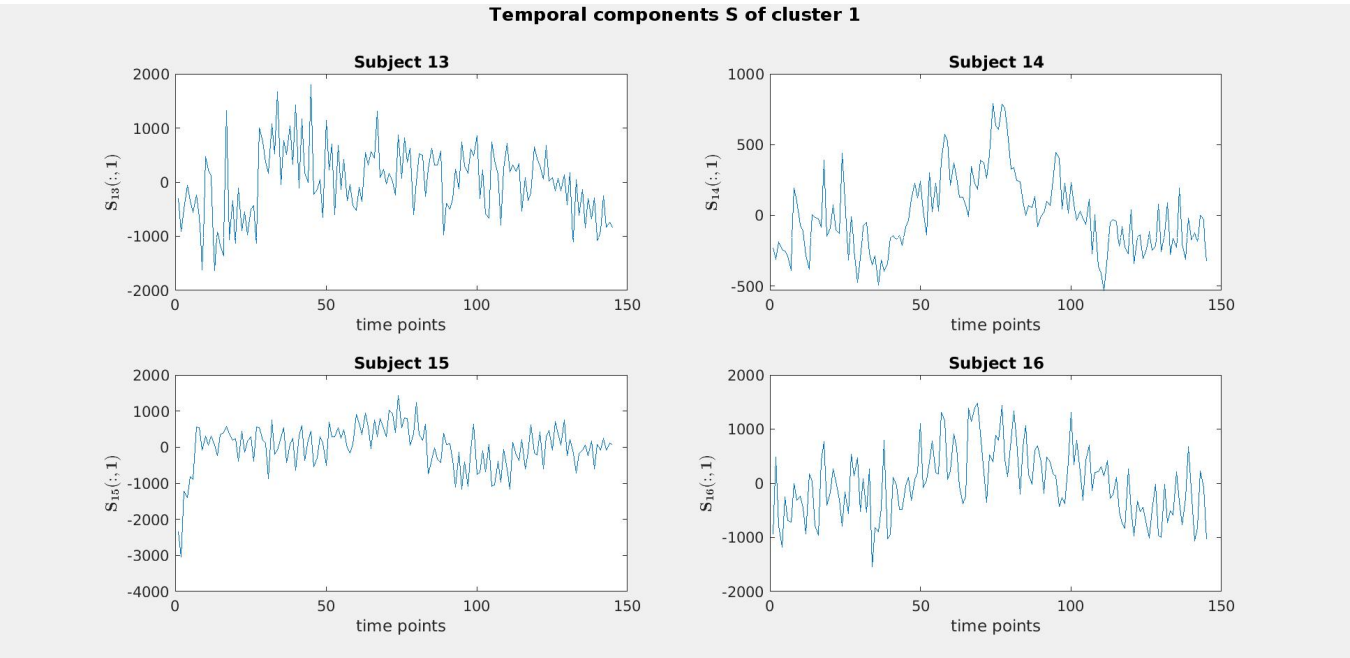


Figure B.239: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 13, \ldots, 16$ and R=27.

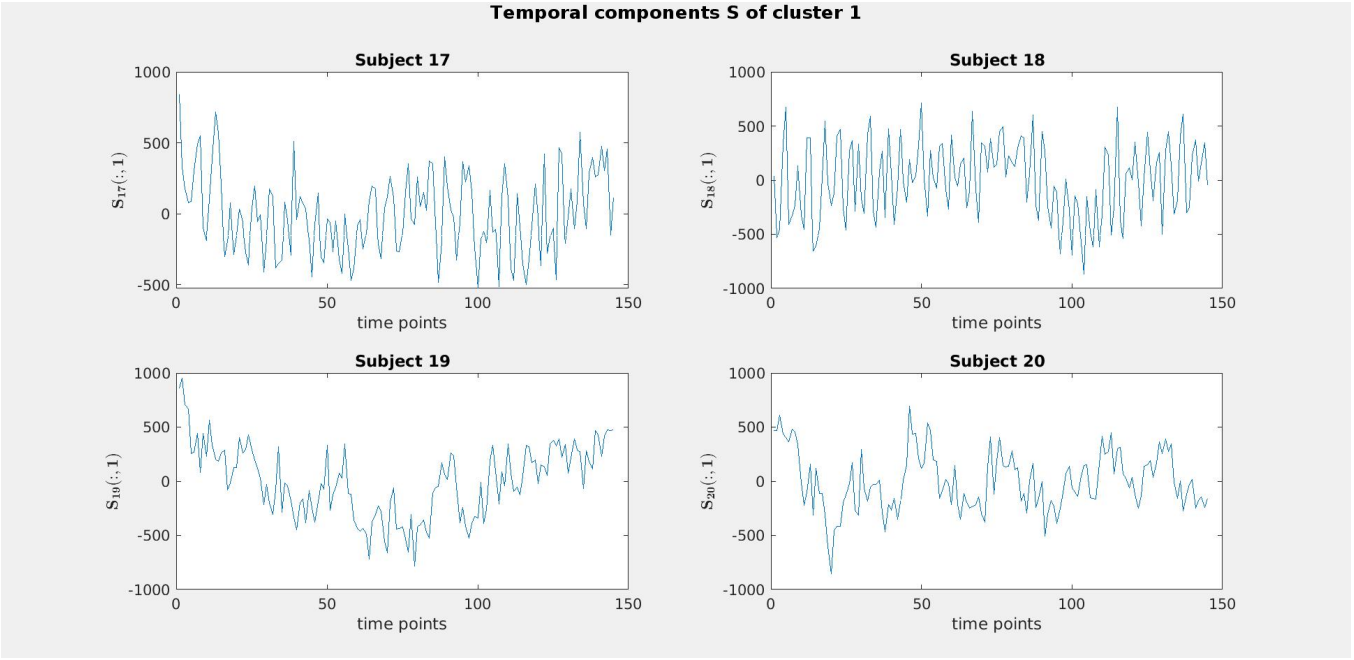


Figure B.240: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 17,\ldots,20$ and R=27.

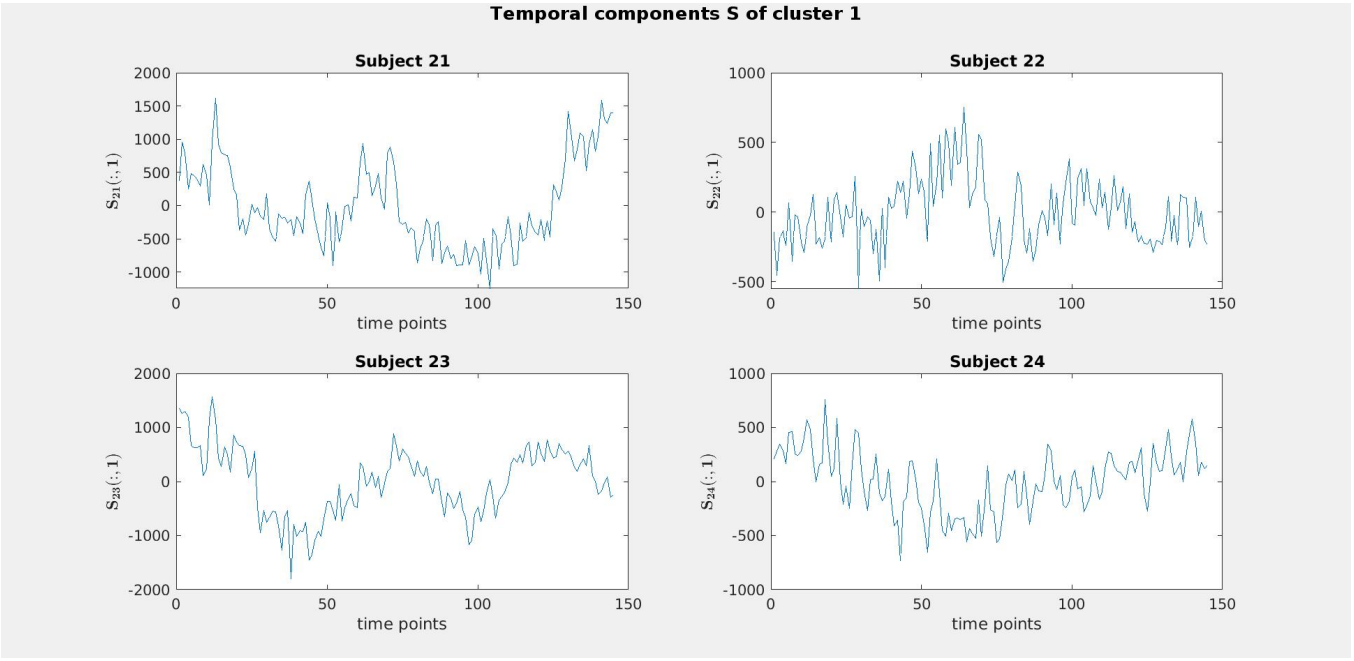


Figure B.241: Waveform of temporal components $\mathbf{S}_k(:,11)$ for $k = 21,\ldots,24$ and R=27.

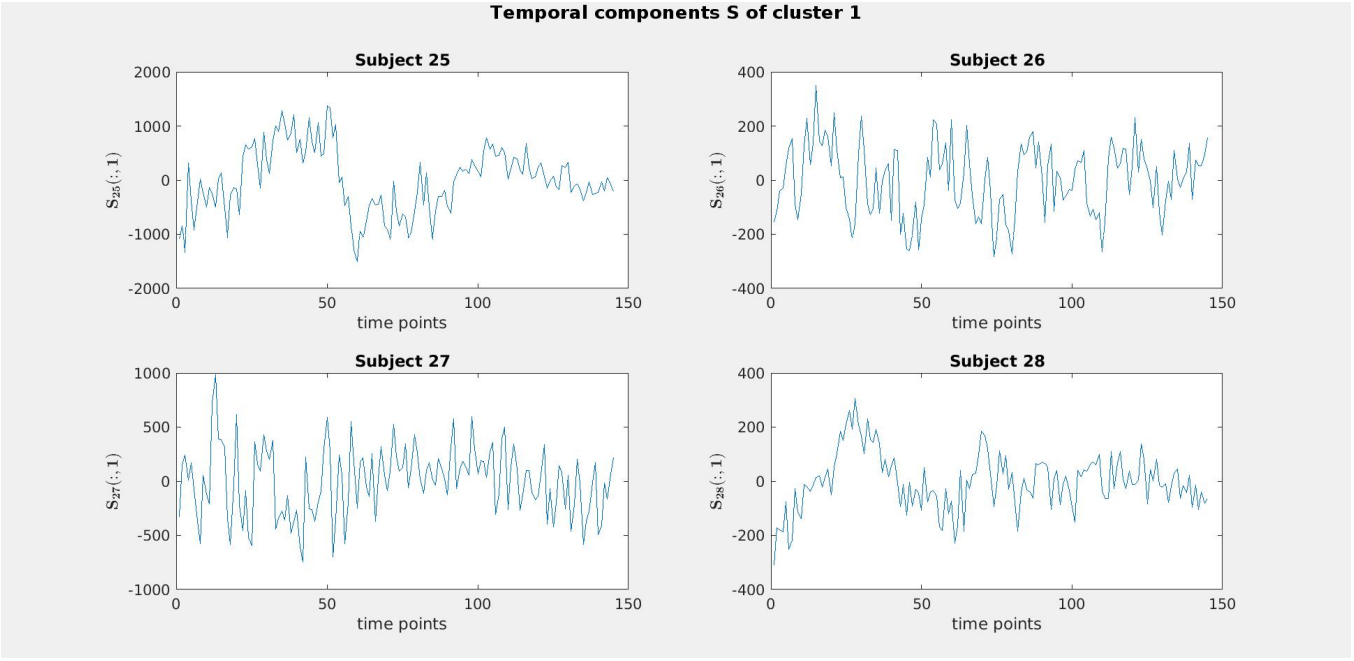


Figure B.242: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 25,\ldots,28$ and R=27.

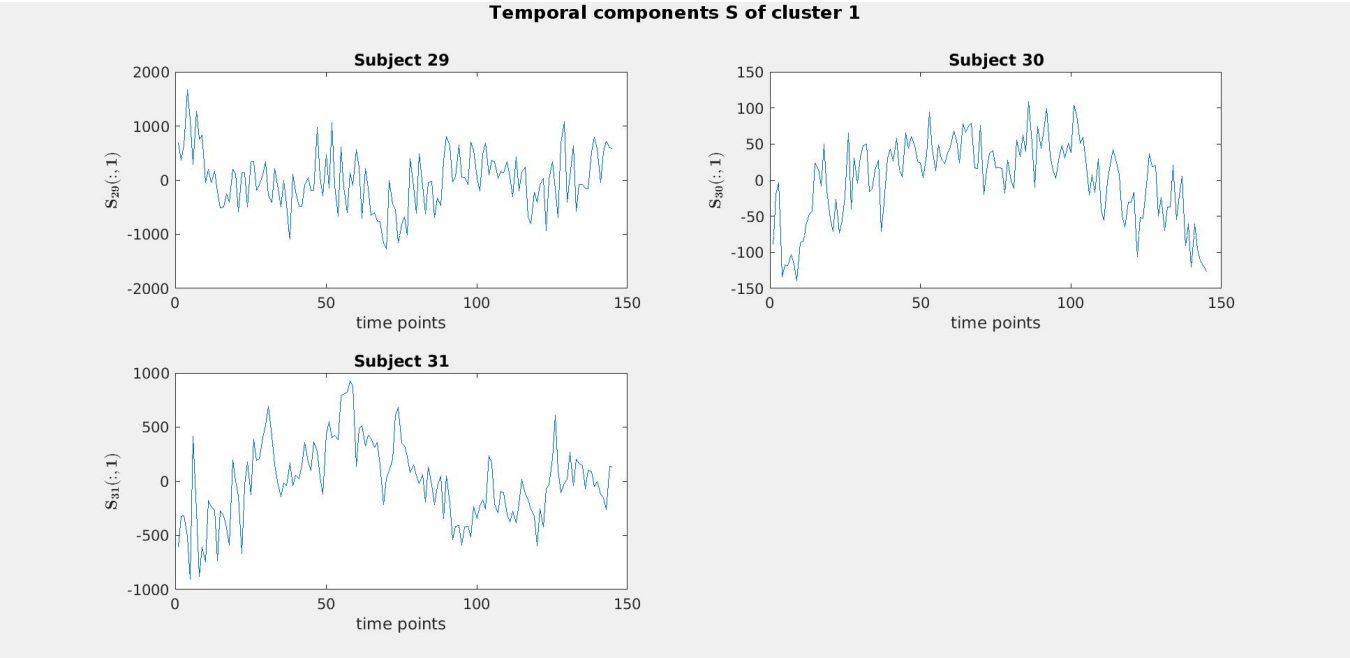


Figure B.243: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 29, \ldots, 31$ and R=27.

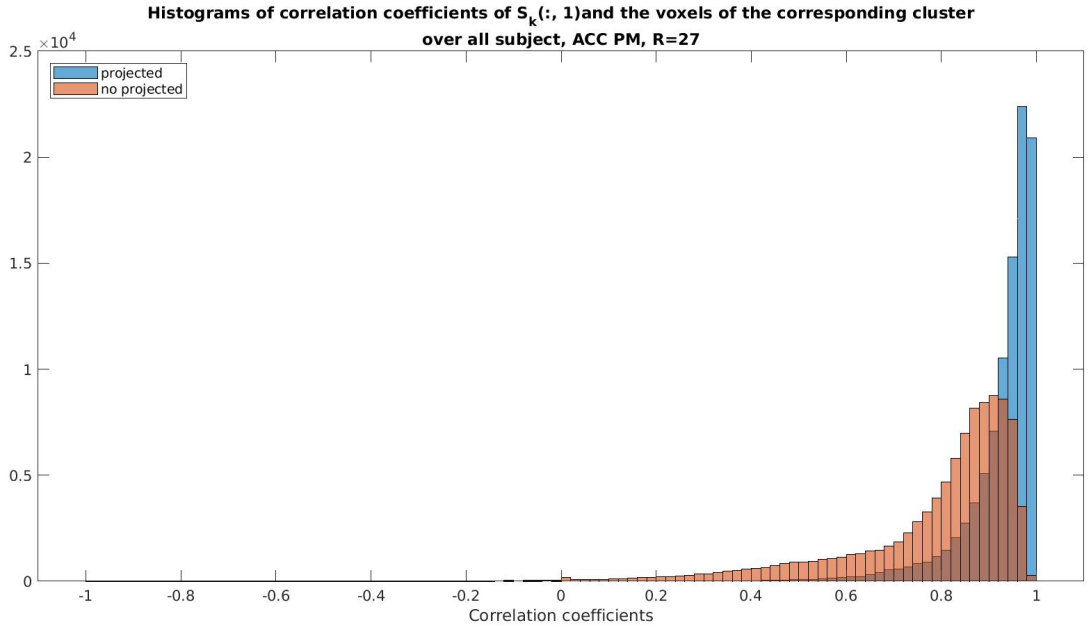


Figure B.244: For the column 1 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,1)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

23. **Cluster 15**

Cluster 15 contains 3595 voxels, which correspond to 1.591074% of the whole brain (background, soft tissue, bone are excluded).

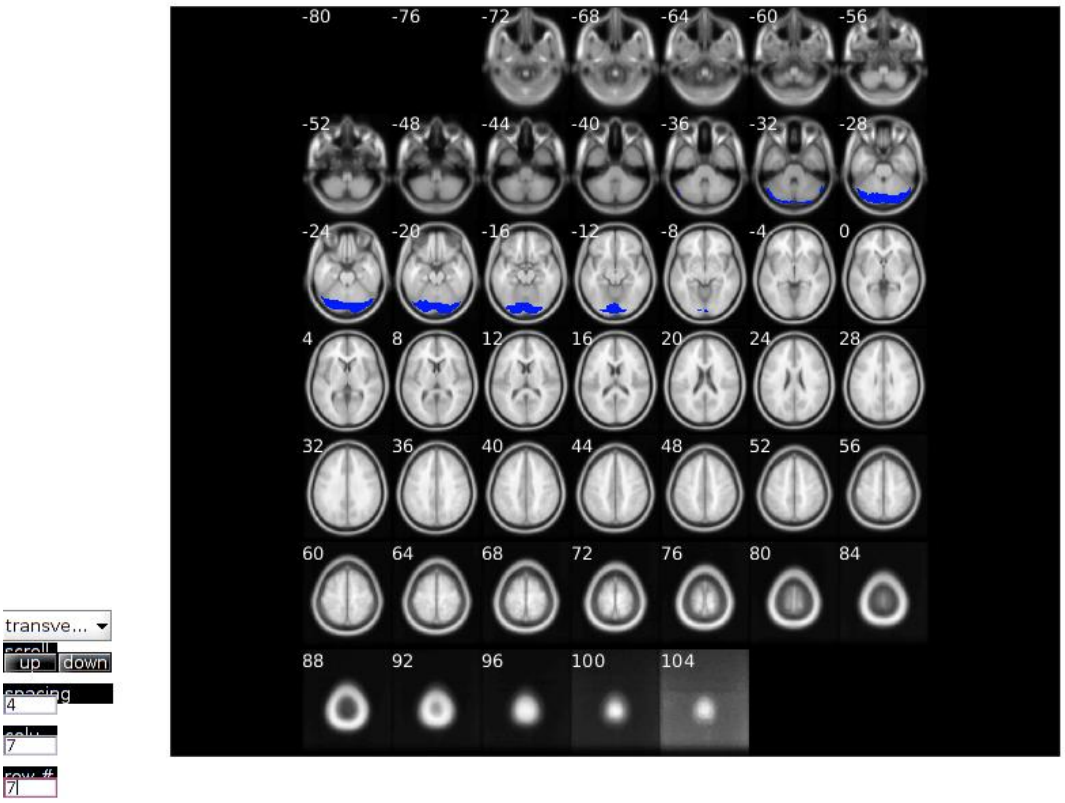


Figure B.245: Cluster 15. Blue areas indicate the voxels which belong to cluster 15.

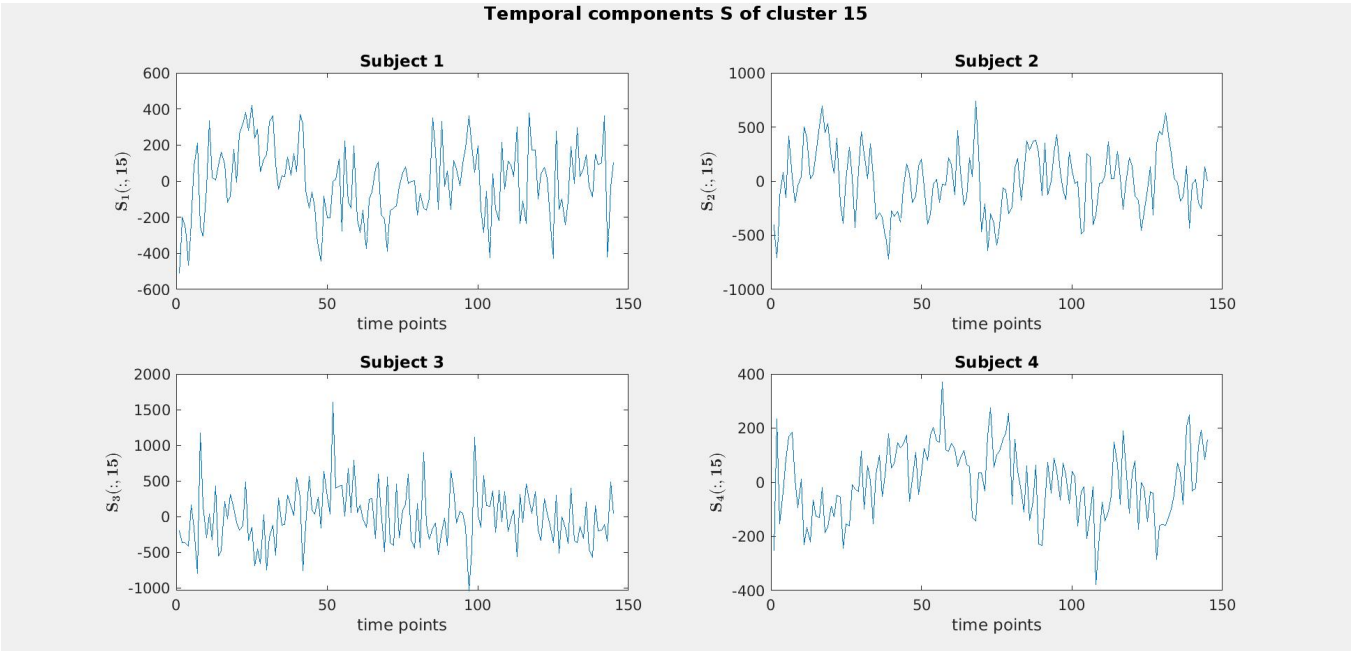


Figure B.246: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 1, \ldots, 4$ and R=27.

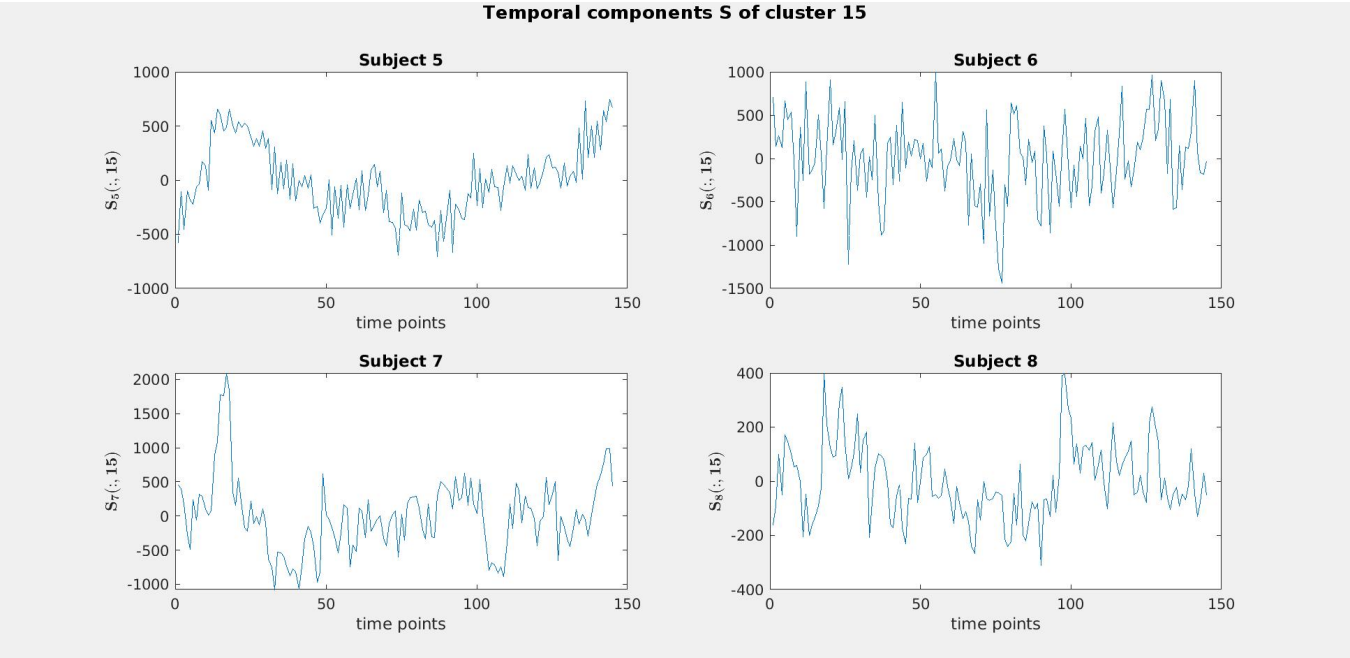


Figure B.247: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 5, \ldots, 8$ and R=27.

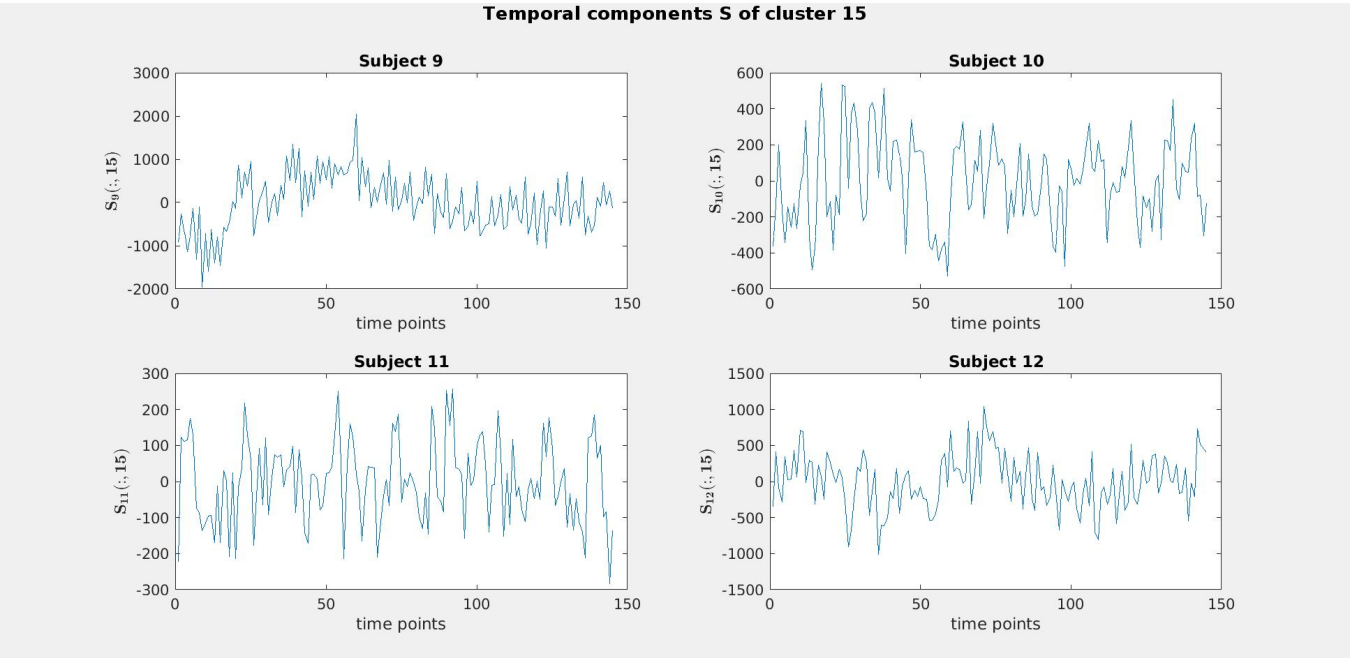


Figure B.248: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 9, \ldots, 12$ and R=27.

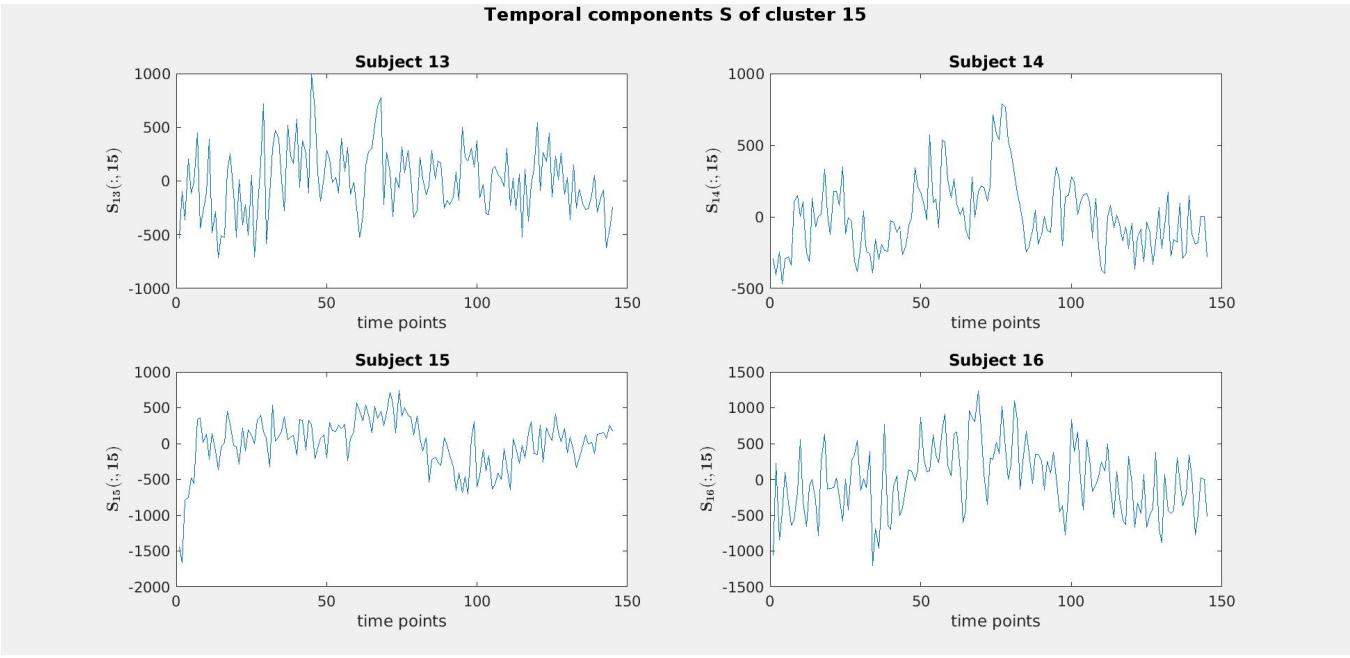


Figure B.249: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 13, \ldots, 16$ and R=27.

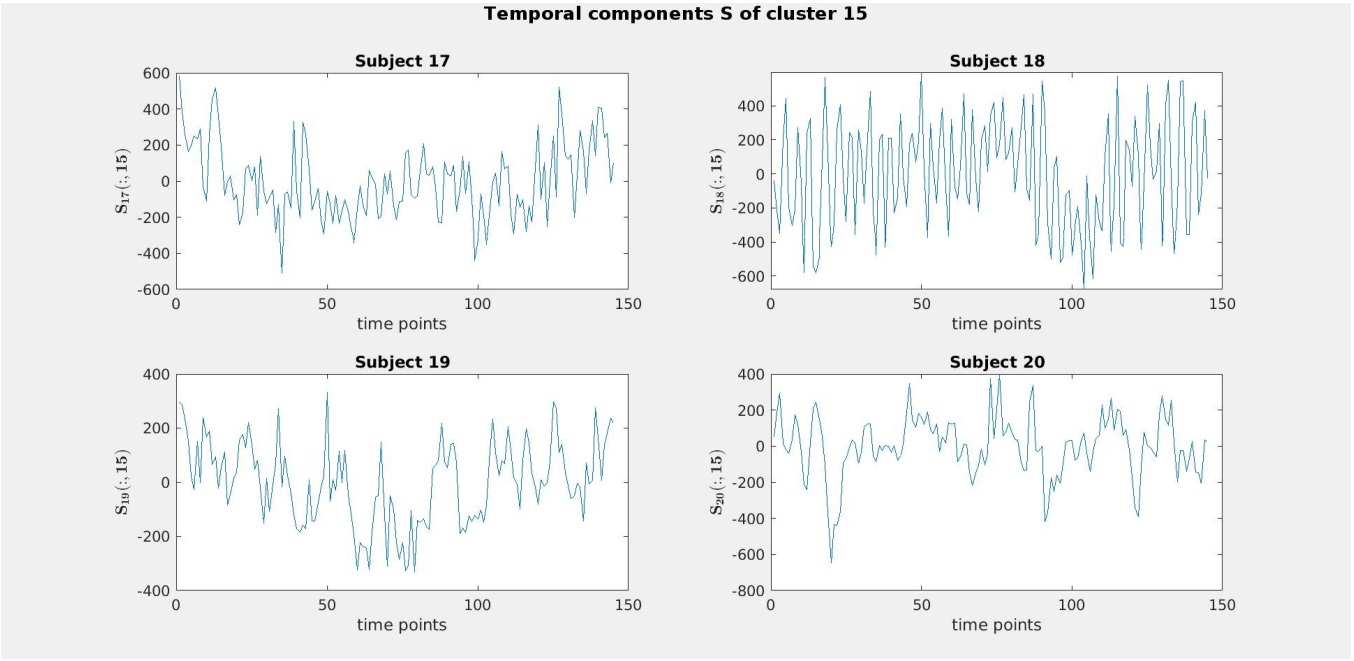


Figure B.250: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 17, \ldots, 20$ and R=27.

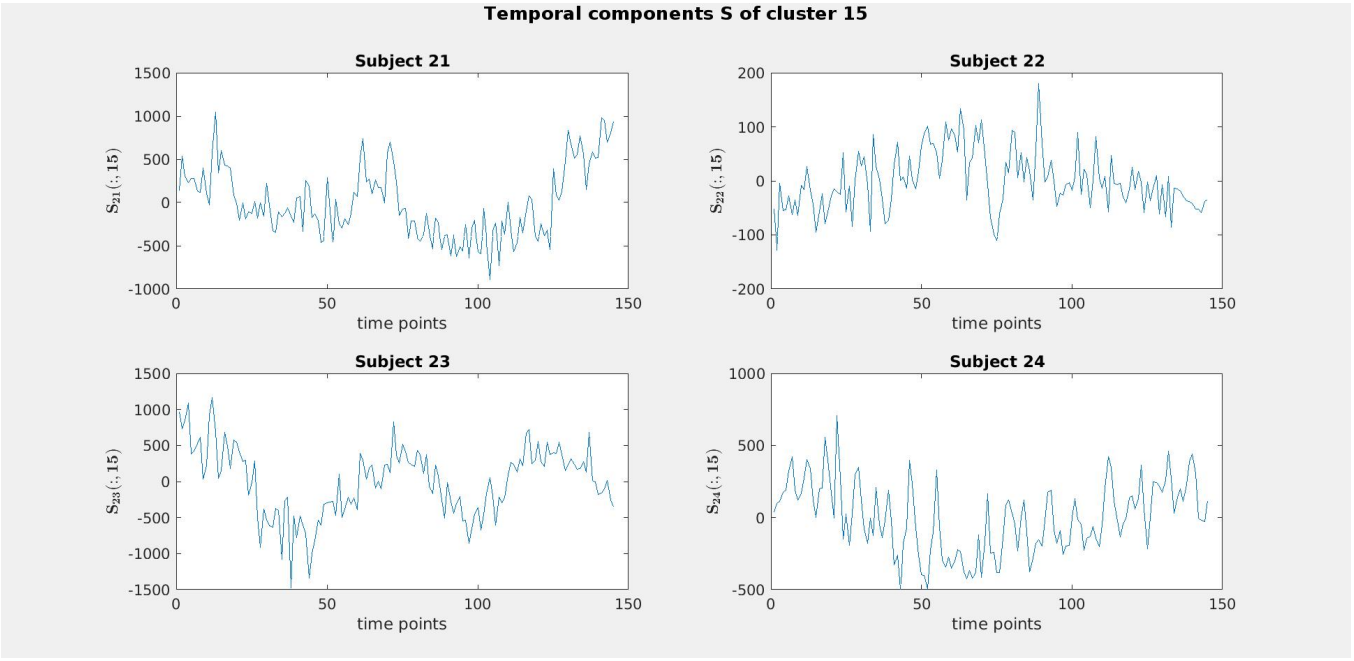


Figure B.251: Waveform of temporal components $\mathbf{S}_k(:, 115)$ for $k = 21, \ldots, 24$ and R=27.

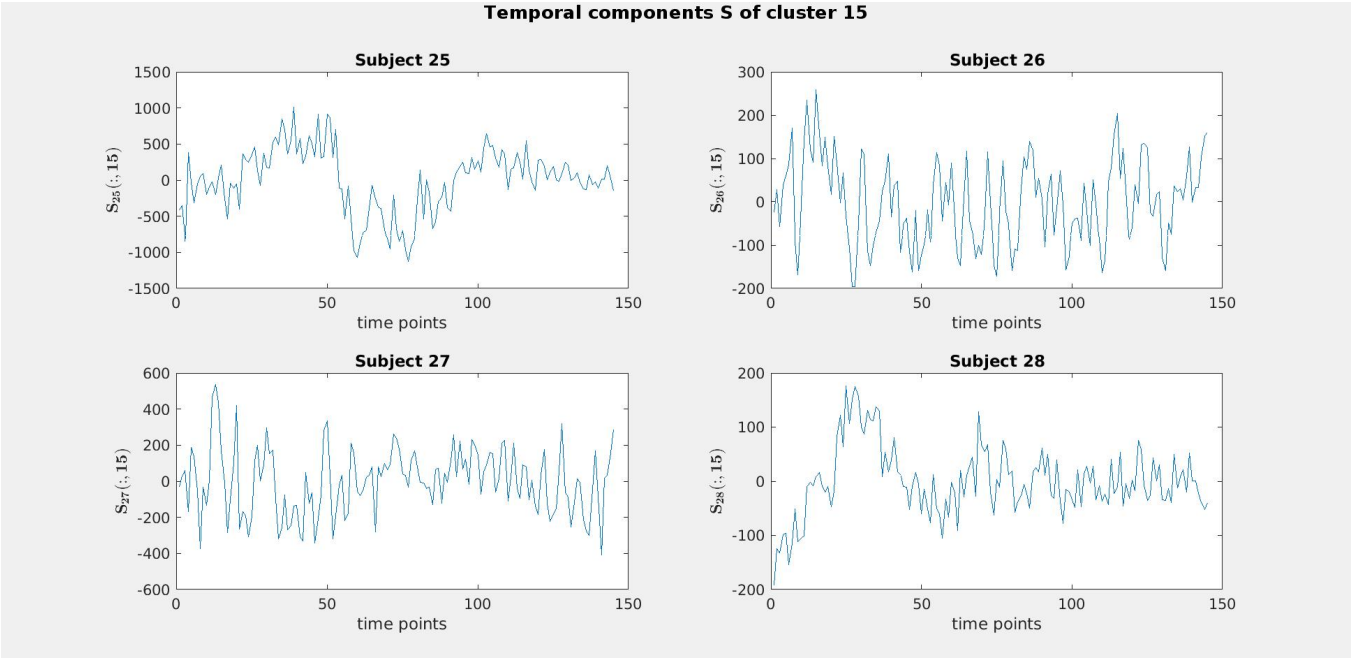


Figure B.252: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 25, \ldots, 28$ and R=27.

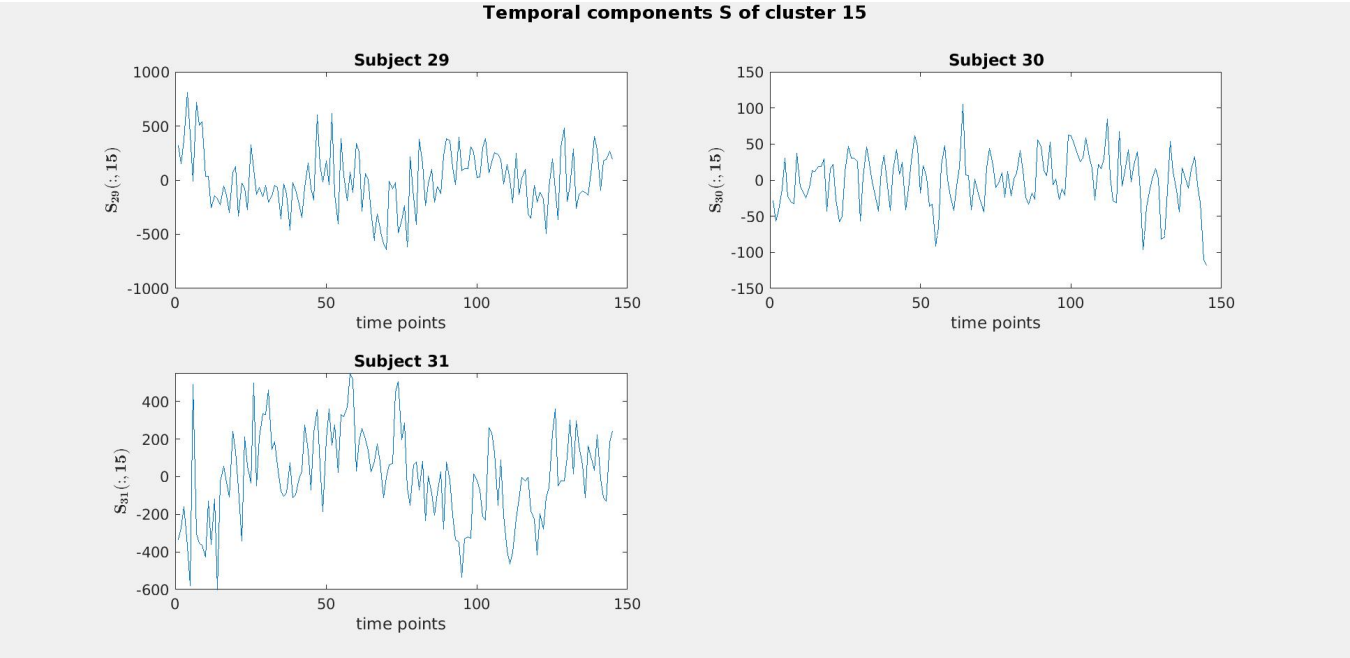


Figure B.253: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 29, \ldots, 31$ and R=27.

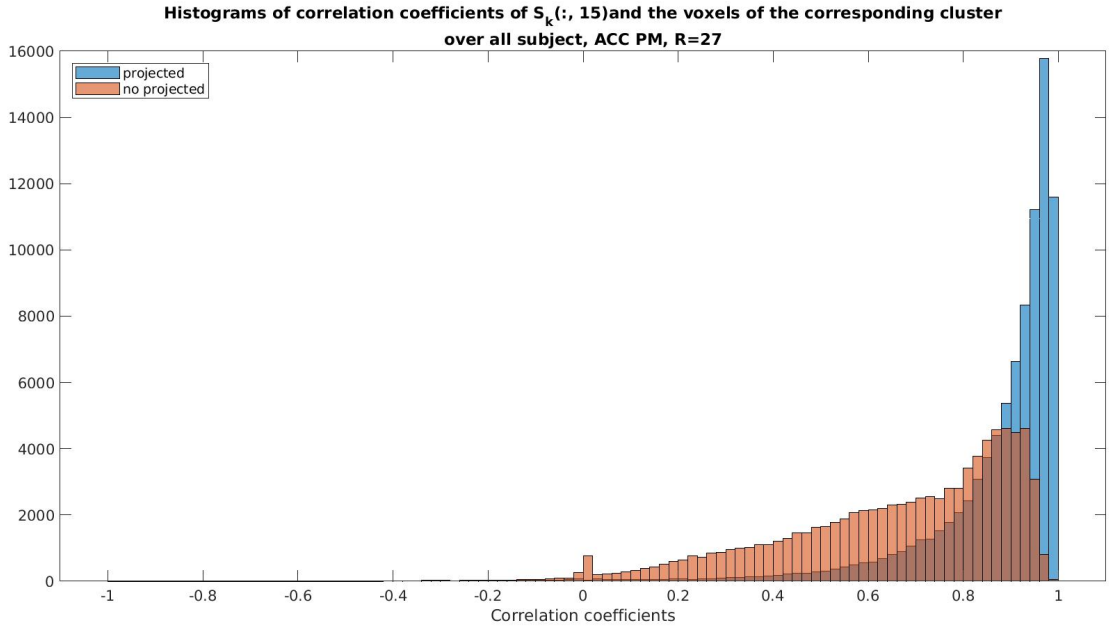


Figure B.254: For the column 15 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 15)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

24. **Cluster 26**

Cluster 26 contains 2442 voxels, which correspond to 1.080780% of the whole brain (background, soft tissue, bone are excluded).

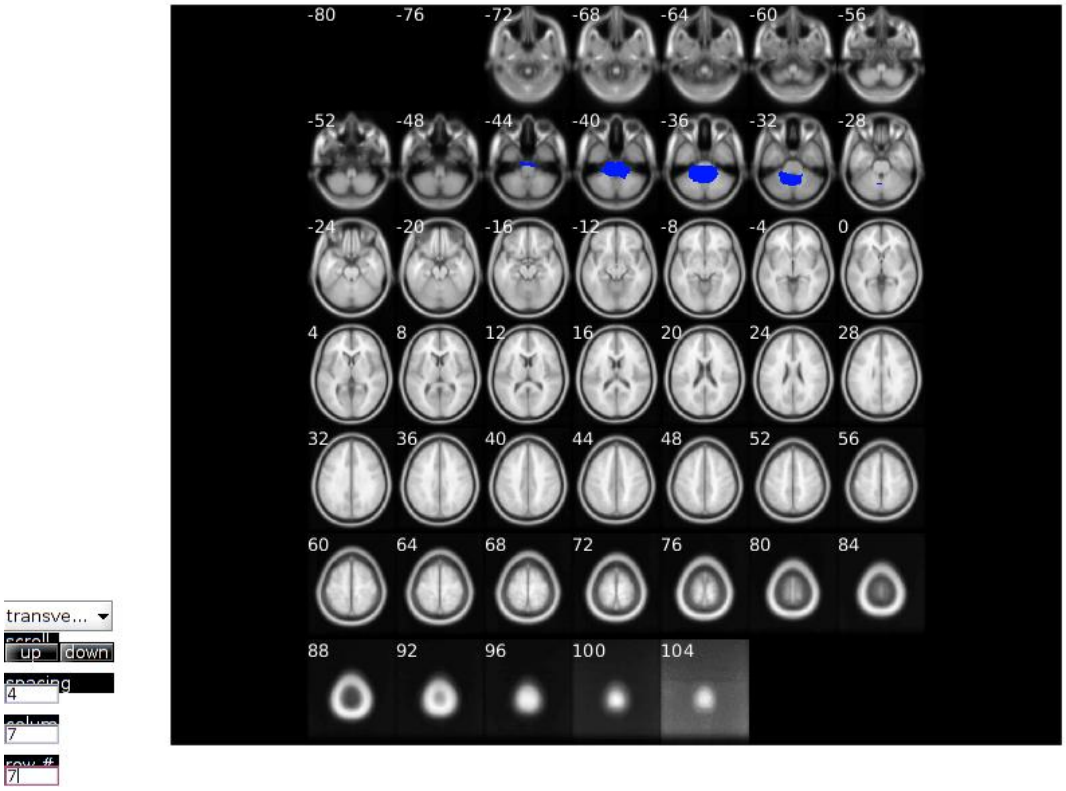


Figure B.255: Cluster 26. Blue areas indicate the voxels which belong to cluster 26.

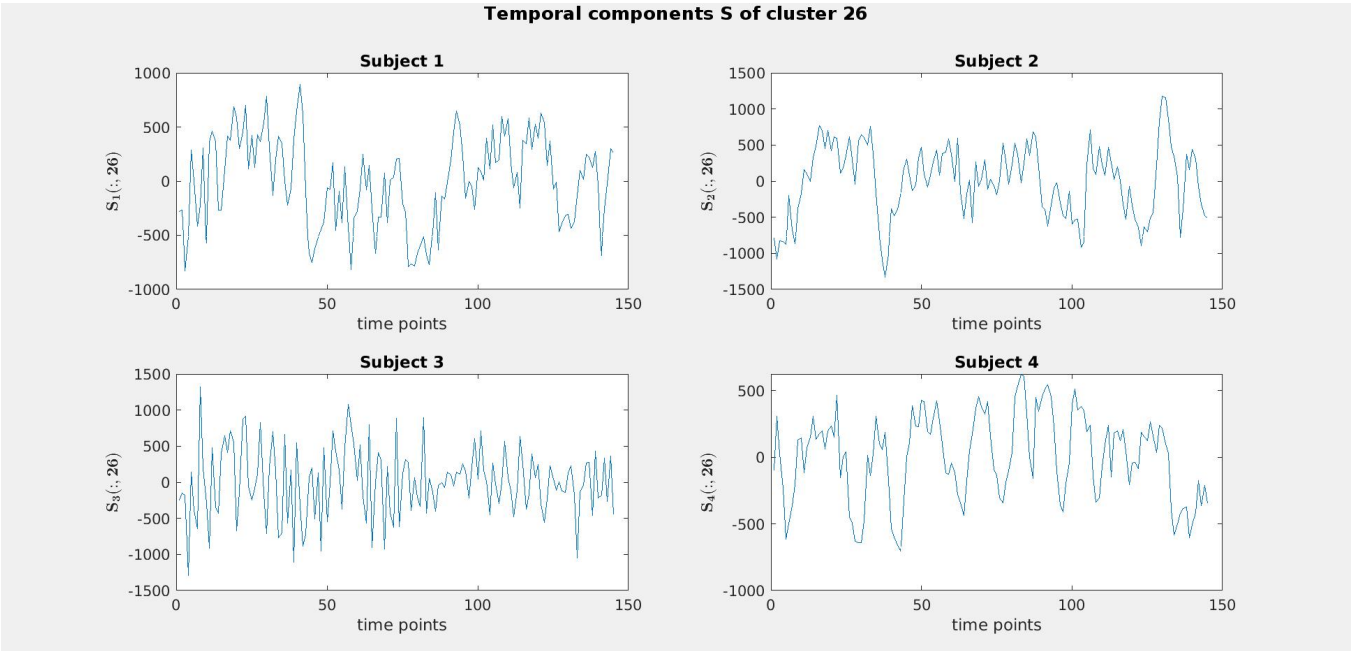


Figure B.256: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 1, \ldots, 4$ and R=27.

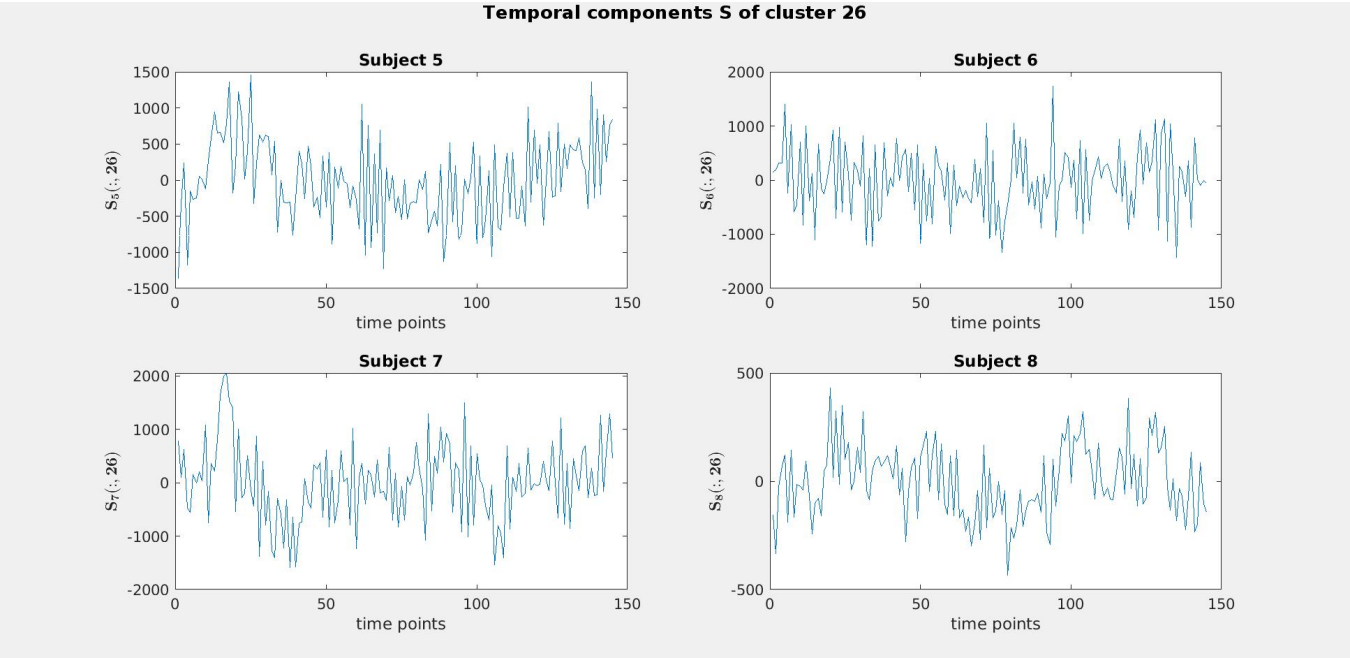


Figure B.257: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 5, \ldots, 8$ and R=27.

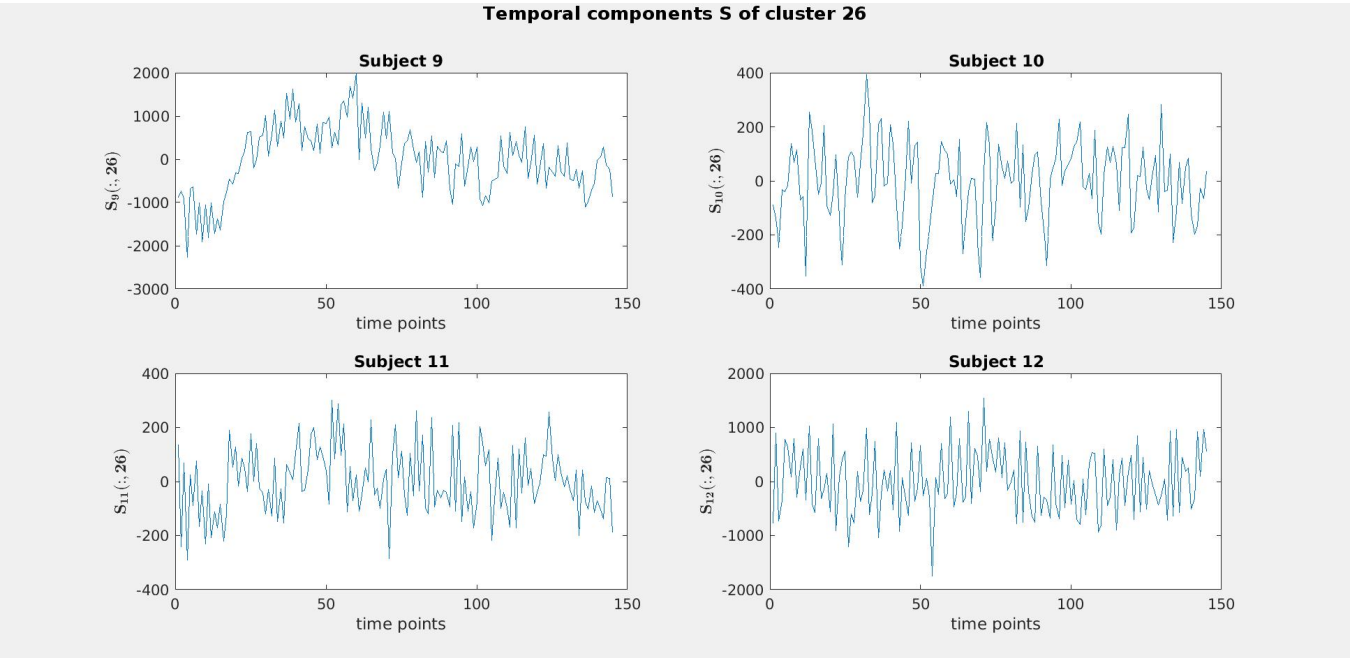


Figure B.258: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 9, \ldots, 12$ and R=27.

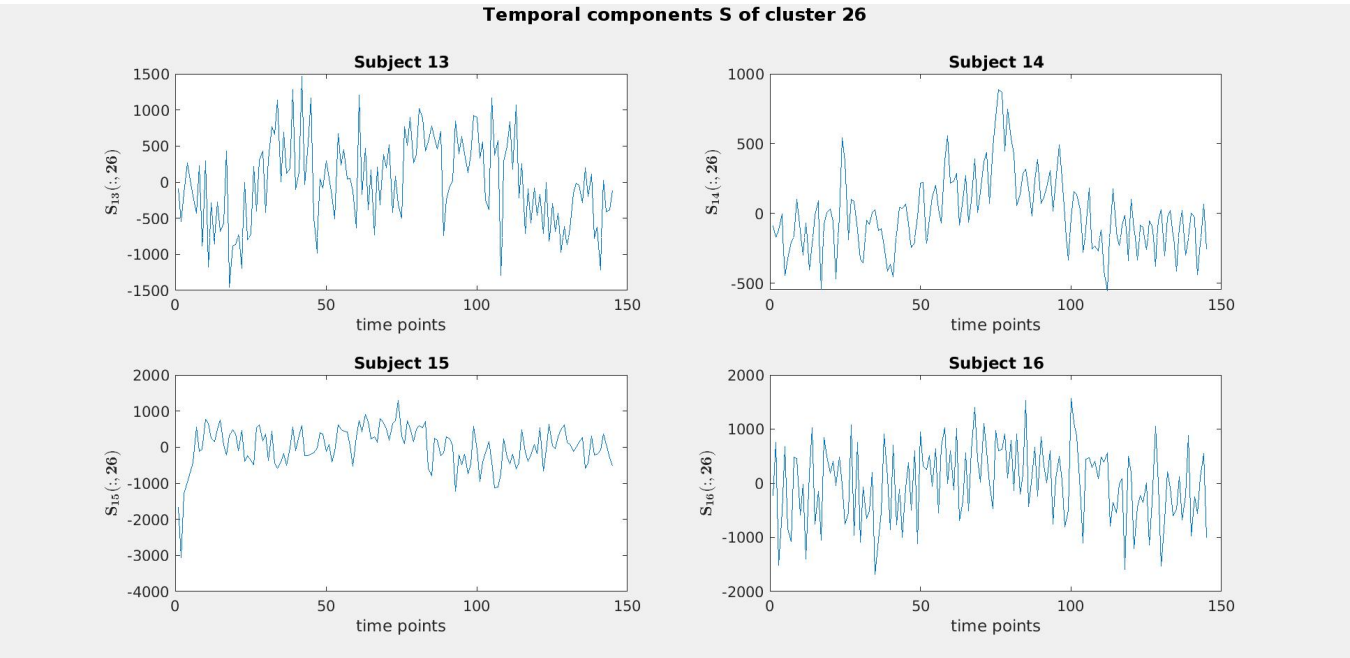


Figure B.259: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 13, \ldots, 16$ and R=27.

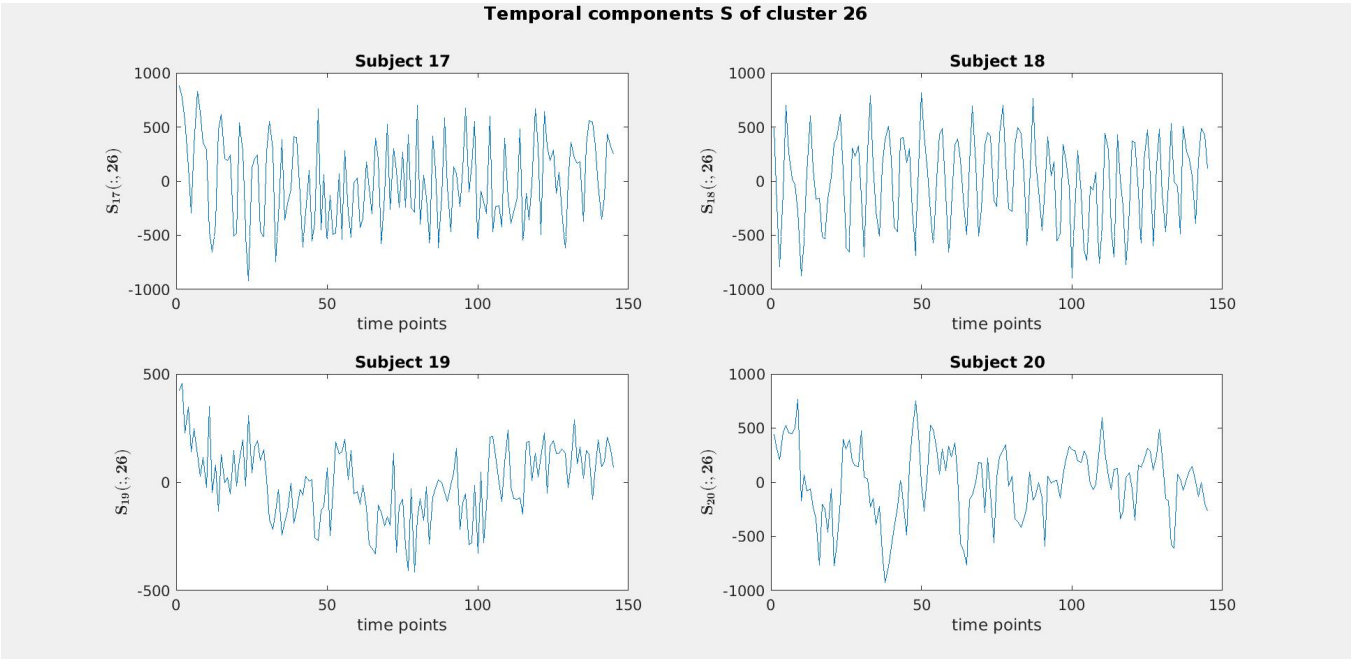


Figure B.260: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 17, \ldots, 20$ and R=27.

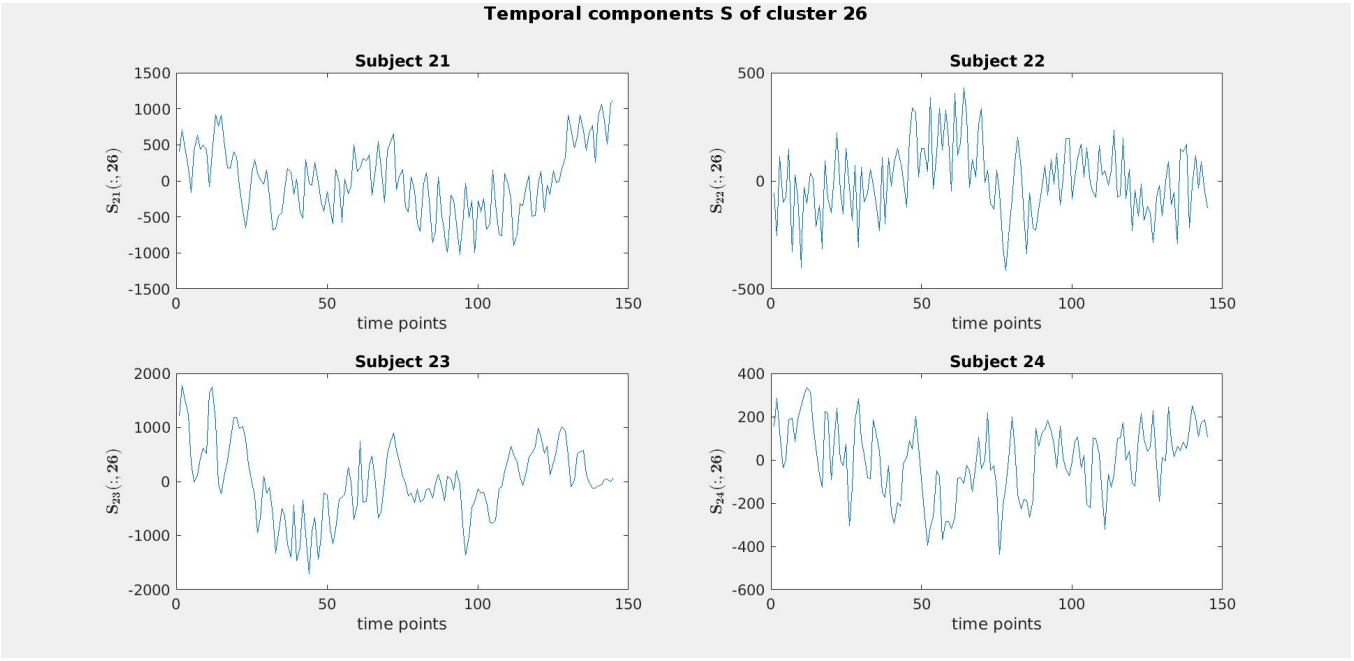


Figure B.261: Waveform of temporal components $\mathbf{S}_k(:, 126)$ for $k = 21, \ldots, 24$ and R=27.

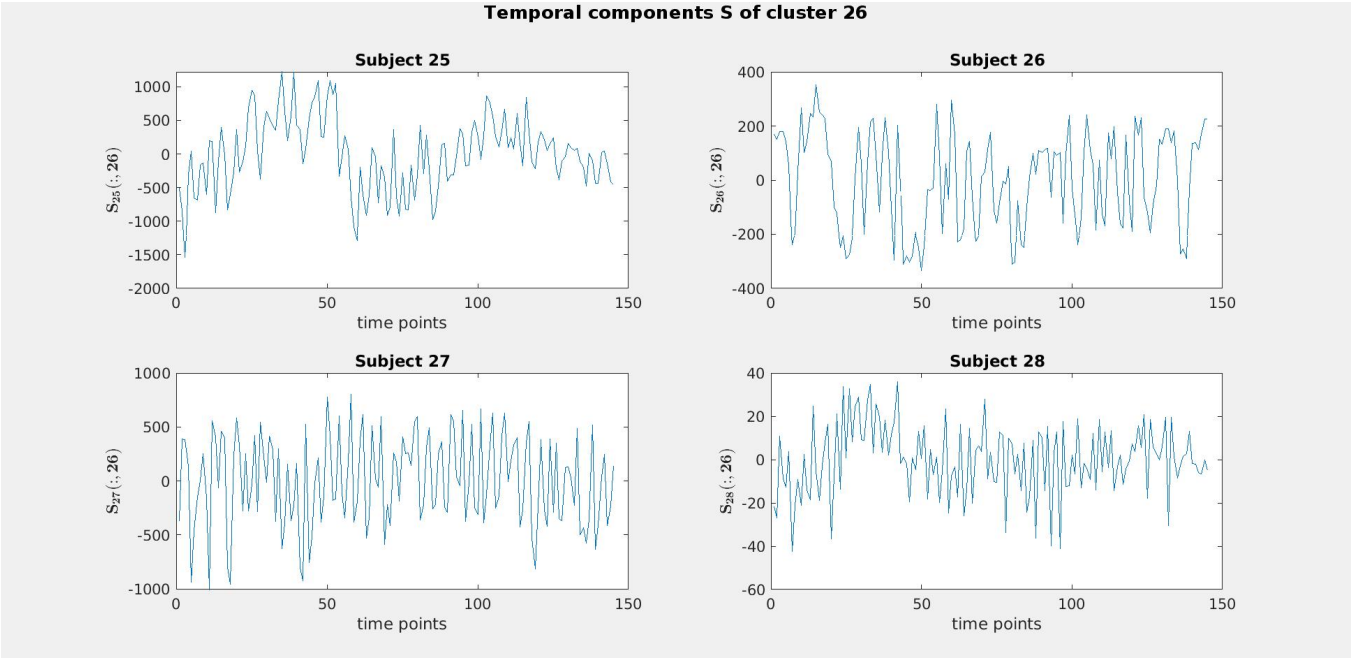


Figure B.262: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 25, \ldots, 28$ and R=27.

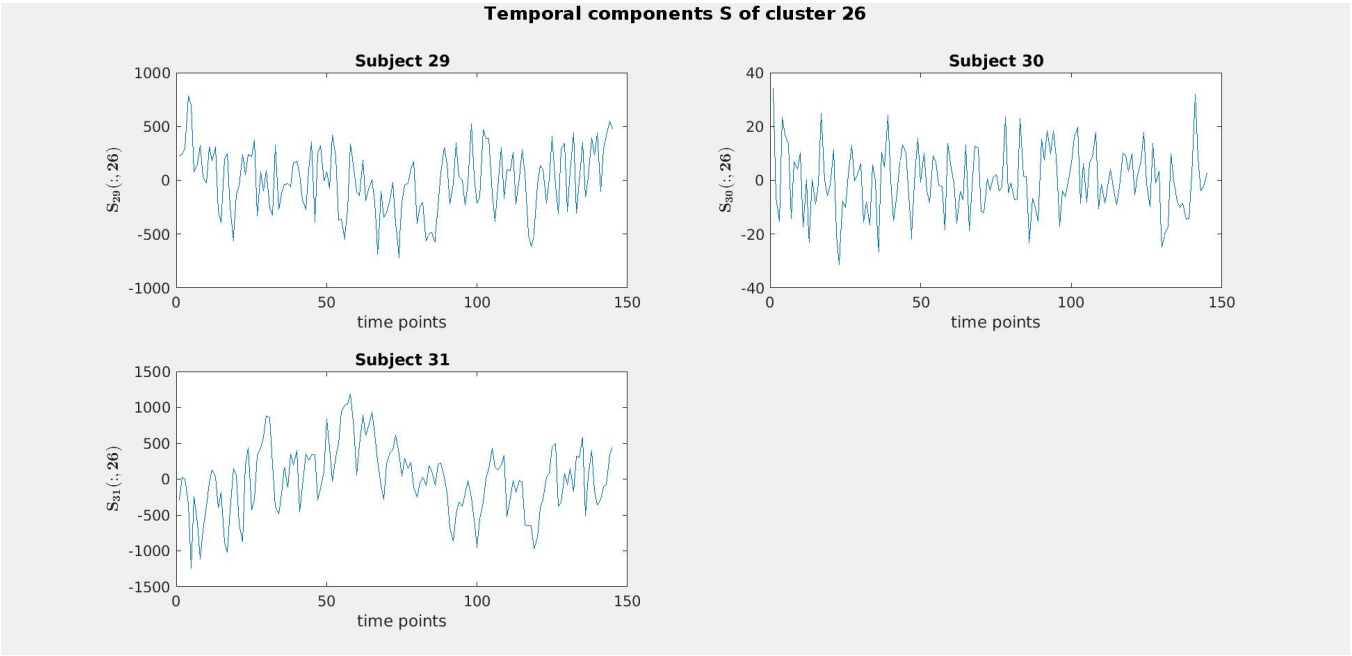


Figure B.263: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 29, \ldots, 31$ and R=27.

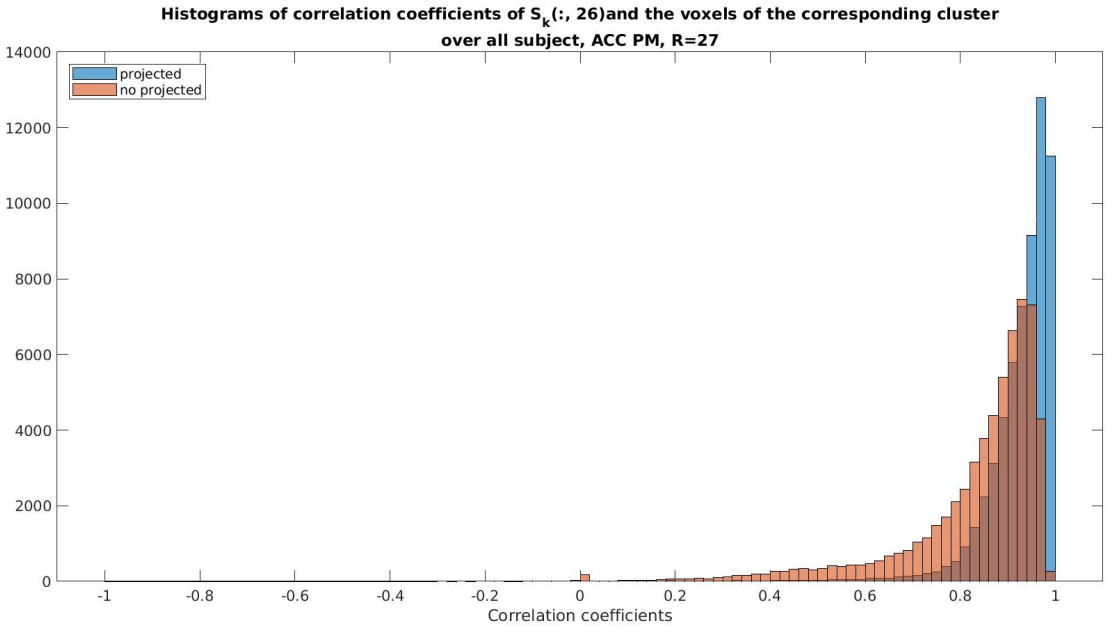


Figure B.264: For the column 26 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 26)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

25. **Cluster 18**

Cluster 18 contains 2392 voxels, which correspond to 1.058651% of the whole brain (background, soft tissue, bone are excluded).

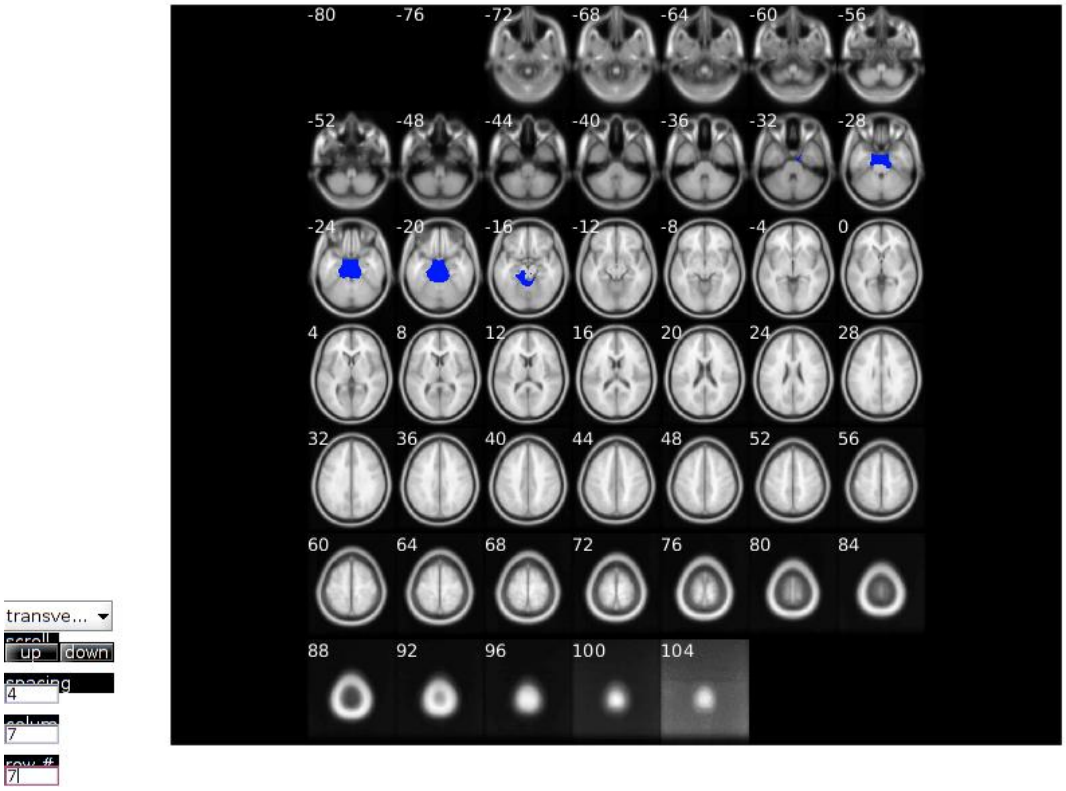


Figure B.265: Cluster 18. Blue areas indicate the voxels which belong to cluster 18.

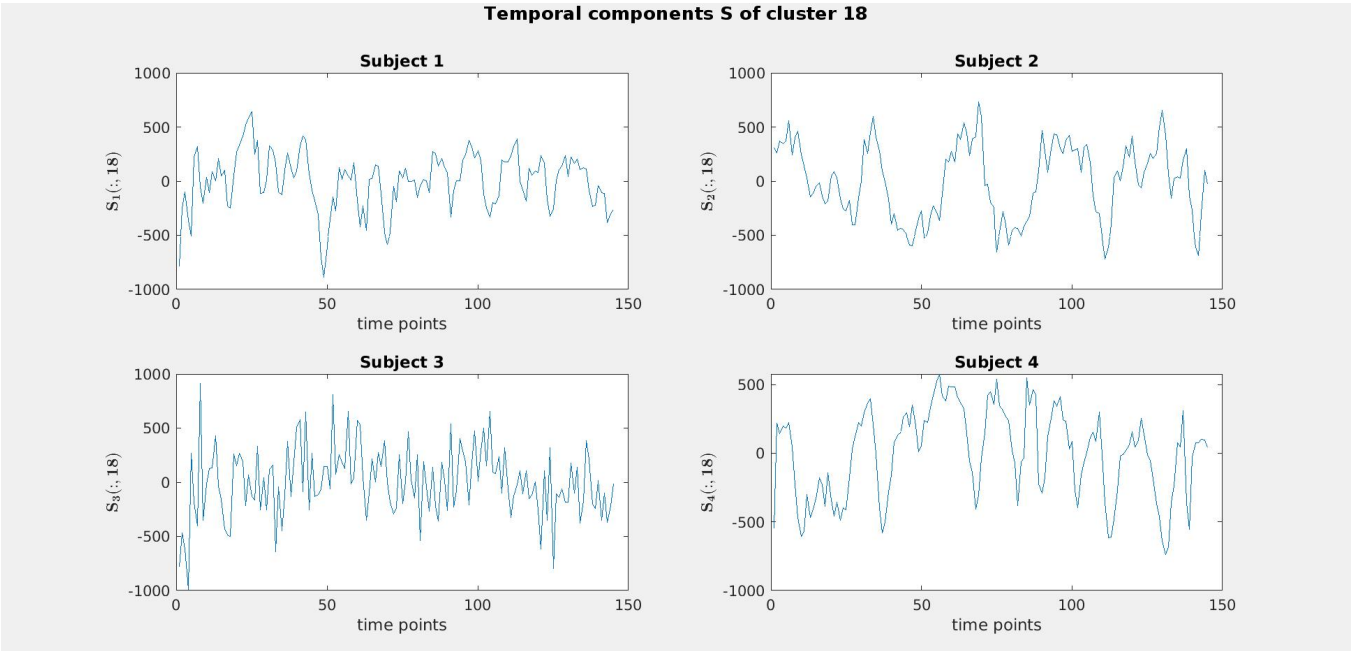


Figure B.266: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 1, \ldots, 4$ and R=27.

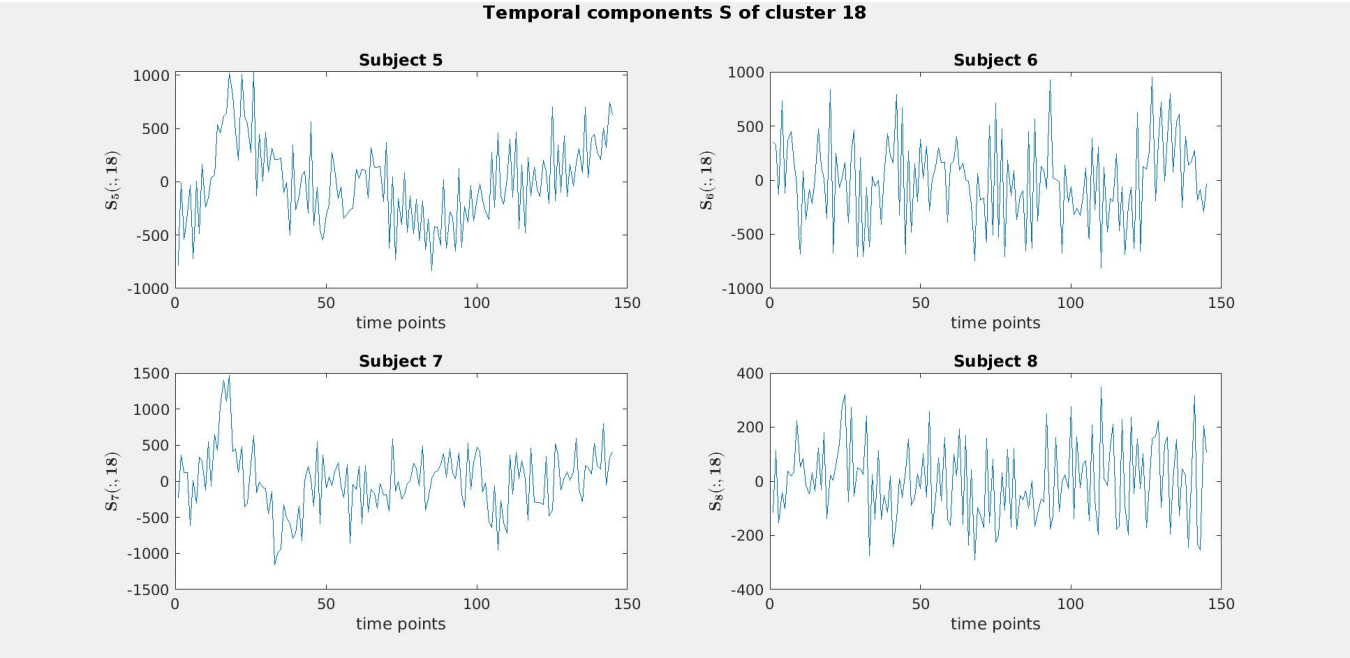


Figure B.267: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 5, \ldots, 8$ and R=27.

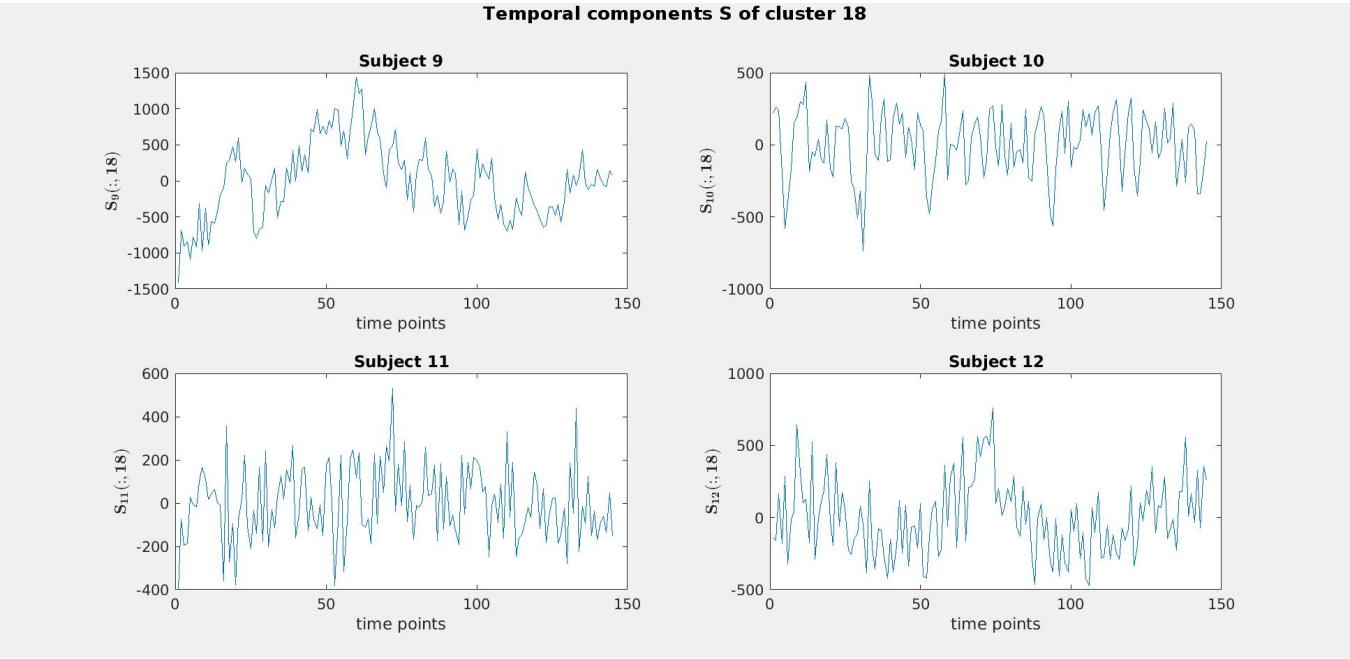


Figure B.268: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 9, \ldots, 12$ and R=27.

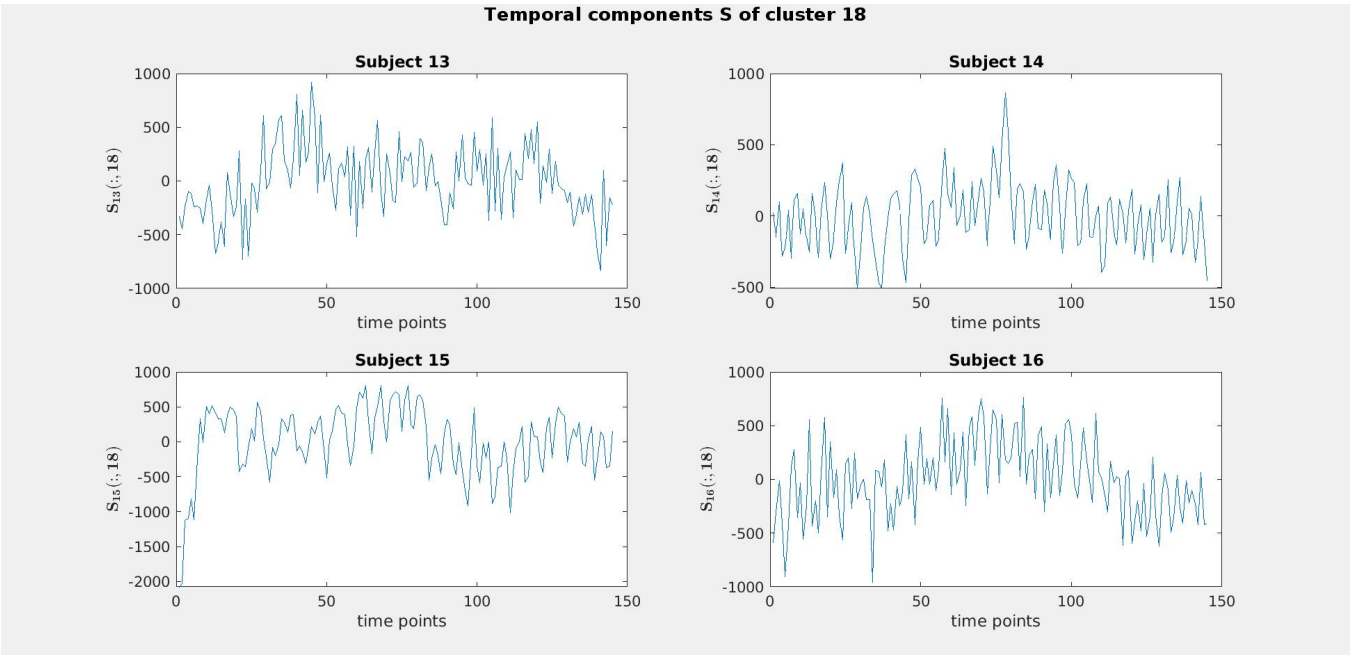


Figure B.269: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 13, \ldots, 16$ and R=27.

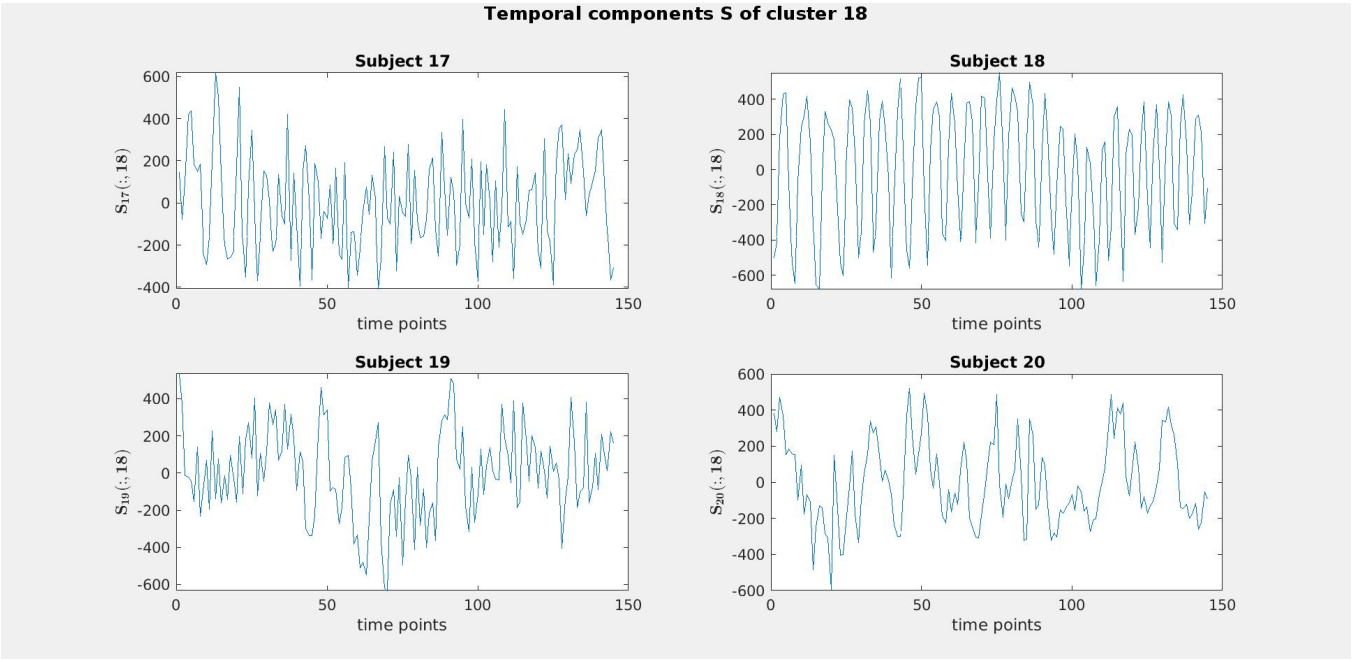


Figure B.270: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 17, \ldots, 20$ and R=27.

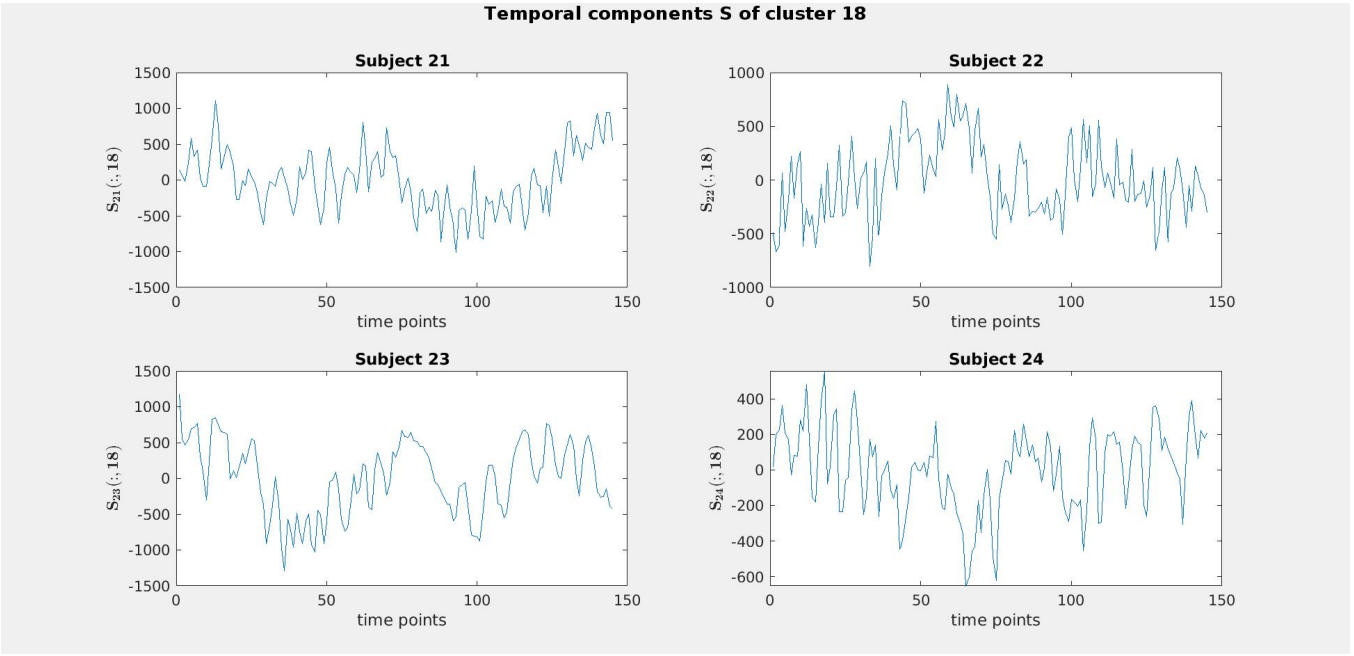


Figure B.271: Waveform of temporal components $\mathbf{S}_k(:, 118)$ for $k = 21, \ldots, 24$ and R=27.

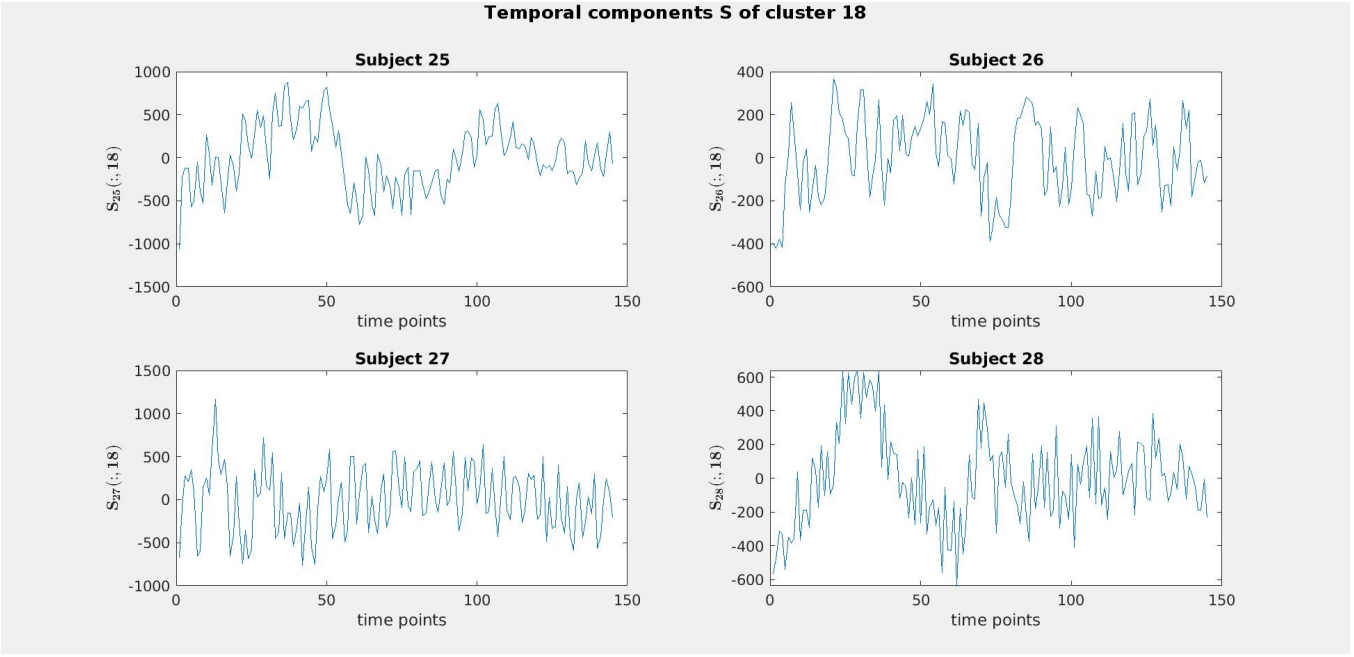


Figure B.272: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 25, \ldots, 28$ and R=27.

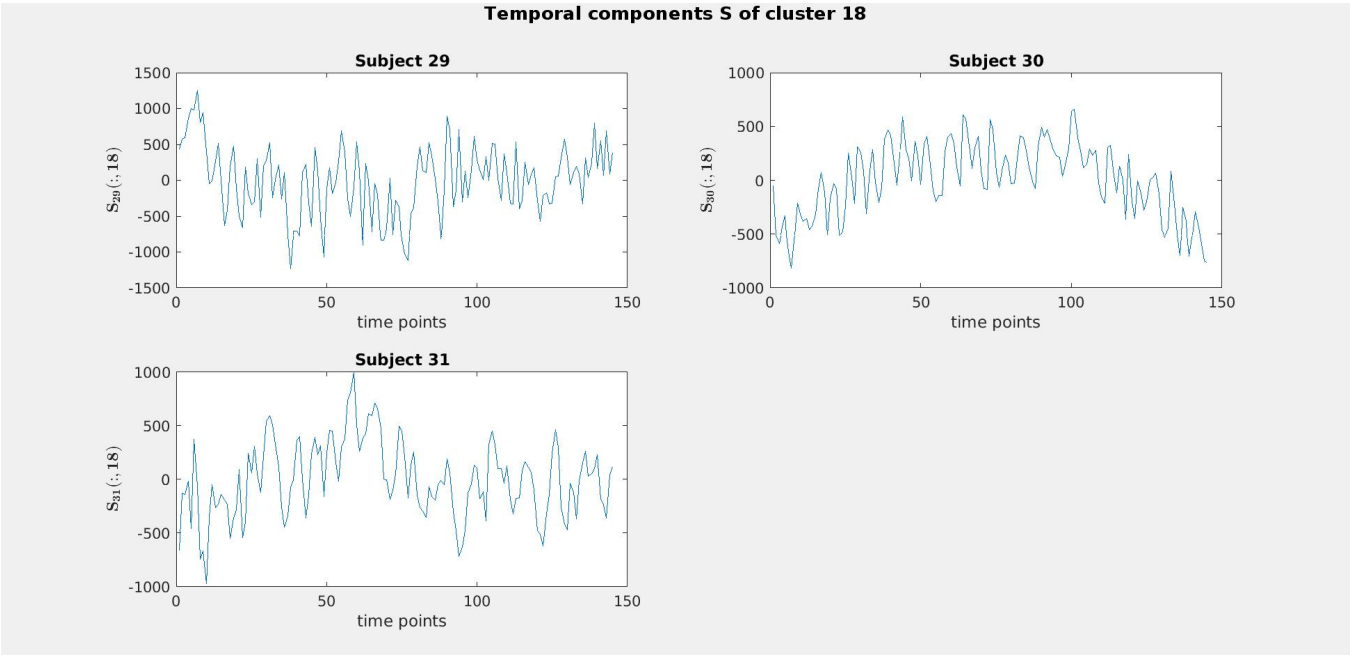


Figure B.273: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 29, \ldots, 31$ and R=27.

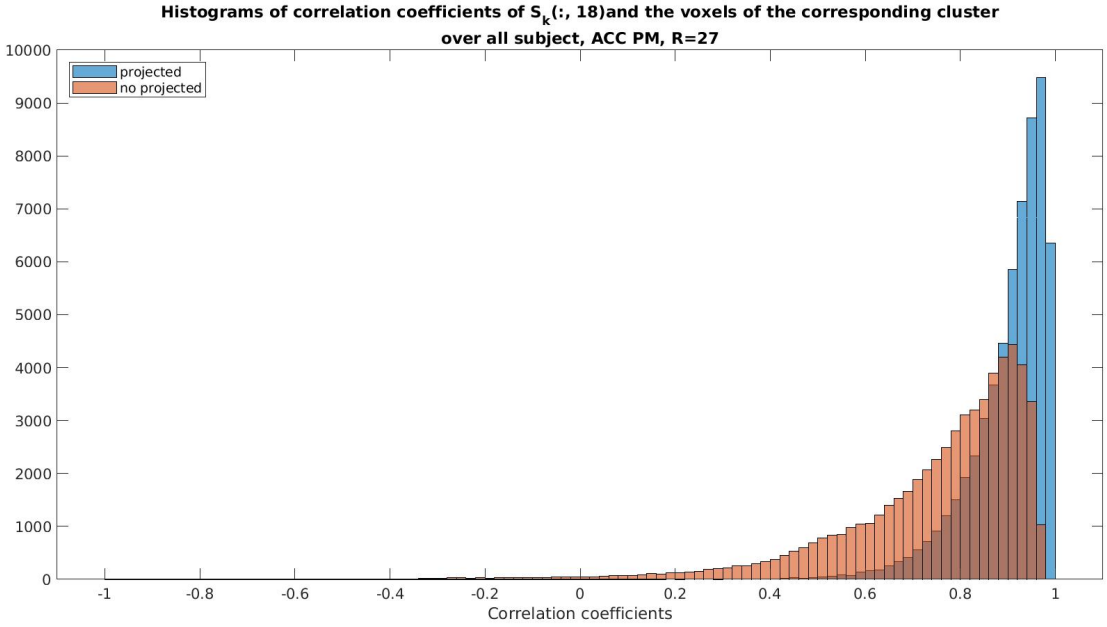


Figure B.274: For the column 18 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 18)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

26. **Cluster 13**

Cluster 13 contains 2264 voxels, which correspond to 1.002000% of the whole brain (background, soft tissue, bone are excluded).

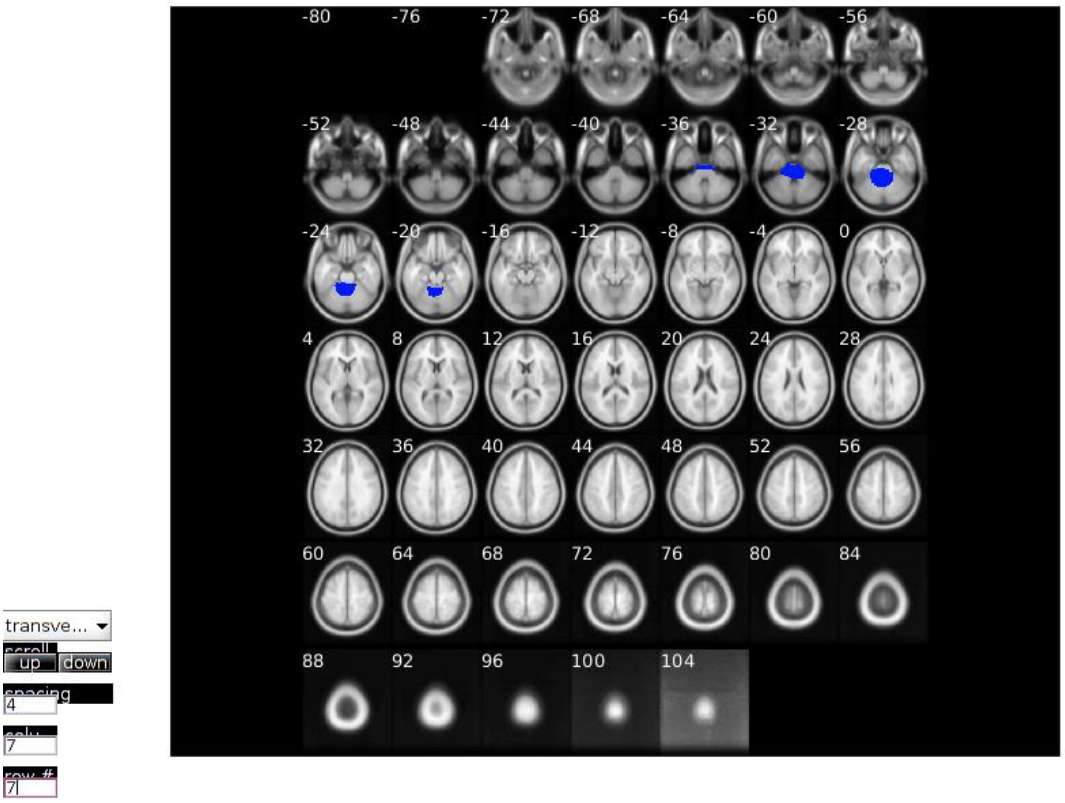


Figure B.275: Cluster 13. Blue areas indicate the voxels which belong to cluster 13.

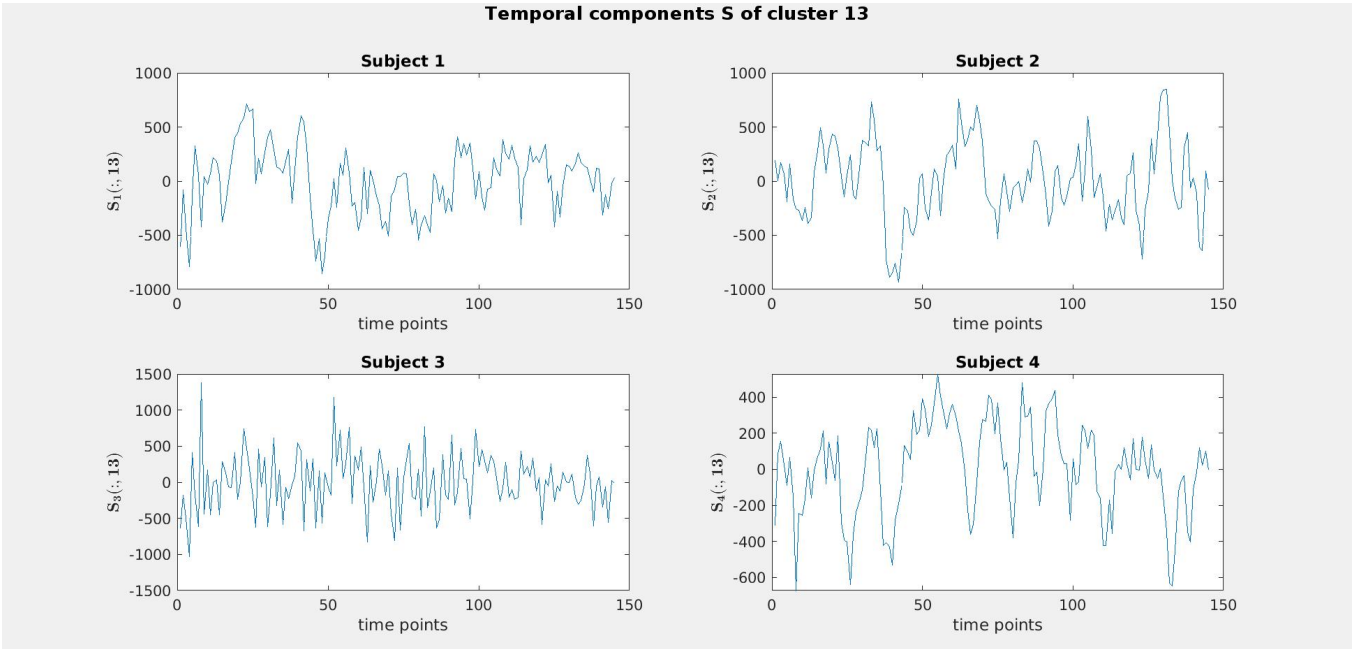


Figure B.276: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 1, \ldots, 4$ and R=27.

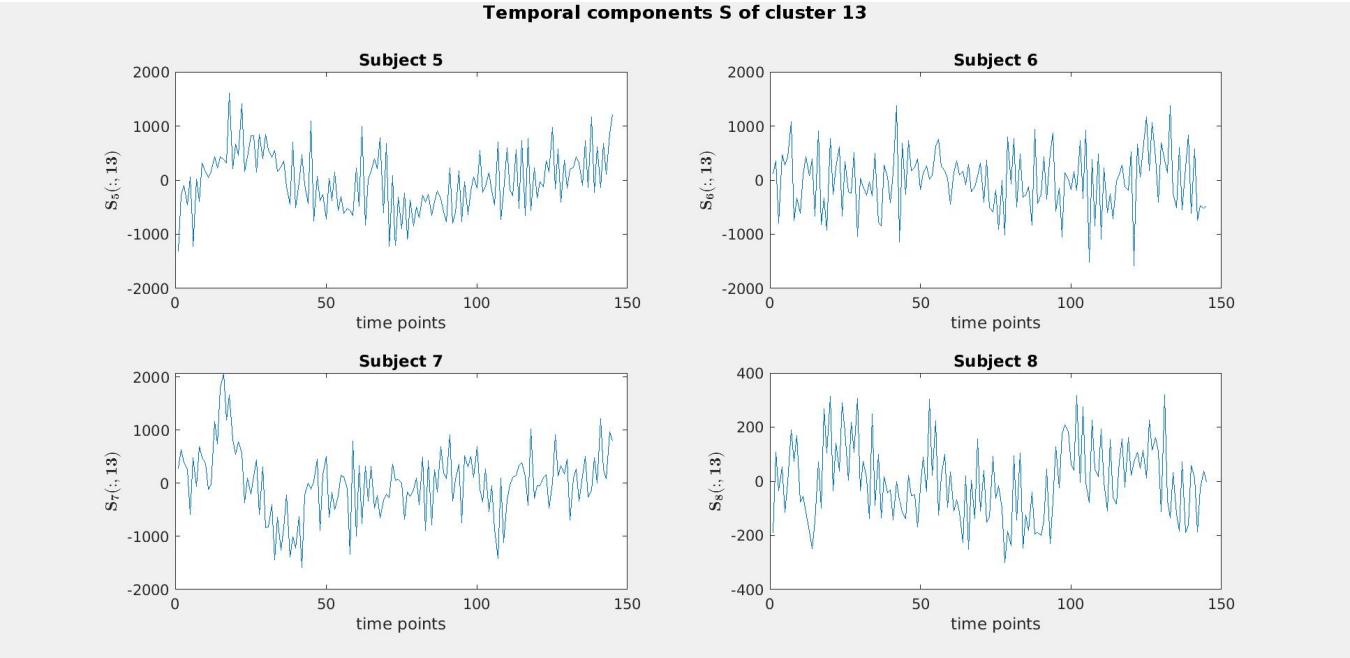


Figure B.277: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 5, \ldots, 8$ and R=27.

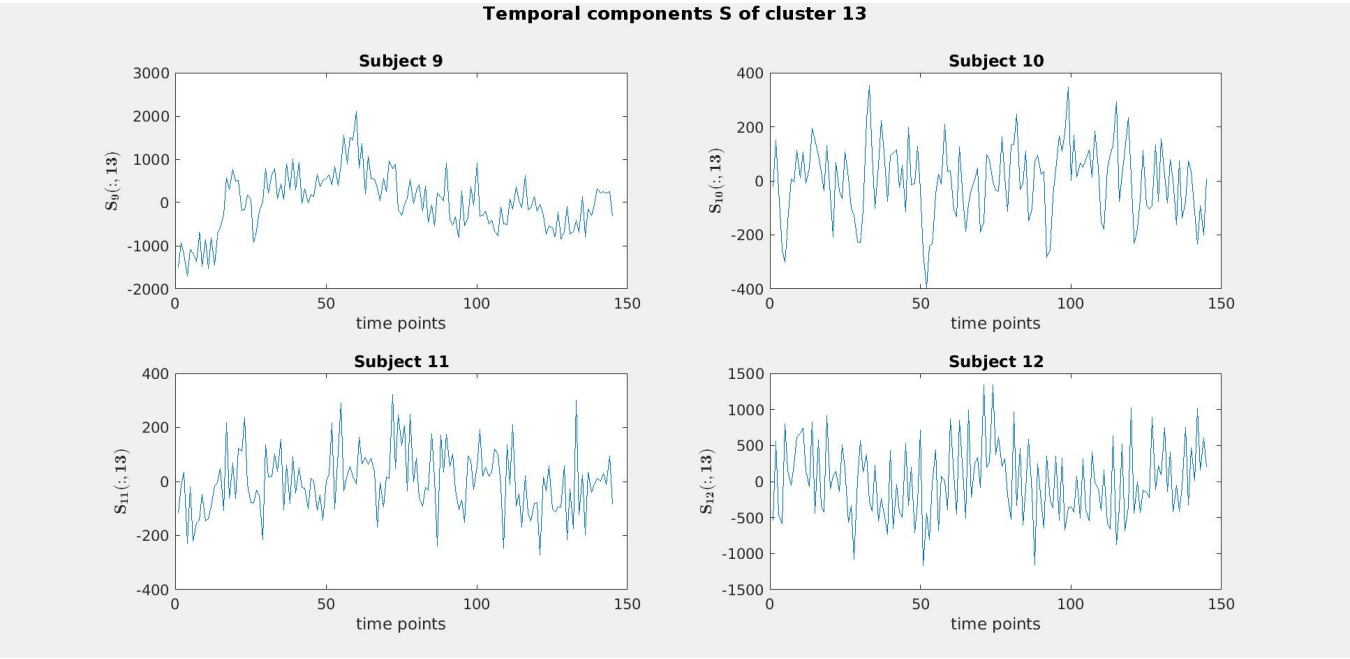


Figure B.278: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 9, \ldots, 12$ and R=27.

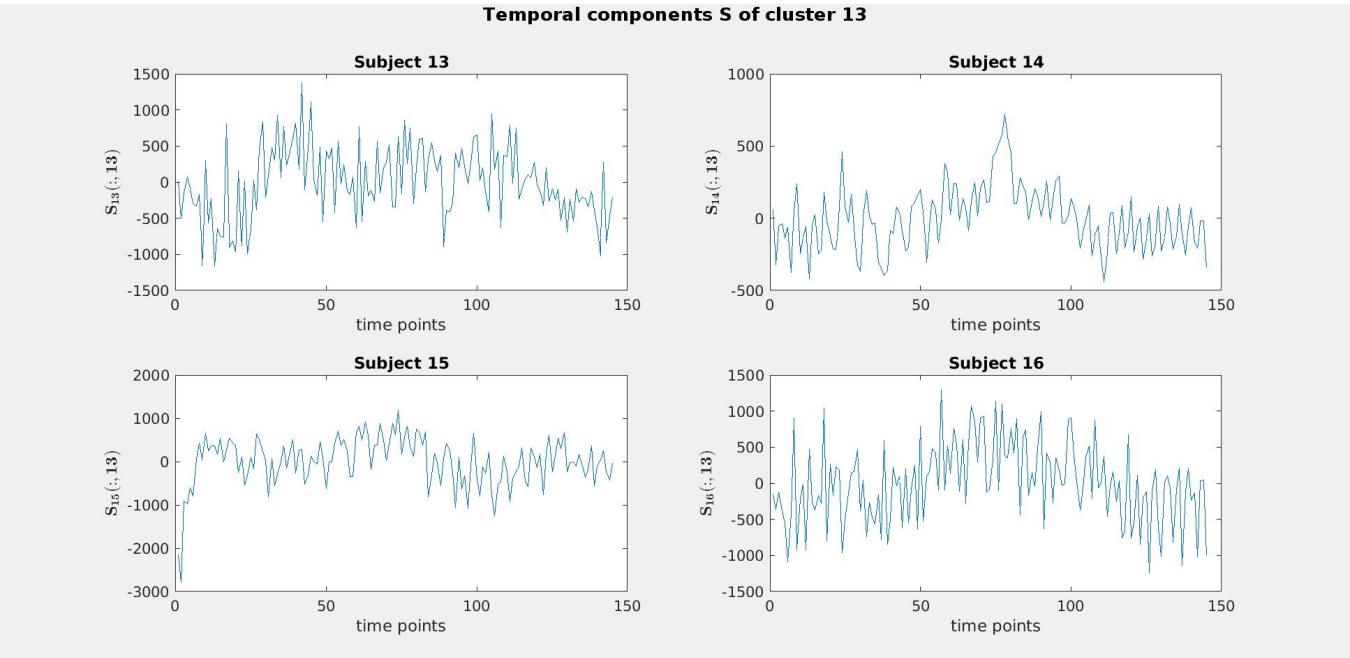


Figure B.279: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 13, \ldots, 16$ and R=27.

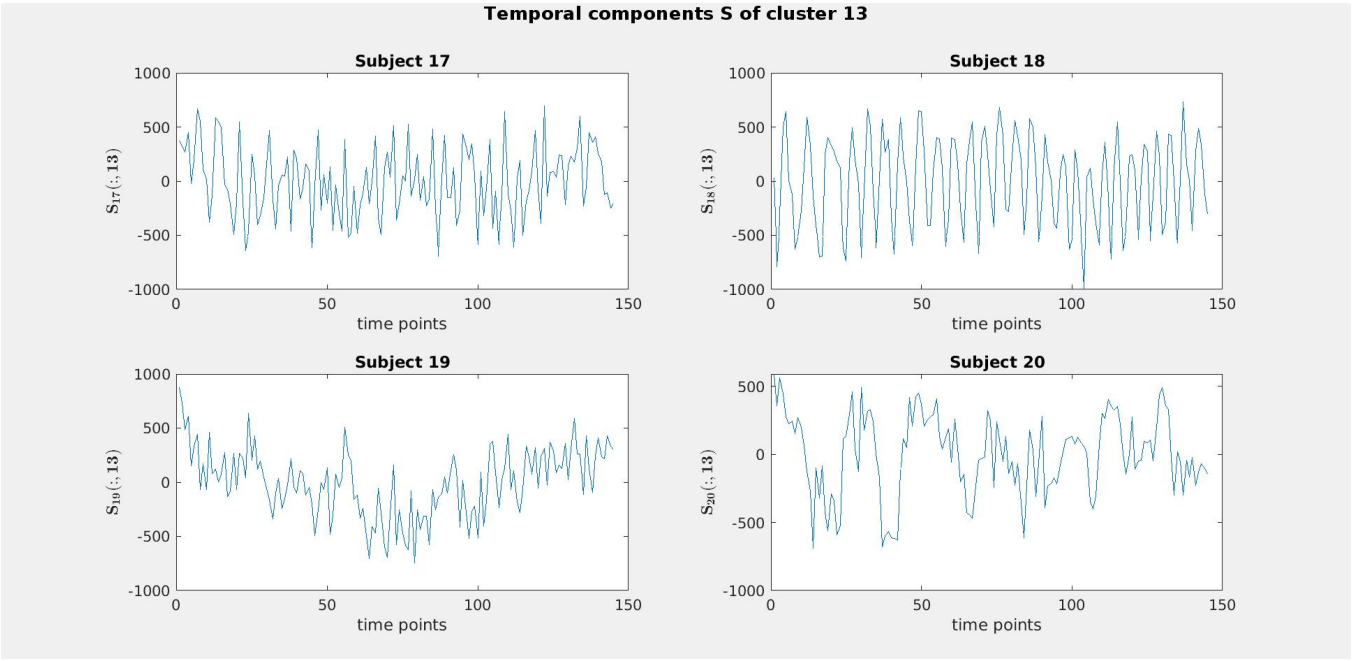


Figure B.280: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 17, \ldots, 20$ and R=27.

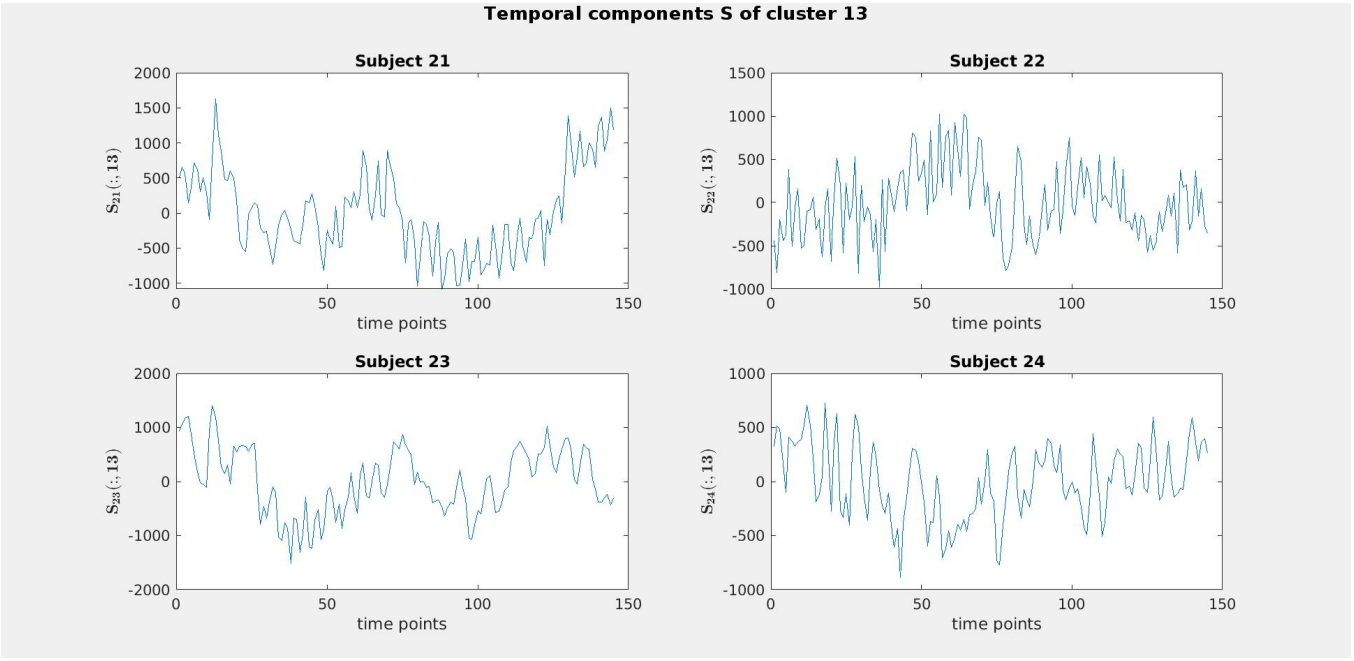


Figure B.281: Waveform of temporal components $\mathbf{S}_k(:, 113)$ for $k = 21, \ldots, 24$ and R=27.

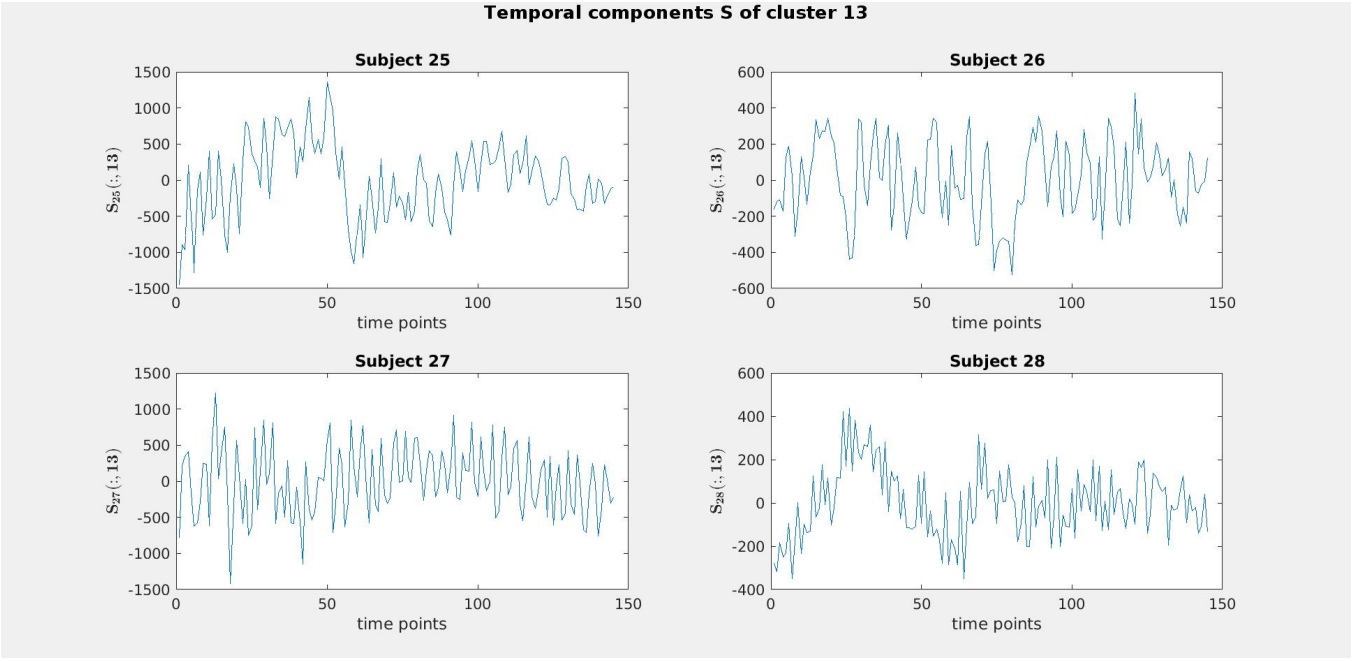


Figure B.282: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 25, \ldots, 28$ and R=27.

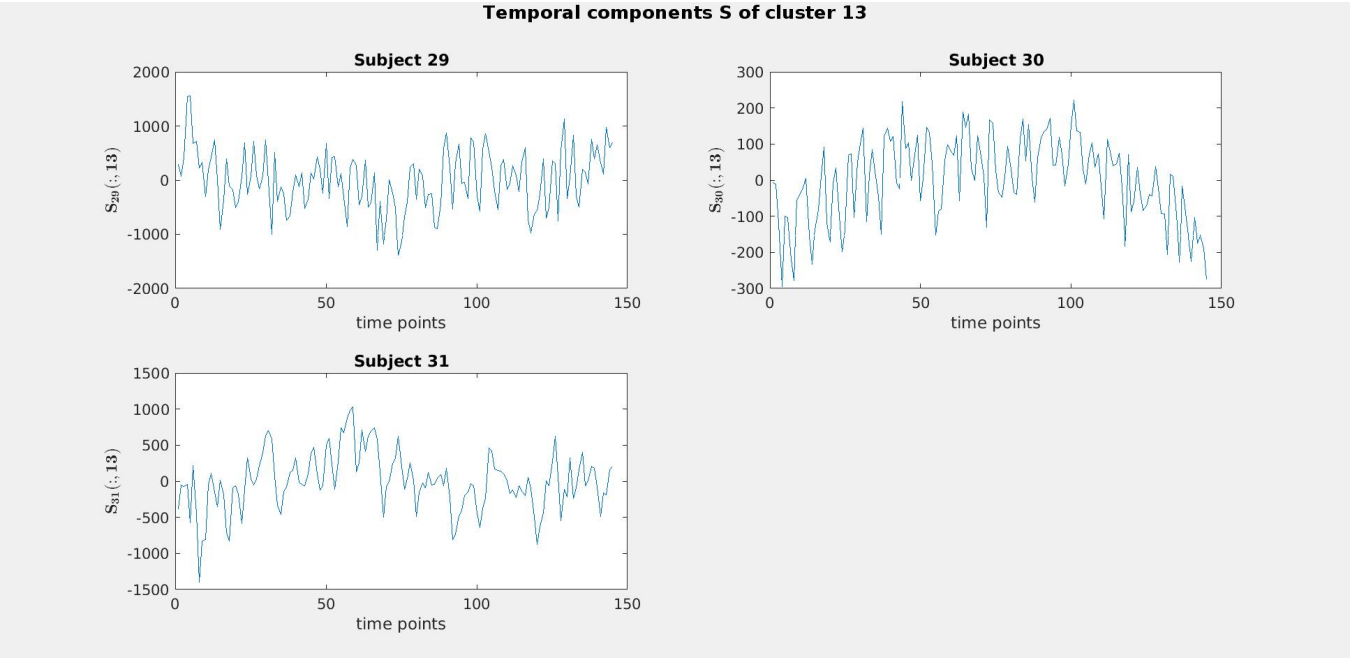


Figure B.283: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 29, \ldots, 31$ and R=27.

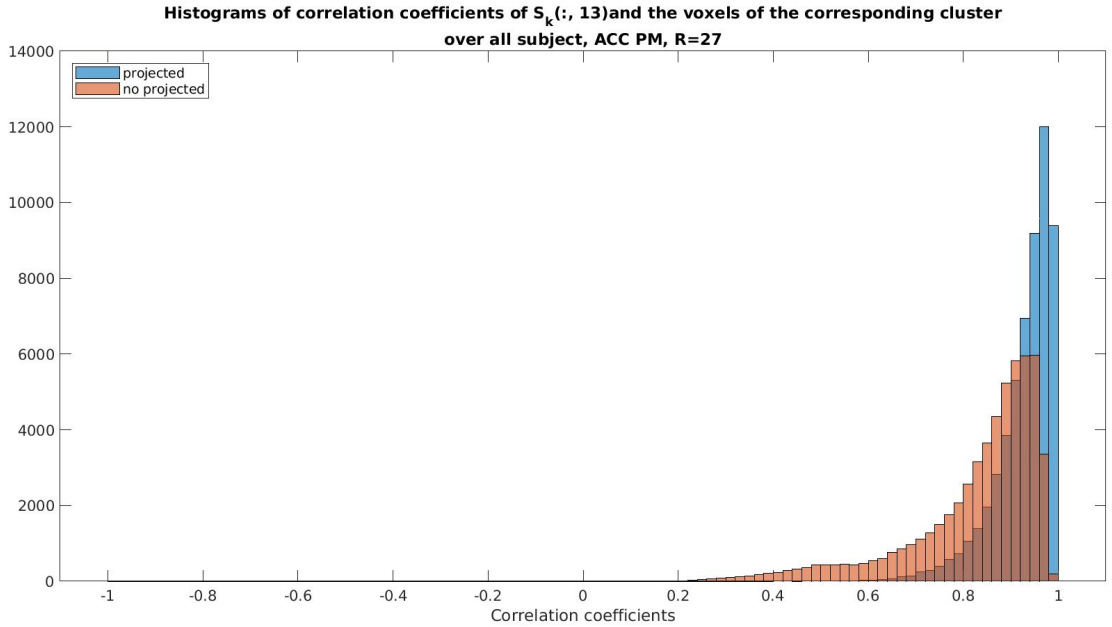


Figure B.284: For the column 13 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 13)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

27. **Cluster 12**

Cluster 12 contains 2201 voxels, which correspond to 0.974118% of the whole brain (background, soft tissue, bone are excluded).

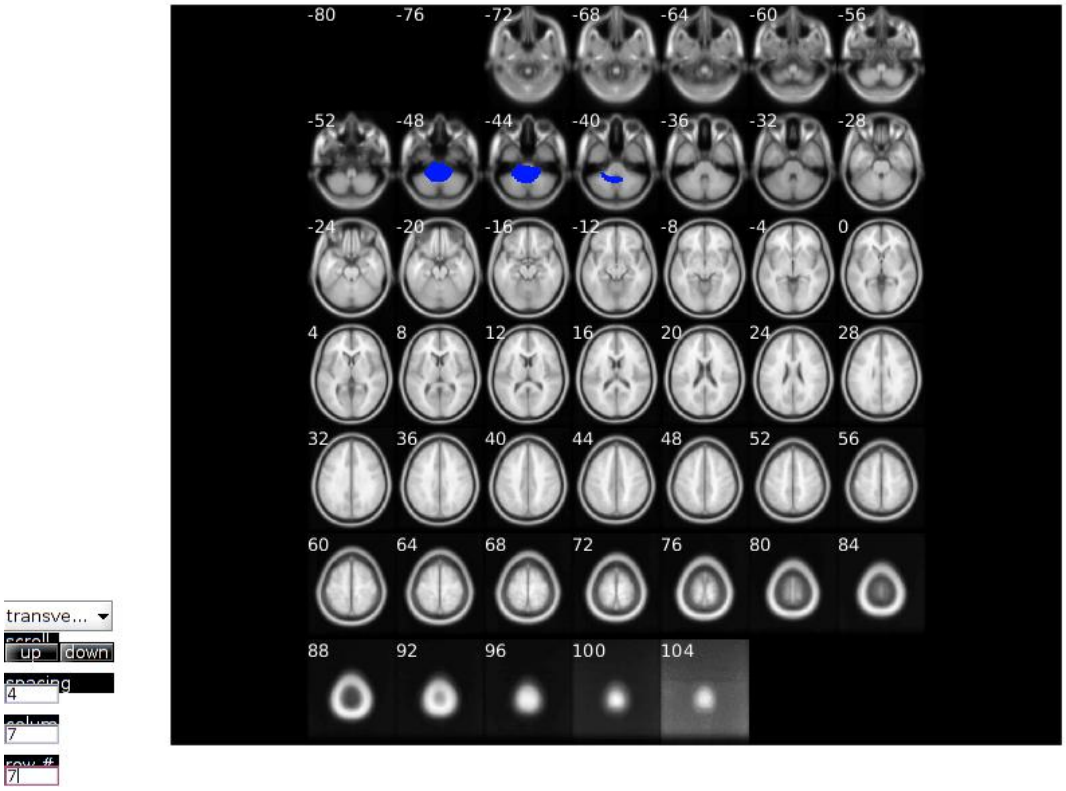


Figure B.285: Cluster 12. Blue areas indicate the voxels which belong to cluster 12.

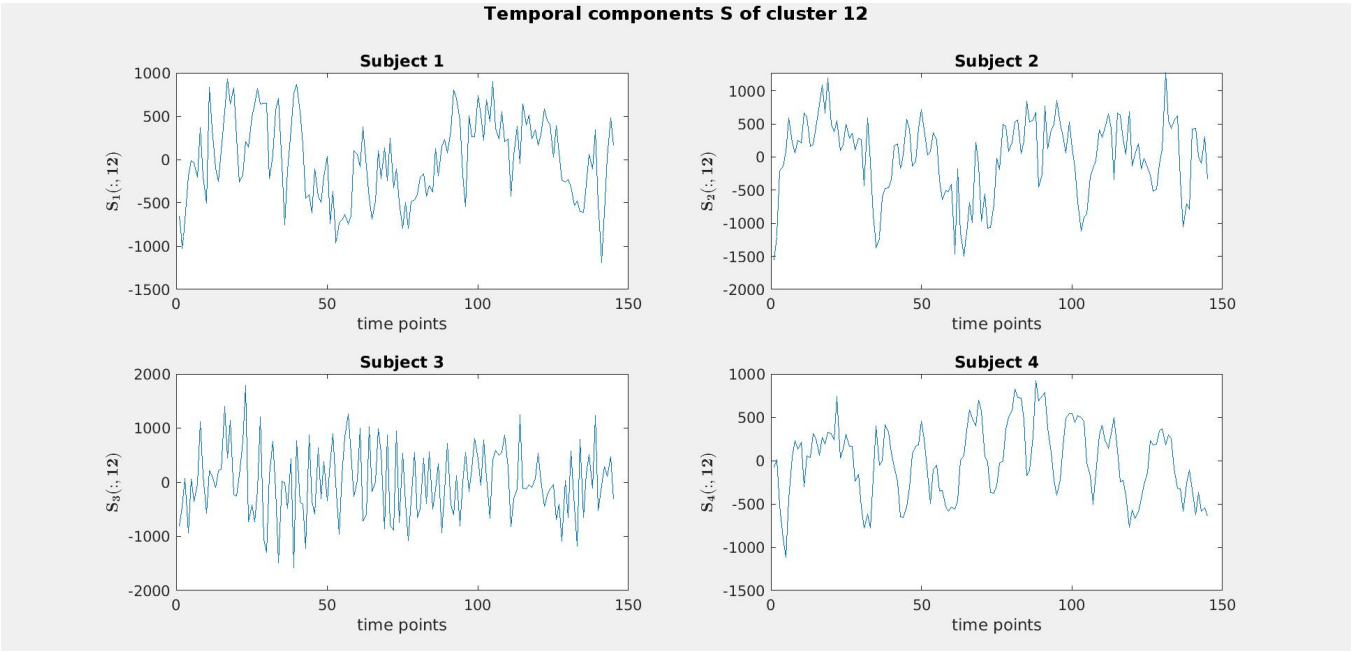


Figure B.286: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 1, \ldots, 4$ and R=27.

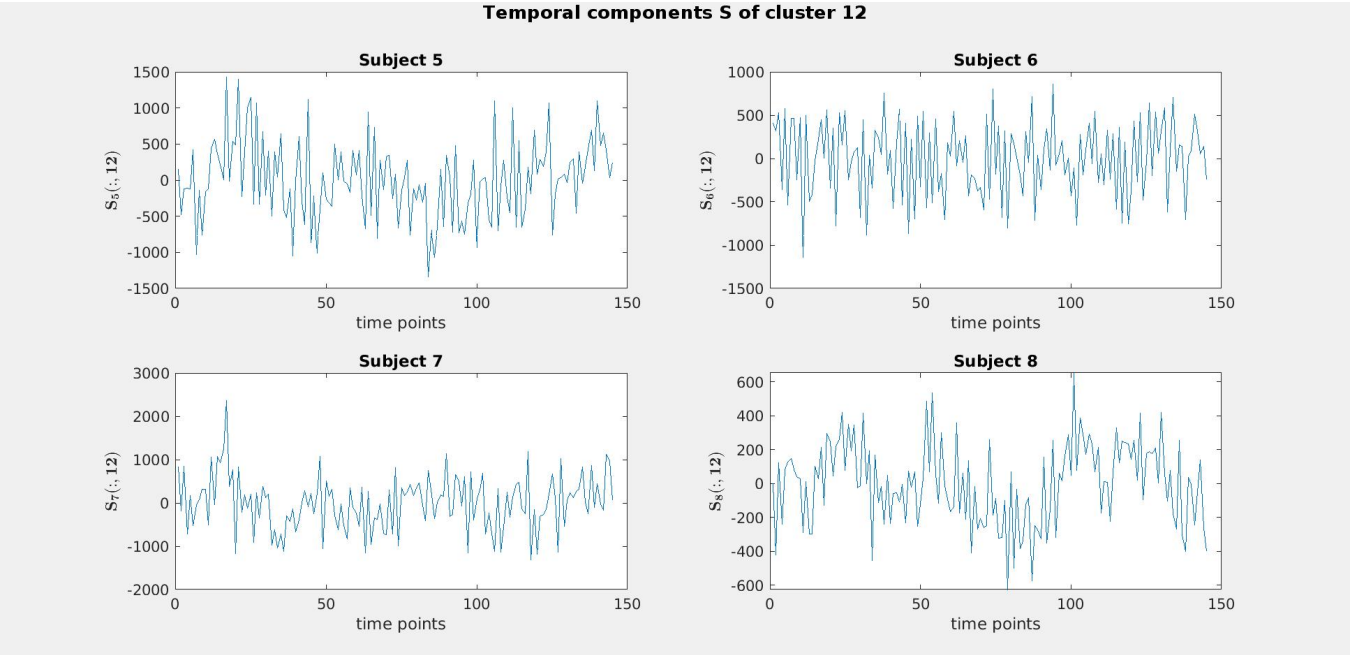


Figure B.287: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 5, \ldots, 8$ and R=27.

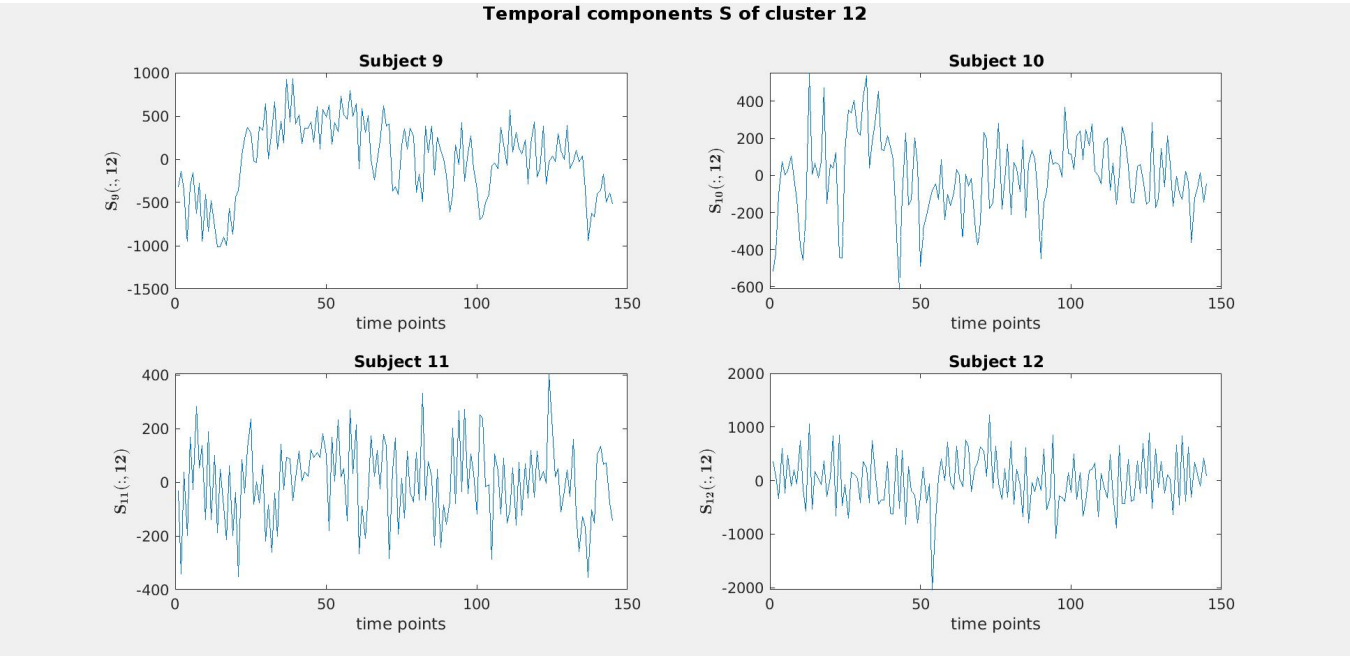


Figure B.288: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 9, \ldots, 12$ and R=27.

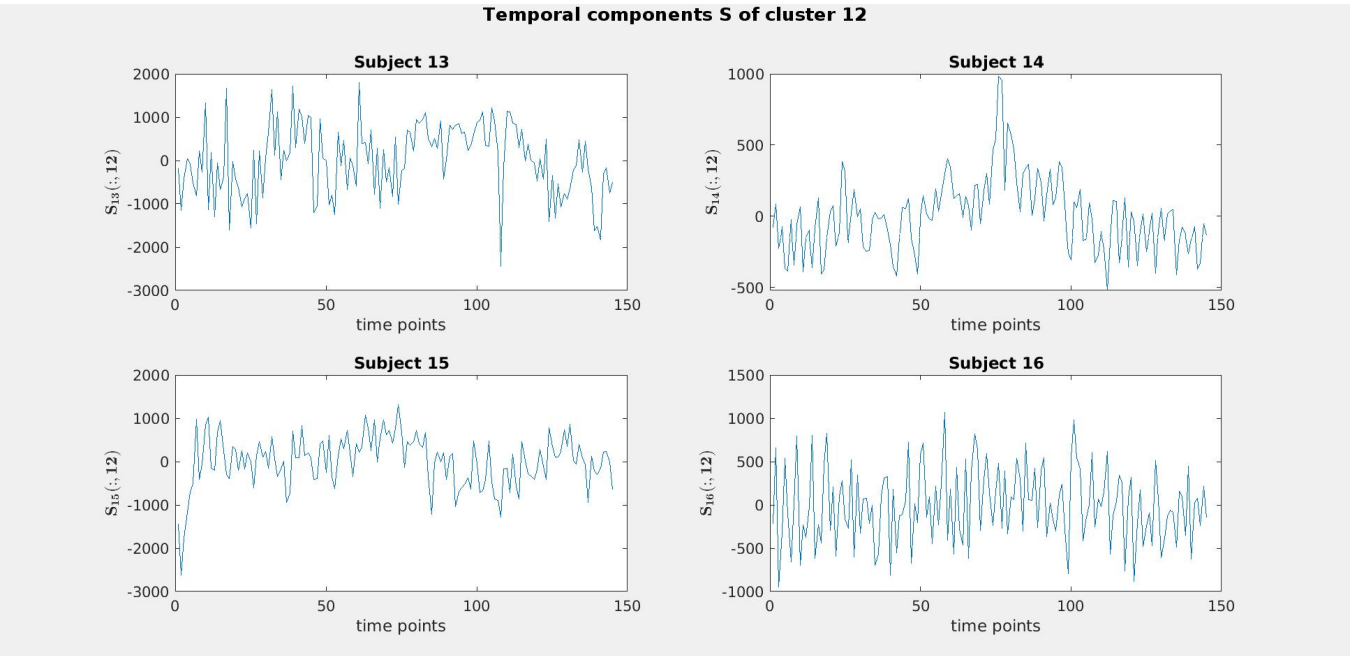


Figure B.289: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 13, \ldots, 16$ and R=27.

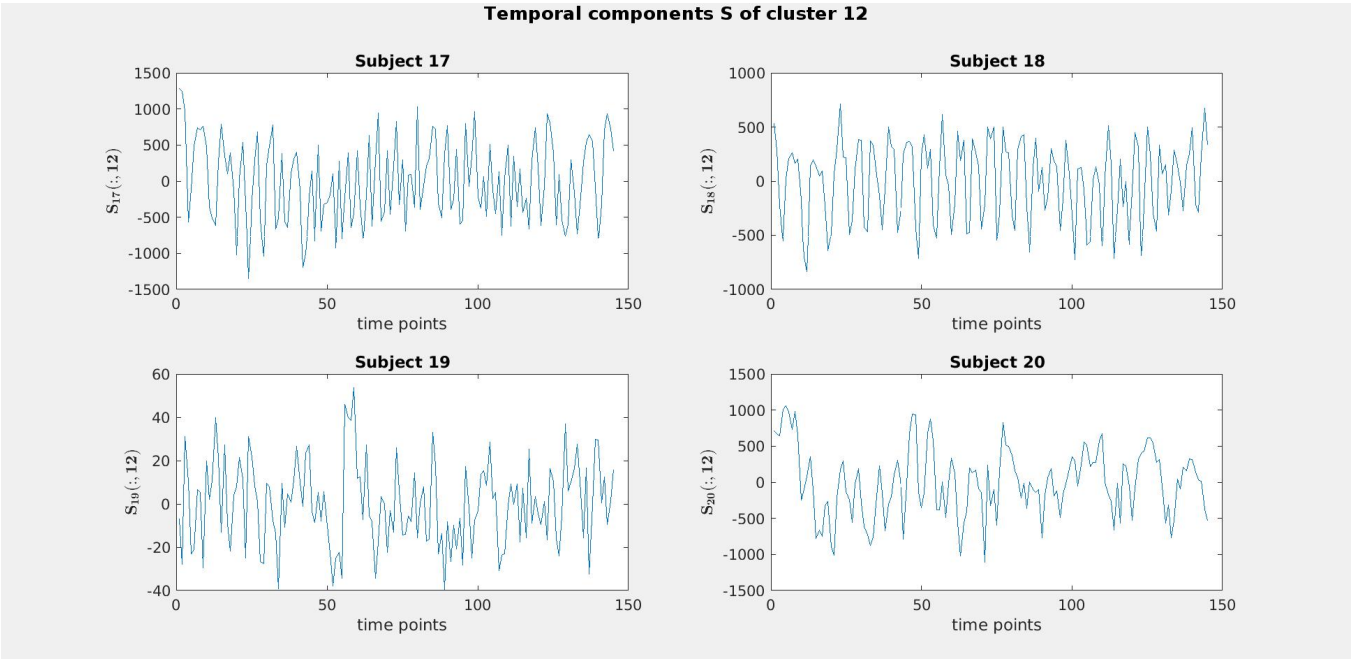


Figure B.290: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 17, \ldots, 20$ and R=27.

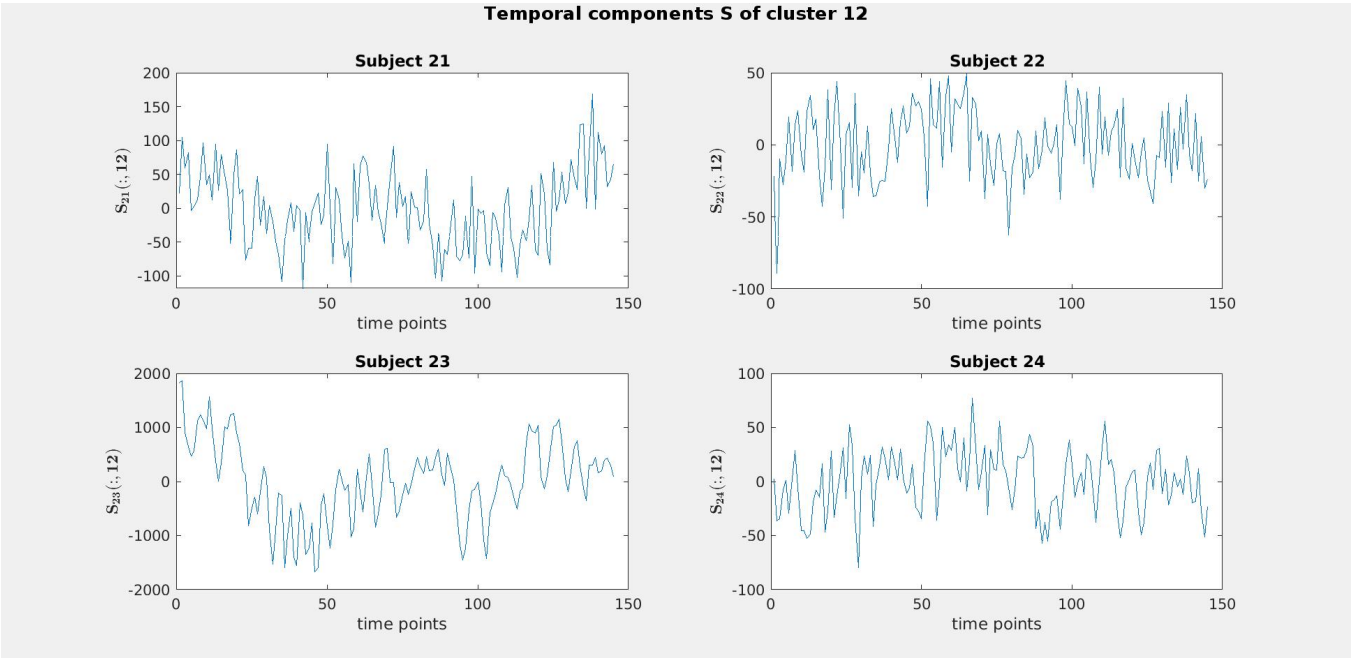


Figure B.291: Waveform of temporal components $\mathbf{S}_k(:, 112)$ for $k = 21, \ldots, 24$ and R=27.

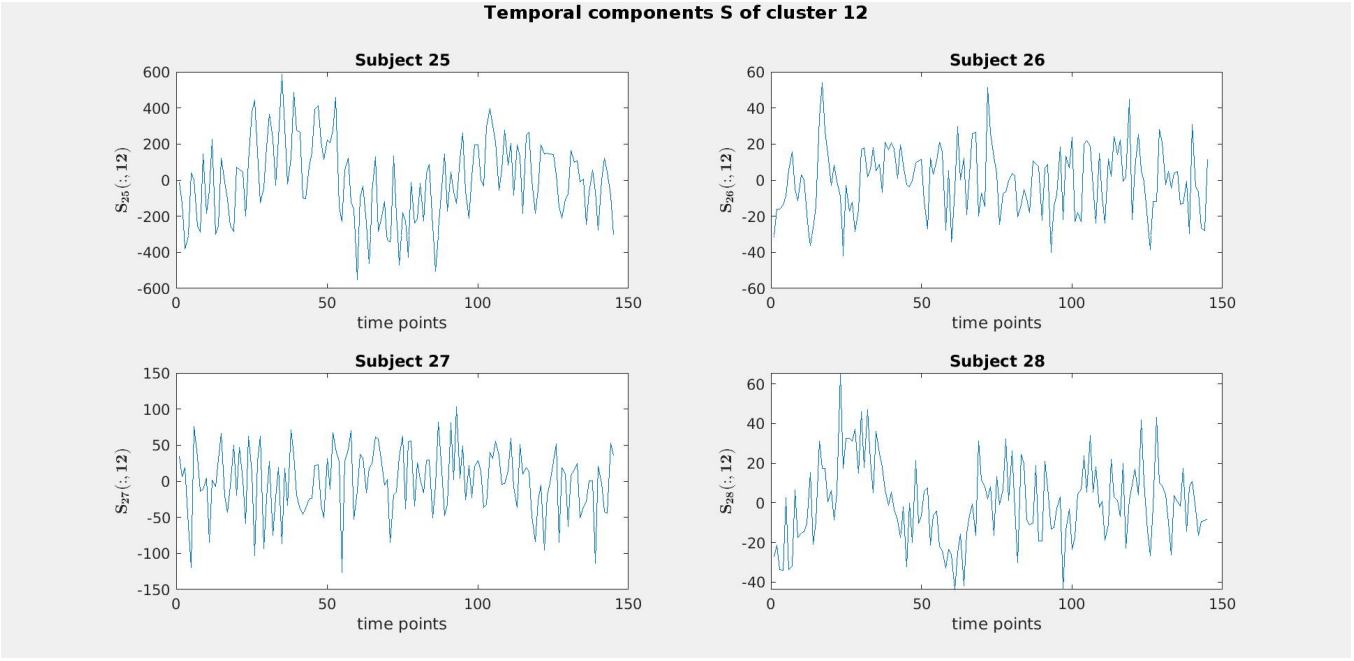


Figure B.292: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 25, \ldots, 28$ and R=27.

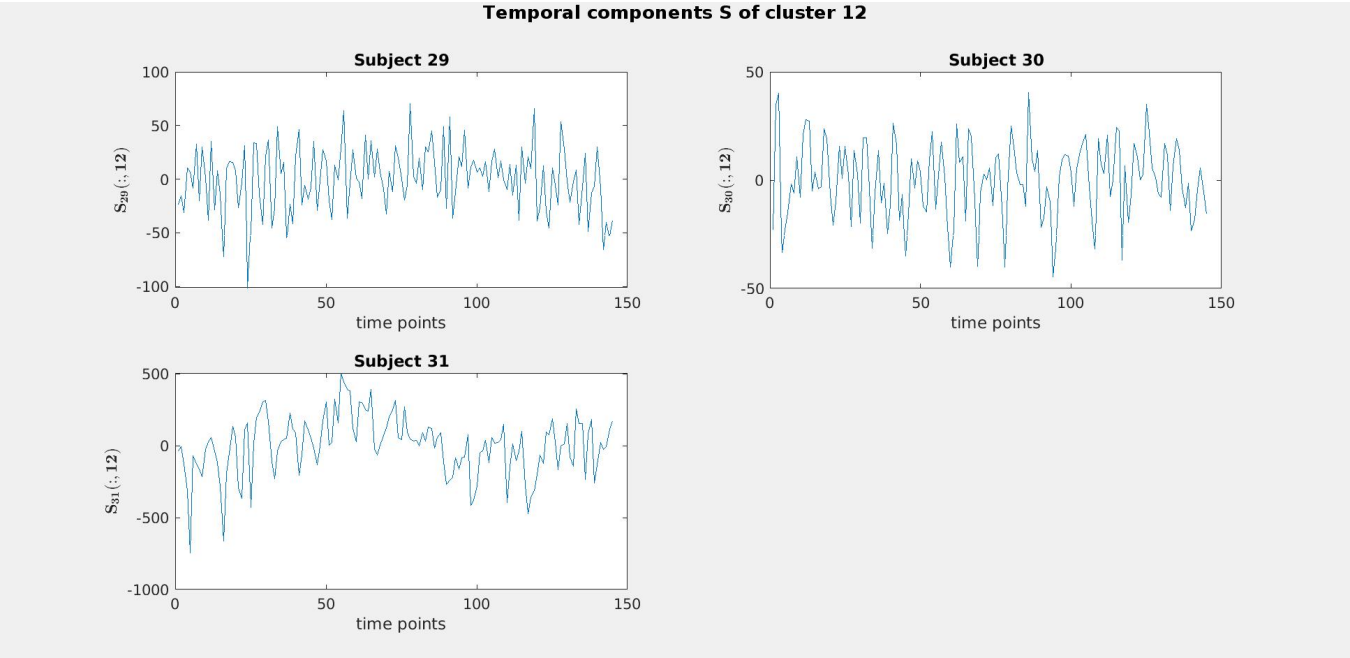


Figure B.293: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 29, \ldots, 31$ and R=27.

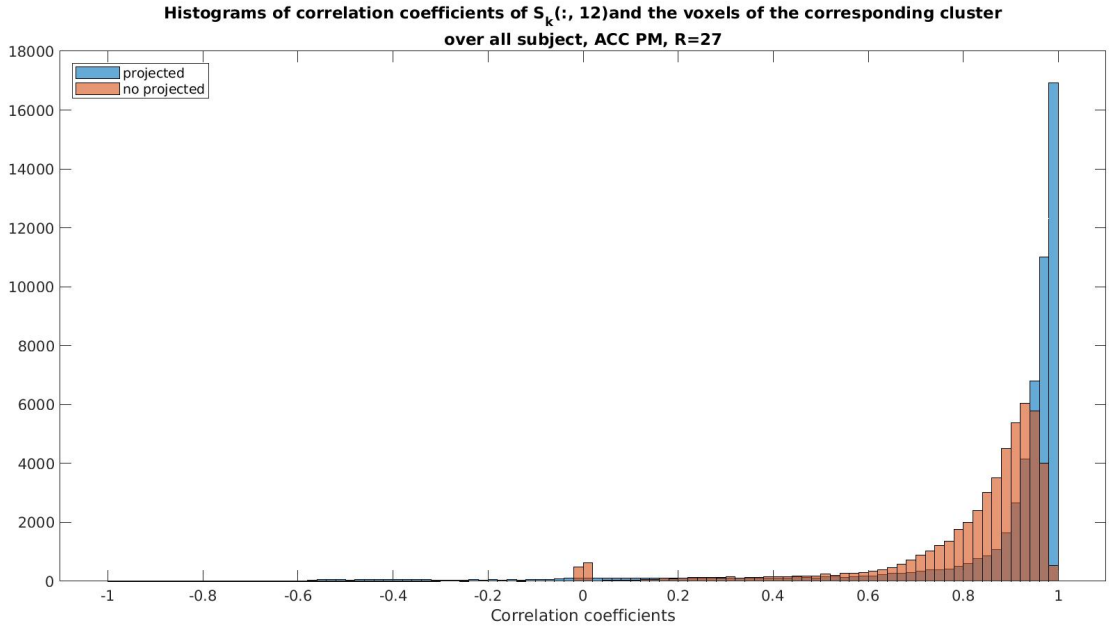


Figure B.294: For the column 12 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 12)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

B.4 Clusters for common subspace dimension equal to 42

1. **Cluster 33**
   Cluster 33 contains 15837 voxels, which correspond to 7.009135% of the whole brain (background, soft tissue, bone are excluded).

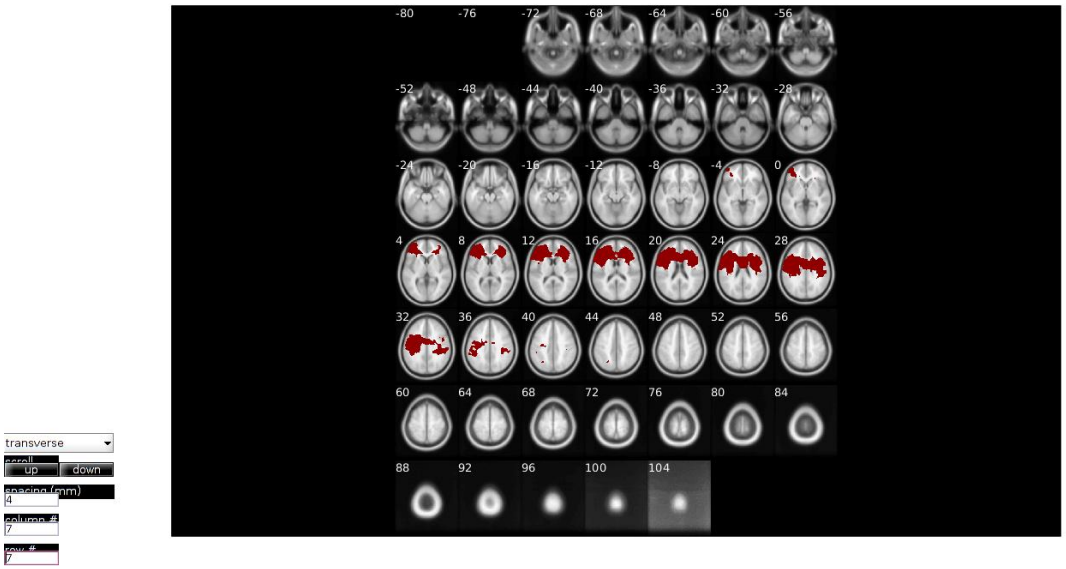


Figure B.295: Cluster 33. Red areas indicate the voxels which belong to cluster 33.

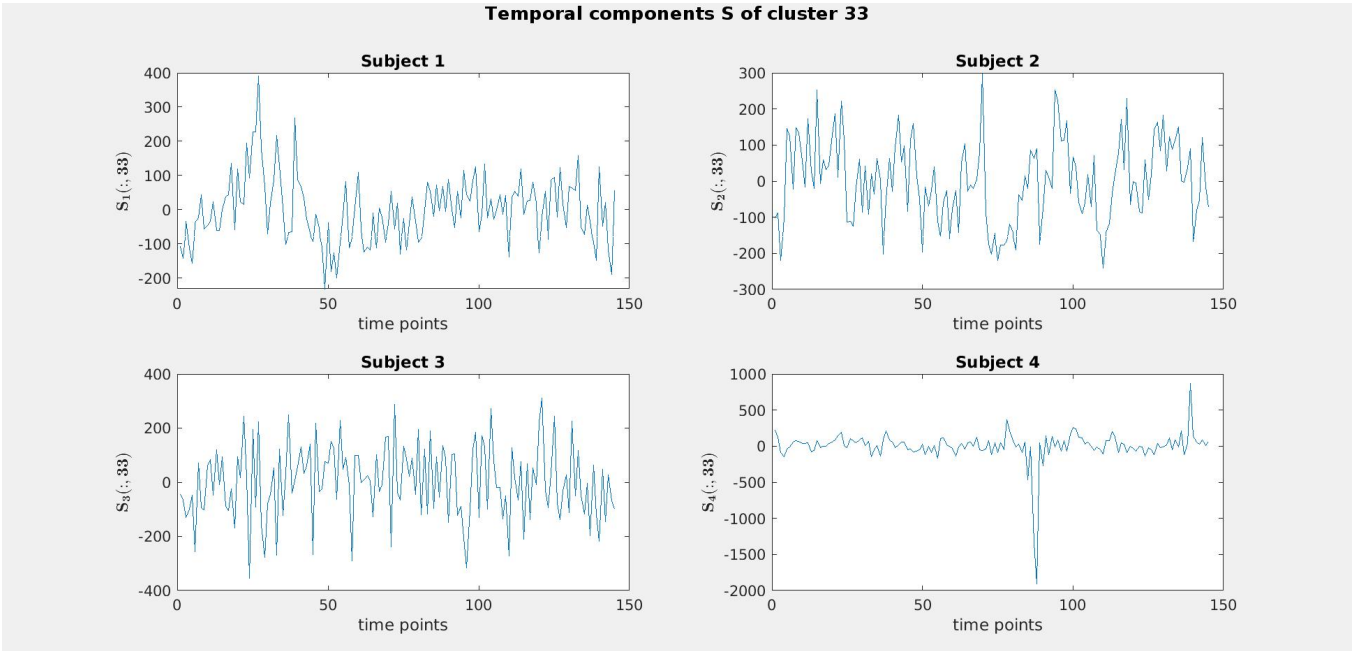


Figure B.296: Waveform of temporal components $\mathbf{S}_k(:, 33)$ for $k = 1, \ldots, 4$ and R=42.

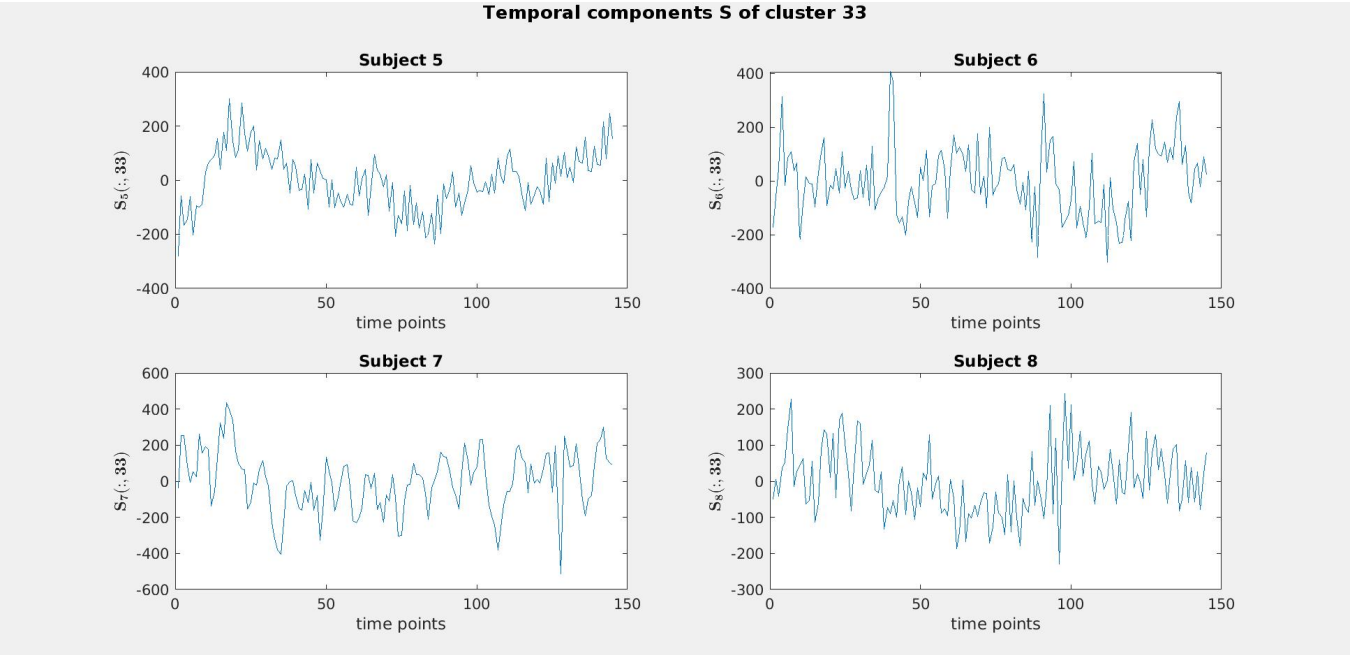


Figure B.297: Waveform of temporal components $\mathbf{S}_k(:, 33)$ for $k = 5, \ldots, 8$ and R=42.

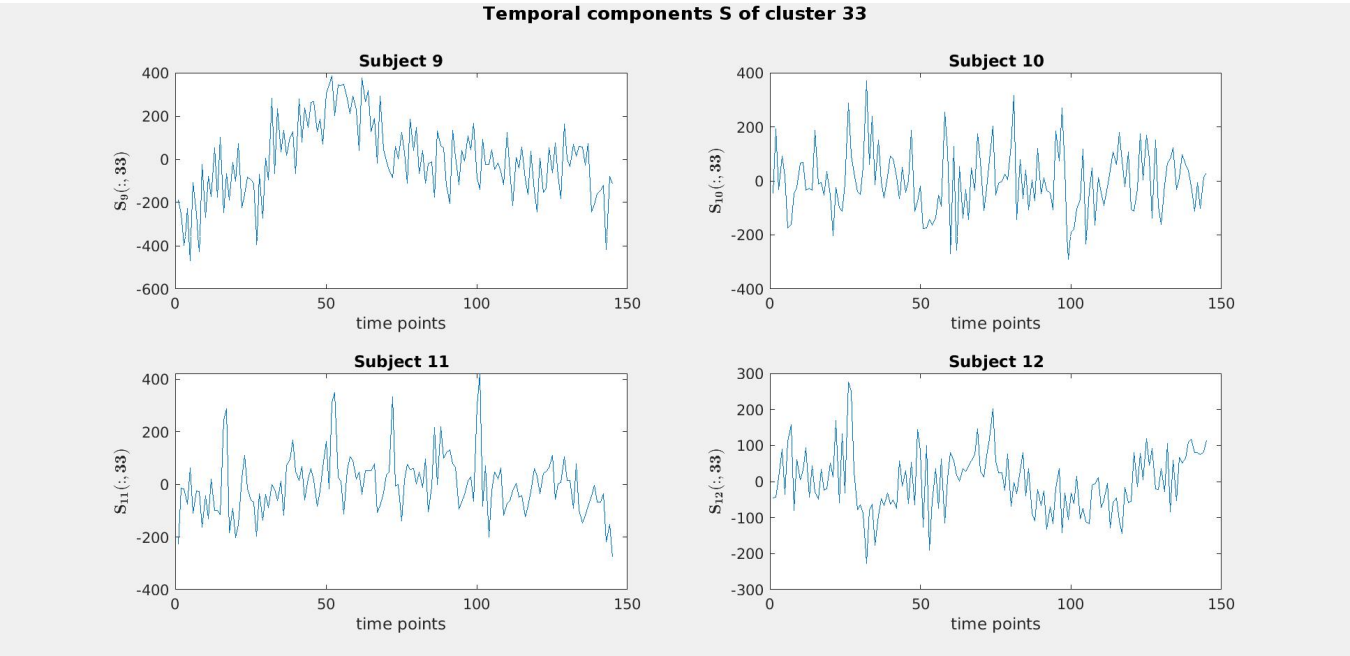


Figure B.298: Waveform of temporal components $\mathbf{S}_k(:, 33)$ for $k = 9, \ldots, 12$ and R=42.

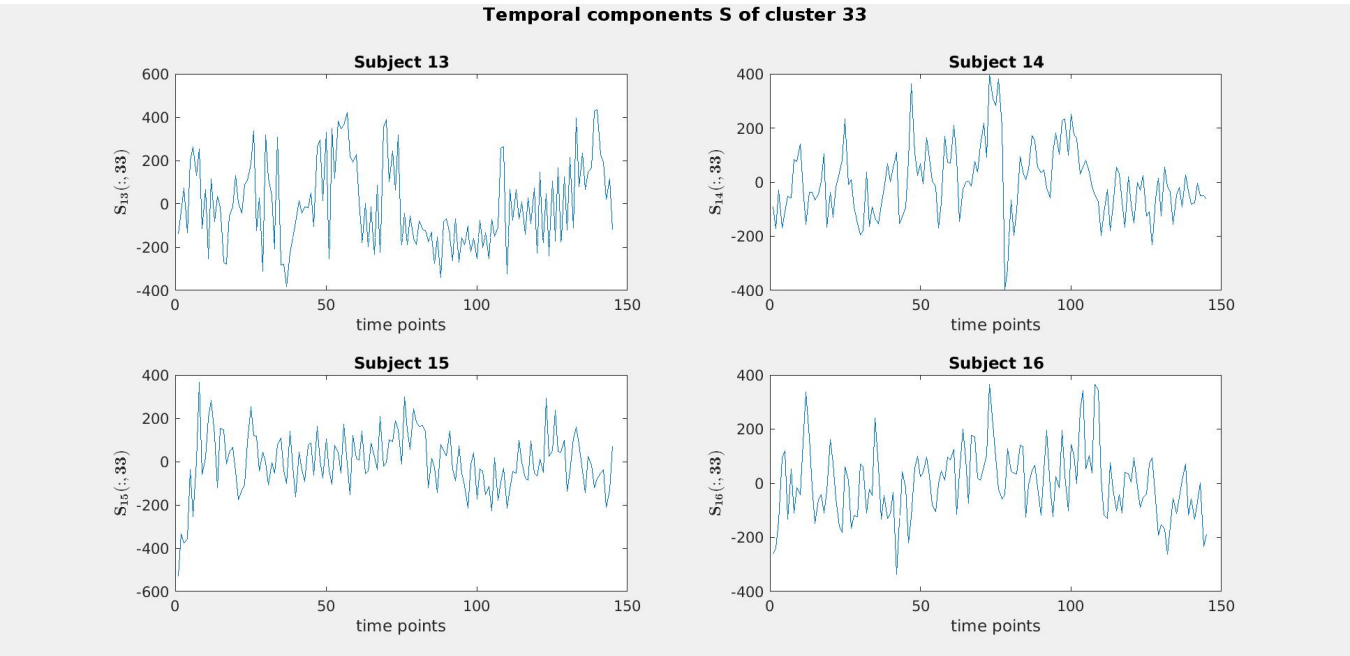


Figure B.299: Waveform of temporal components $\mathbf{S}_k(:, 33)$ for $k = 13, \ldots, 16$ and R=42.

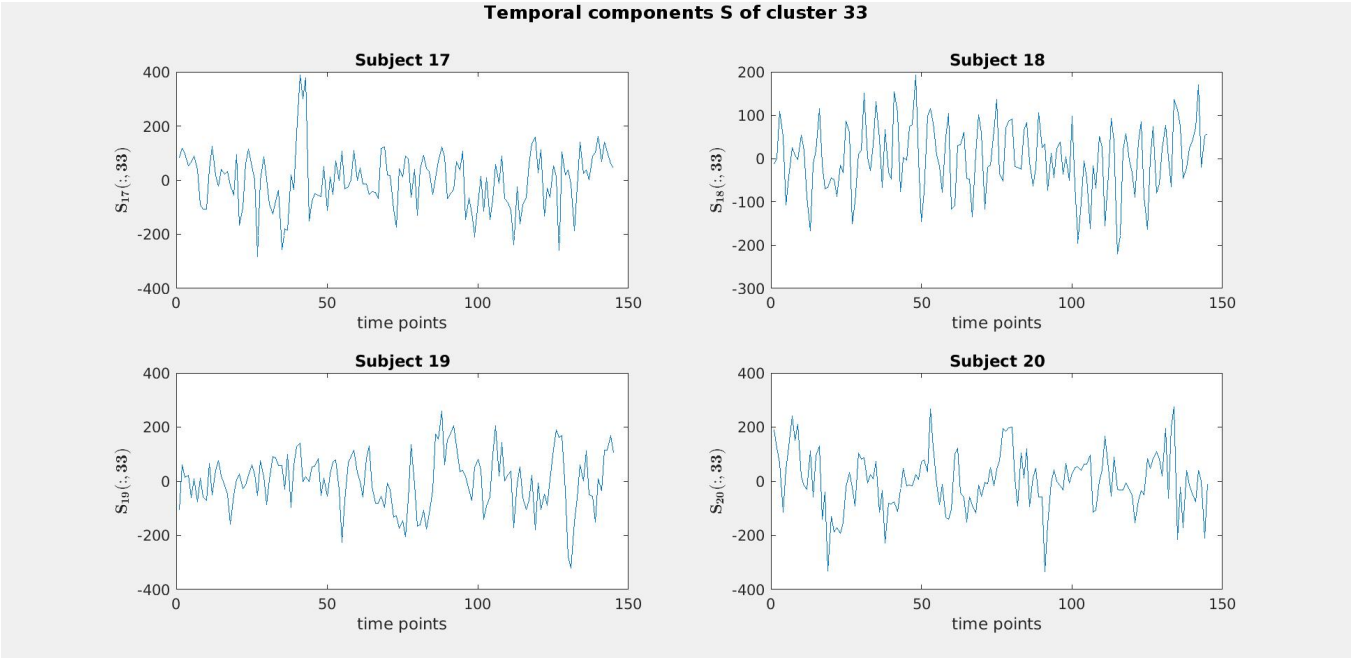


Figure B.300: Waveform of temporal components $\mathbf{S}_k(:, 33)$ for $k = 17, \ldots, 20$ and R=42.

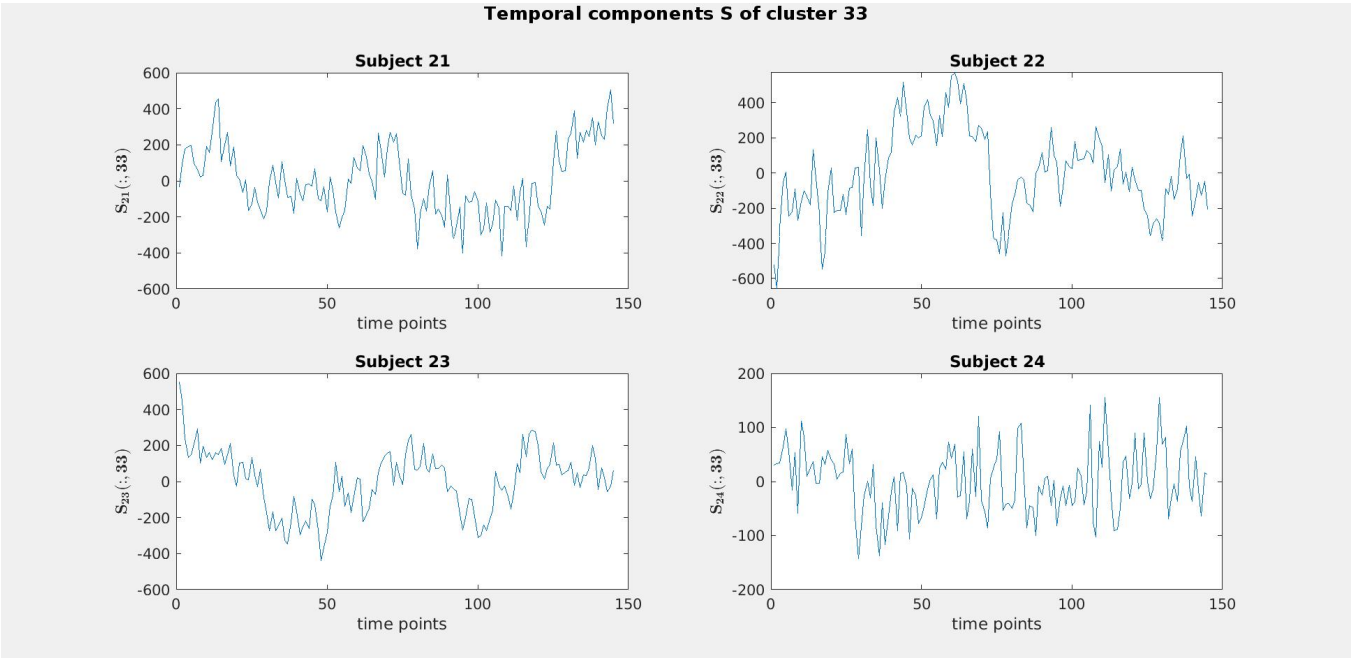


Figure B.301: Waveform of temporal components $\mathbf{S}_k(:, 133)$ for $k = 21, \ldots, 24$ and R=42.

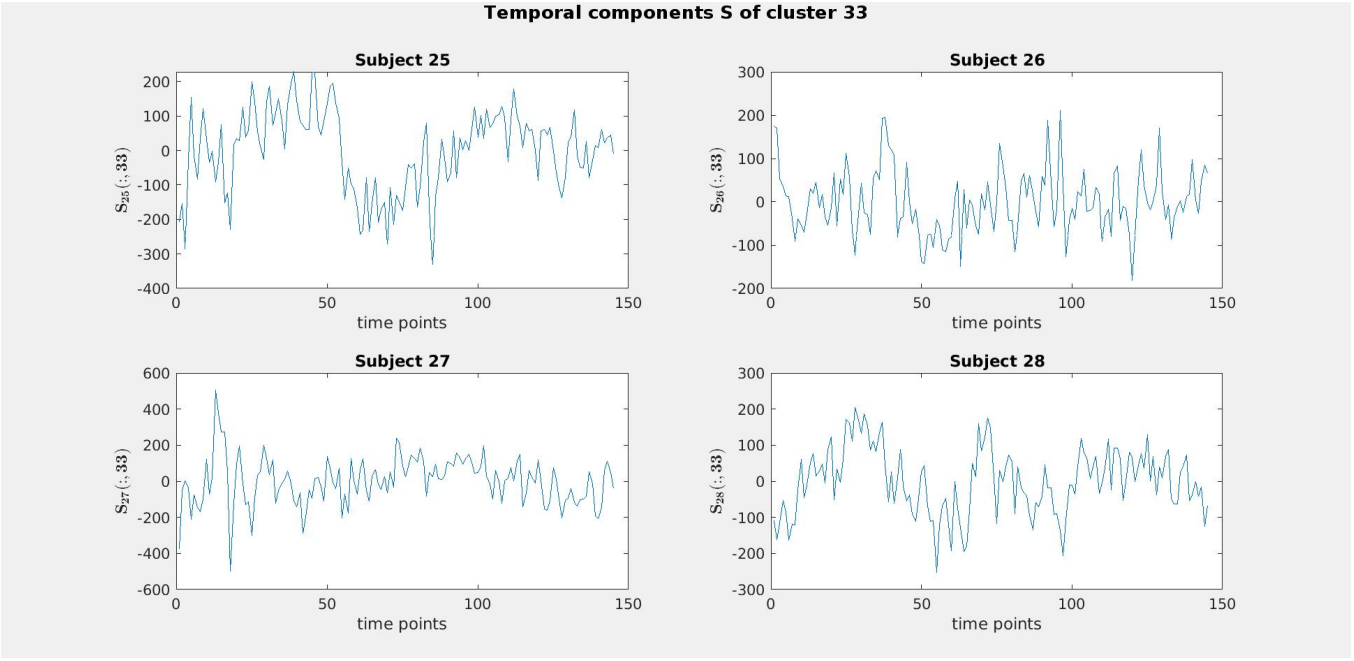


Figure B.302: Waveform of temporal components $\mathbf{S}_k(:, 33)$ for $k = 25, \ldots, 28$ and R=42.

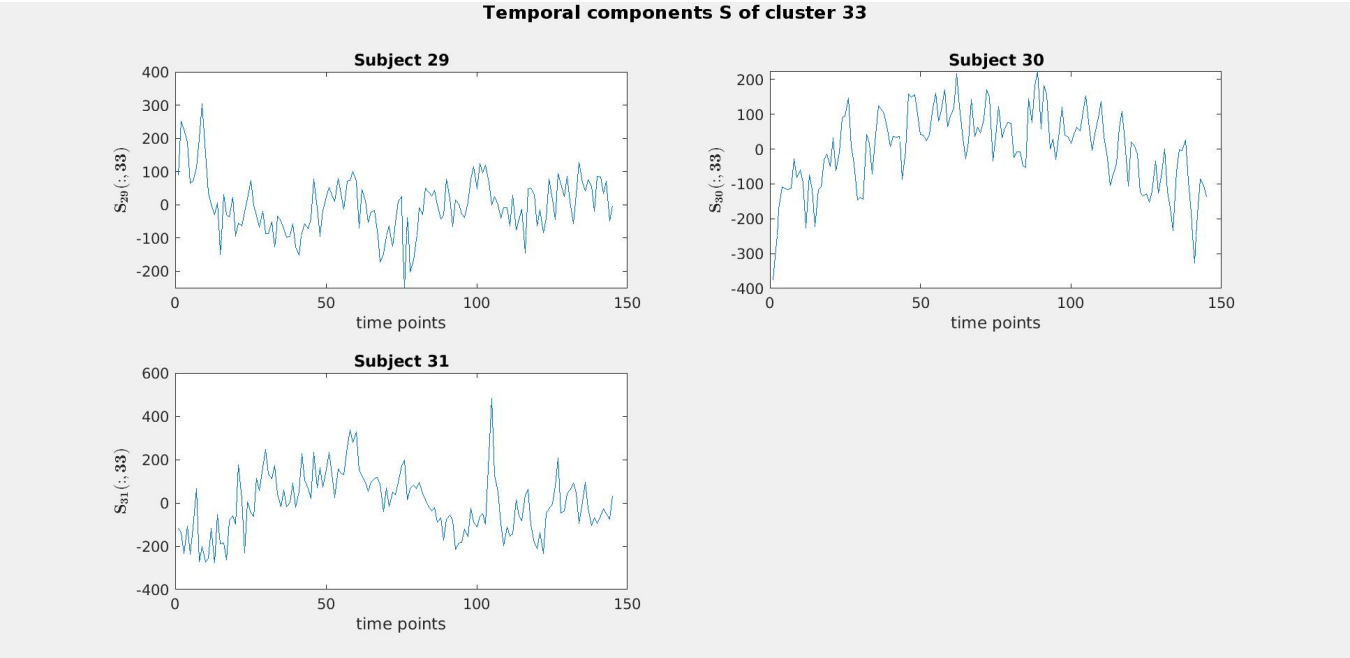


Figure B.303: Waveform of temporal components $\mathbf{S}_k(:, 33)$ for $k = 29, \ldots, 31$ and R=42.

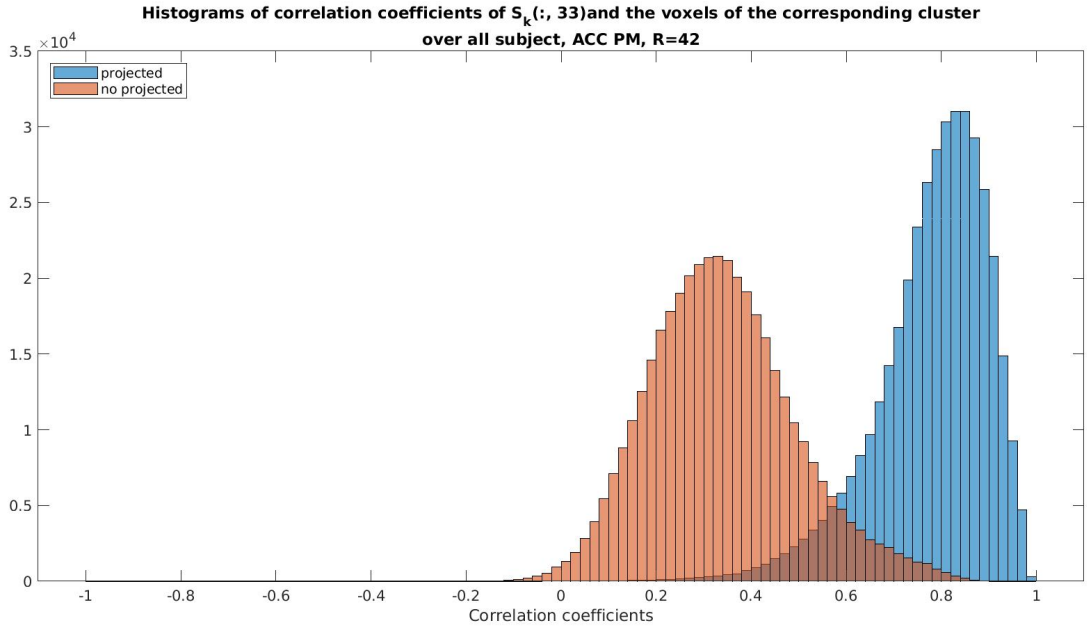


Figure B.304: For the column 33 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 33)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

2. **Cluster 5**
   Cluster 5 contains 11207 voxels, which correspond to 4.959991% of the whole brain (background, soft tissue, bone are excluded).

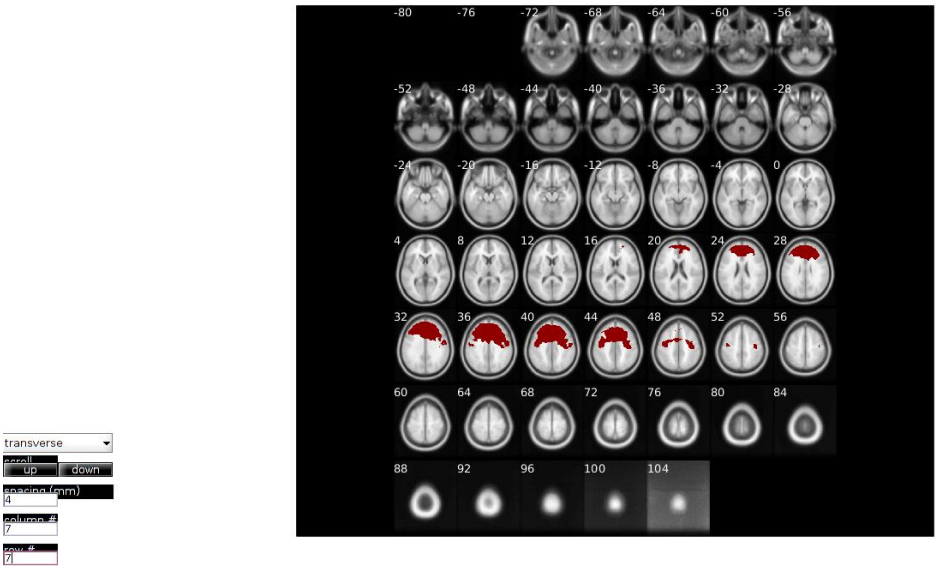


Figure B.305: Cluster 5. Red areas indicate the voxels which belong to cluster 5.

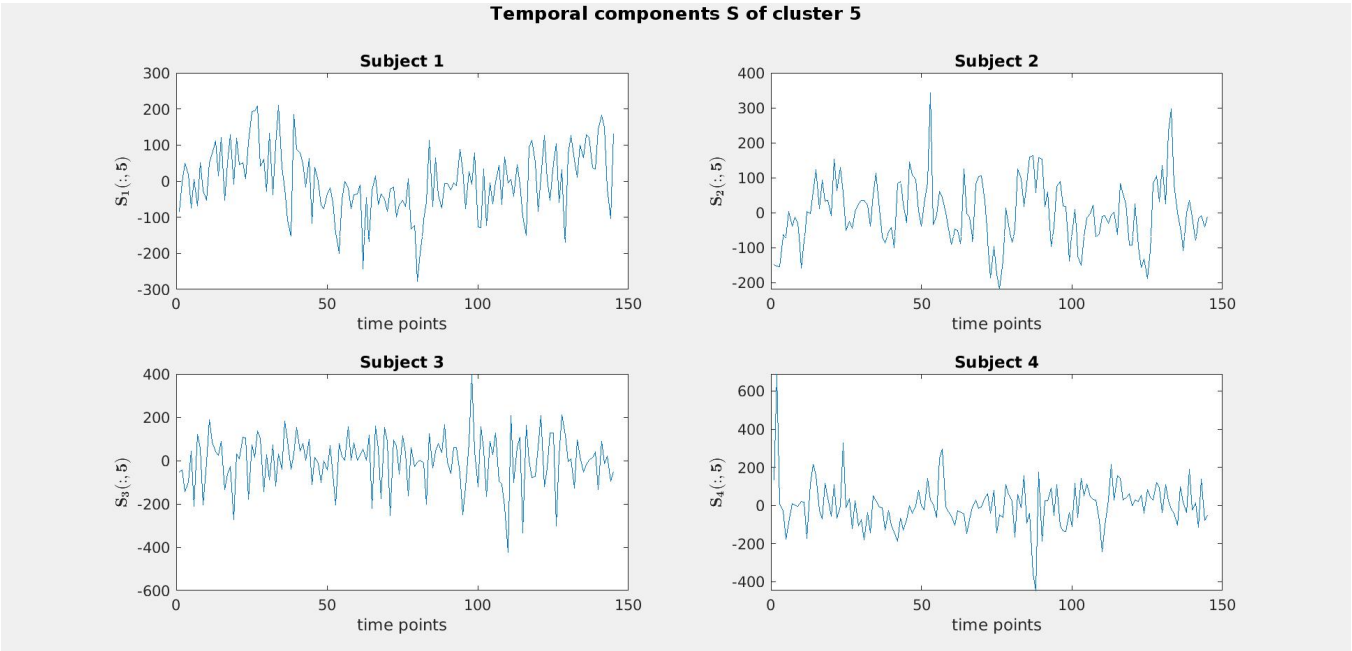


Figure B.306: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 1,\ldots,4$ and R=42.

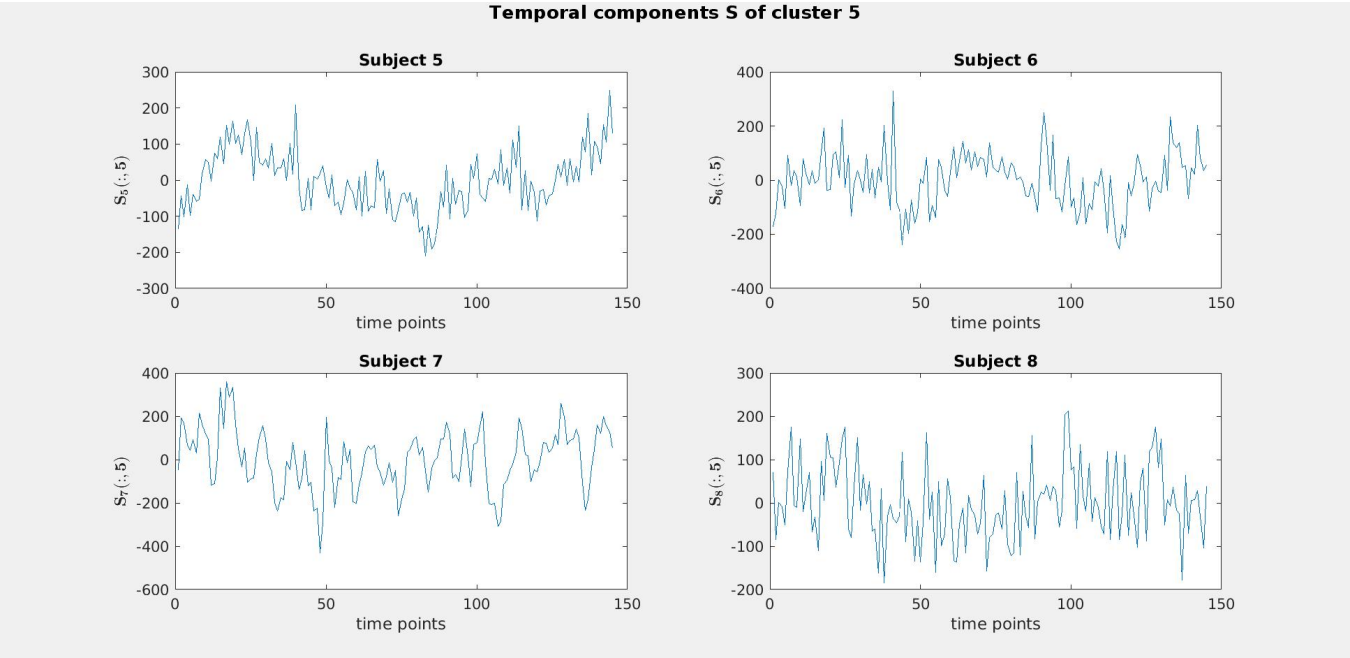


Figure B.307: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 5,\ldots,8$ and R=42.

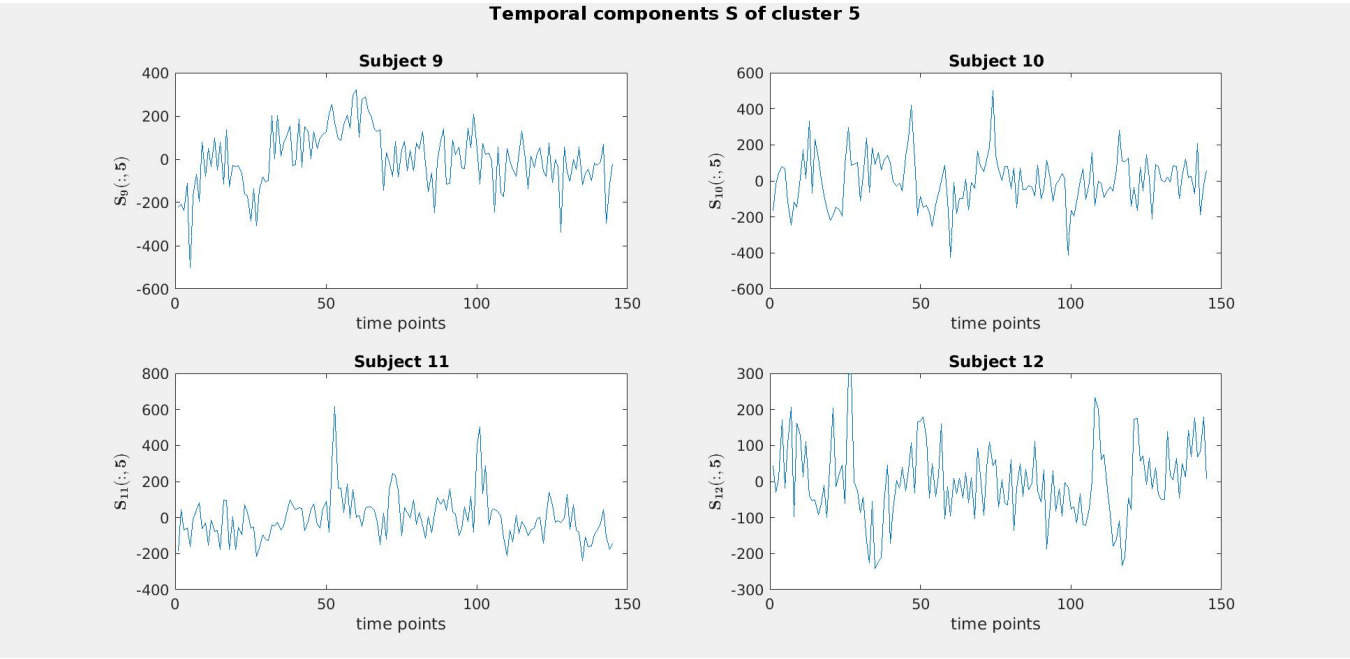


Figure B.308: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 9,\ldots,12$ and R=42.

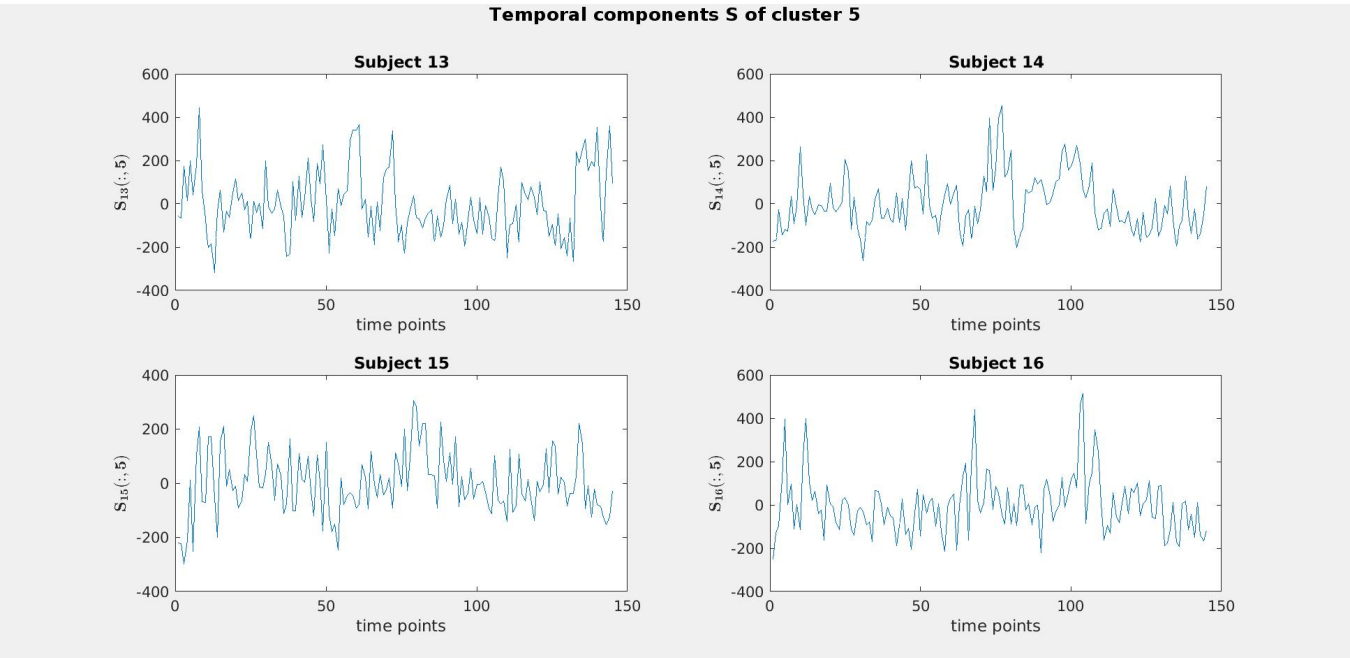


Figure B.309: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 13,\ldots,16$ and R=42.

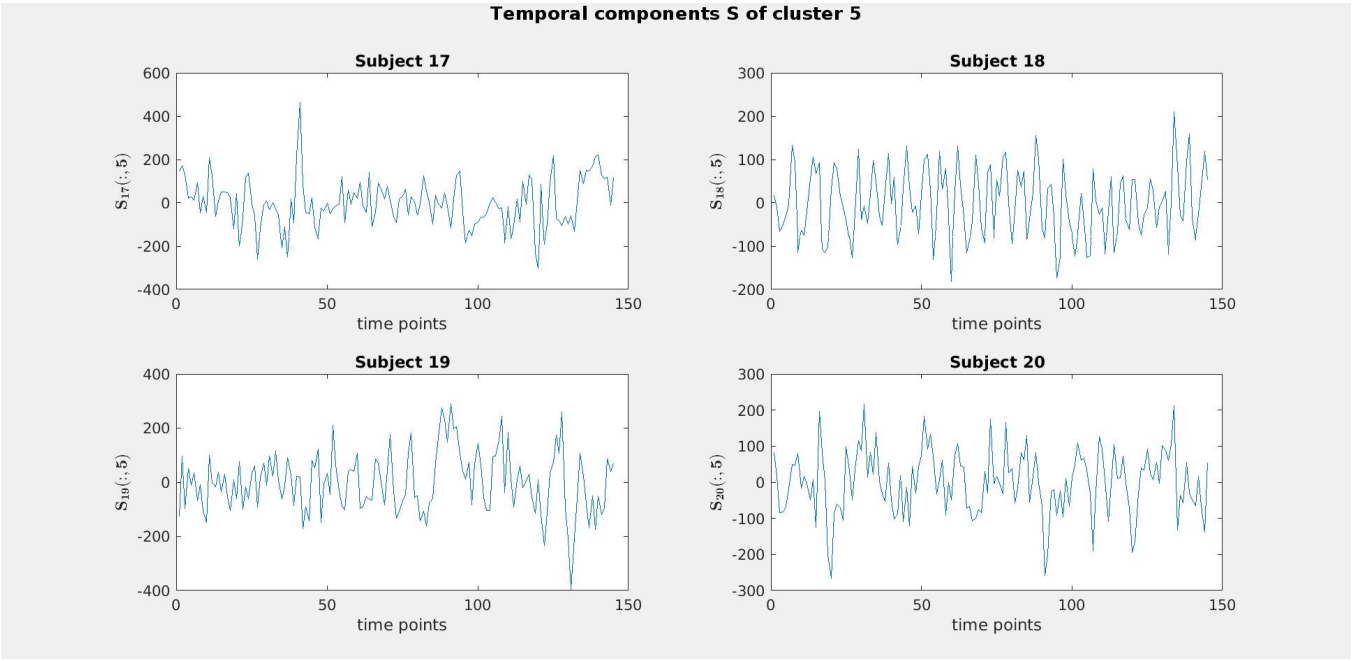


Figure B.310: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 17,\ldots,20$ and R=42.

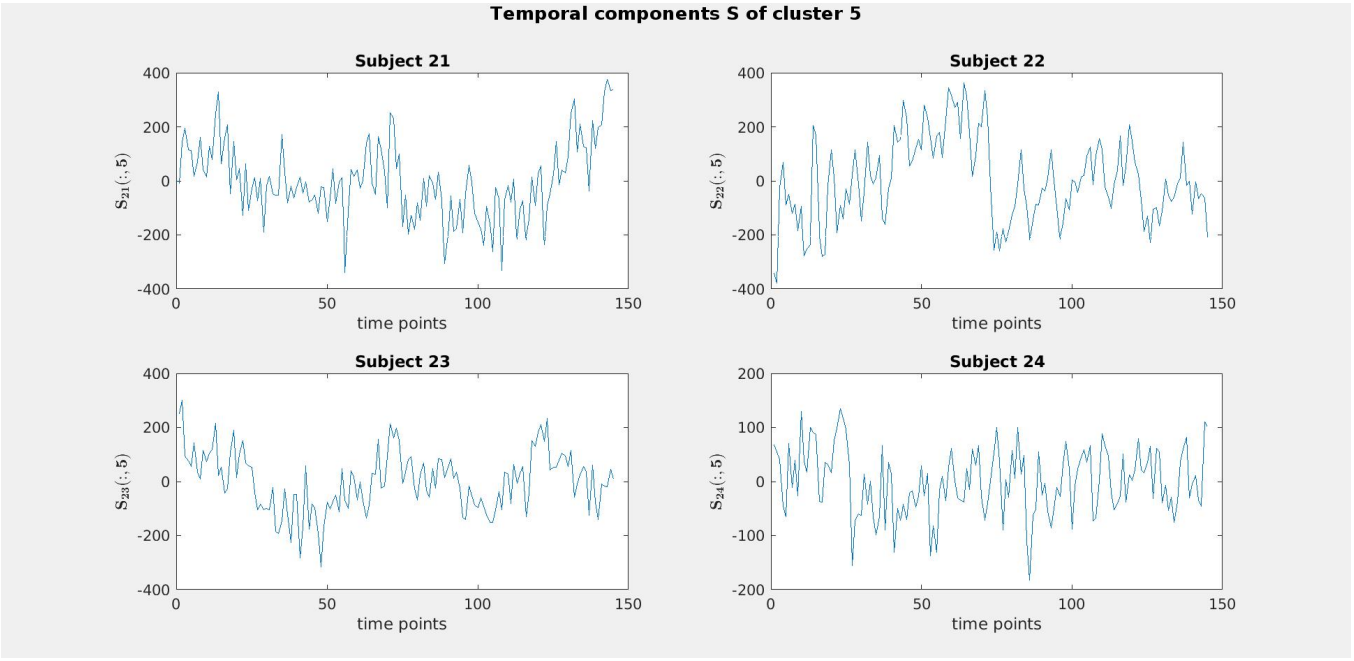


Figure B.311: Waveform of temporal components $\mathbf{S}_k(:,15)$ for $k = 21,\ldots,24$ and R=42.

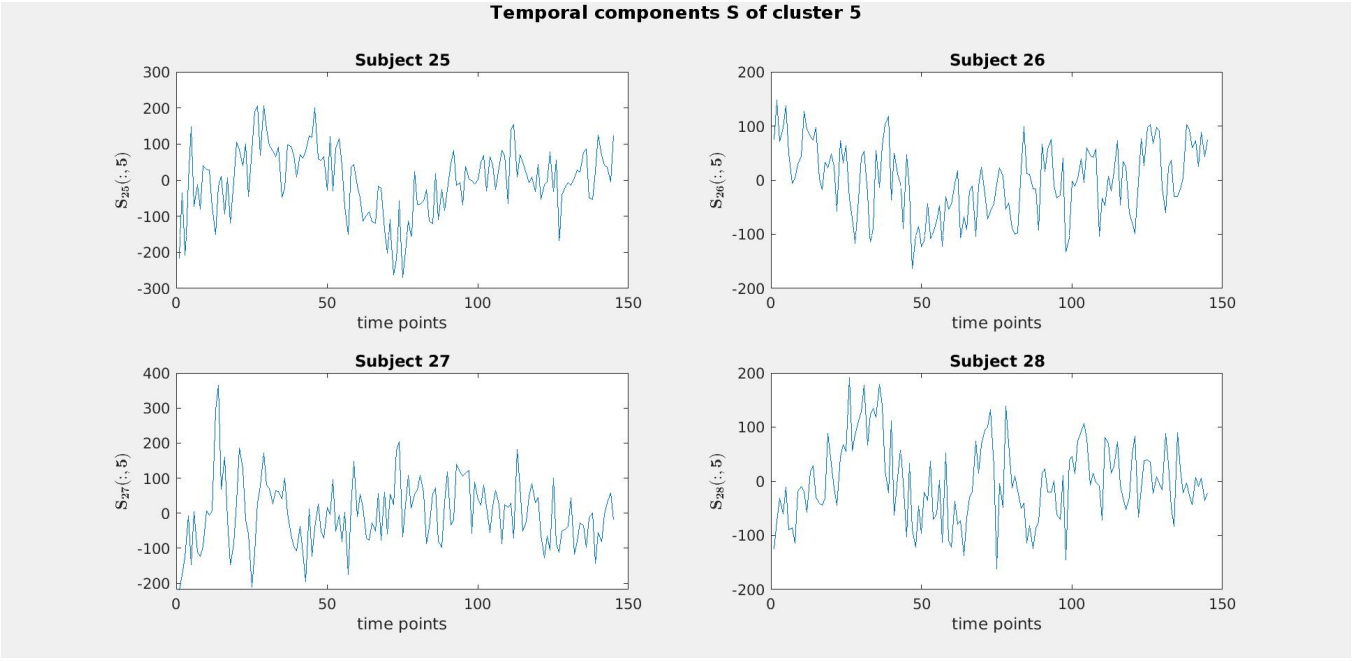


Figure B.312: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 25,\ldots,28$ and R=42.

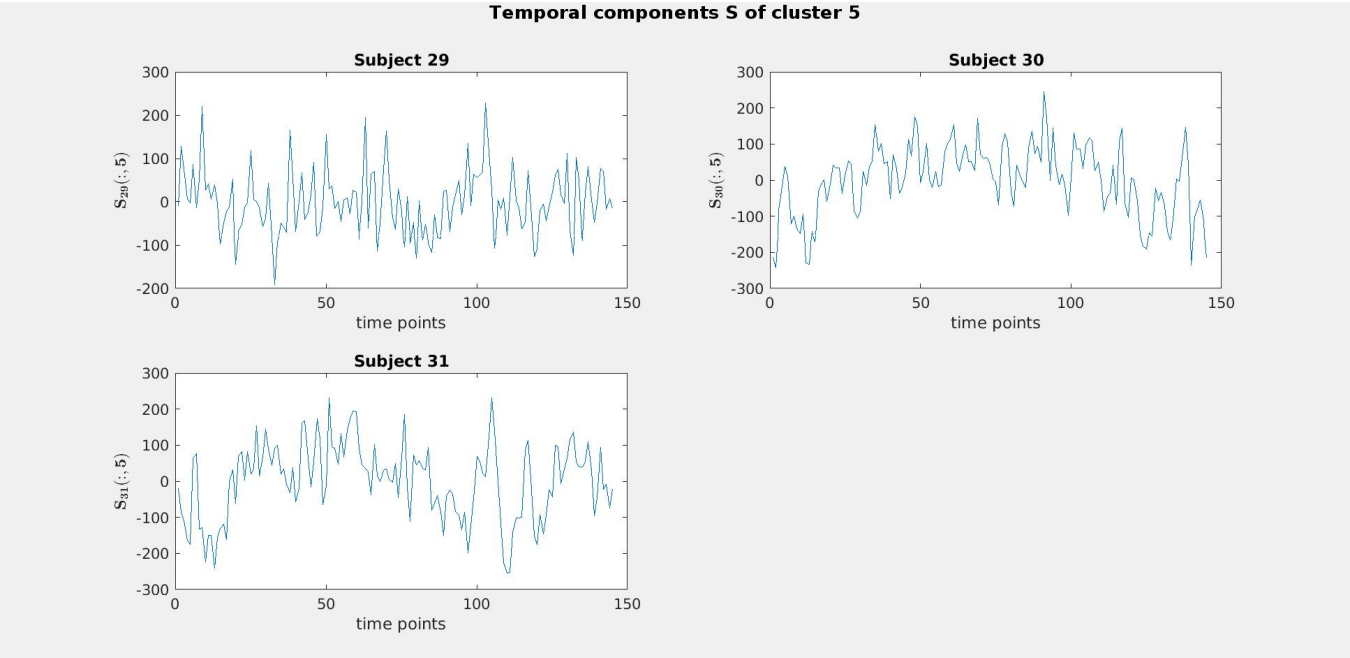


Figure B.313: Waveform of temporal components $\mathbf{S}_k(:,5)$ for $k = 29, \ldots, 31$ and R=42.

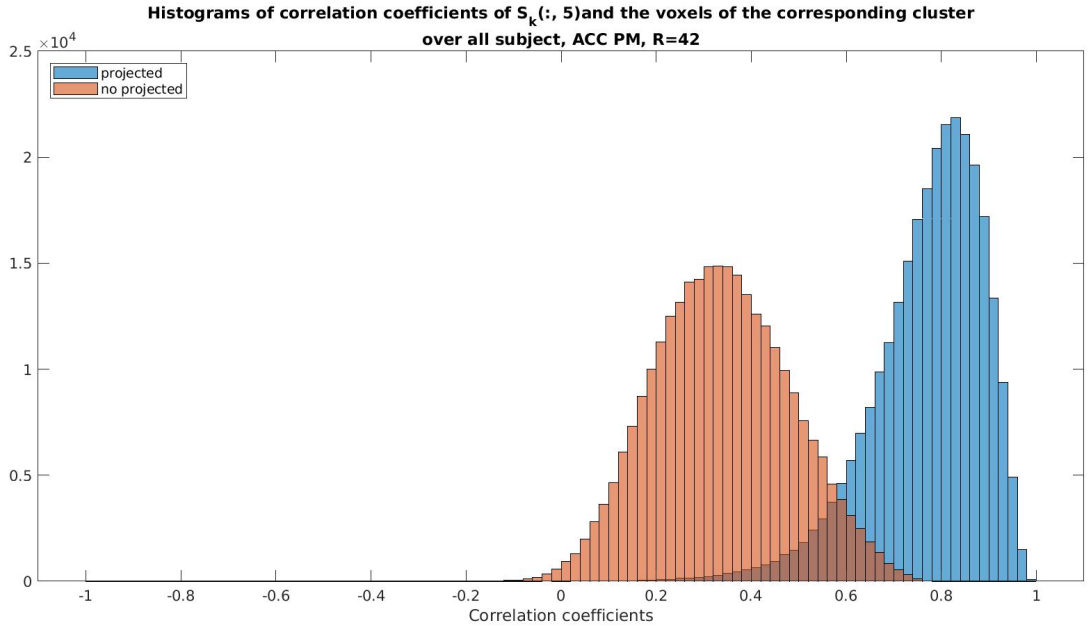


Figure B.314: For the column 5 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,5)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

3. **Cluster 42**
   Cluster 42 contains 10756 voxels, which correspond to 4.760387% of the whole brain (background, soft tissue, bone are excluded).

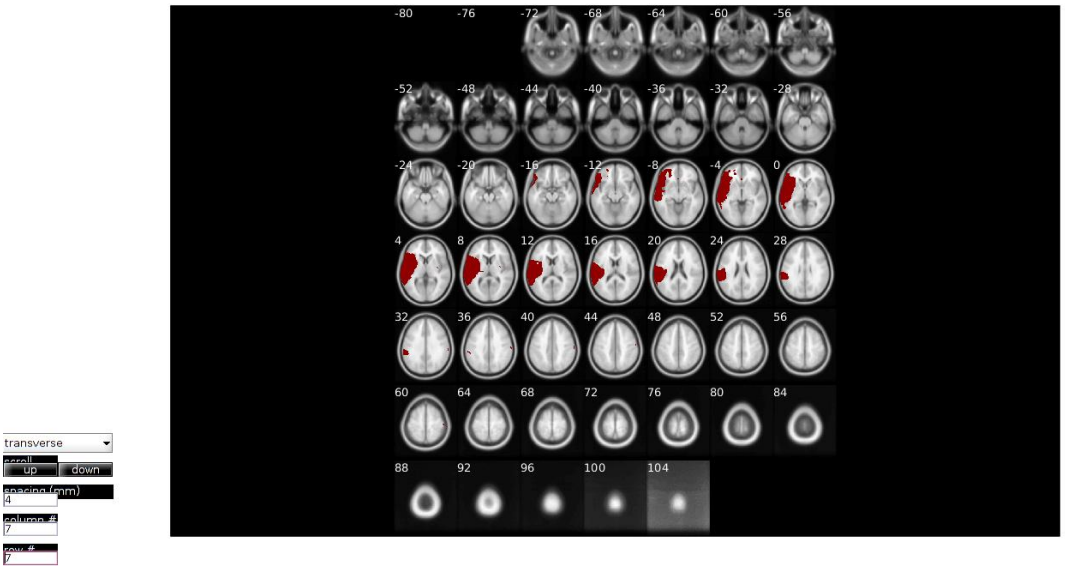


Figure B.315: Cluster 42. Red areas indicate the voxels which belong to cluster 42.

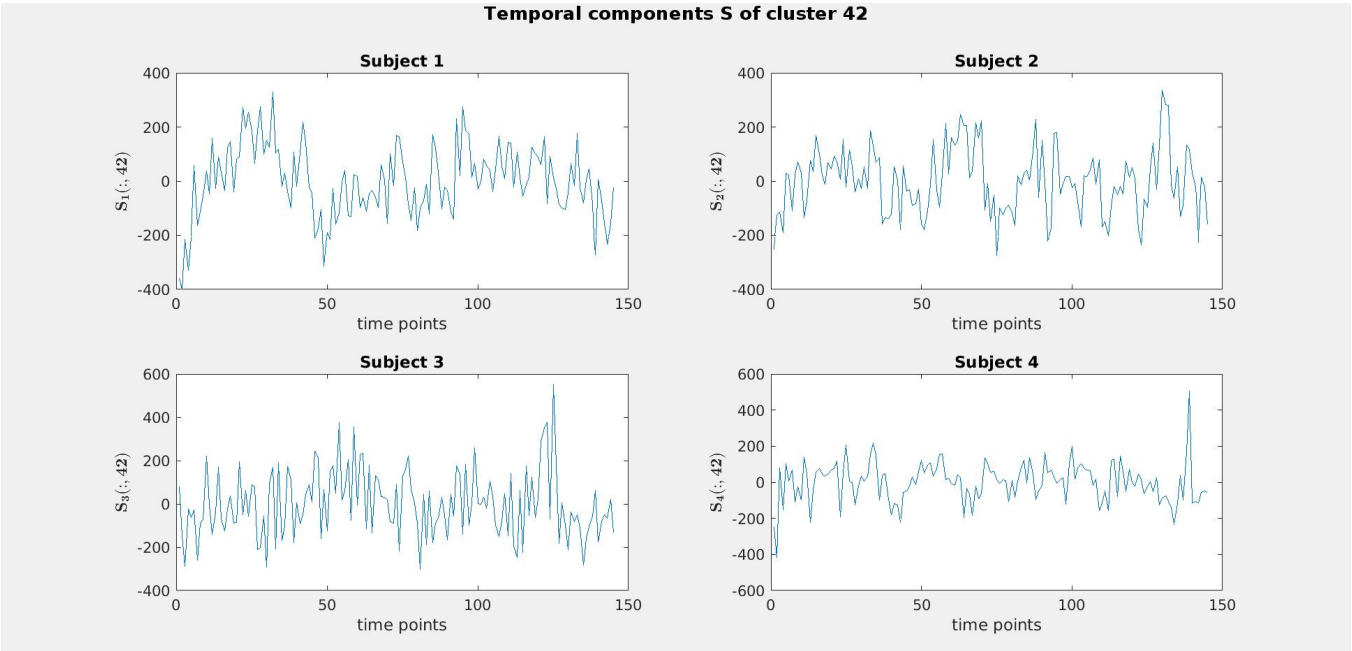


Figure B.316: Waveform of temporal components $\mathbf{S}_k(:, 42)$ for $k = 1, \ldots, 4$ and R=42.

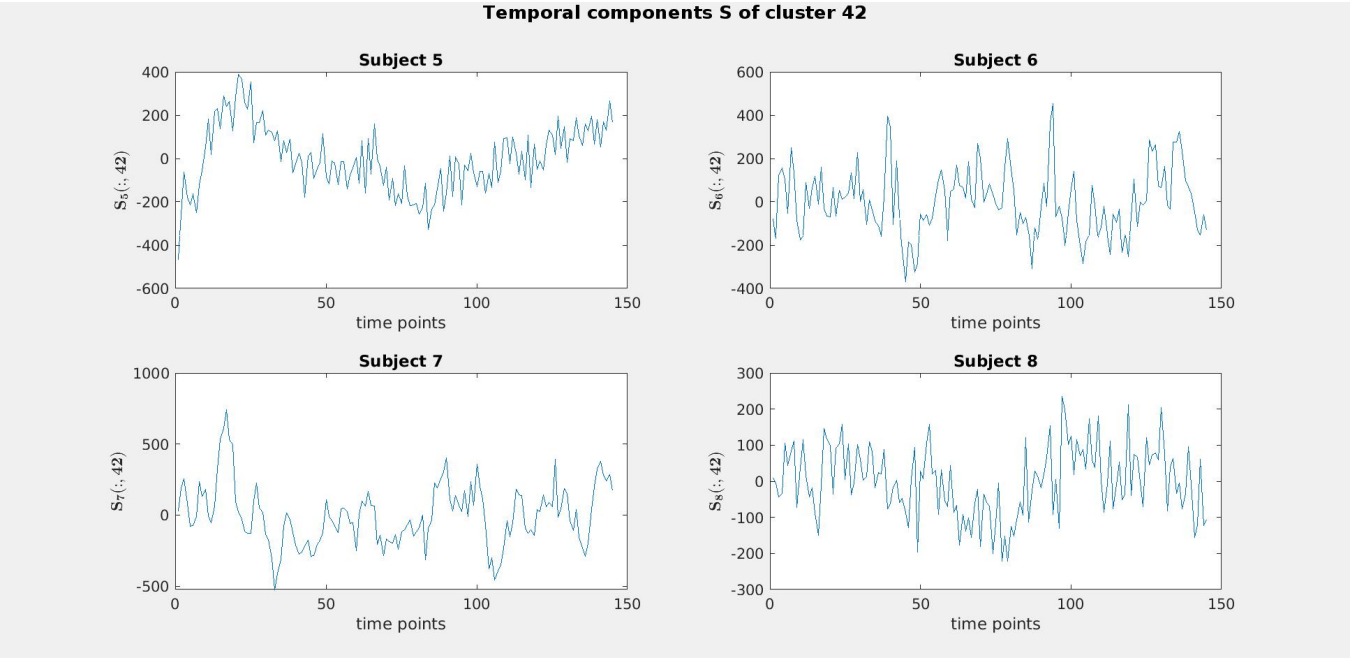


Figure B.317: Waveform of temporal components $\mathbf{S}_k(:, 42)$ for $k = 5, \ldots, 8$ and R=42.

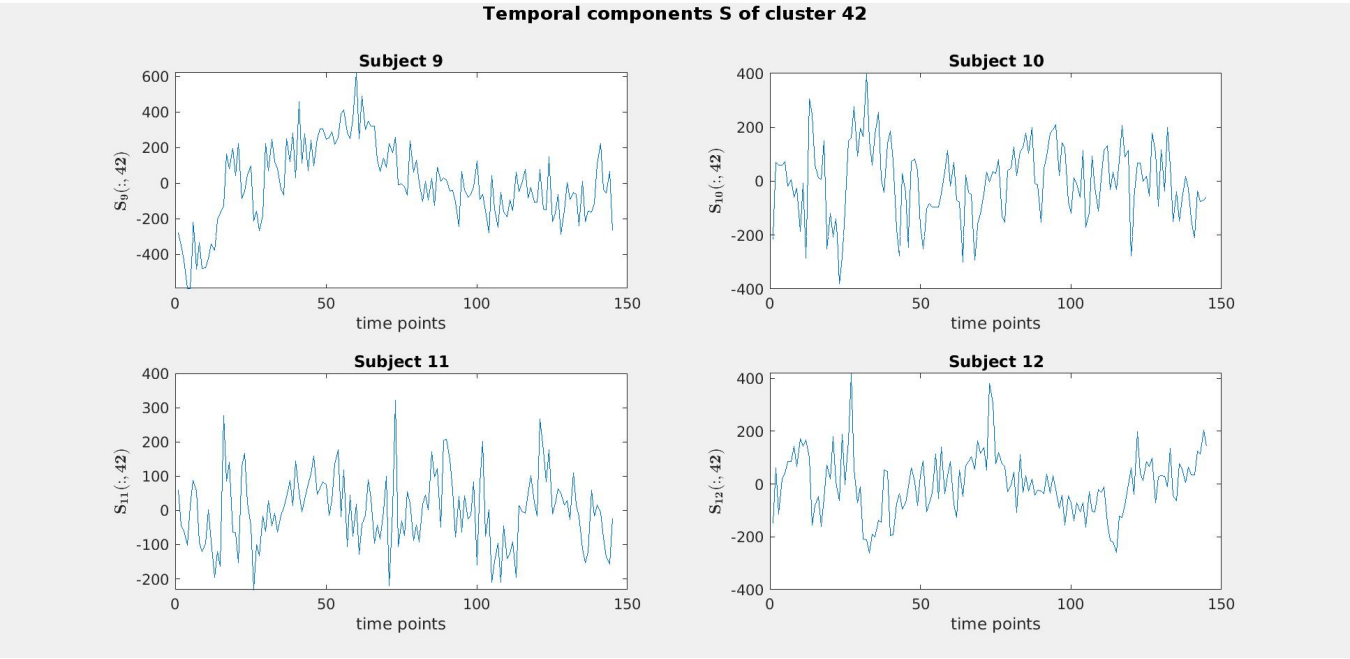


Figure B.318: Waveform of temporal components $\mathbf{S}_k(:, 42)$ for $k = 9, \ldots, 12$ and R=42.

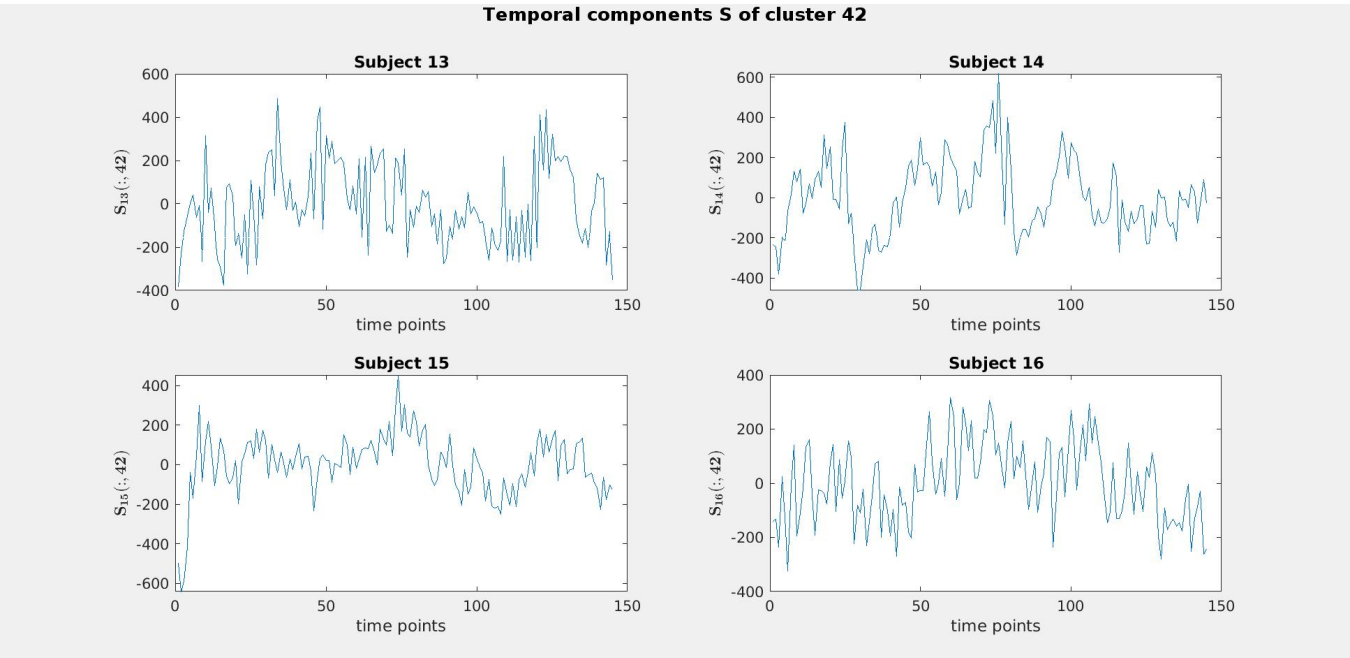


Figure B.319: Waveform of temporal components $\mathbf{S}_k(:, 42)$ for $k = 13, \ldots, 16$ and R=42.

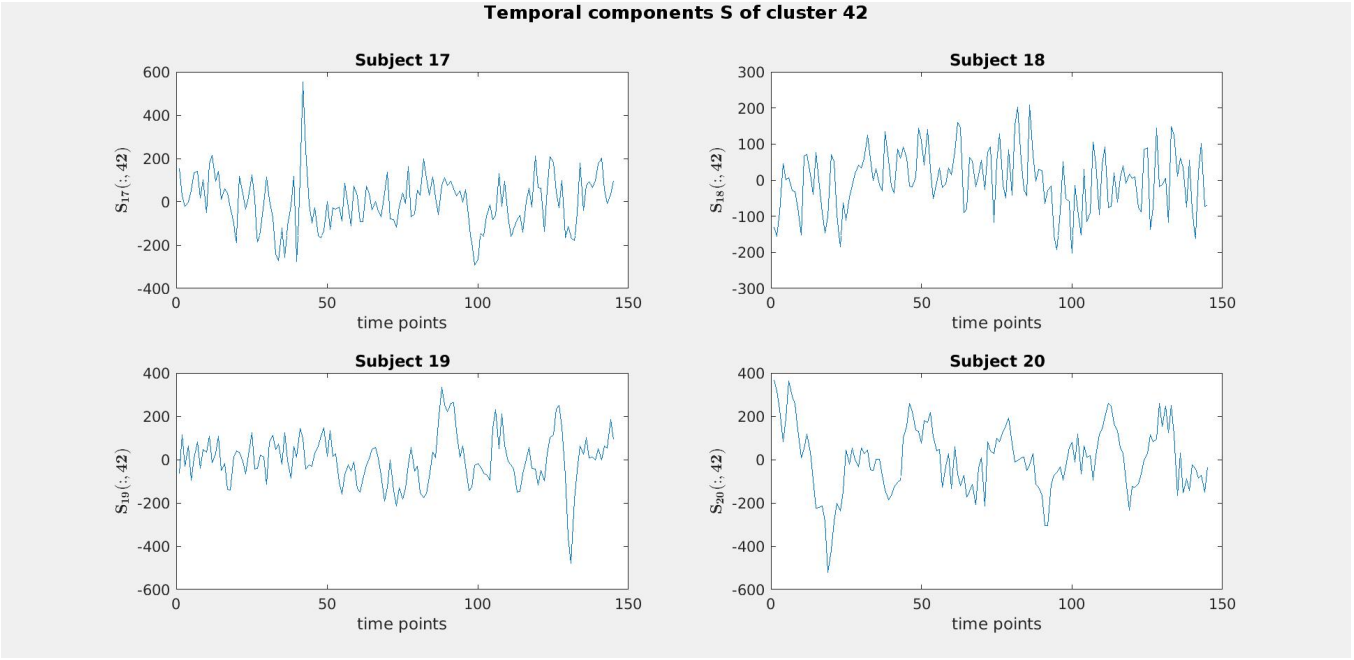


Figure B.320: Waveform of temporal components $\mathbf{S}_k(:, 42)$ for $k = 17, \ldots, 20$ and R=42.

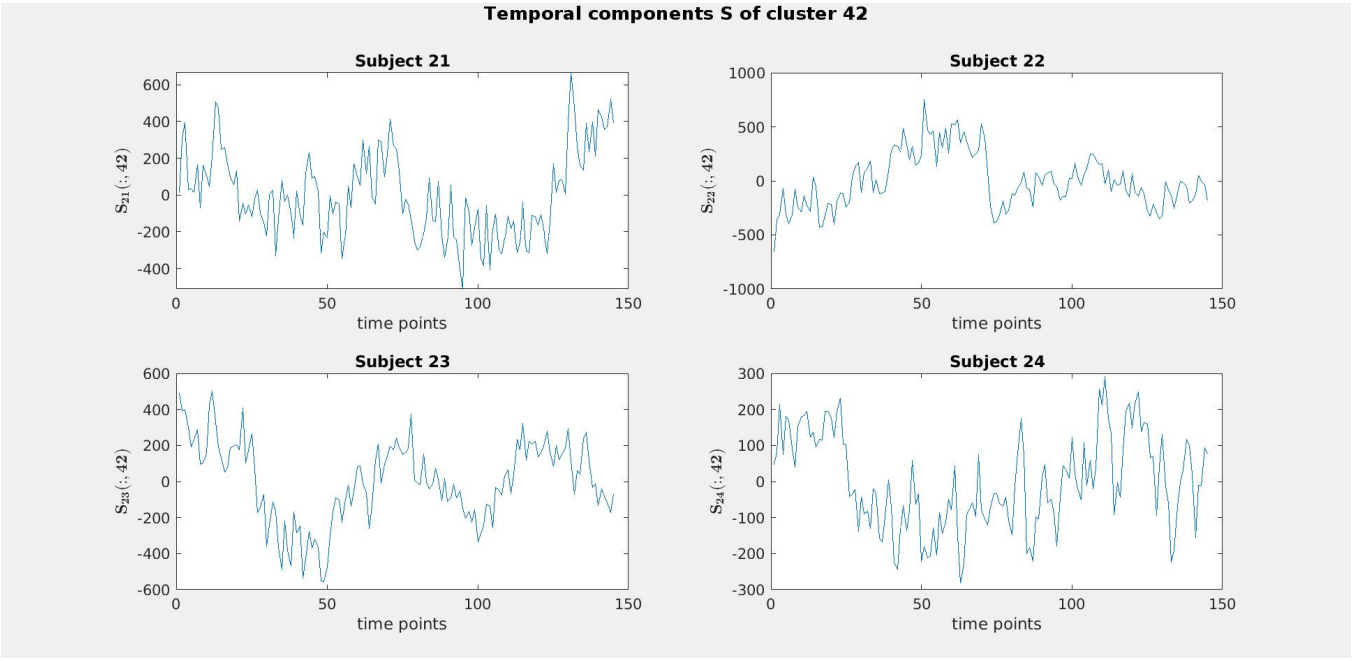


Figure B.321: Waveform of temporal components $\mathbf{S}_k(:, 142)$ for $k = 21, \ldots, 24$ and R=42.

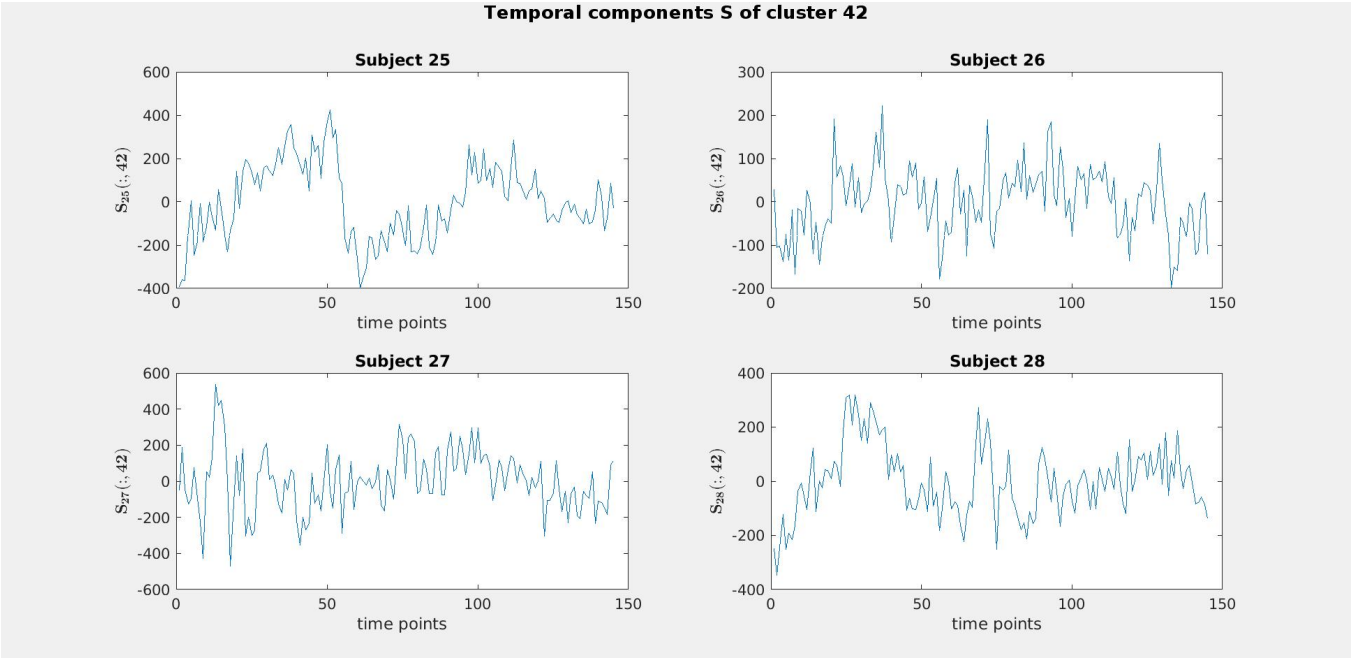


Figure B.322: Waveform of temporal components $\mathbf{S}_k(:, 42)$ for $k = 25, \ldots, 28$ and R=42.

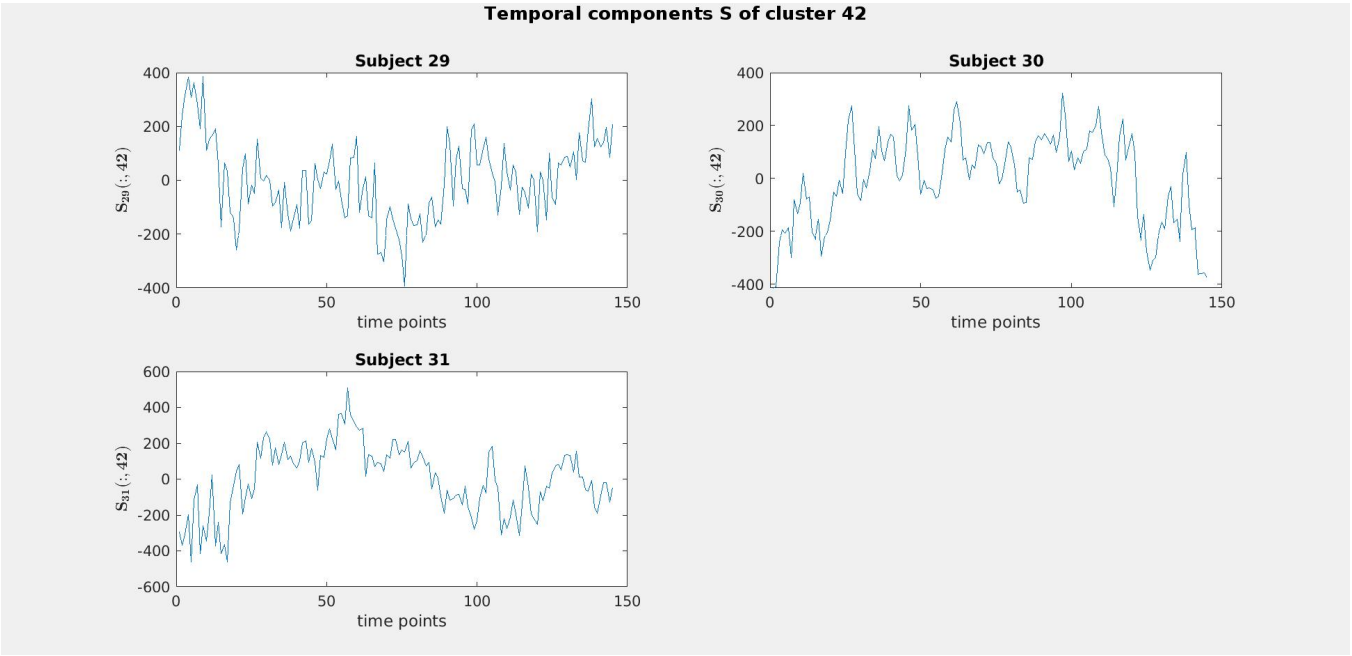


Figure B.323: Waveform of temporal components $\mathbf{S}_k(:, 42)$ for $k = 29, \ldots, 31$ and R=42.

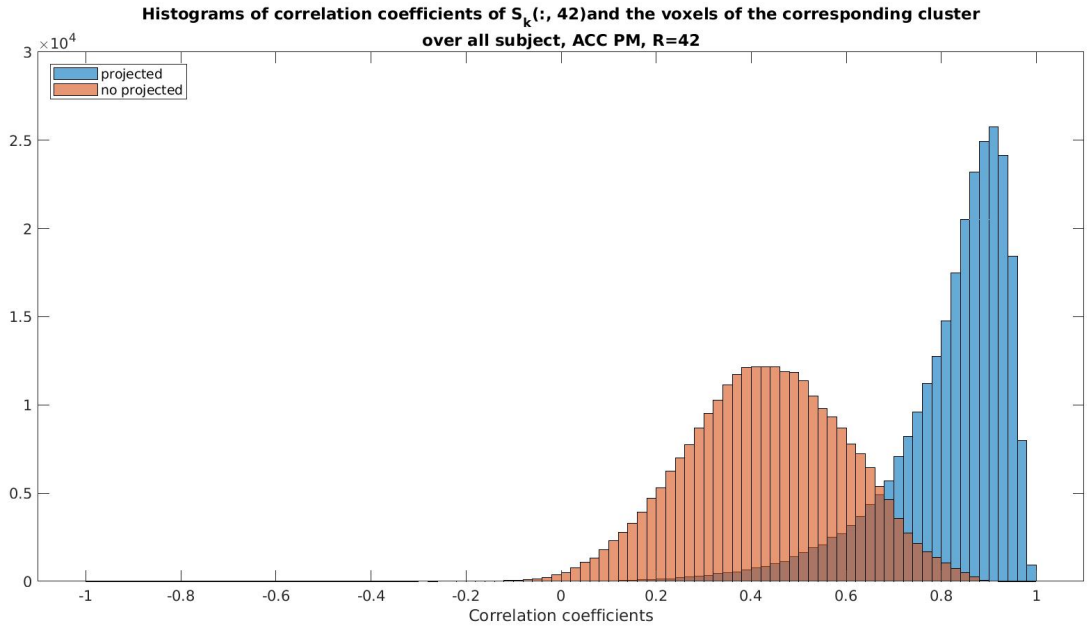


Figure B.324: For the column 42 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 42)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

4. **Cluster 3**
   Cluster 3 contains 10479 voxels, which correspond to 4.637793% of the whole brain (background, soft tissue, bone are excluded).

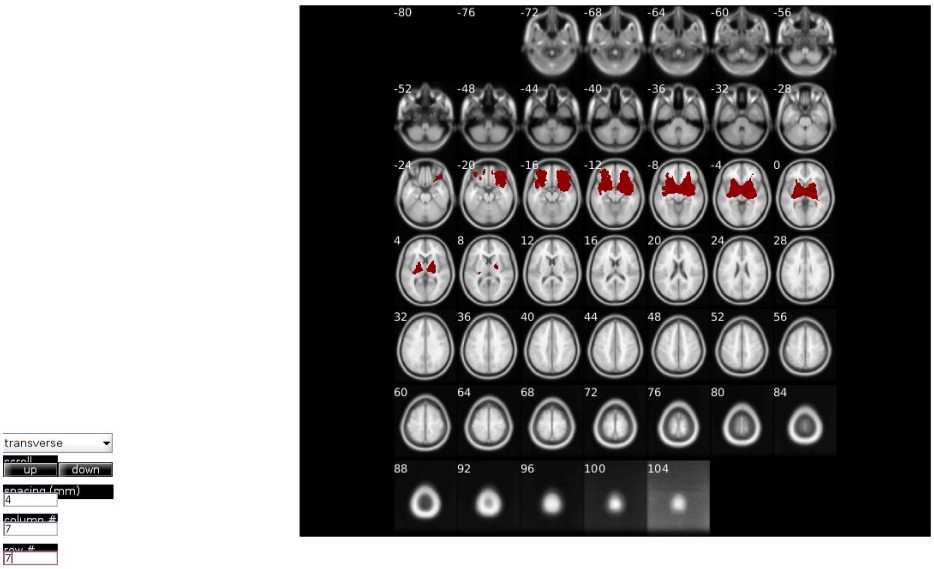


Figure B.325: Cluster 3. Red areas indicate the voxels which belong to cluster 3.

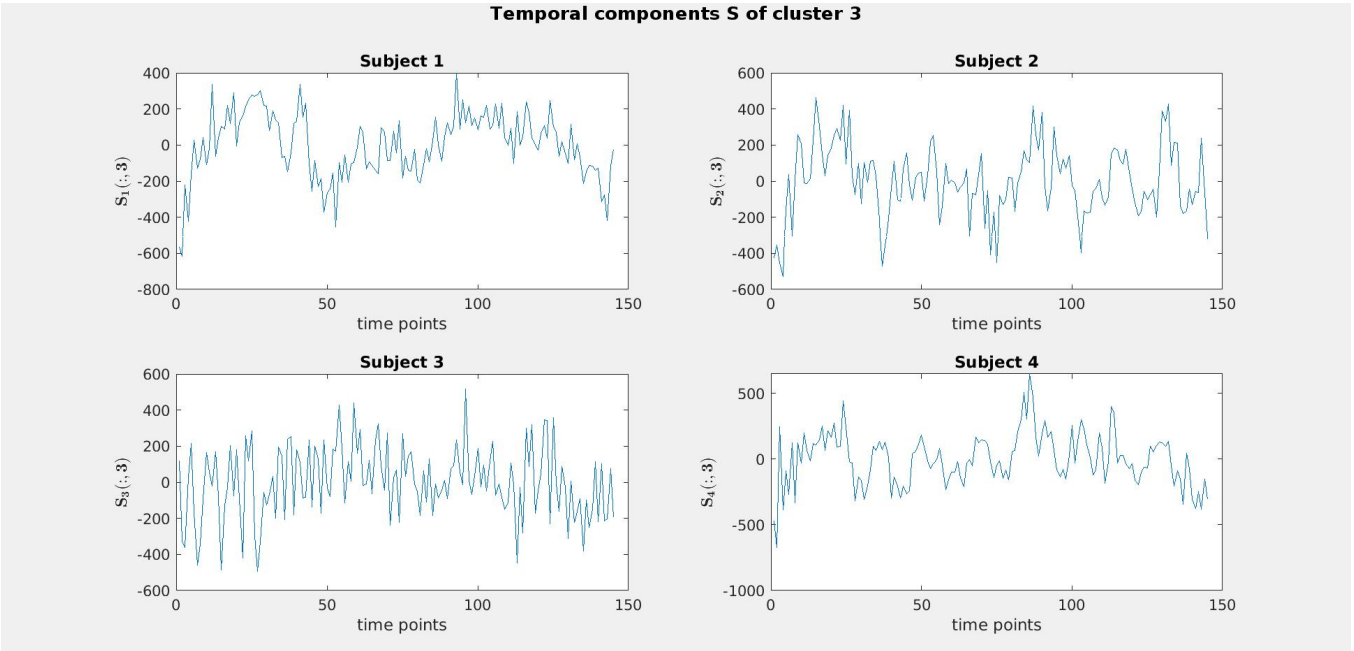


Figure B.326: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 1,\ldots,4$ and R=42.

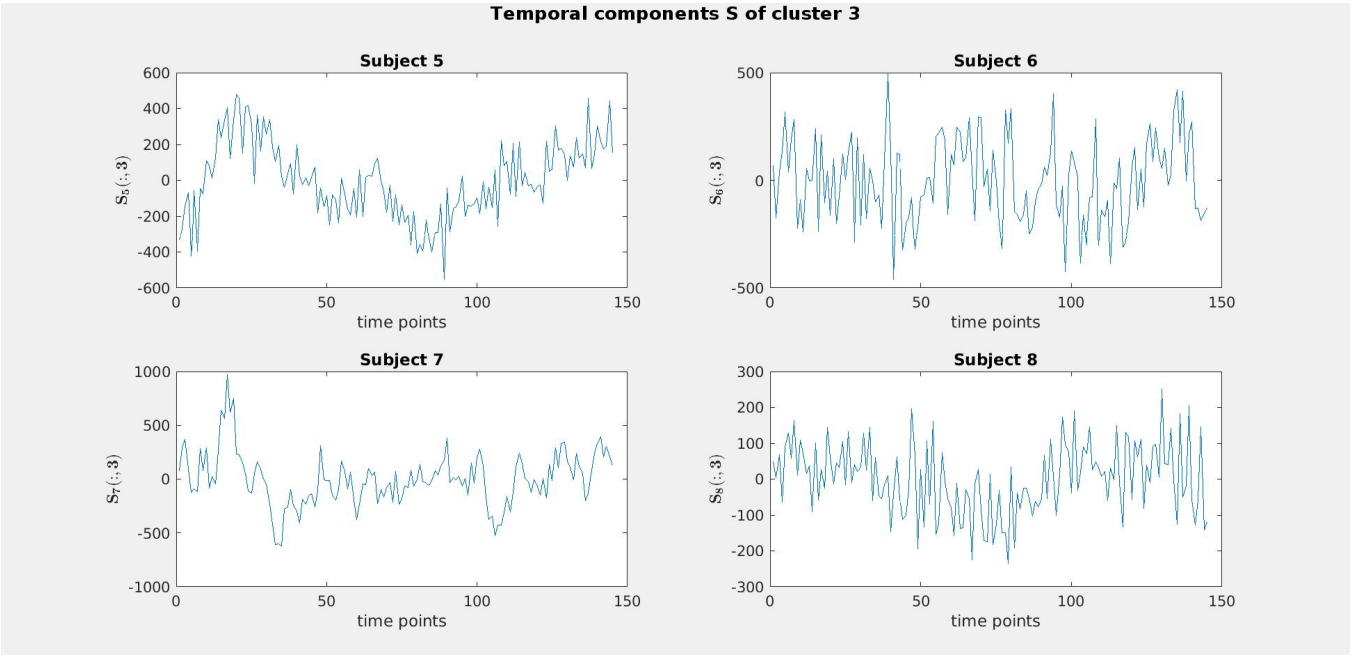


Figure B.327: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 5,\ldots,8$ and R=42.

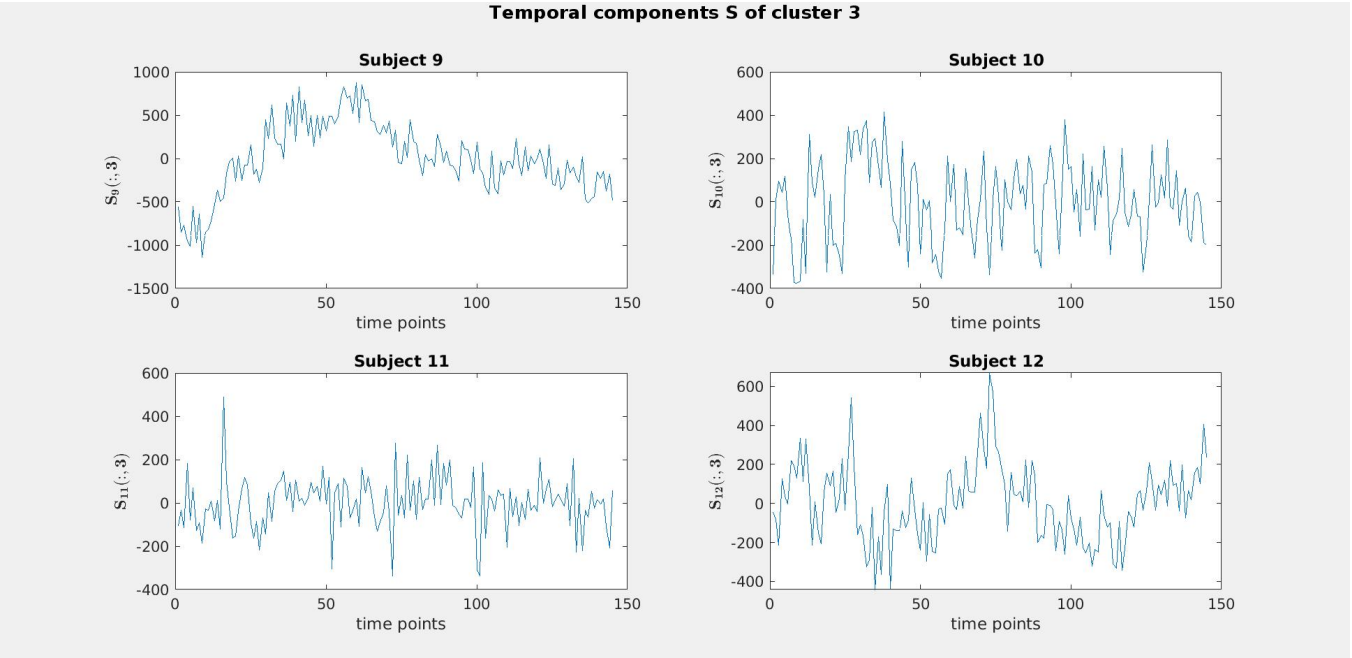


Figure B.328: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 9,\ldots,12$ and R=42.

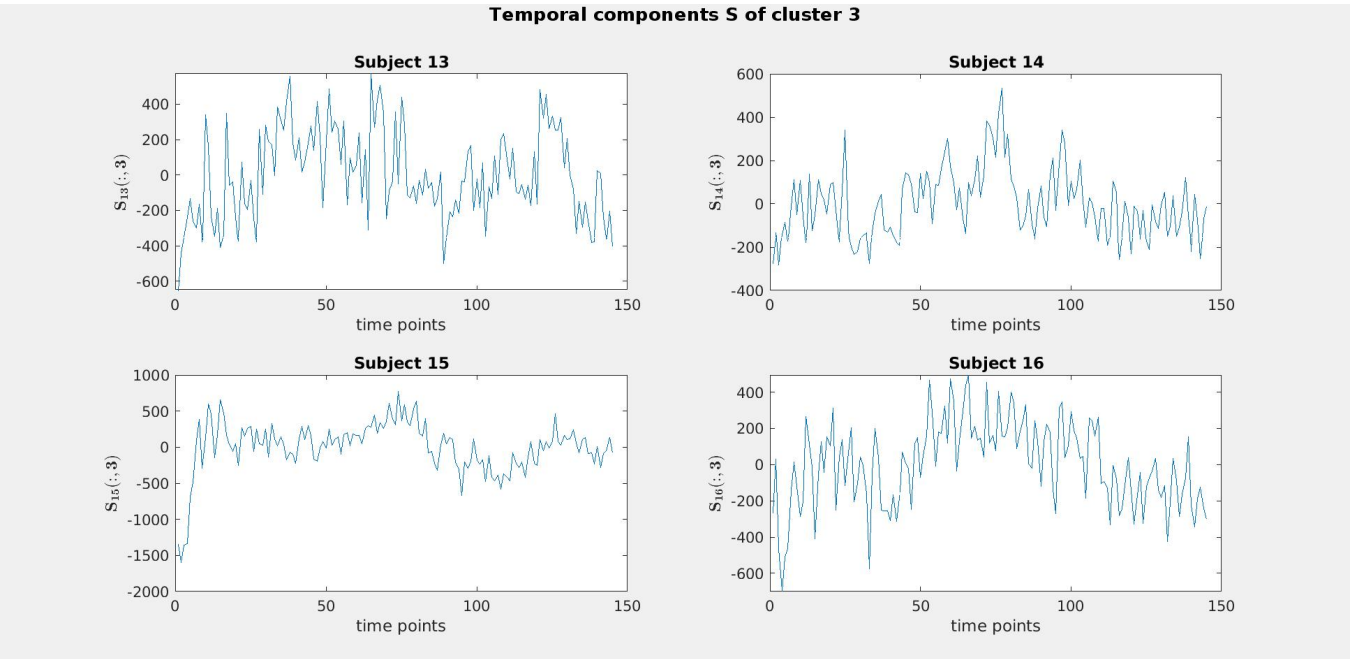


Figure B.329: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 13,\ldots,16$ and R=42.

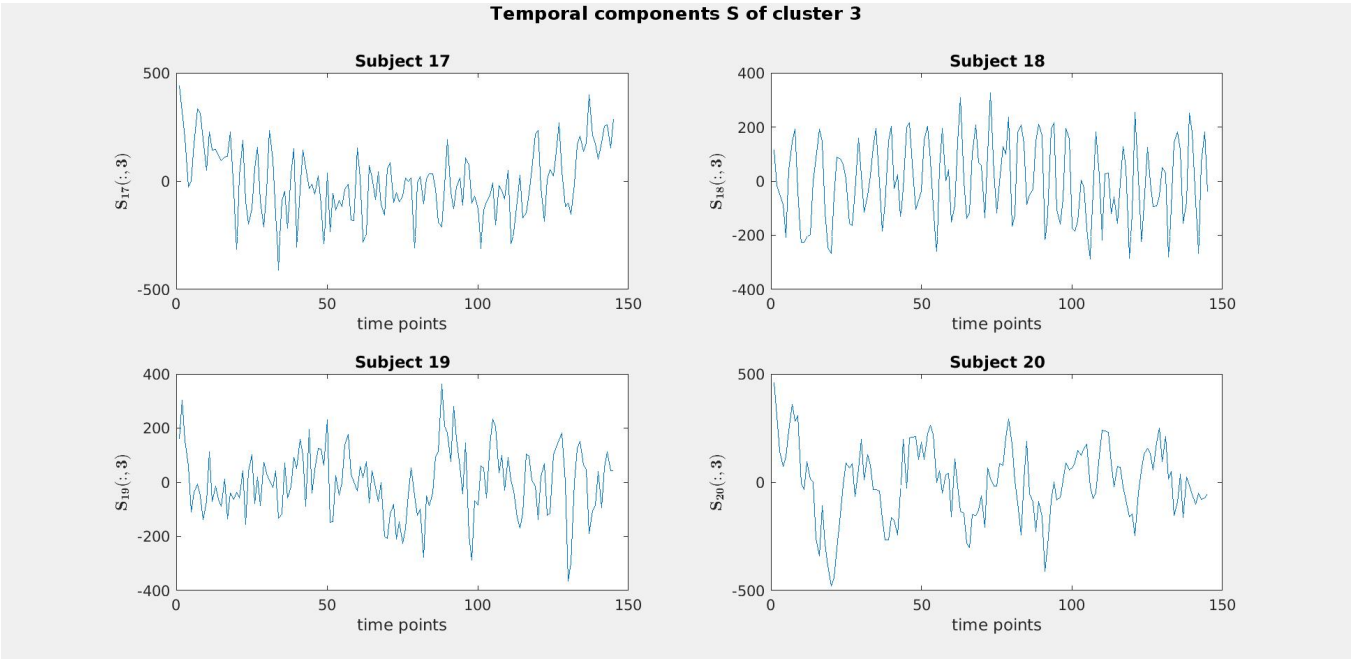


Figure B.330: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 17,\ldots,20$ and R=42.

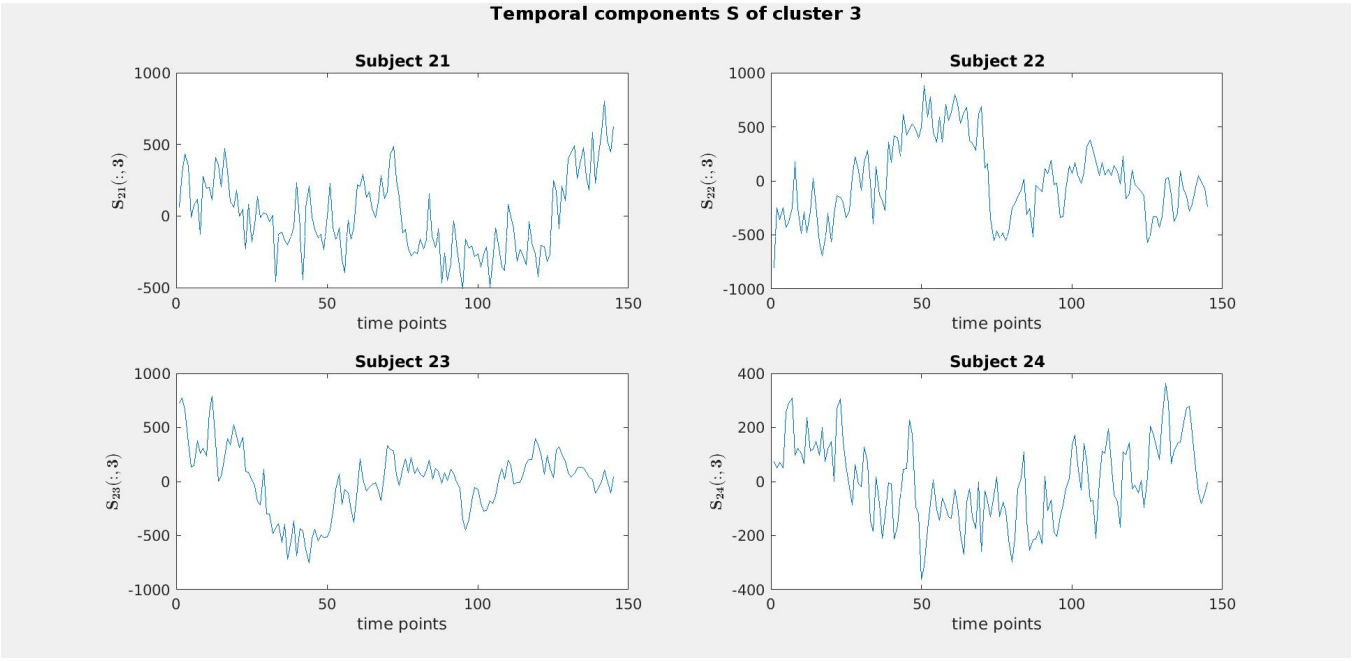


Figure B.331: Waveform of temporal components $\mathbf{S}_k(:,13)$ for $k = 21,\ldots,24$ and R=42.

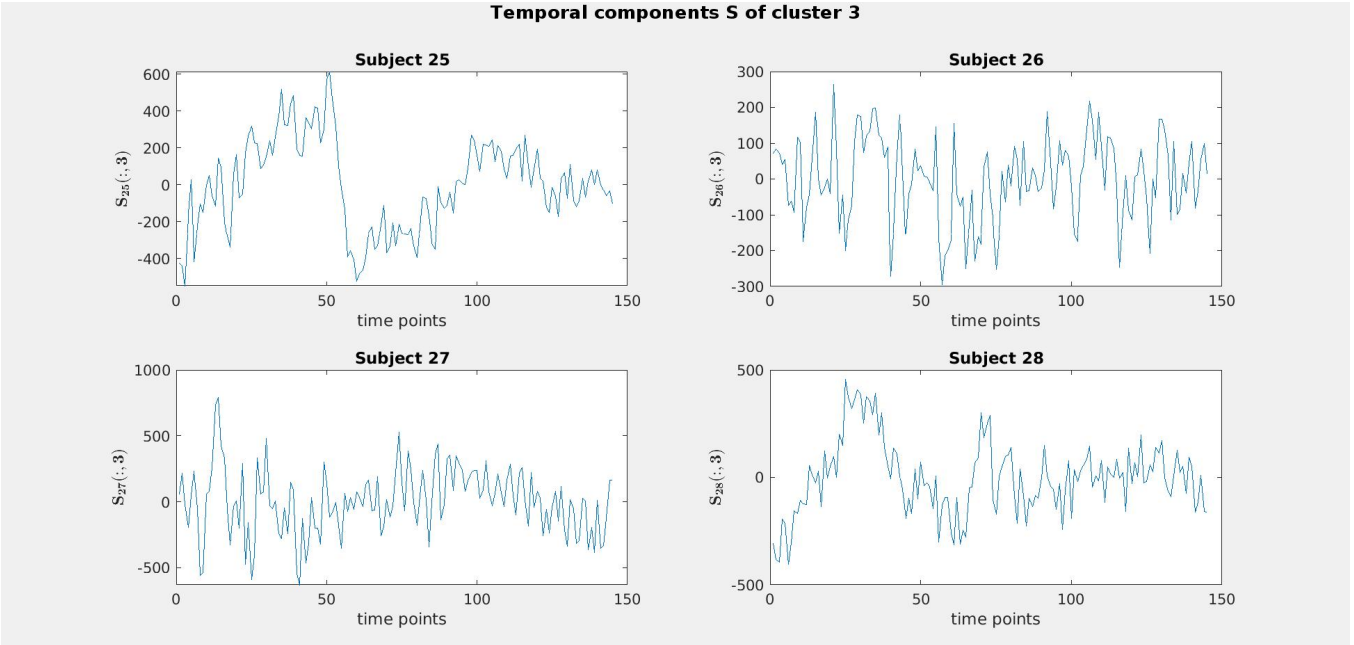


Figure B.332: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 25,\ldots,28$ and R=42.

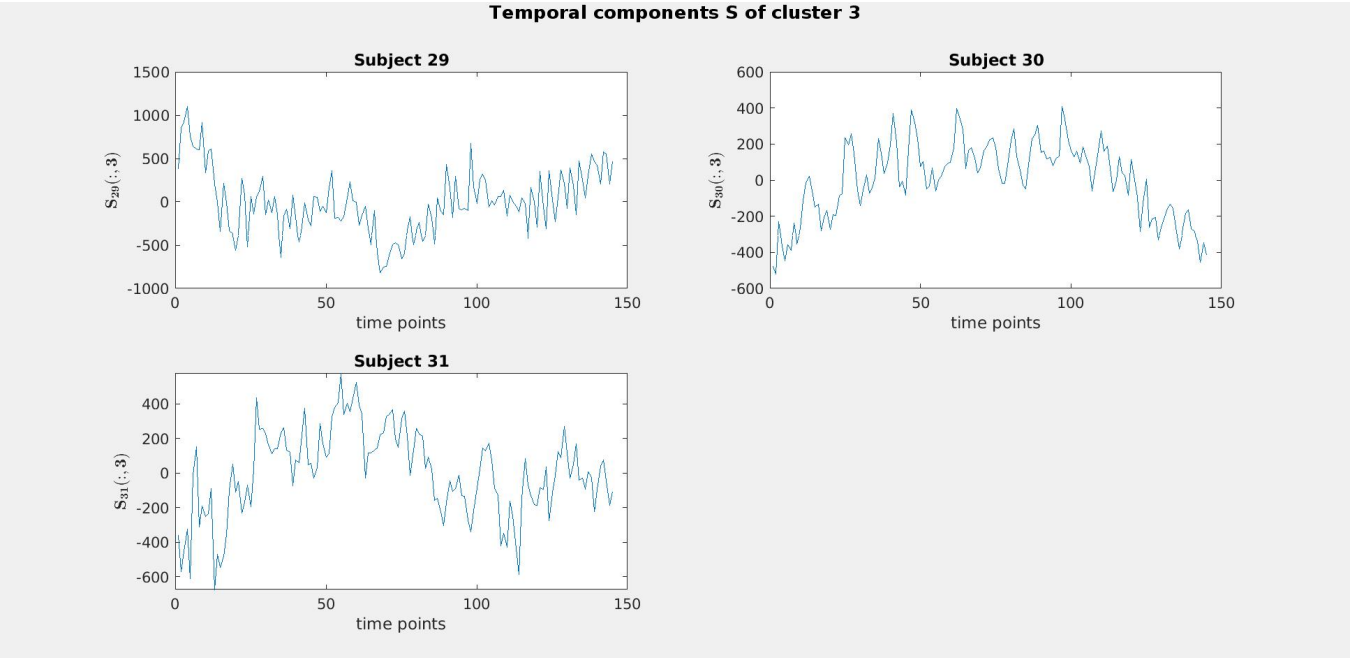


Figure B.333: Waveform of temporal components $\mathbf{S}_k(:,3)$ for $k = 29, \ldots, 31$ and R=42.

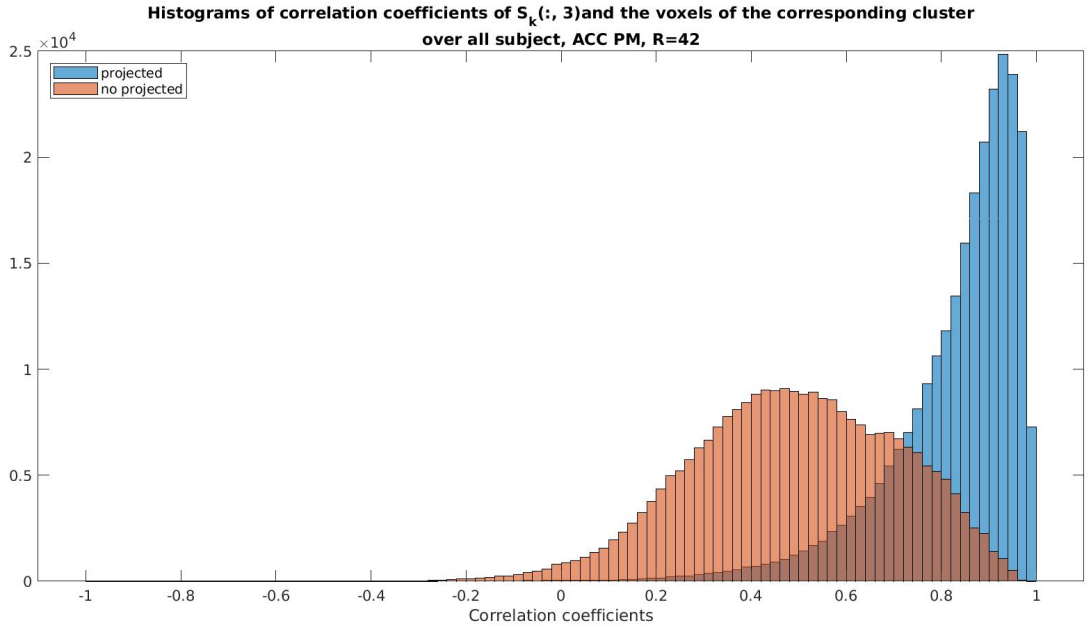


Figure B.334: For the column 3 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,3)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

5. **Cluster 25**
   Cluster 25 contains 9604 voxels, which correspond to 4.250536% of the whole brain (background, soft tissue, bone are excluded).

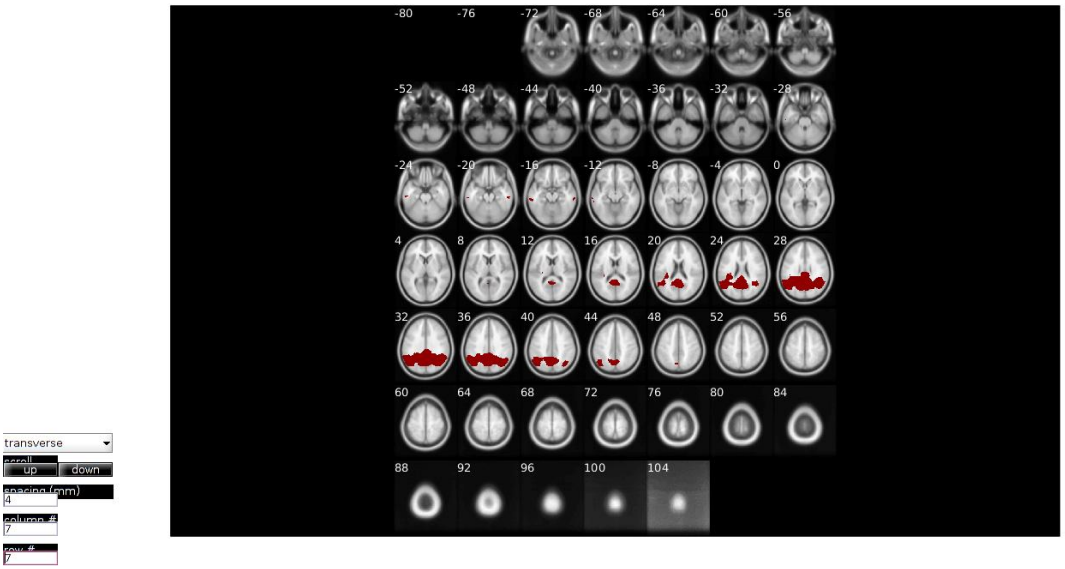


Figure B.335: Cluster 25. Red areas indicate the voxels which belong to cluster 25.

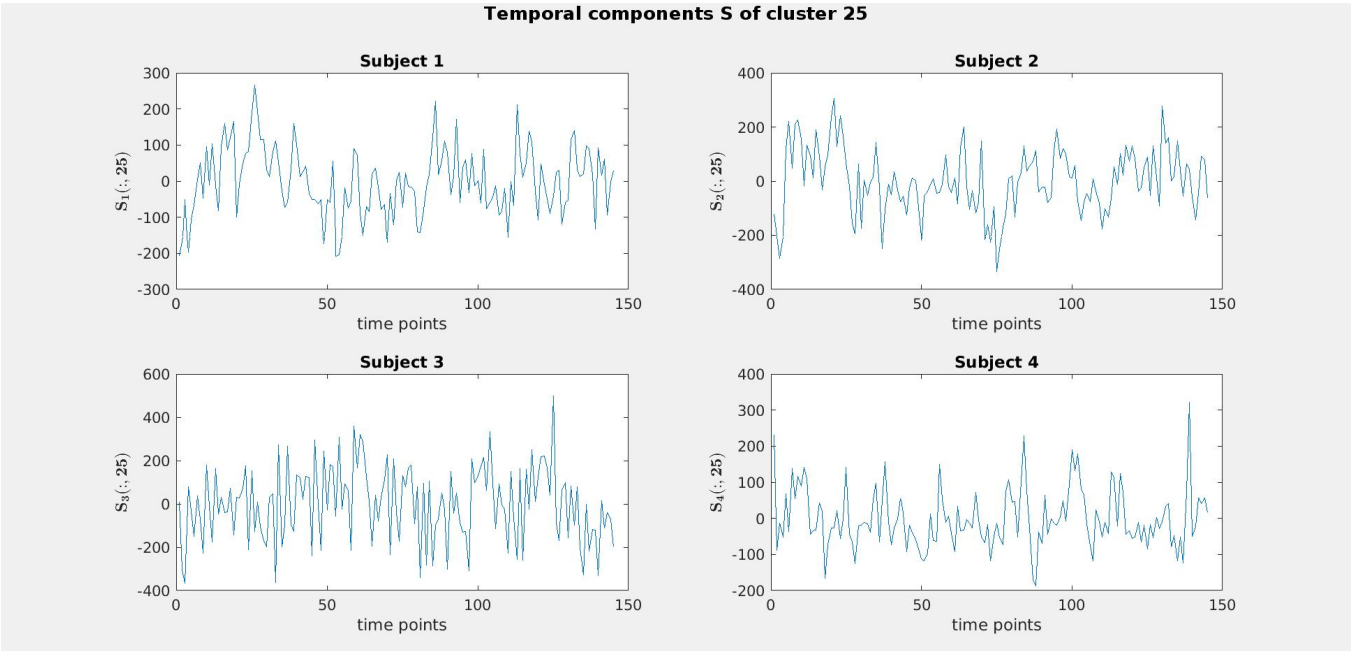


Figure B.336: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 1, \ldots, 4$ and R=42.

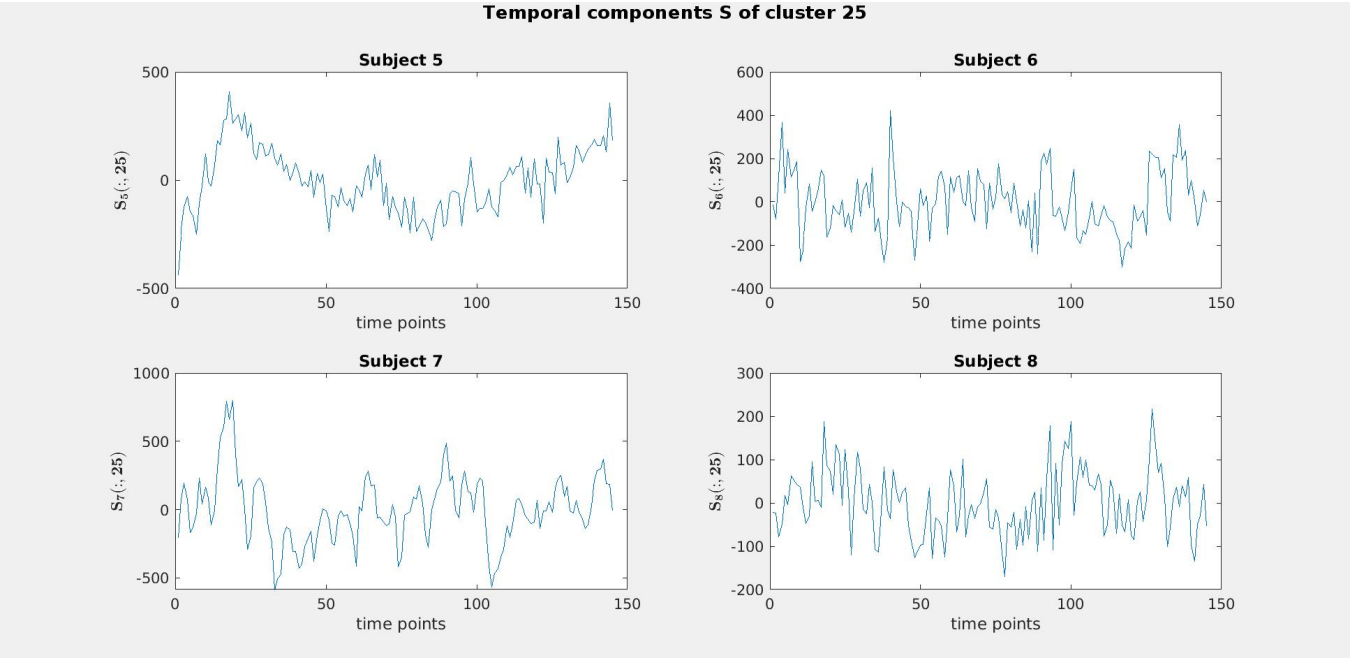


Figure B.337: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 5, \ldots, 8$ and R=42.

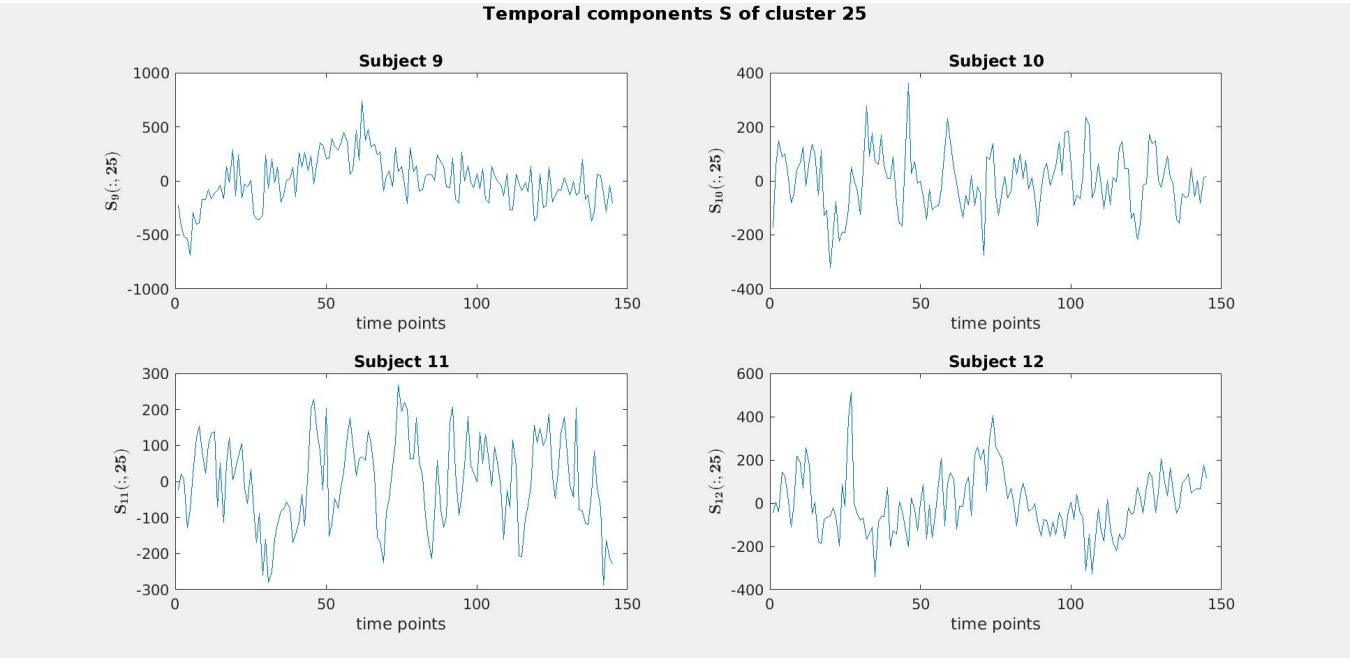


Figure B.338: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 9, \ldots, 12$ and R=42.

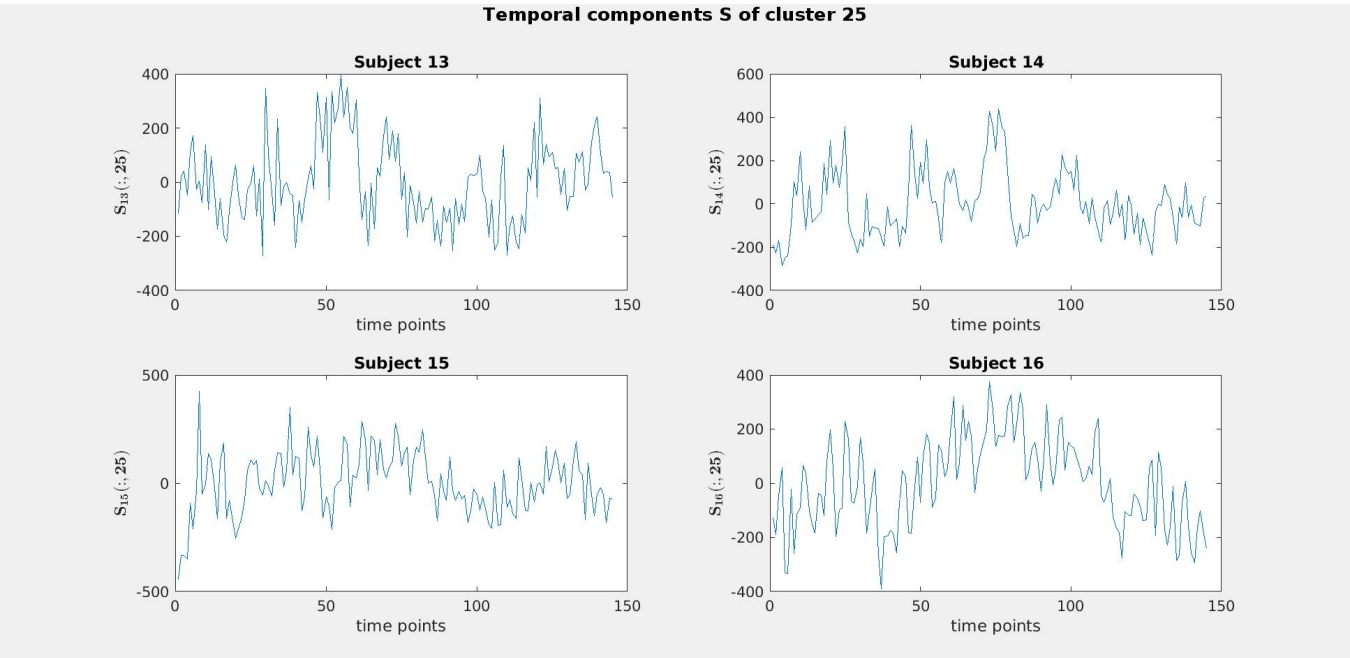


Figure B.339: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 13, \ldots, 16$ and R=42.

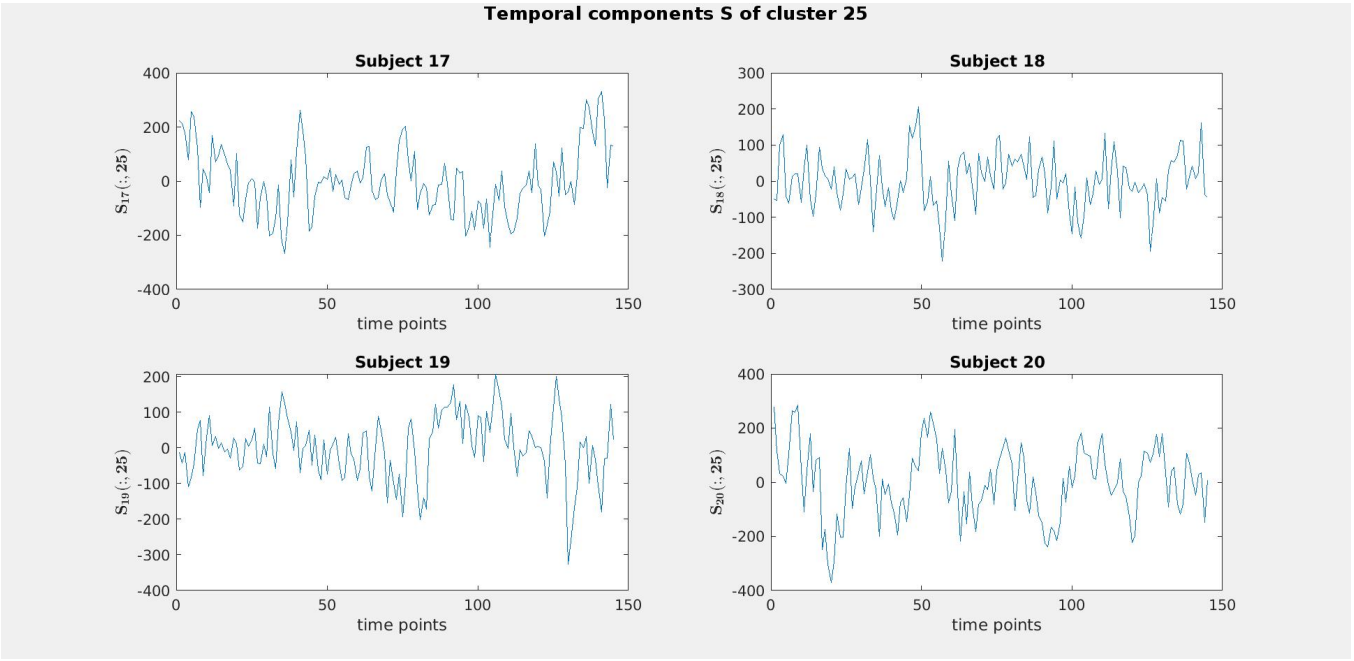


Figure B.340: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 17, \ldots, 20$ and R=42.

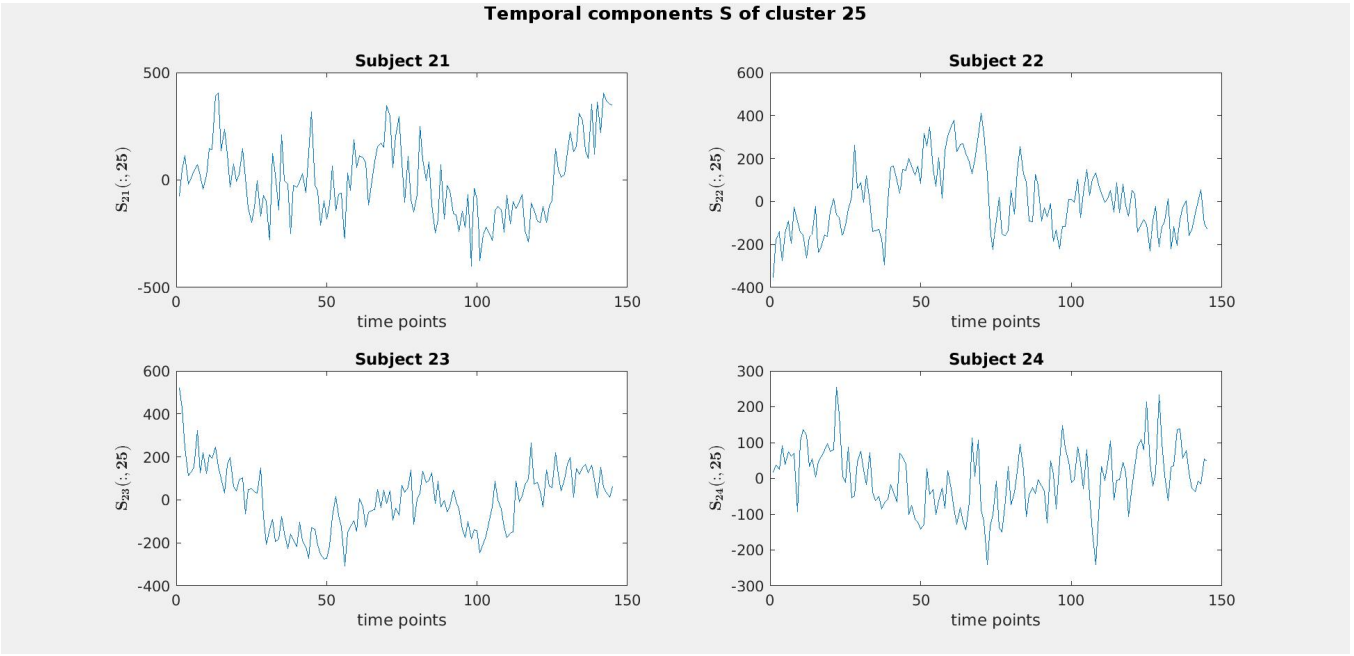


Figure B.341: Waveform of temporal components $\mathbf{S}_k(:, 125)$ for $k = 21, \ldots, 24$ and R=42.

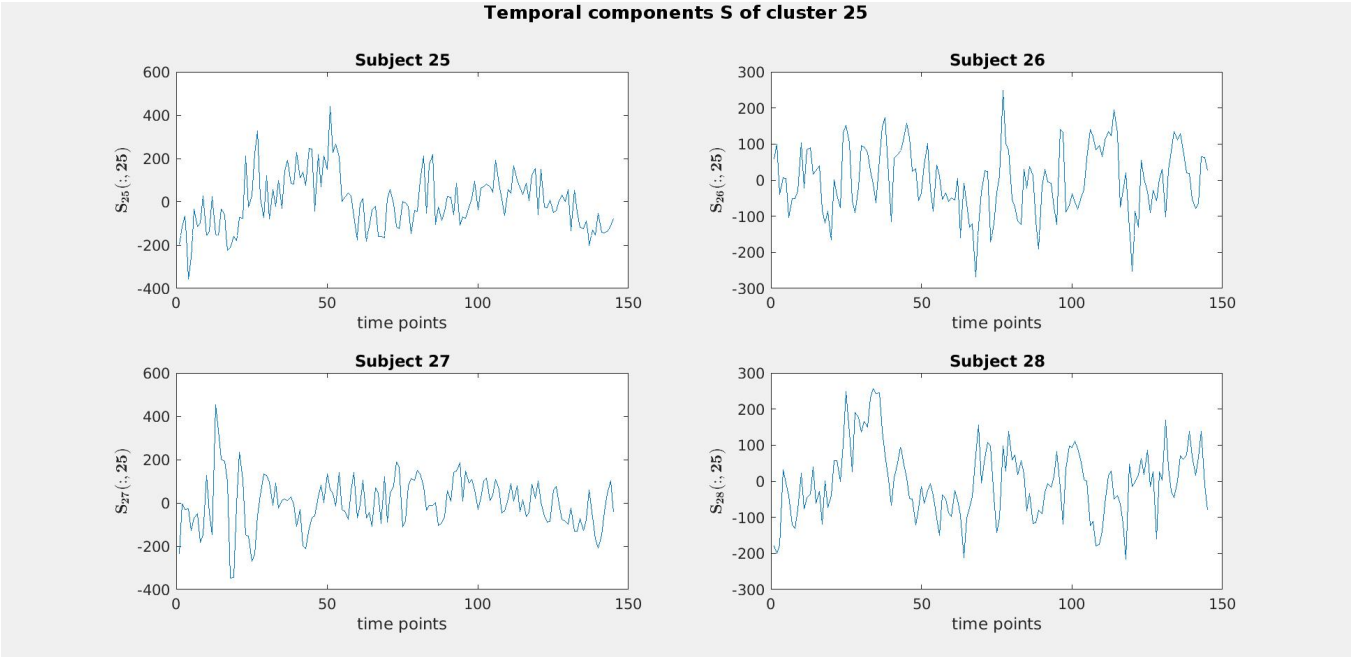


Figure B.342: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 25, \ldots, 28$ and R=42.

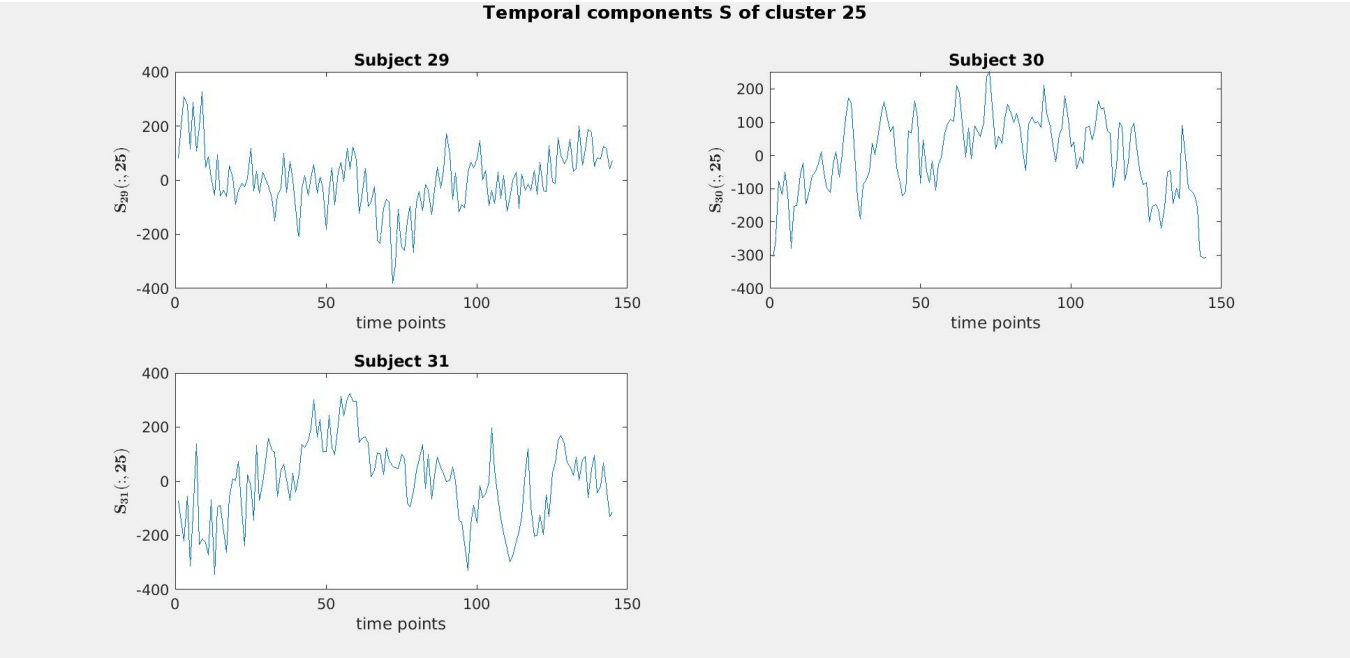


Figure B.343: Waveform of temporal components $\mathbf{S}_k(:, 25)$ for $k = 29, \ldots, 31$ and R=42.

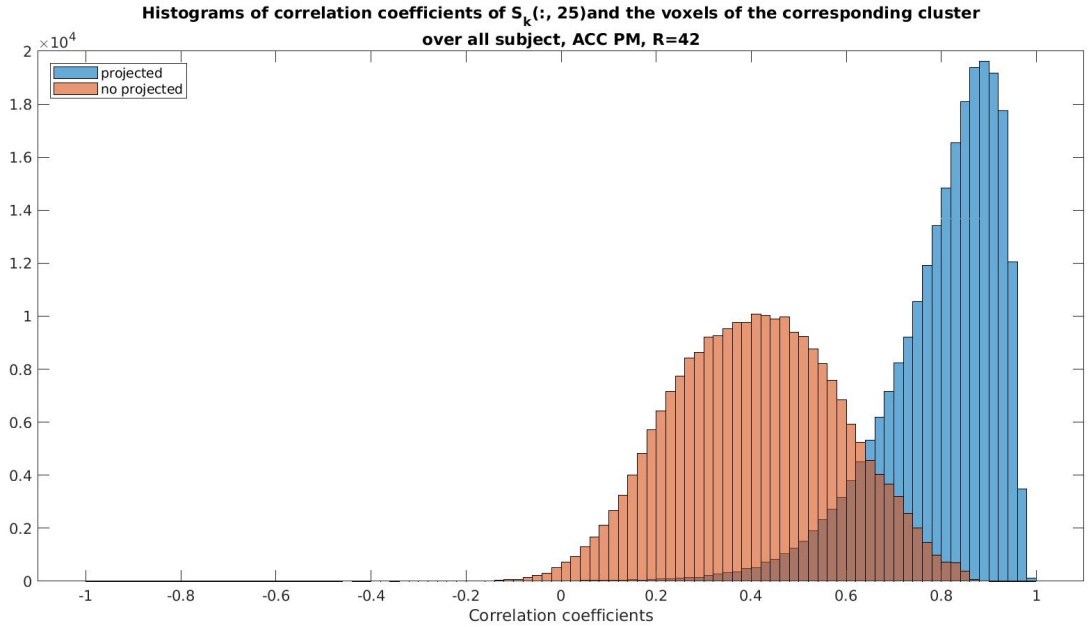


Figure B.344: For the column 25 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 25)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

6. **Cluster 13**
   Cluster 13 contains 9560 voxels, which correspond to 4.231062% of the whole brain (background, soft tissue, bone are excluded).

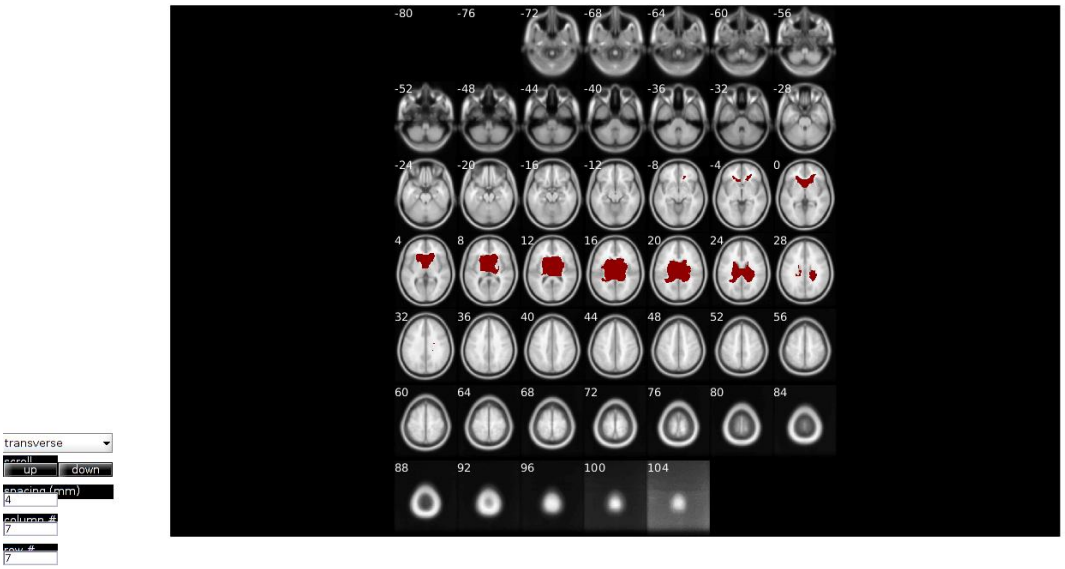


Figure B.345: Cluster 13. Red areas indicate the voxels which belong to cluster 13.

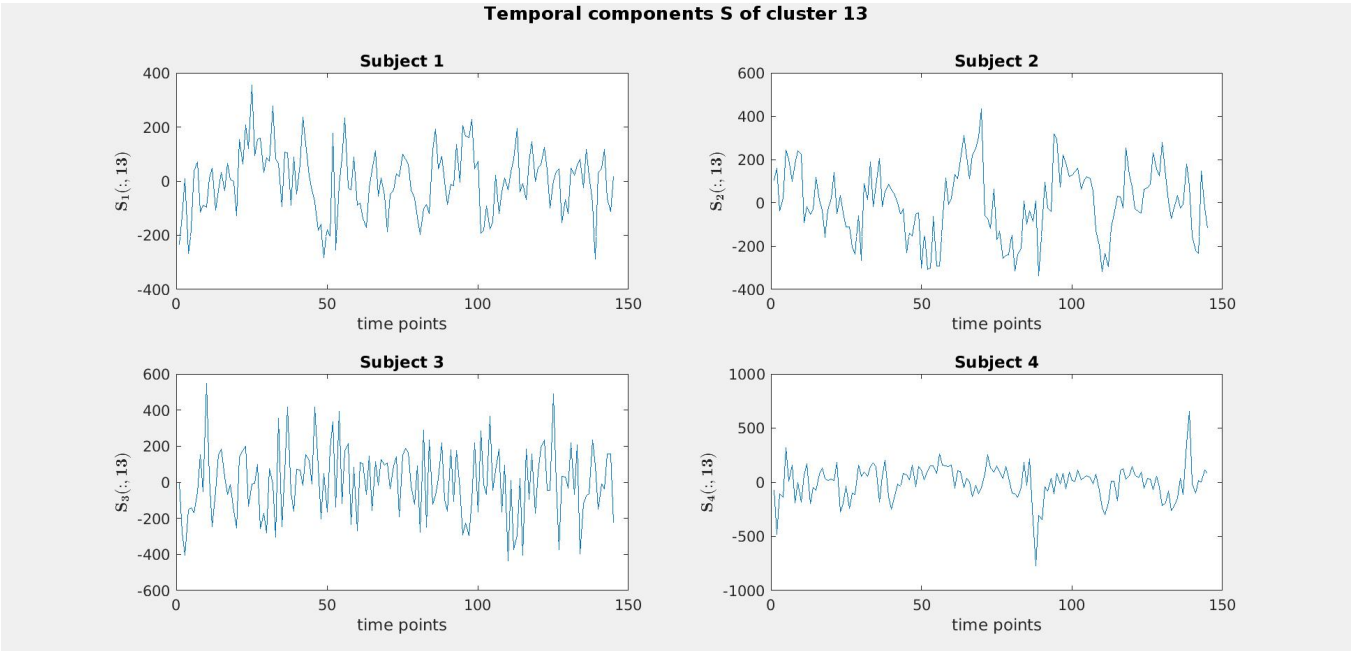


Figure B.346: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 1, \ldots, 4$ and R=42.

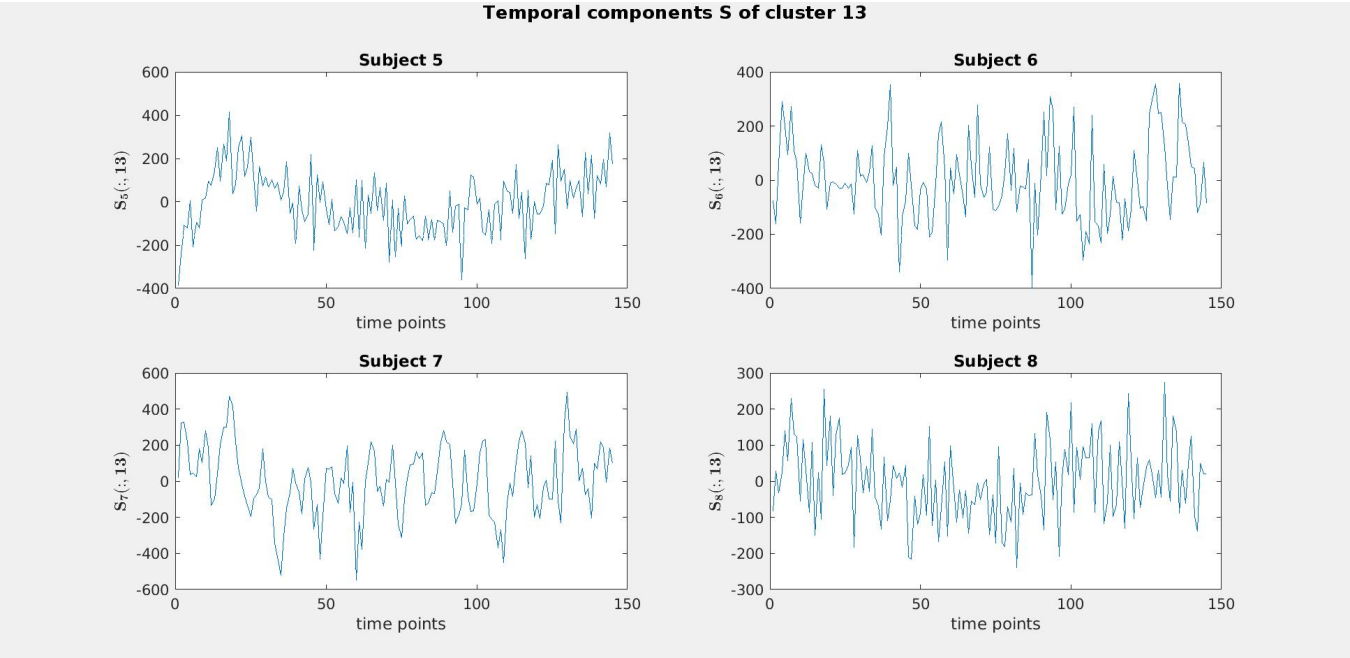


Figure B.347: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 5, \ldots, 8$ and R=42.

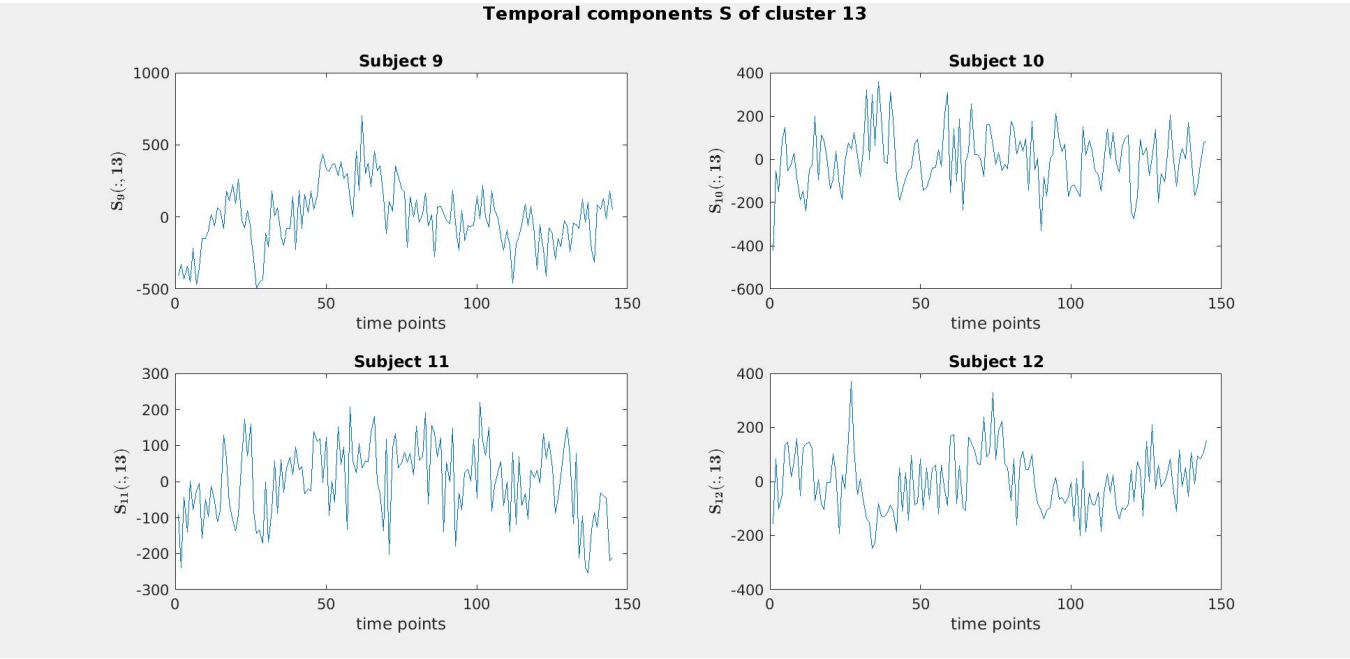


Figure B.348: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 9, \ldots, 12$ and R=42.

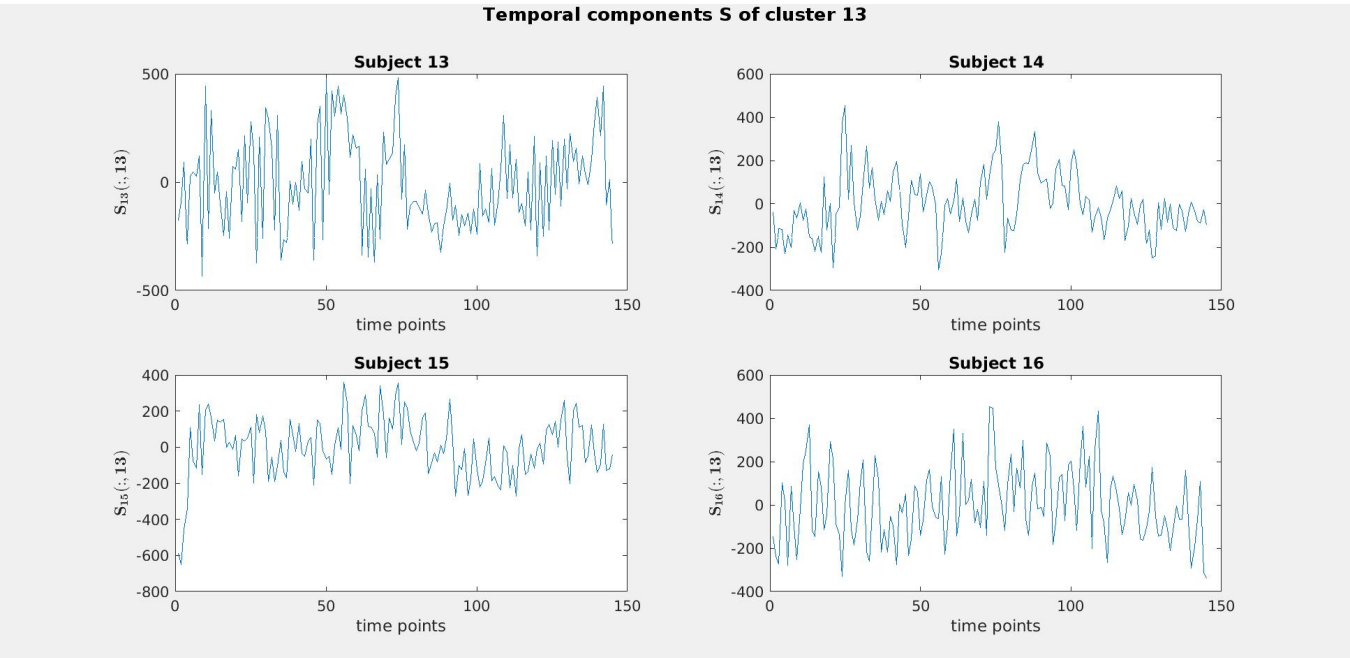


Figure B.349: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 13, \ldots, 16$ and R=42.

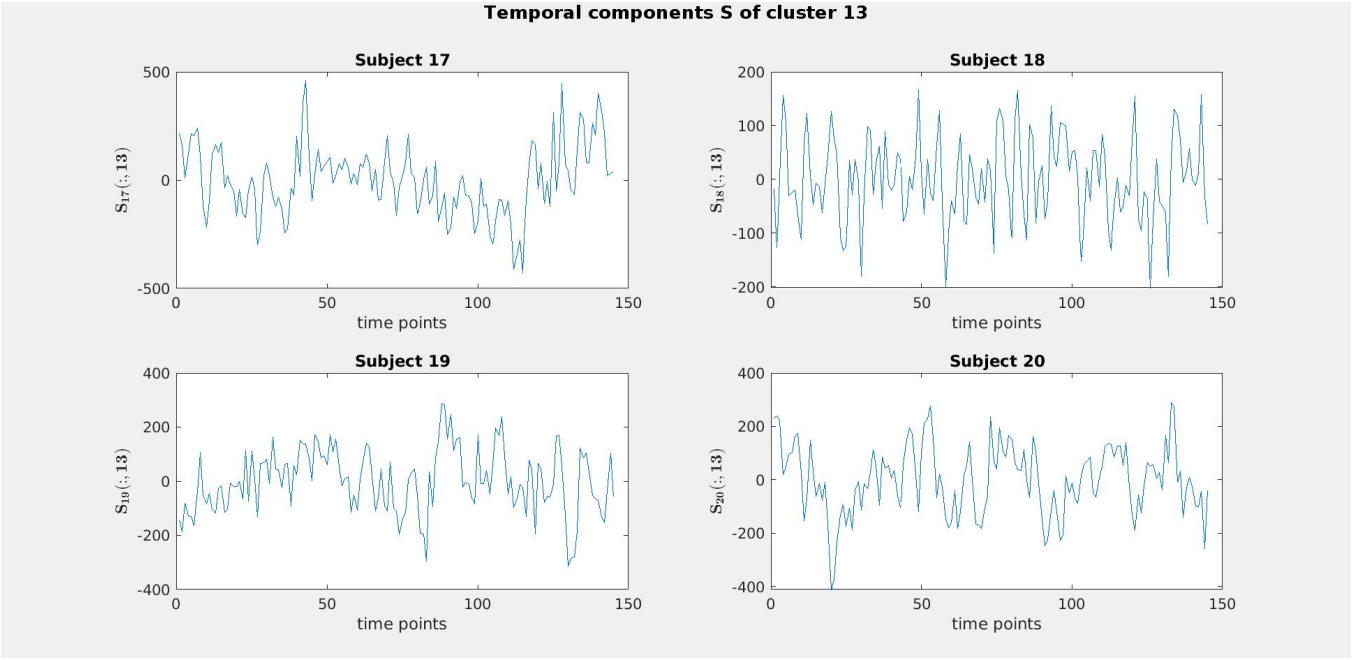


Figure B.350: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 17, \ldots, 20$ and R=42.

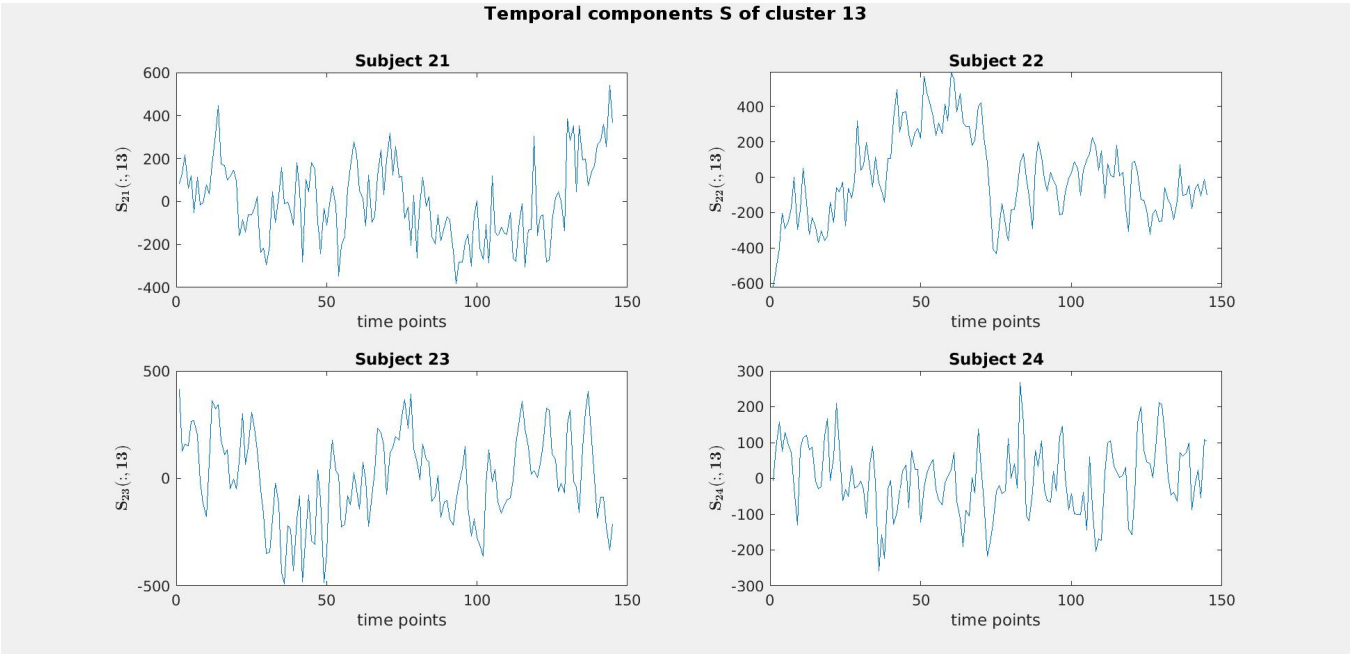


Figure B.351: Waveform of temporal components $\mathbf{S}_k(:, 113)$ for $k = 21, \ldots, 24$ and R=42.

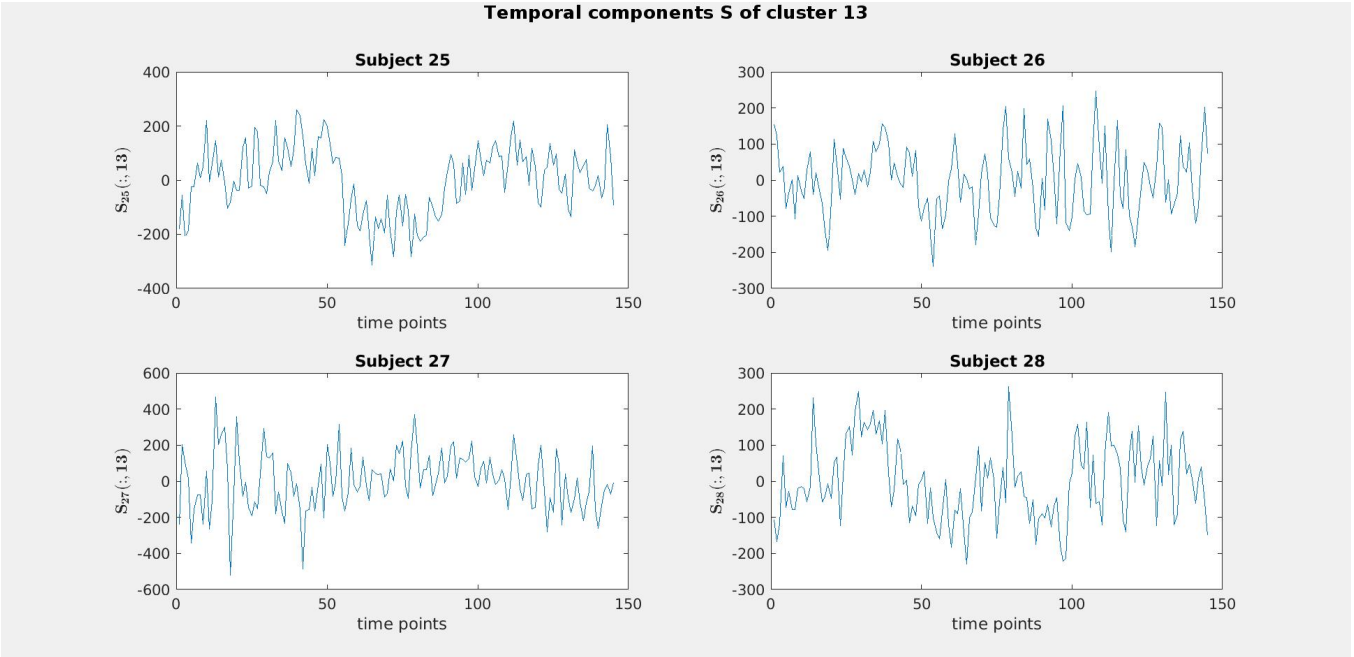


Figure B.352: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 25, \ldots, 28$ and R=42.

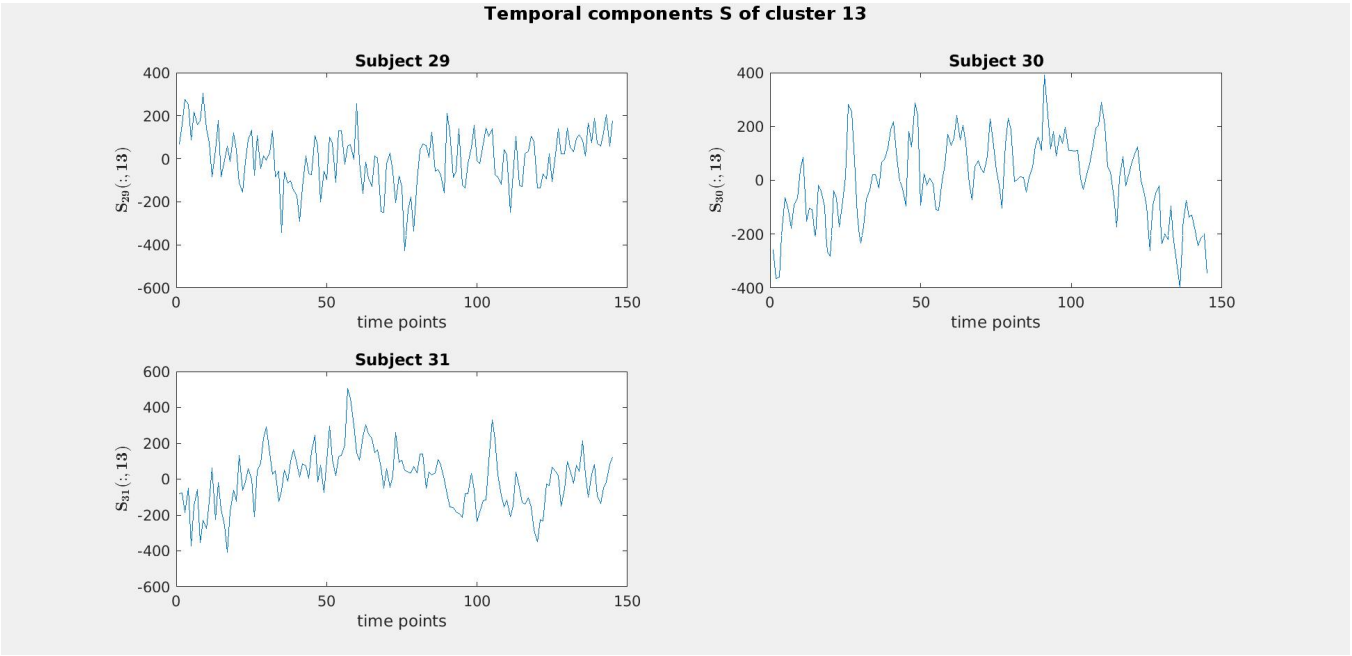


Figure B.353: Waveform of temporal components $\mathbf{S}_k(:, 13)$ for $k = 29, \ldots, 31$ and R=42.

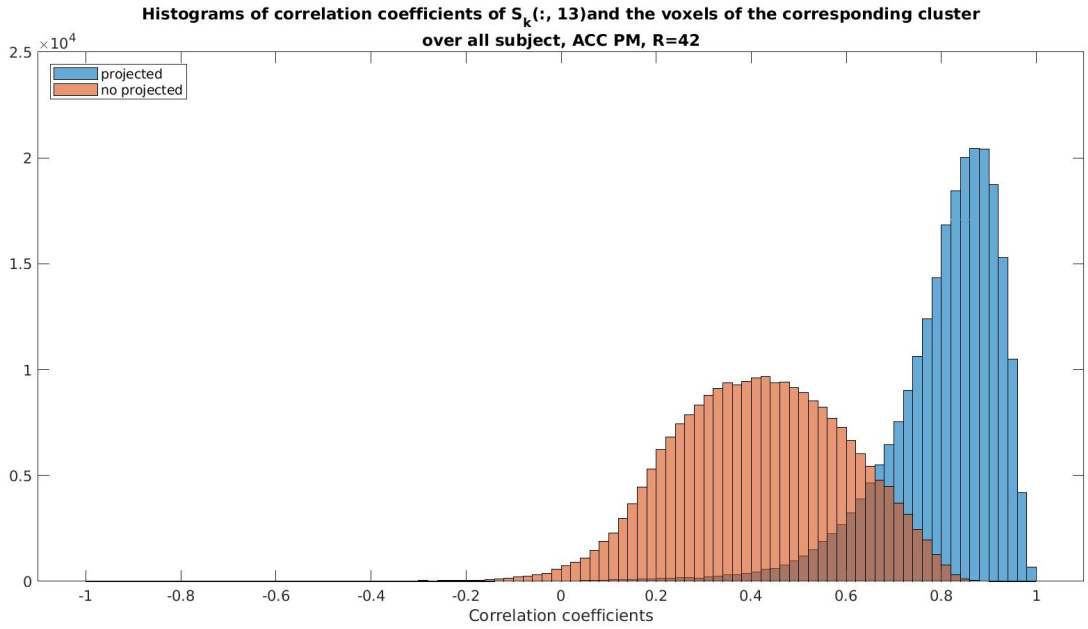


Figure B.354: For the column 13 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 13)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

7. **Cluster 20**

   Cluster 20 contains 9198 voxels, which correspond to 4.070848% of the whole brain (background, soft tissue, bone are excluded).

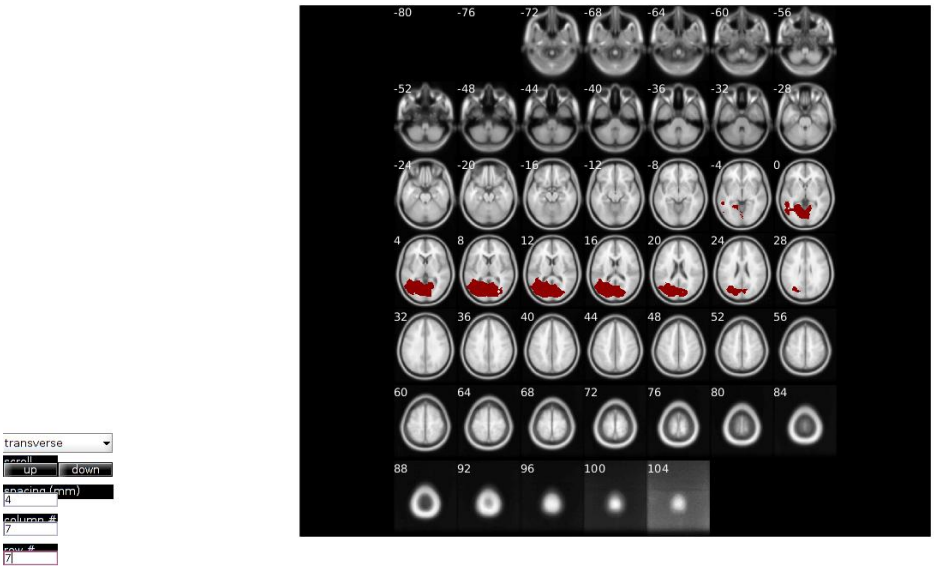


Figure B.355: Cluster 20. Red areas indicate the voxels which belong to cluster 20.

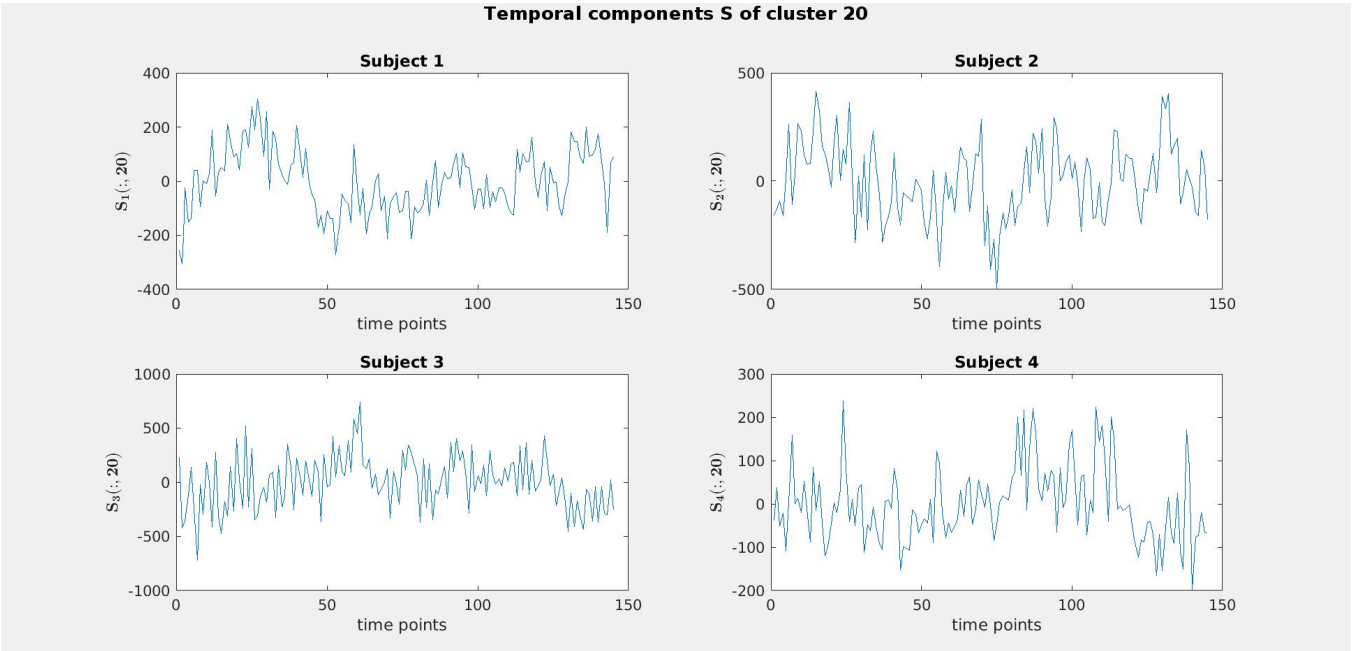


Figure B.356: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 1, \ldots, 4$ and R=42.

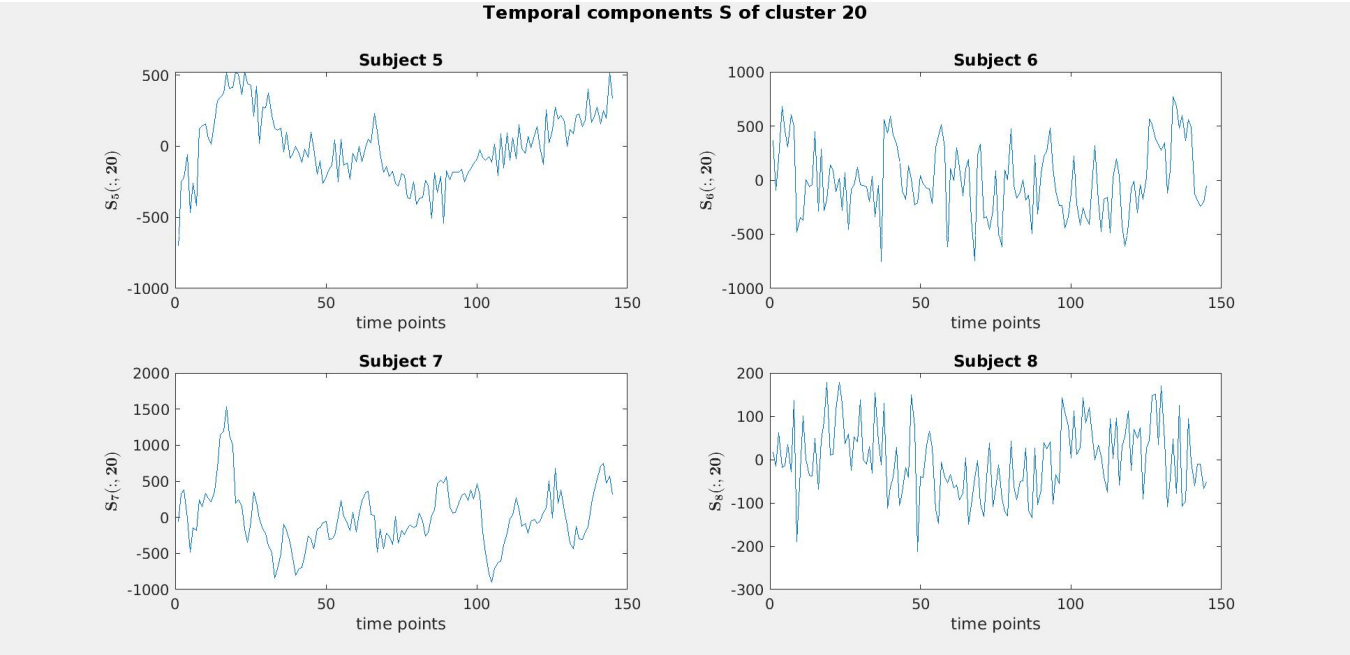


Figure B.357: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 5, \ldots, 8$ and R=42.

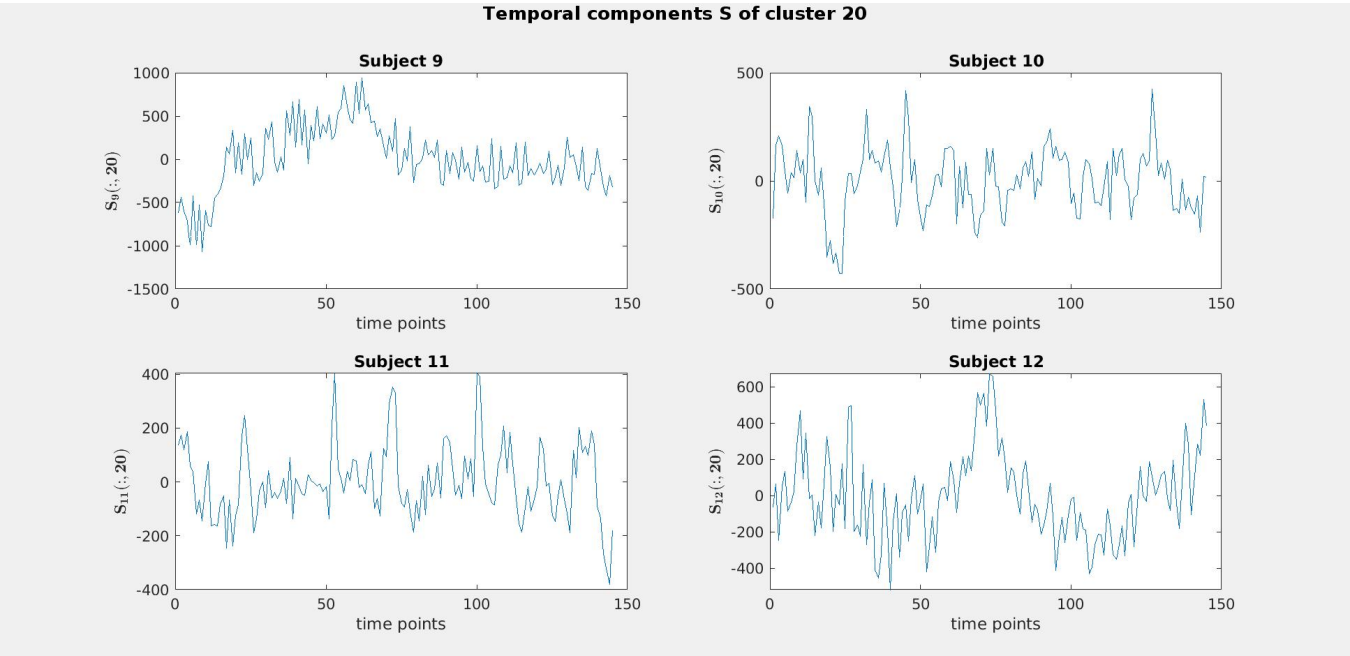


Figure B.358: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 9, \ldots, 12$ and R=42.

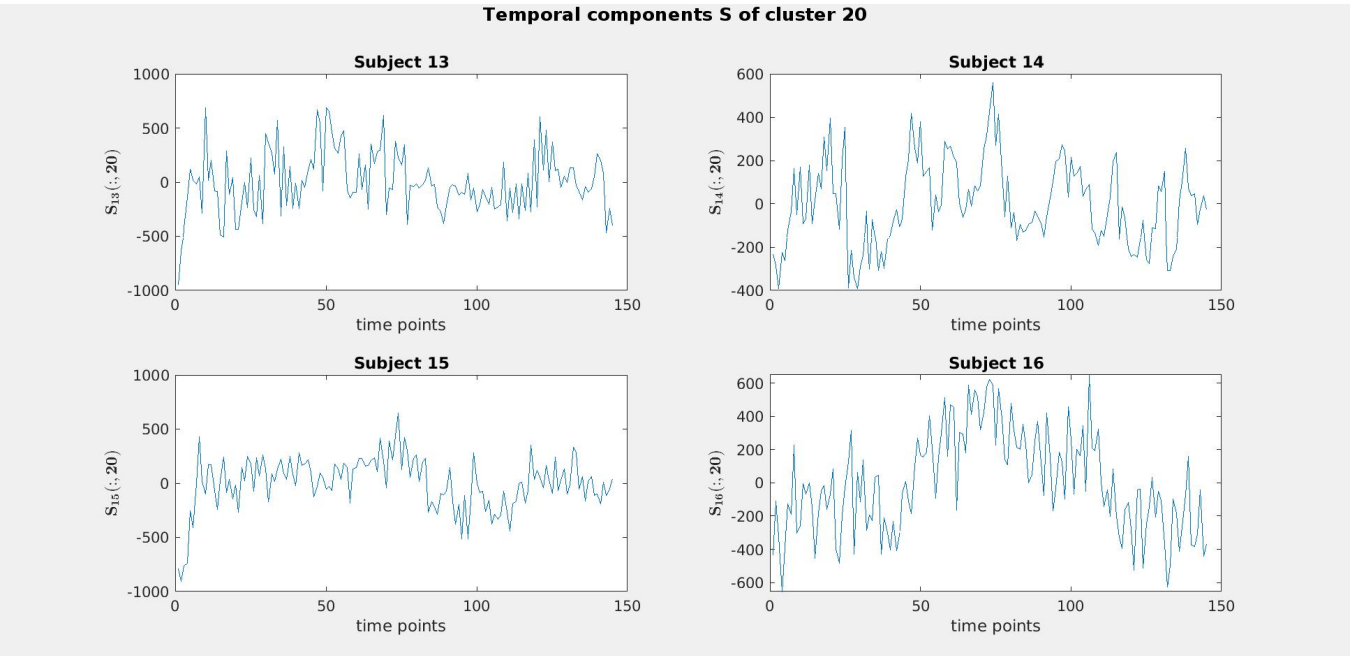


Figure B.359: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 13, \ldots, 16$ and R=42.

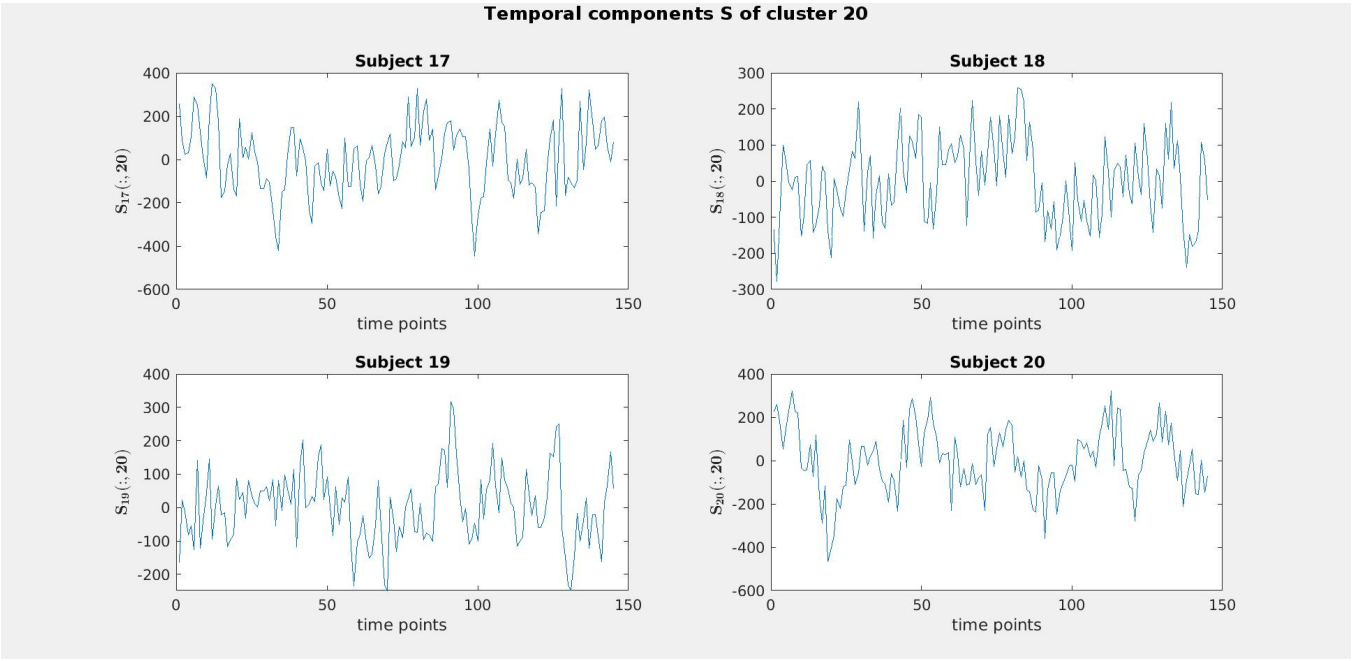


Figure B.360: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 17, \ldots, 20$ and R=42.

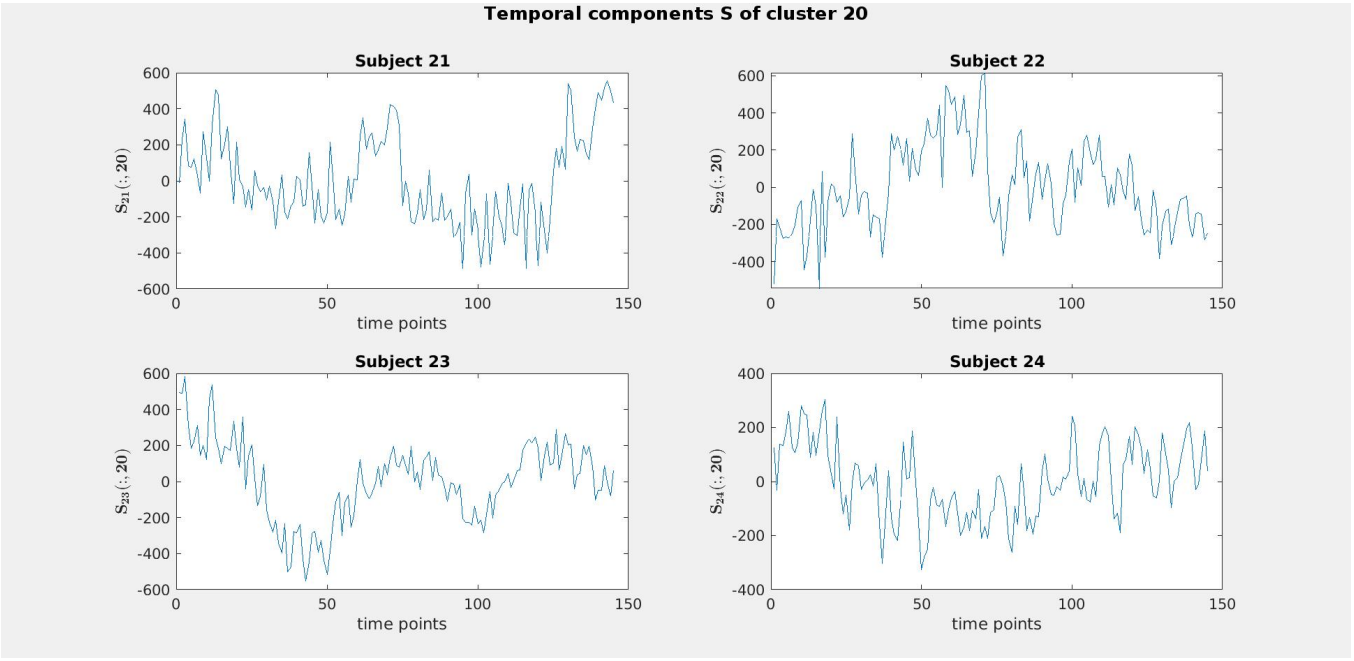


Figure B.361: Waveform of temporal components $\mathbf{S}_k(:, 120)$ for $k = 21, \ldots, 24$ and R=42.

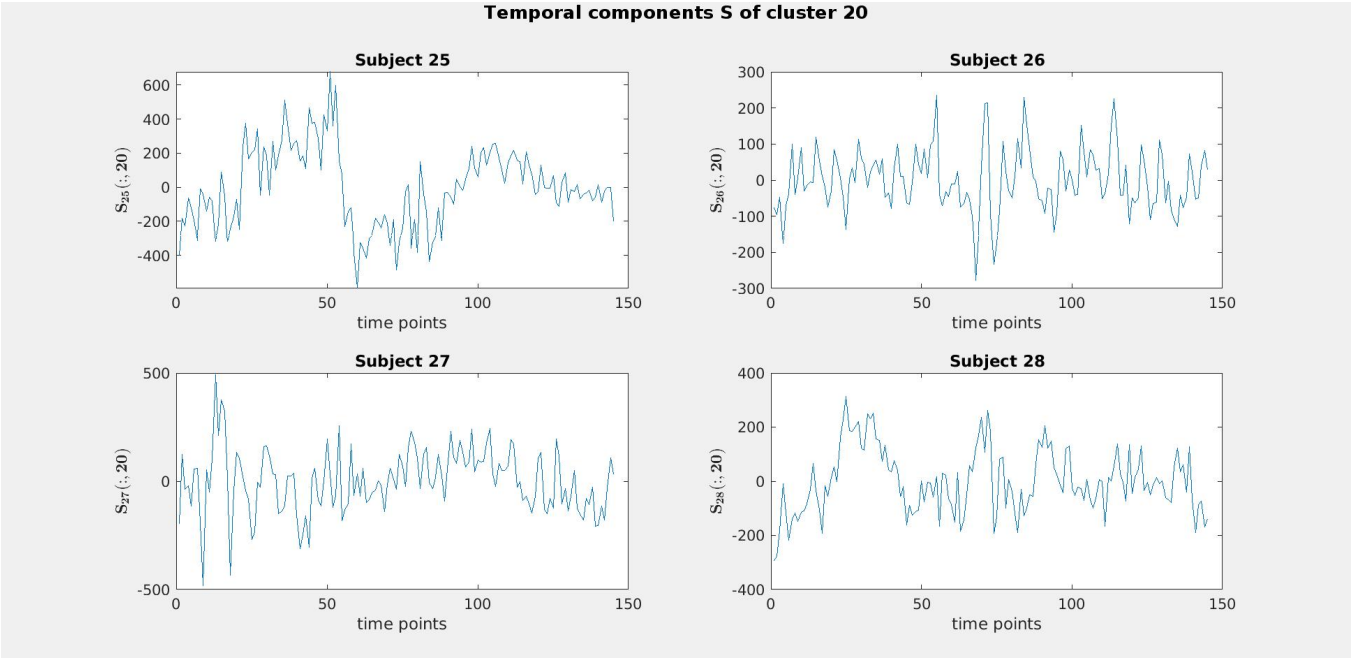


Figure B.362: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 25, \ldots, 28$ and R=42.

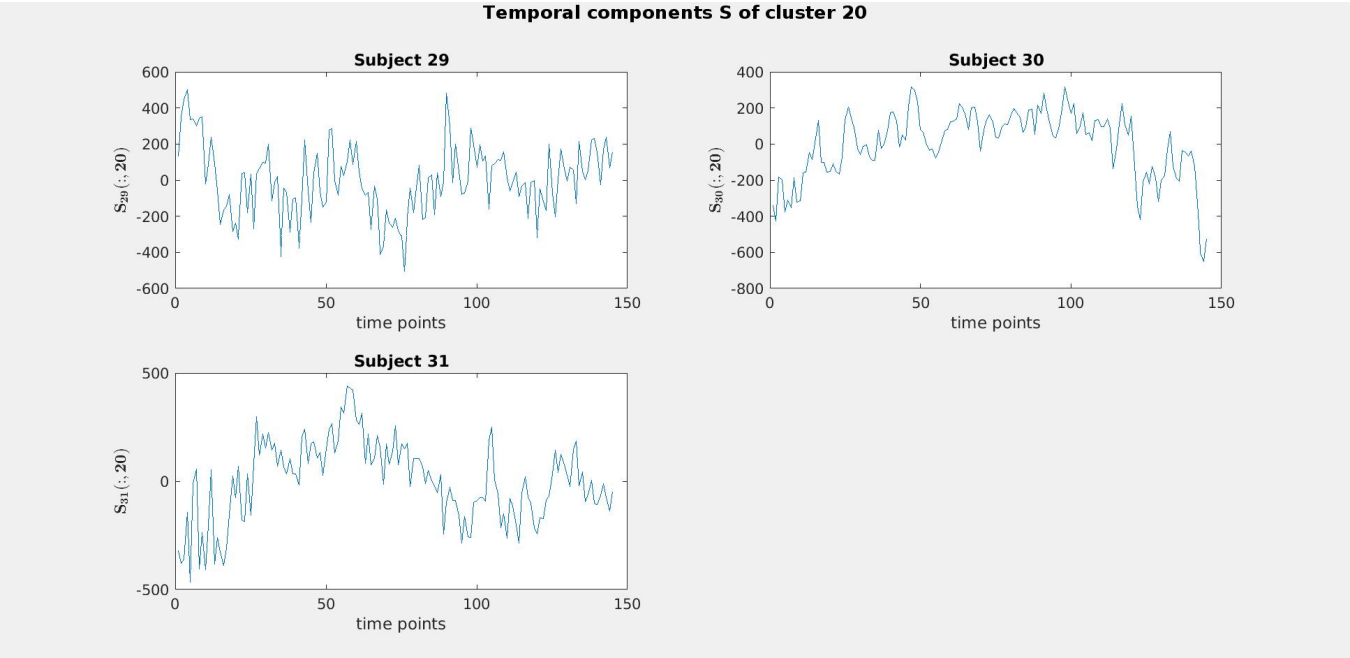


Figure B.363: Waveform of temporal components $\mathbf{S}_k(:, 20)$ for $k = 29, \ldots, 31$ and R=42.

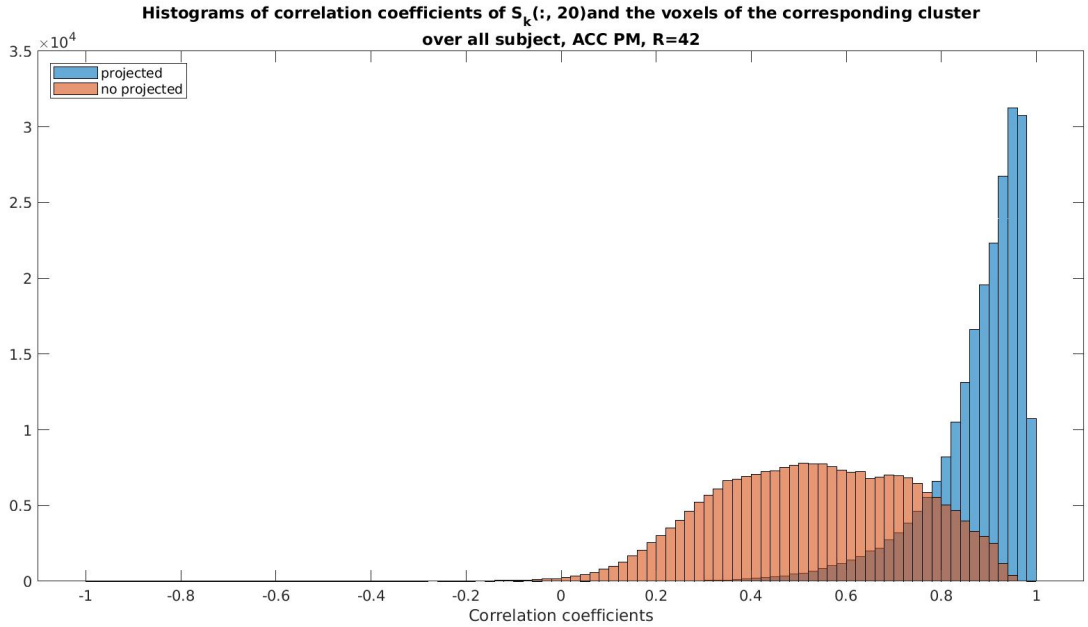


Figure B.364: For the column 20 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 20)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

8. **Cluster 29**
   Cluster 29 contains 8347 voxels, which correspond to 3.694213% of the whole brain (background, soft tissue, bone are excluded).

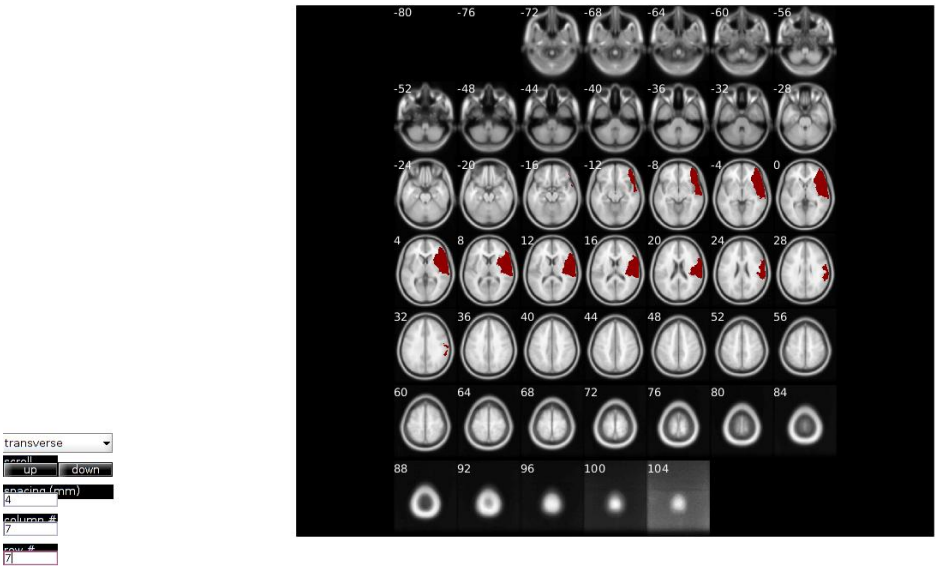


Figure B.365: Cluster 29. Red areas indicate the voxels which belong to cluster 29.

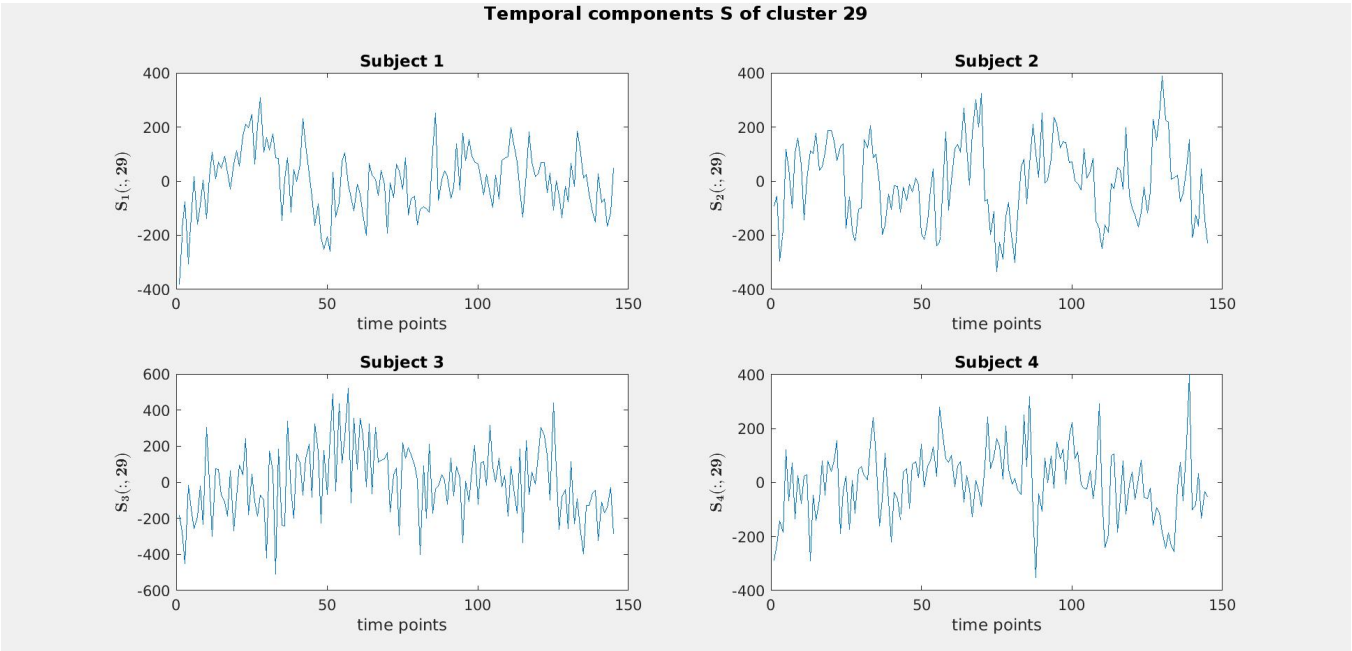


Figure B.366: Waveform of temporal components $\mathbf{S}_k(:, 29)$ for $k = 1, \ldots, 4$ and R=42.

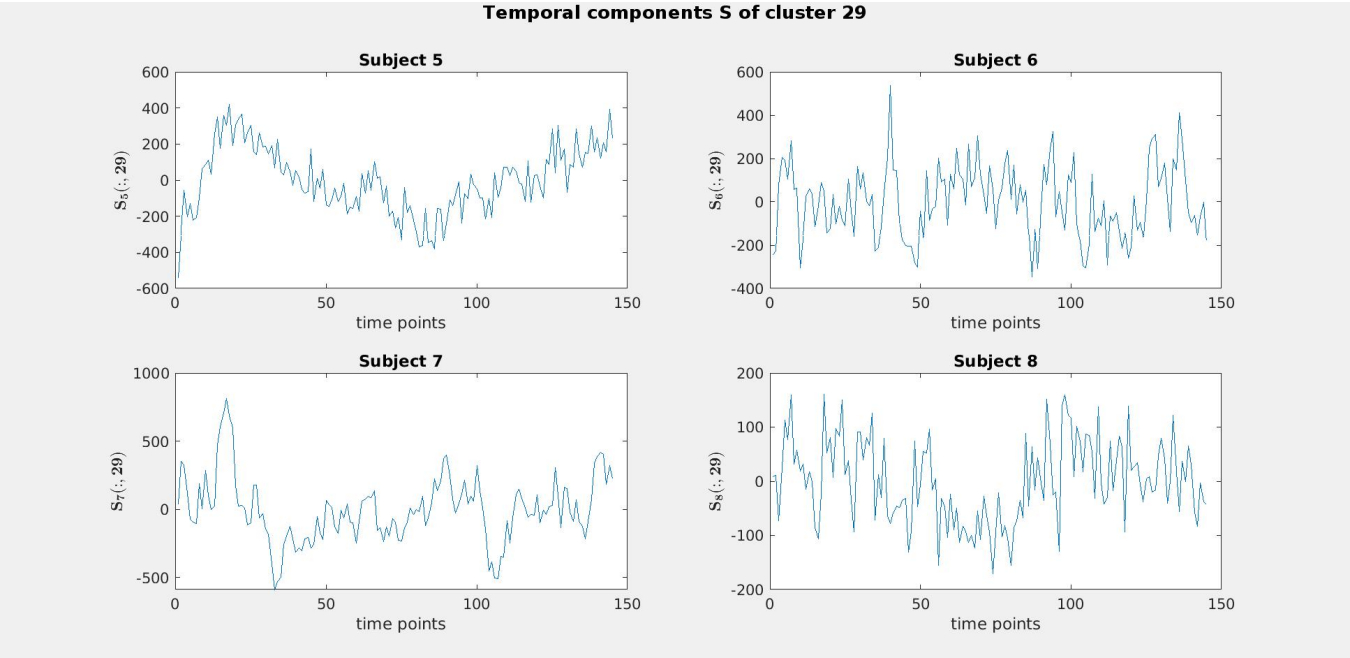


Figure B.367: Waveform of temporal components $\mathbf{S}_k(:, 29)$ for $k = 5, \ldots, 8$ and R=42.

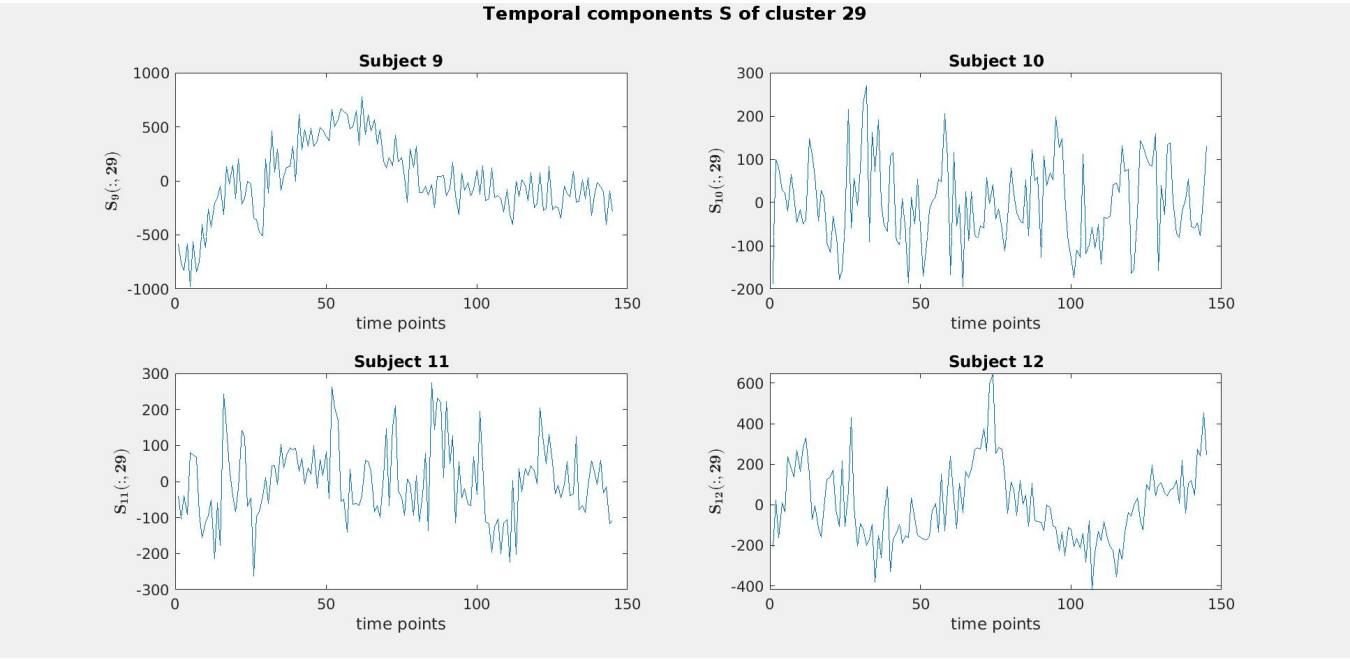


Figure B.368: Waveform of temporal components $\mathbf{S}_k(:, 29)$ for $k = 9, \ldots, 12$ and R=42.

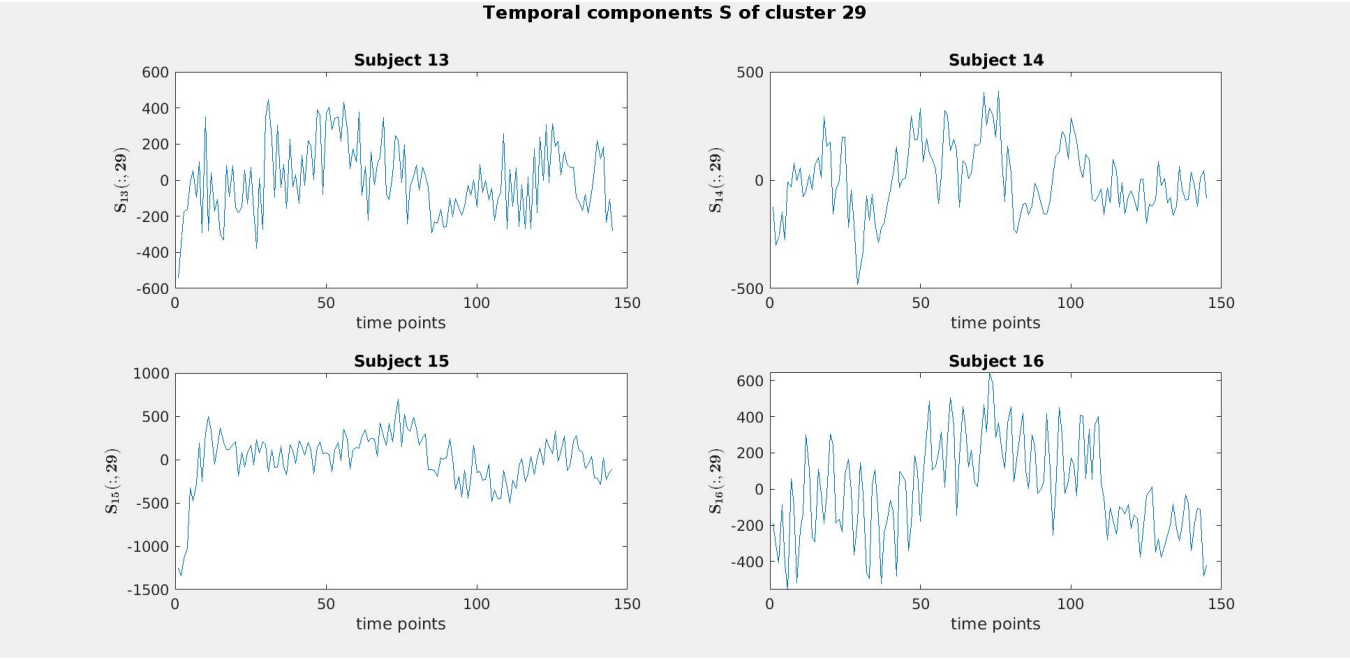


Figure B.369: Waveform of temporal components $\mathbf{S}_k(:, 29)$ for $k = 13, \ldots, 16$ and R=42.

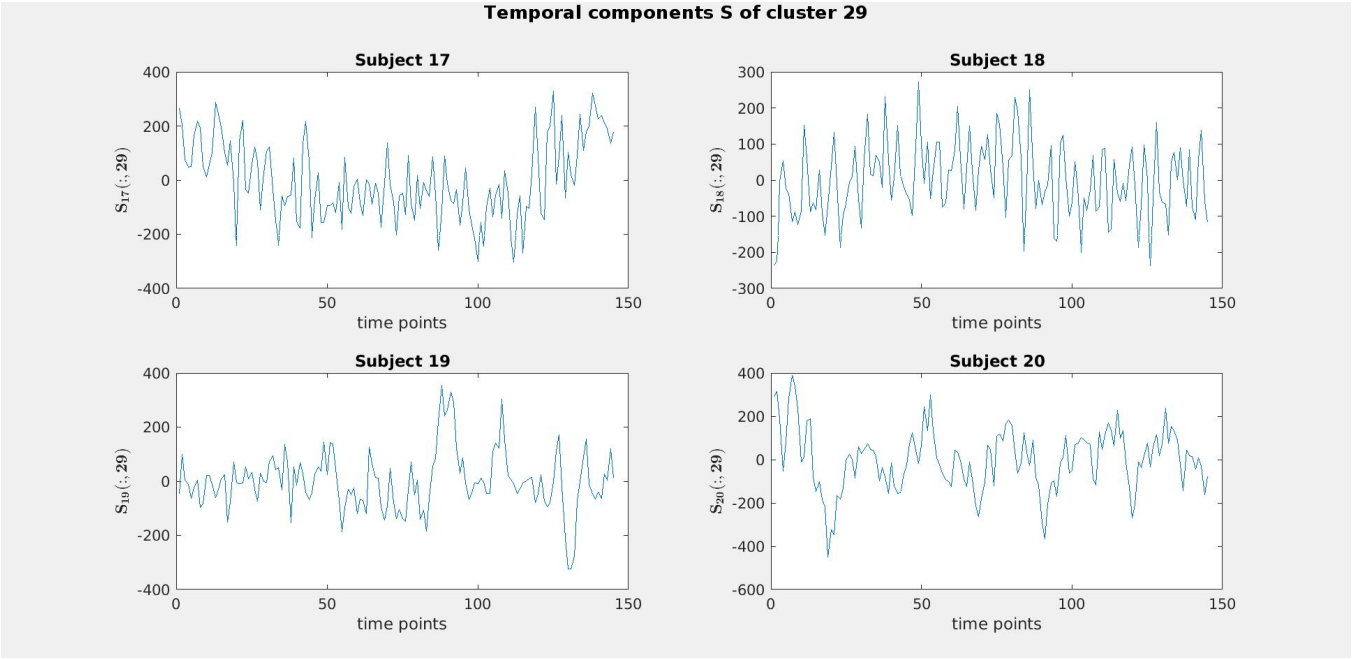


Figure B.370: Waveform of temporal components $\mathbf{S}_k(:, 29)$ for $k = 17, \ldots, 20$ and R=42.

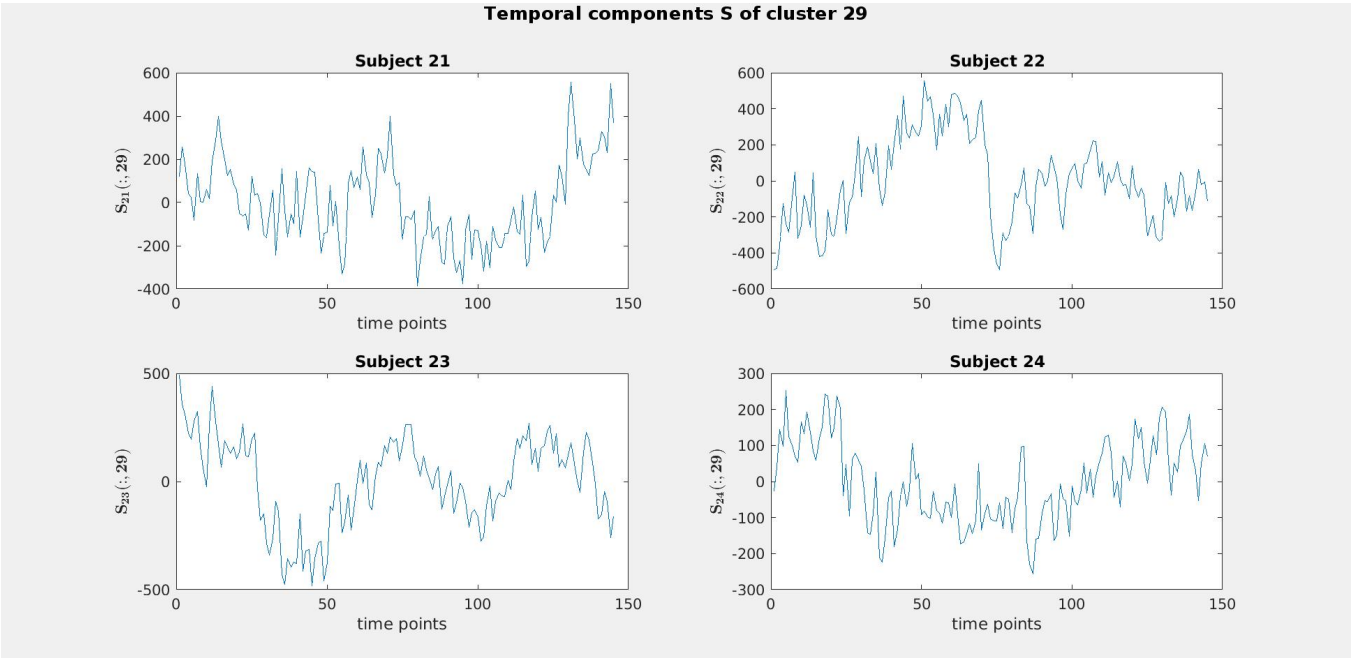


Figure B.371: Waveform of temporal components $\mathbf{S}_k(:, 129)$ for $k = 21, \ldots, 24$ and R=42.

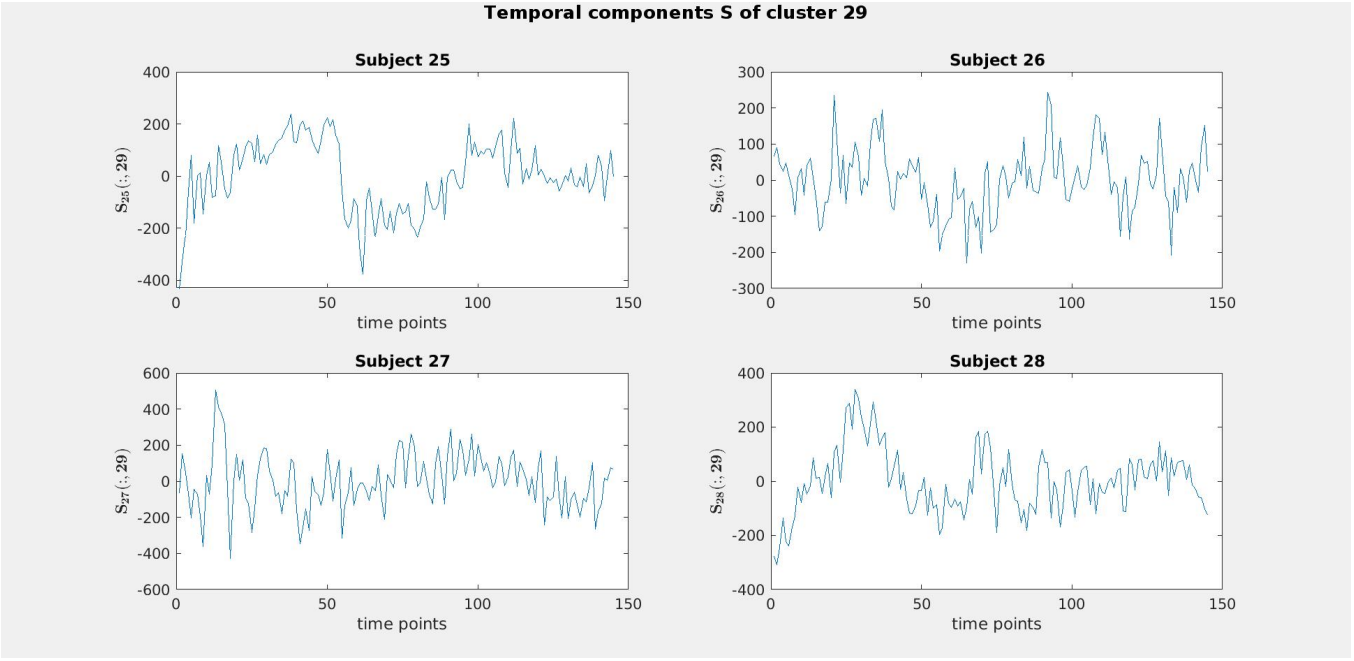


Figure B.372: Waveform of temporal components $\mathbf{S}_k(:, 29)$ for $k = 25, \ldots, 28$ and R=42.

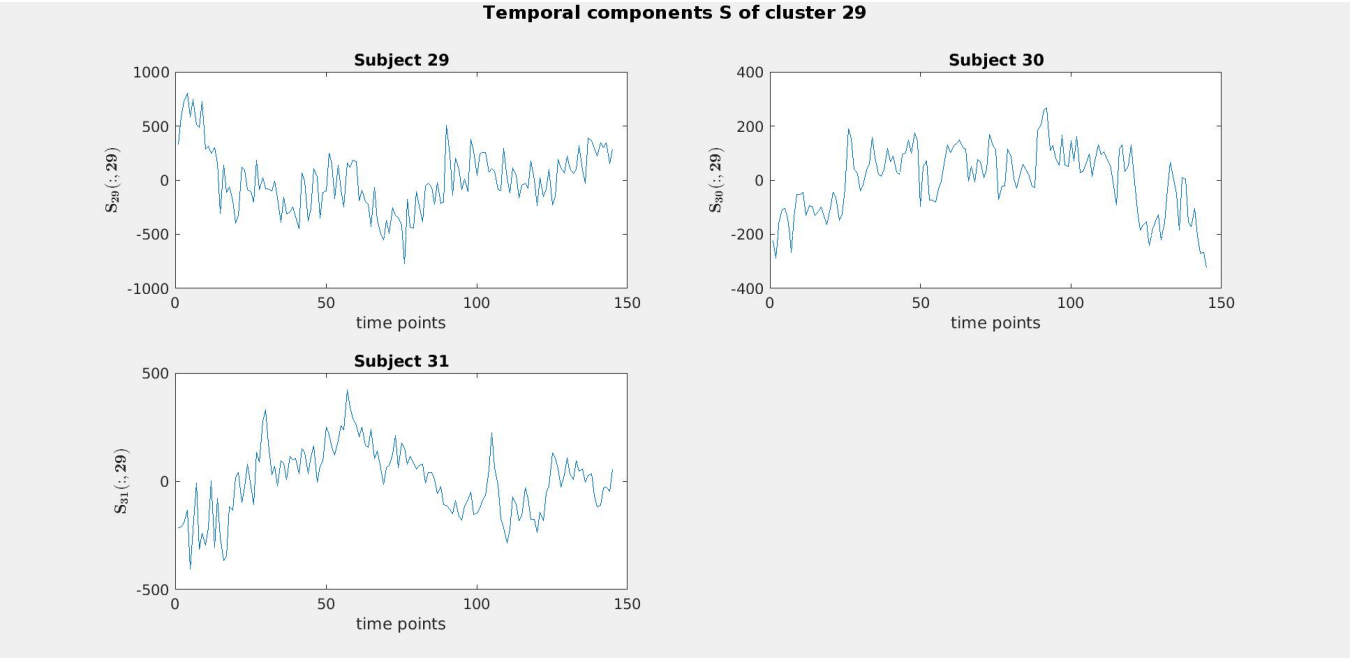


Figure B.373: Waveform of temporal components $\mathbf{S}_k(:, 29)$ for $k = 29, \ldots, 31$ and R=42.

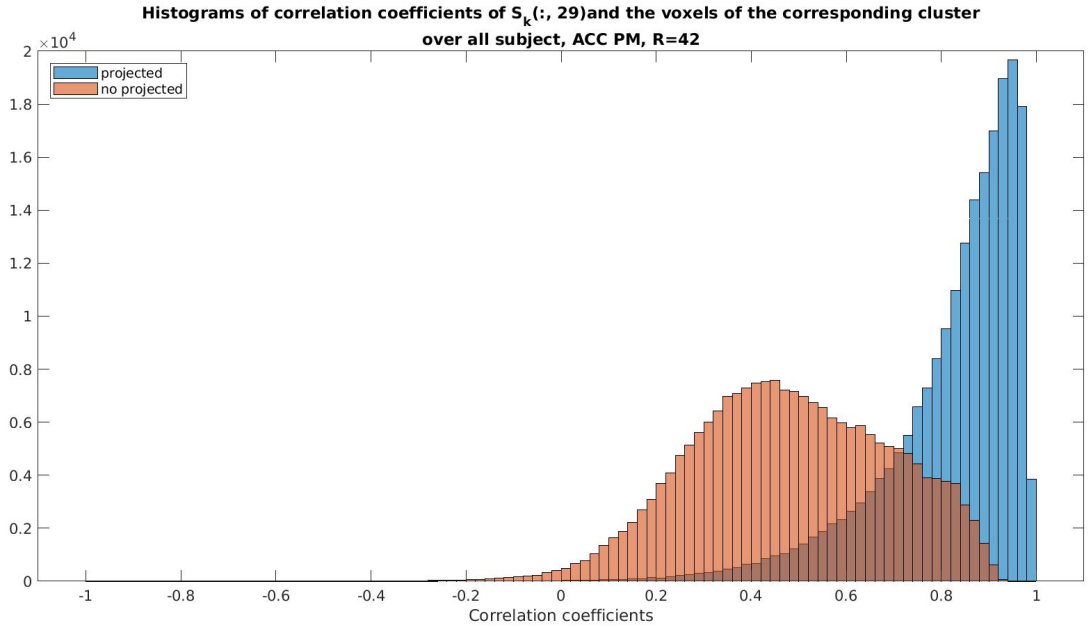


Figure B.374: For the column 29 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 29)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

9. **Cluster 41**

   Cluster 41 contains 7618 voxels, which correspond to 3.371572% of the whole brain (background, soft tissue, bone are excluded).

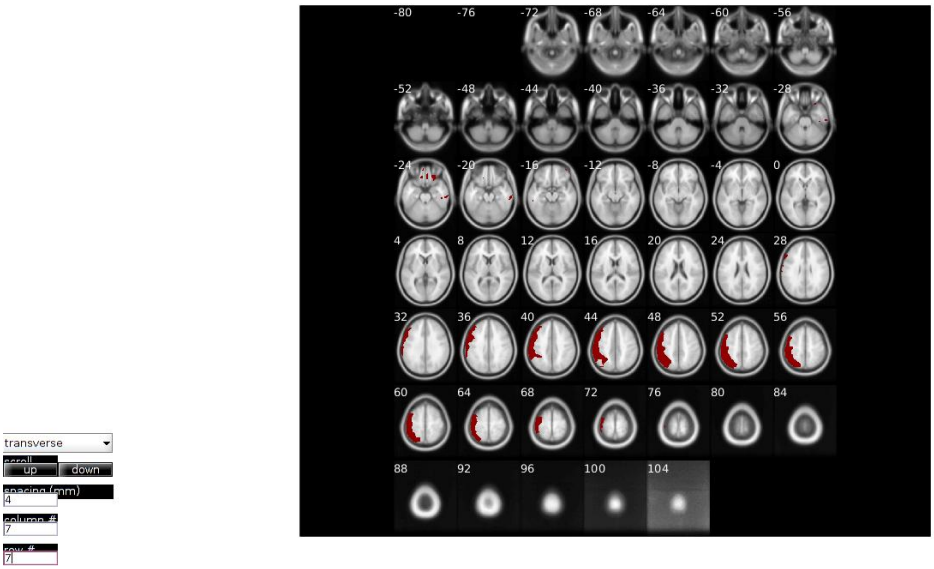


Figure B.375: Cluster 41. Red areas indicate the voxels which belong to cluster 41.

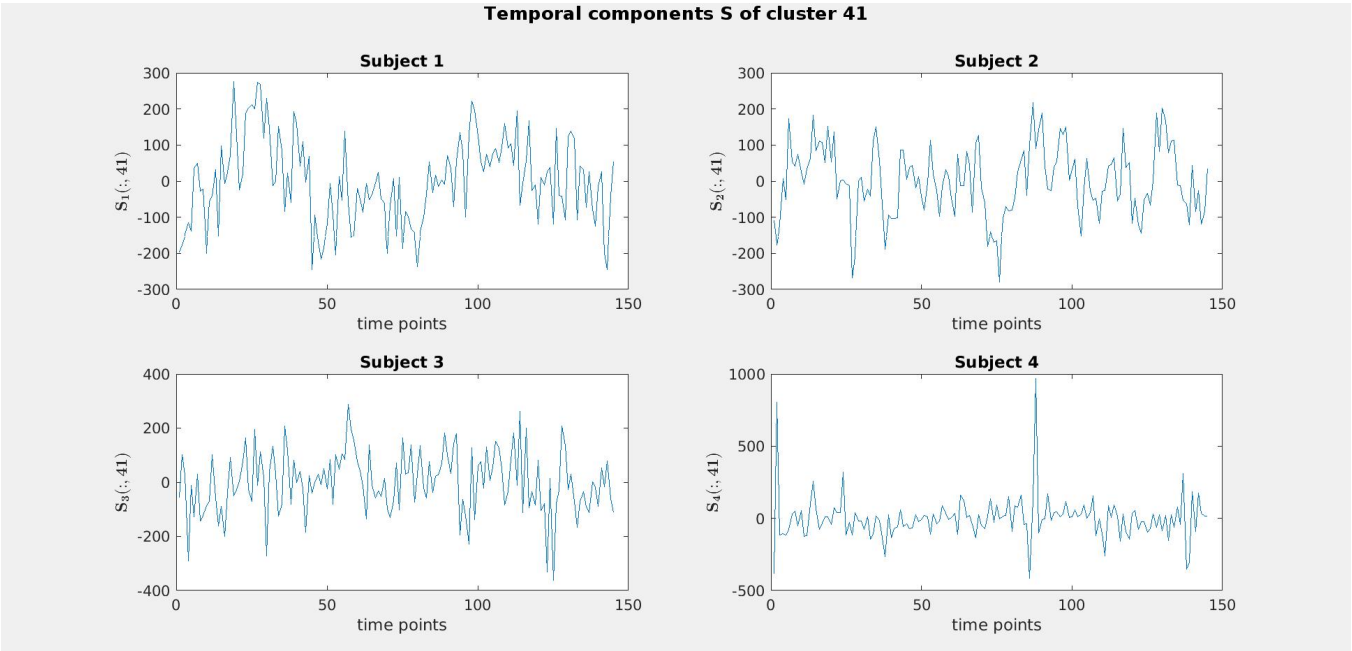


Figure B.376: Waveform of temporal components $\mathbf{S}_k(:, 41)$ for $k = 1, \ldots, 4$ and R=42.

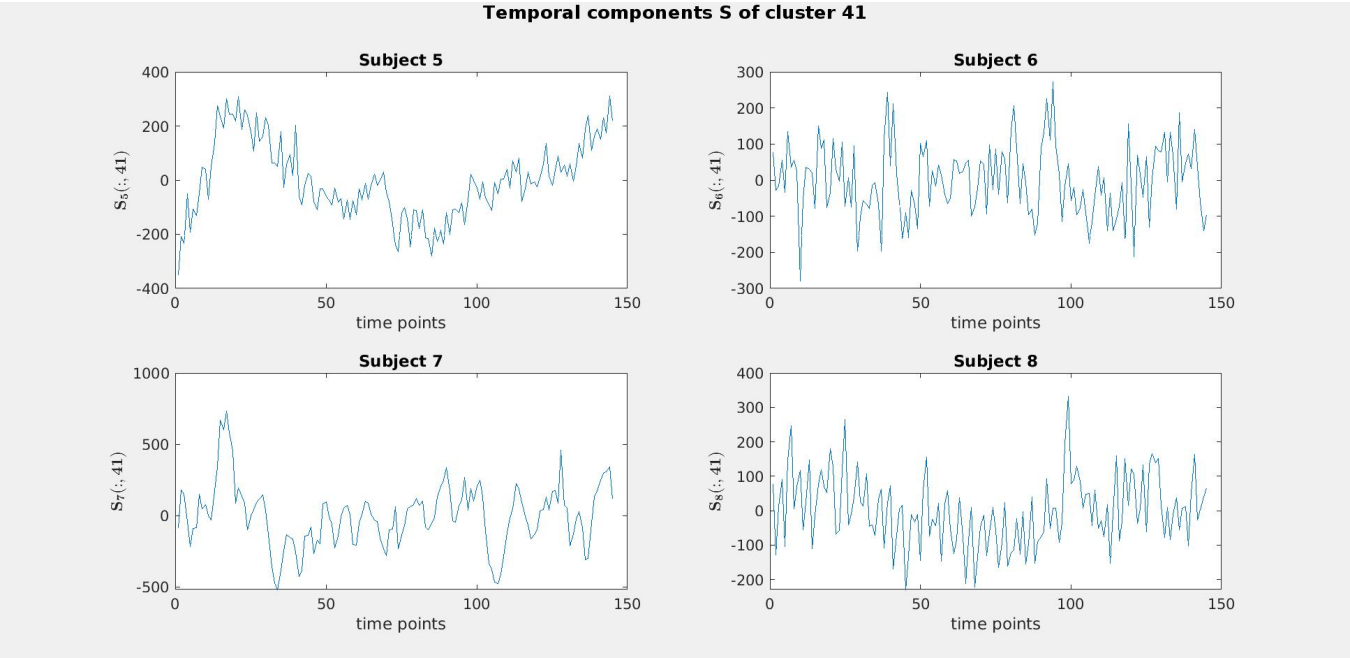


Figure B.377: Waveform of temporal components $\mathbf{S}_k(:, 41)$ for $k = 5, \ldots, 8$ and R=42.

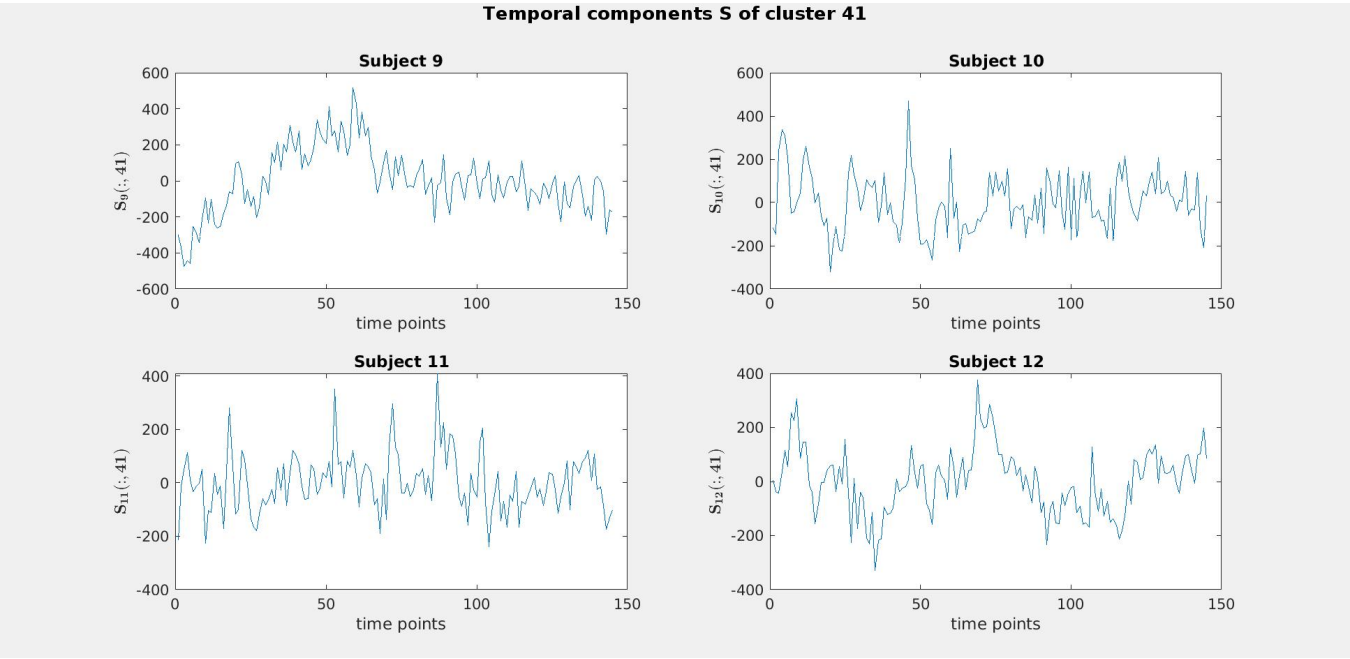


Figure B.378: Waveform of temporal components $\mathbf{S}_k(:, 41)$ for $k = 9, \ldots, 12$ and R=42.

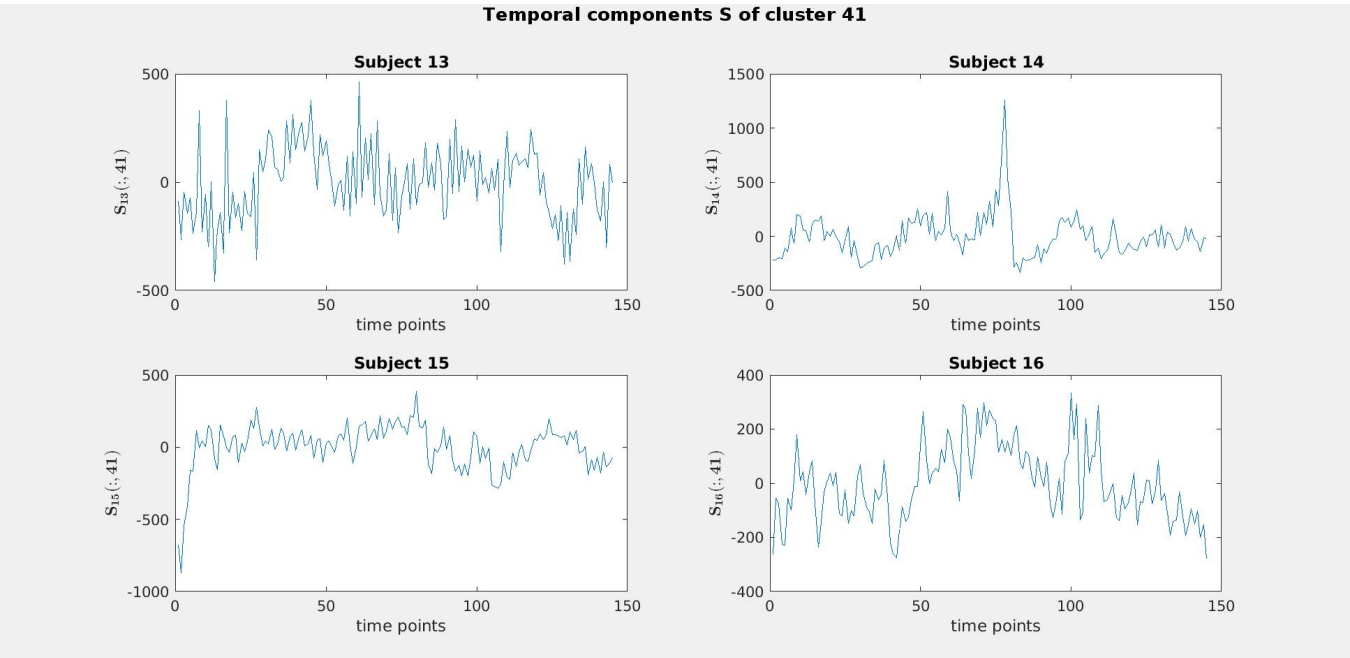


Figure B.379: Waveform of temporal components $\mathbf{S}_k(:, 41)$ for $k = 13, \ldots, 16$ and R=42.

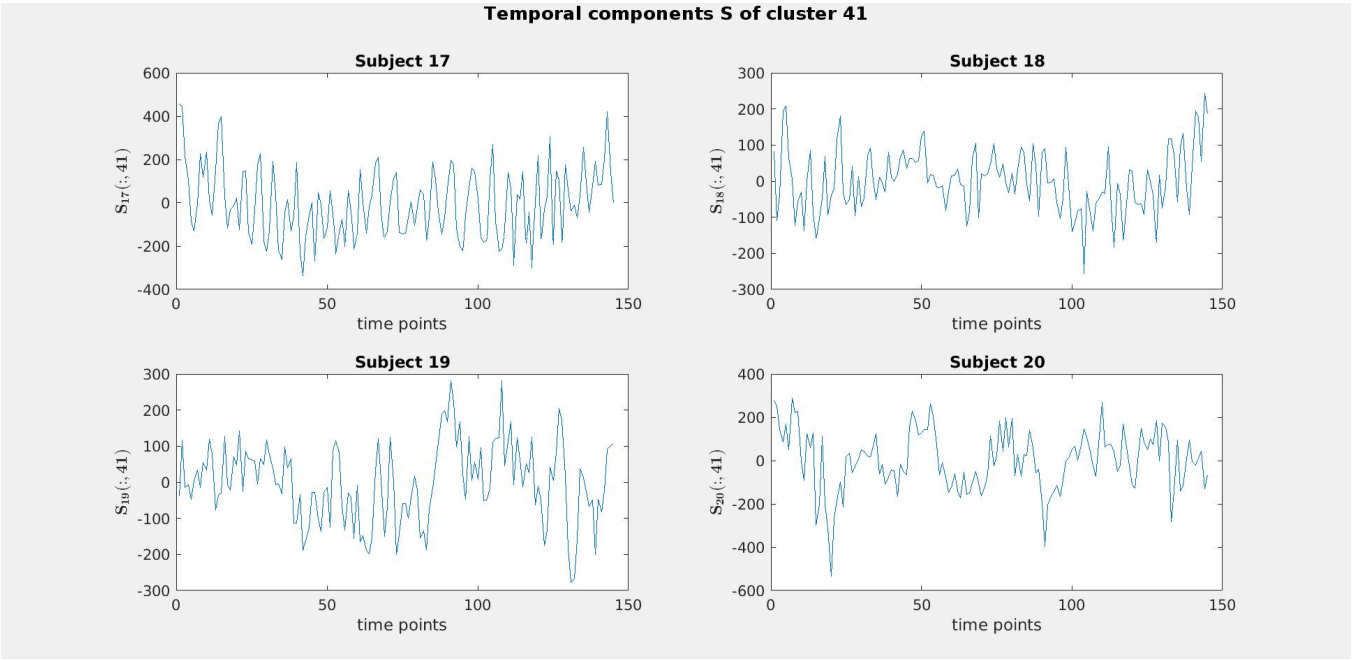


Figure B.380: Waveform of temporal components $\mathbf{S}_k(:, 41)$ for $k = 17, \ldots, 20$ and R=42.

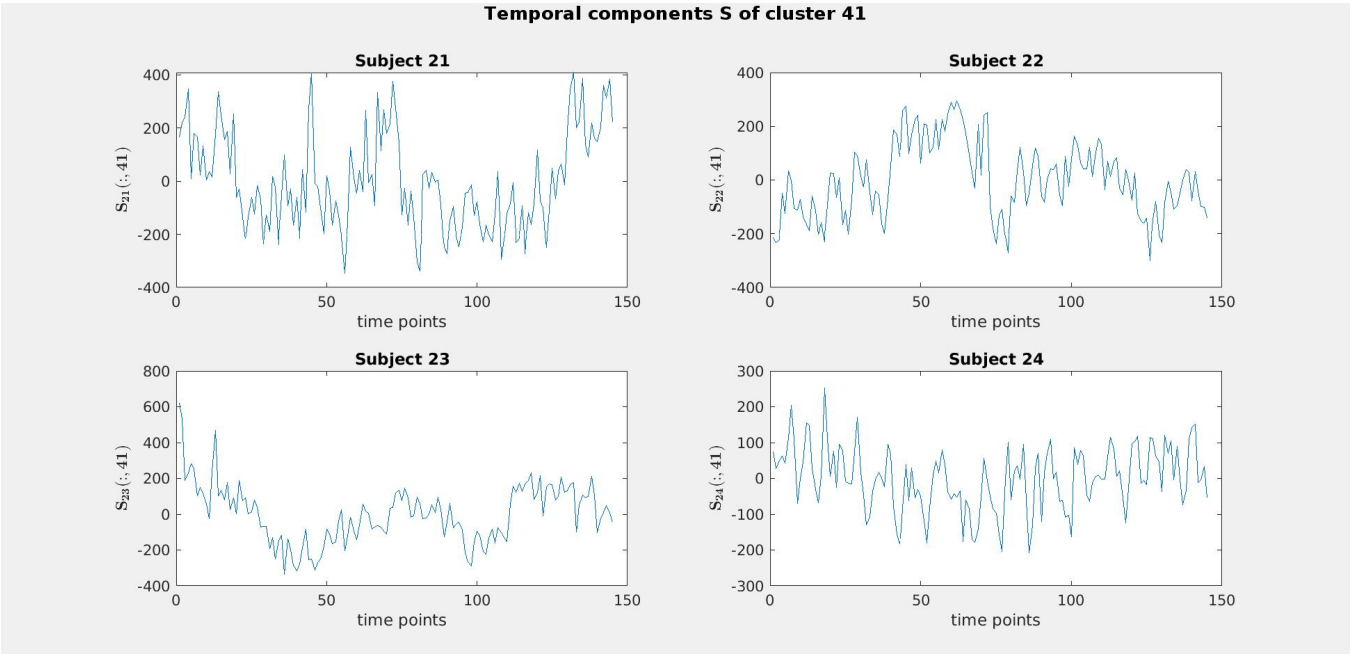


Figure B.381: Waveform of temporal components $\mathbf{S}_k(:, 141)$ for $k = 21, \ldots, 24$ and R=42.

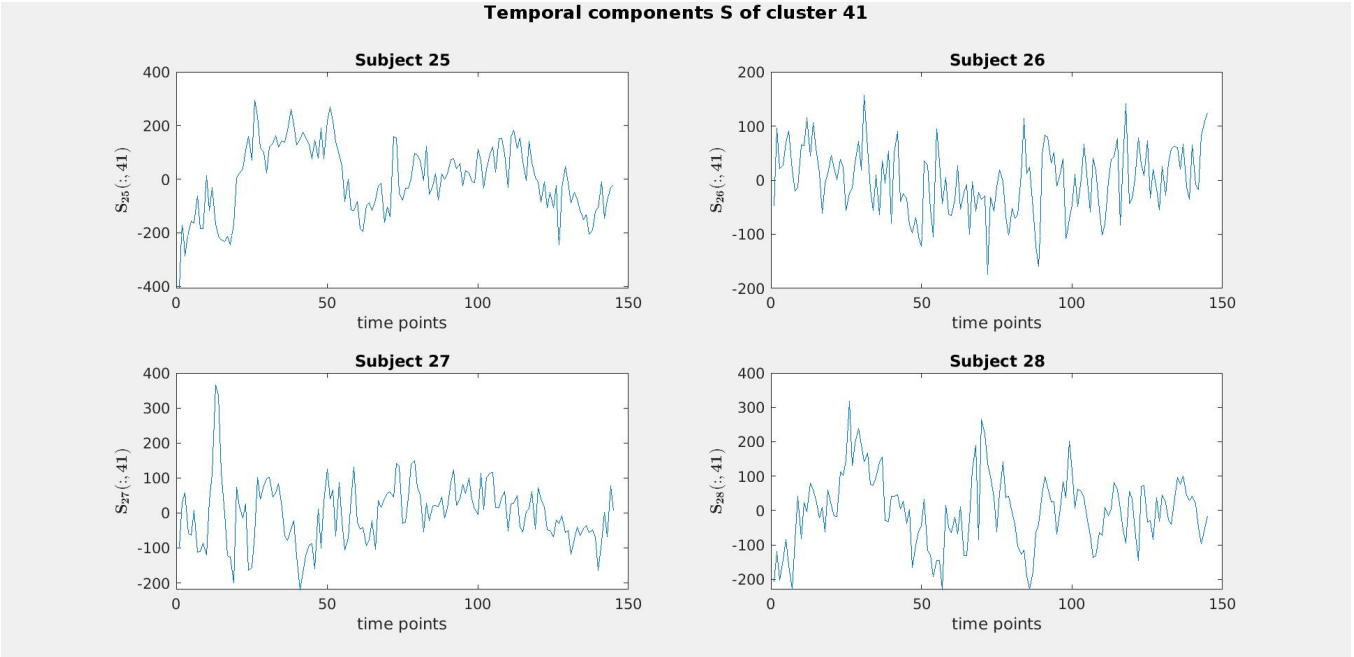


Figure B.382: Waveform of temporal components $\mathbf{S}_k(:, 41)$ for $k = 25, \ldots, 28$ and R=42.

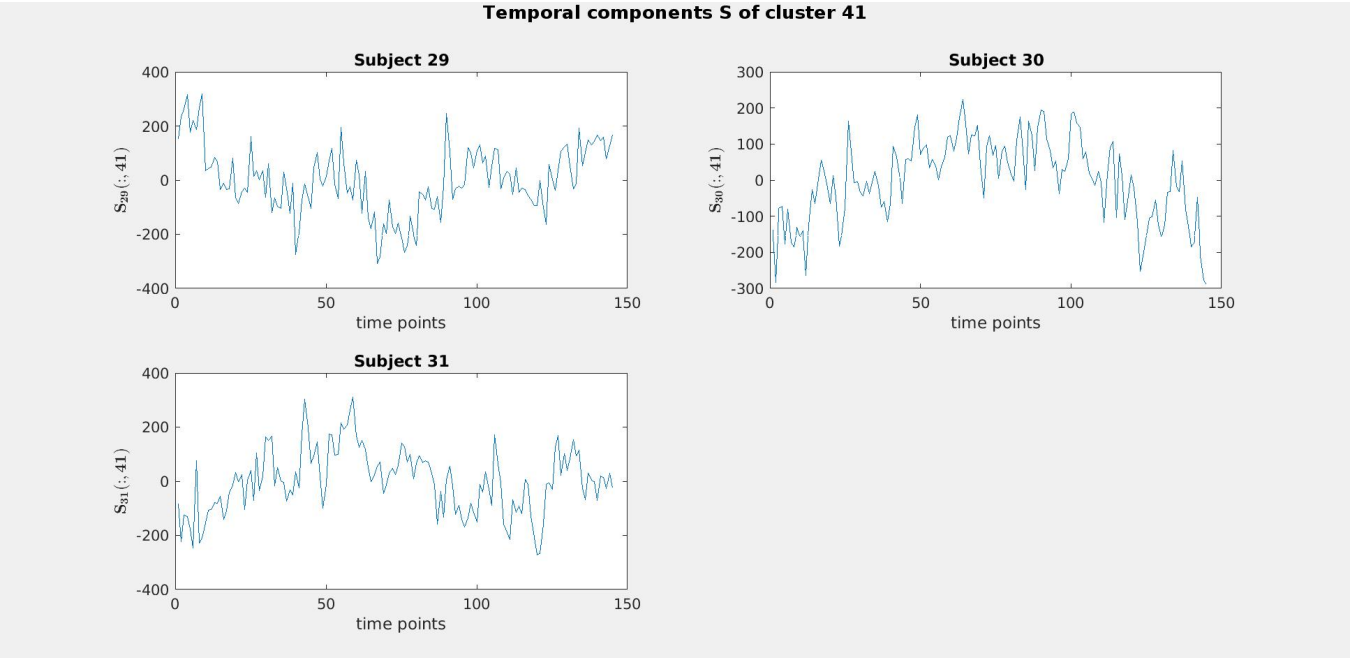


Figure B.383: Waveform of temporal components $\mathbf{S}_k(:, 41)$ for $k = 29, \ldots, 31$ and R=42.

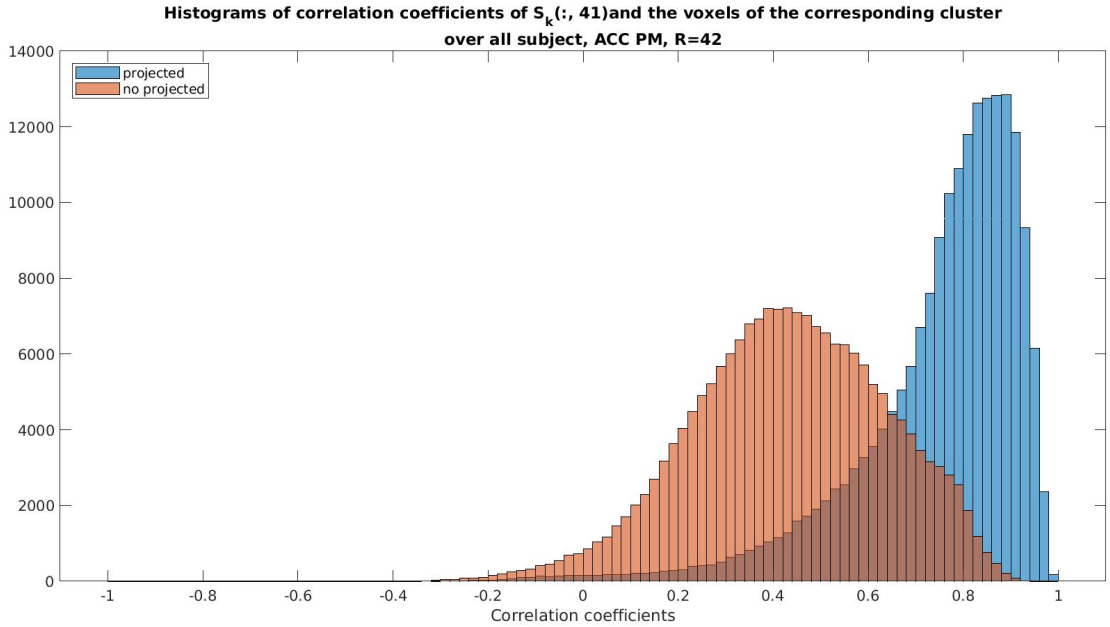


Figure B.384: For the column 41 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 41)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

10. **Cluster 23**

    Cluster 23 contains 7456 voxels, which correspond to 3.299874% of the whole brain (background, soft tissue, bone are excluded).

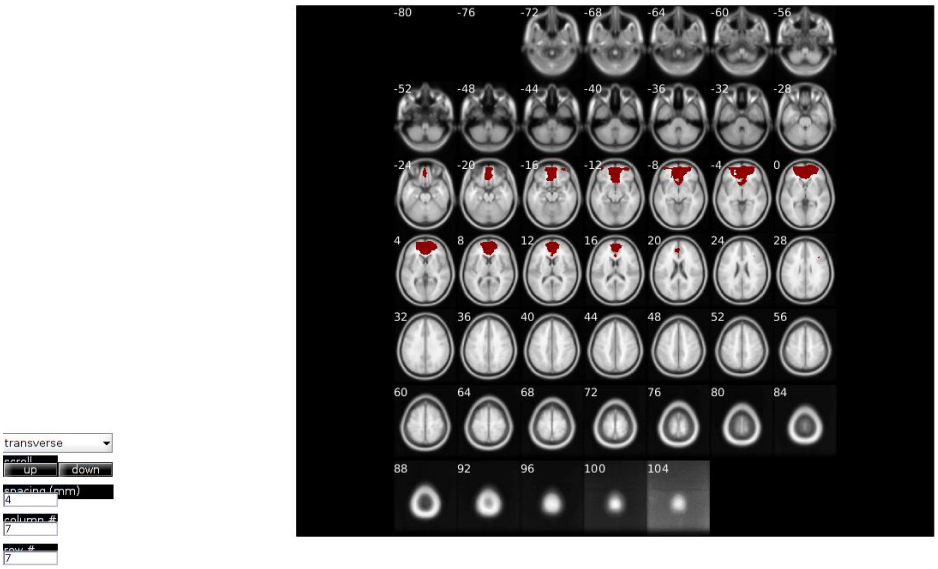


Figure B.385: Cluster 23. Red areas indicate the voxels which belong to cluster 23.

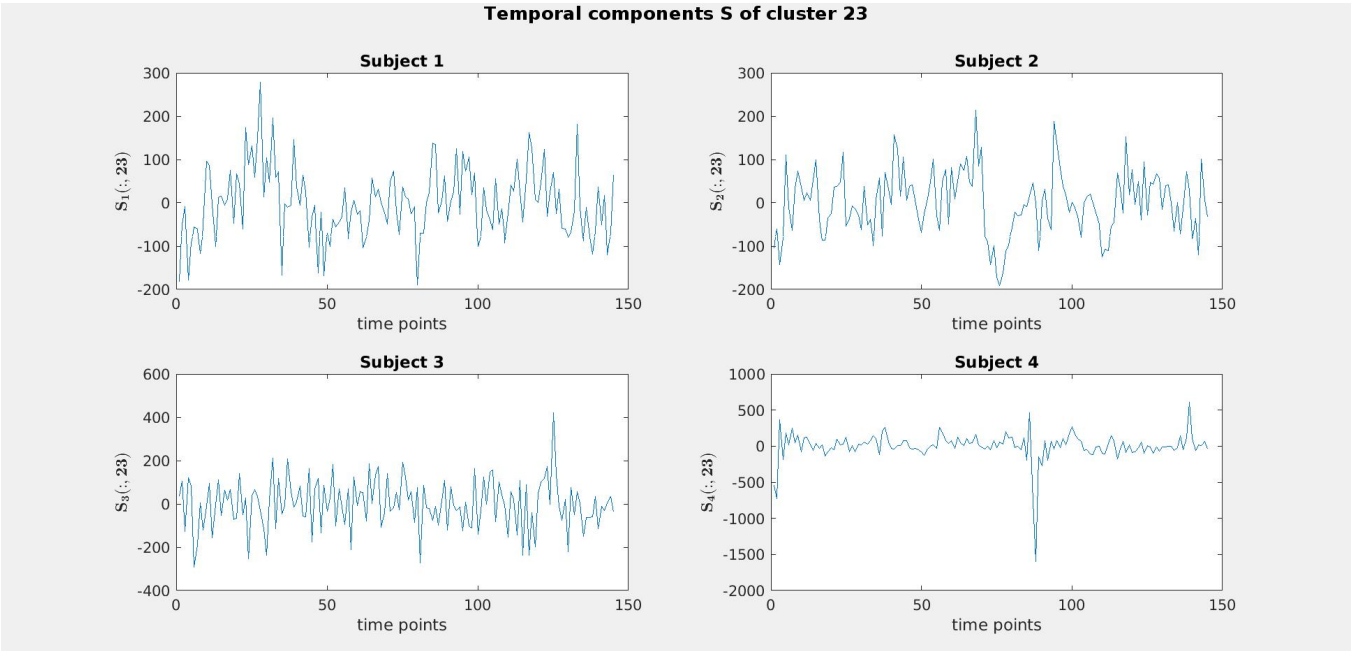


Figure B.386: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 1, \ldots, 4$ and R=42.

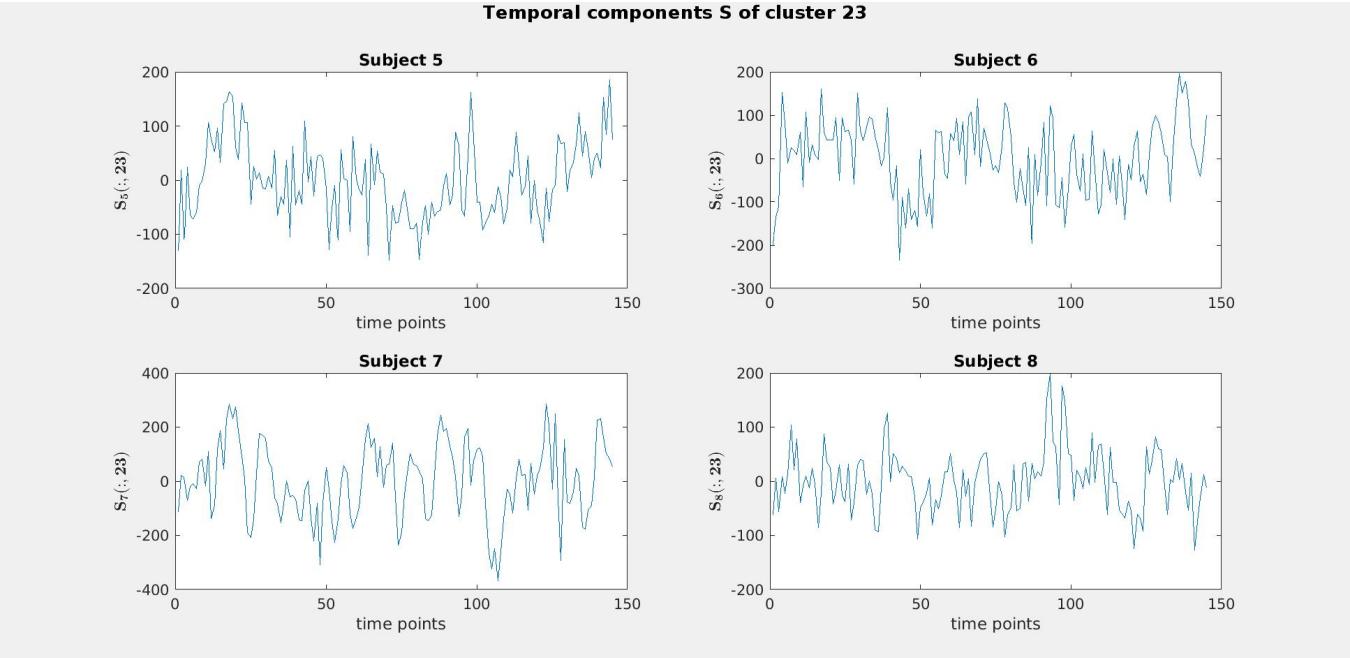


Figure B.387: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 5, \ldots, 8$ and R=42.

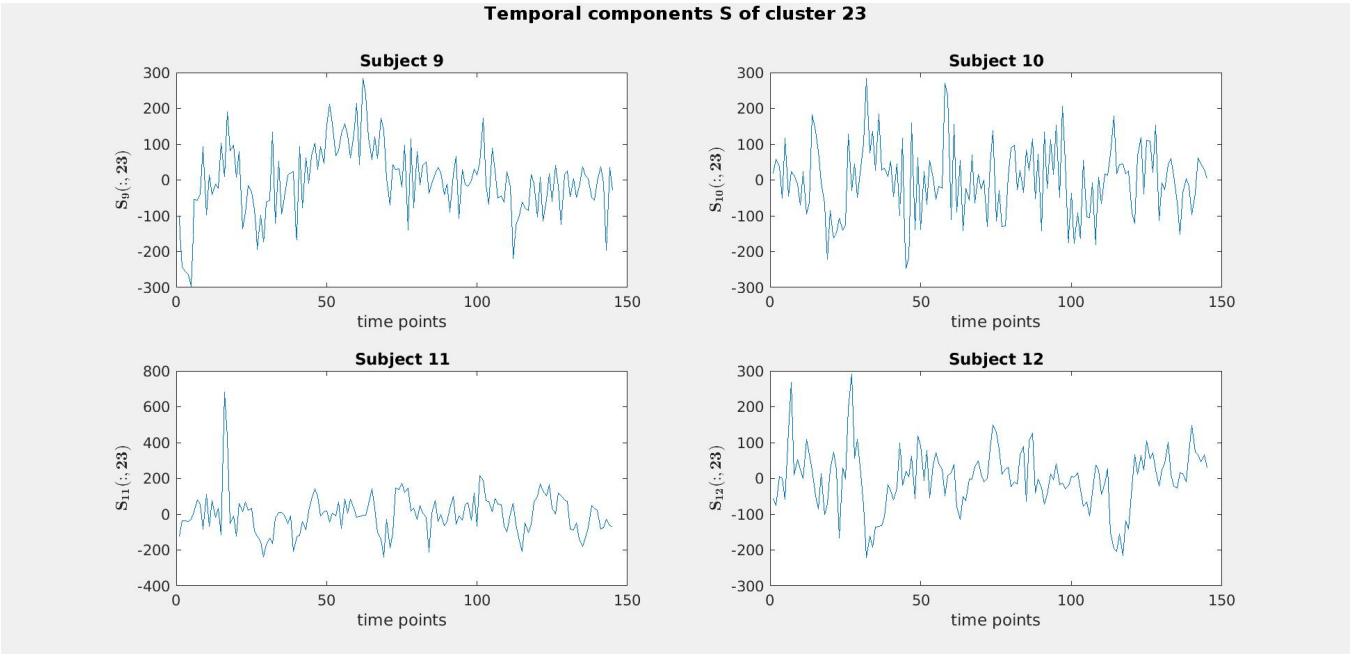


Figure B.388: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 9, \ldots, 12$ and R=42.

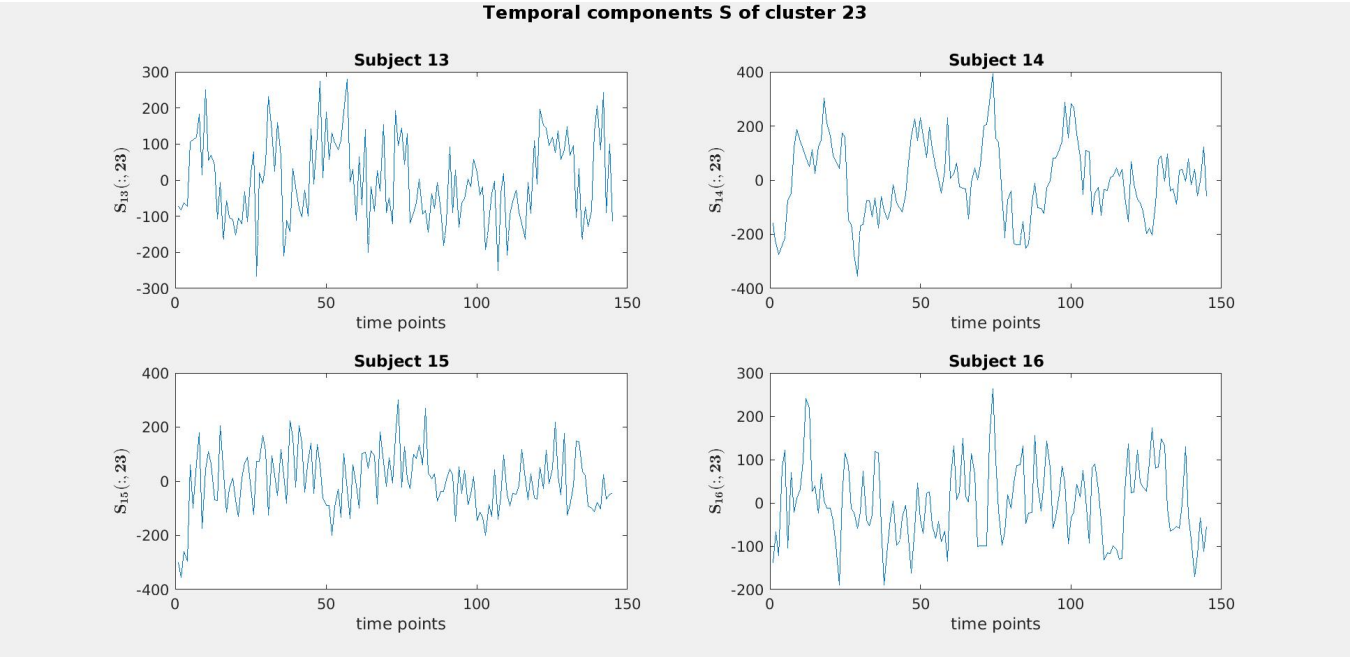


Figure B.389: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 13, \ldots, 16$ and R=42.

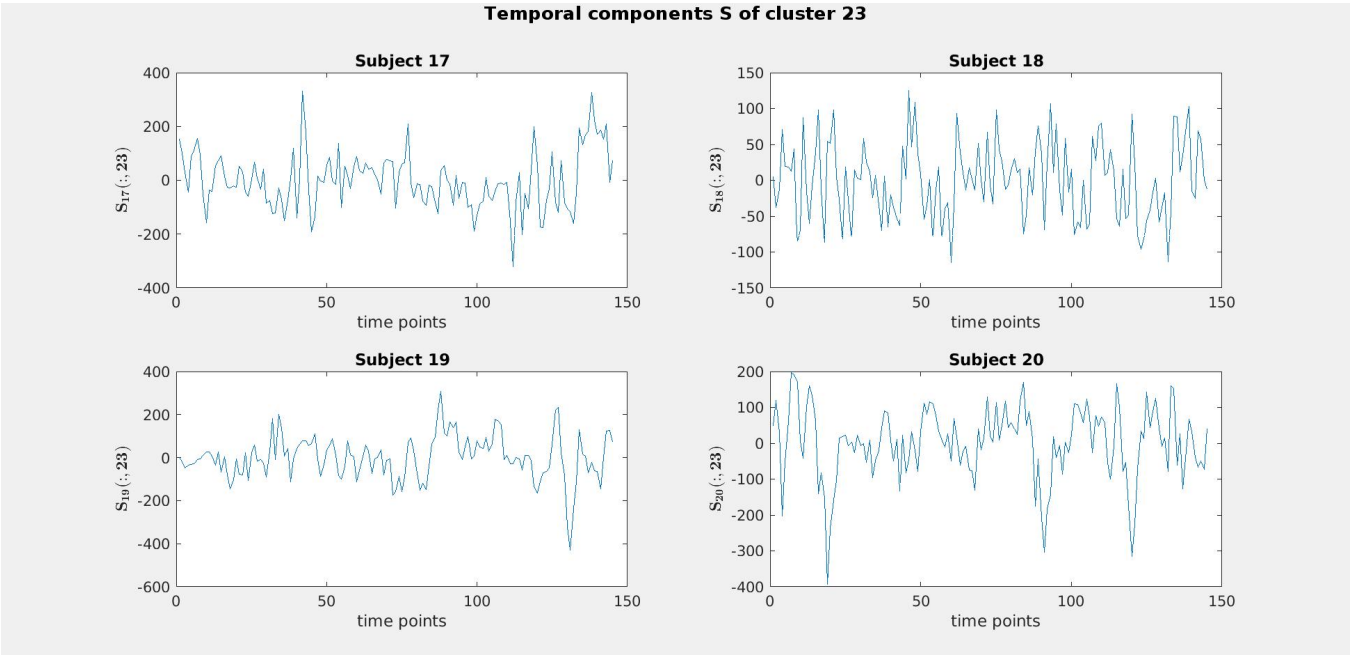


Figure B.390: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 17, \ldots, 20$ and R=42.

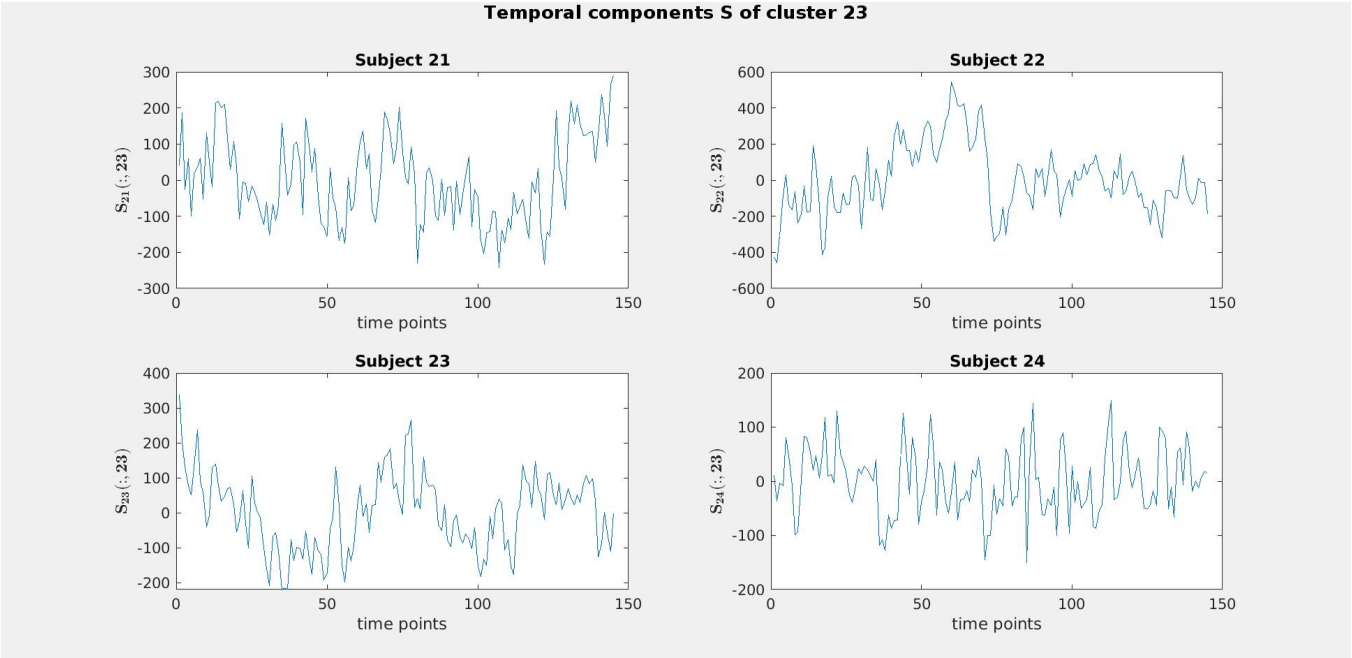


Figure B.391: Waveform of temporal components $\mathbf{S}_k(:, 123)$ for $k = 21, \ldots, 24$ and R=42.

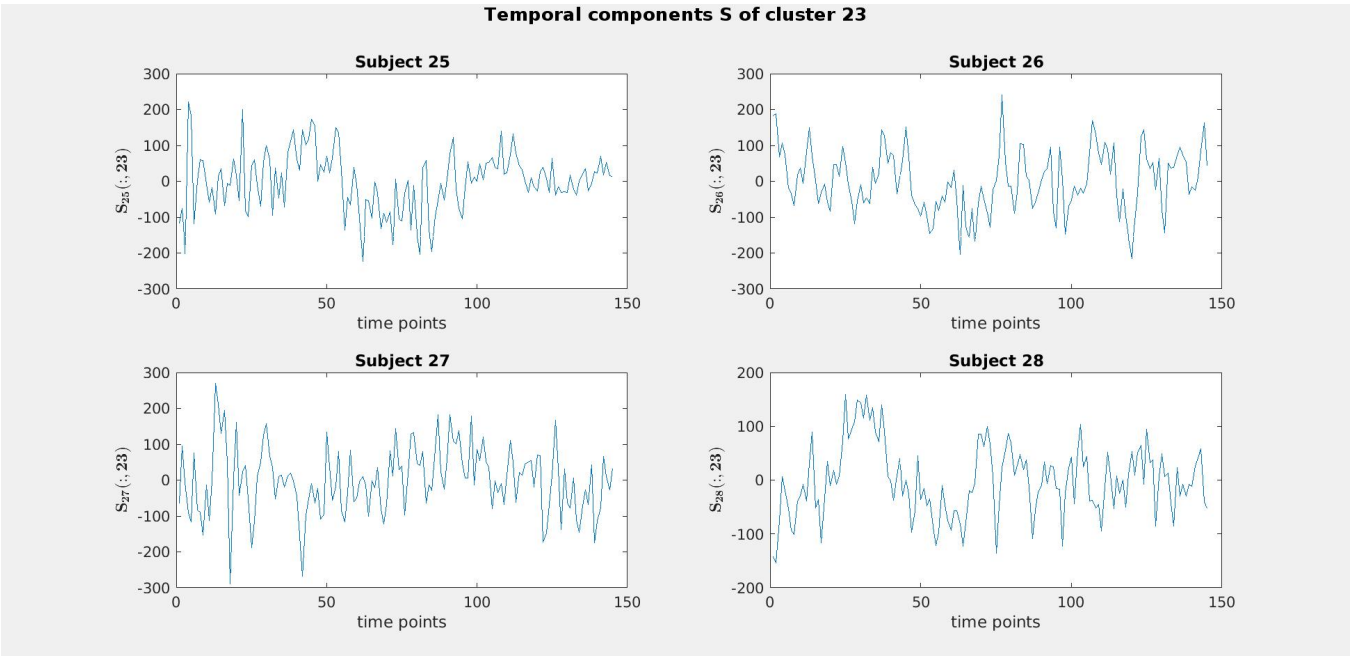


Figure B.392: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 25, \ldots, 28$ and R=42.

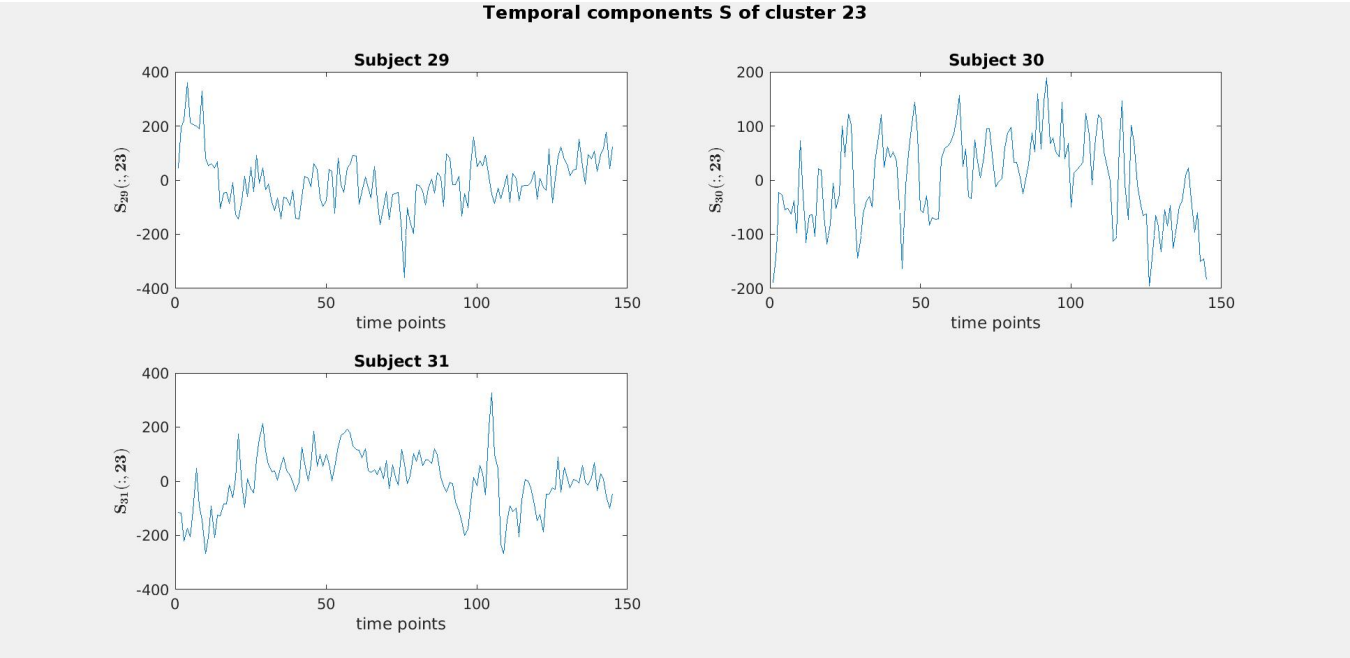


Figure B.393: Waveform of temporal components $\mathbf{S}_k(:, 23)$ for $k = 29, \ldots, 31$ and R=42.

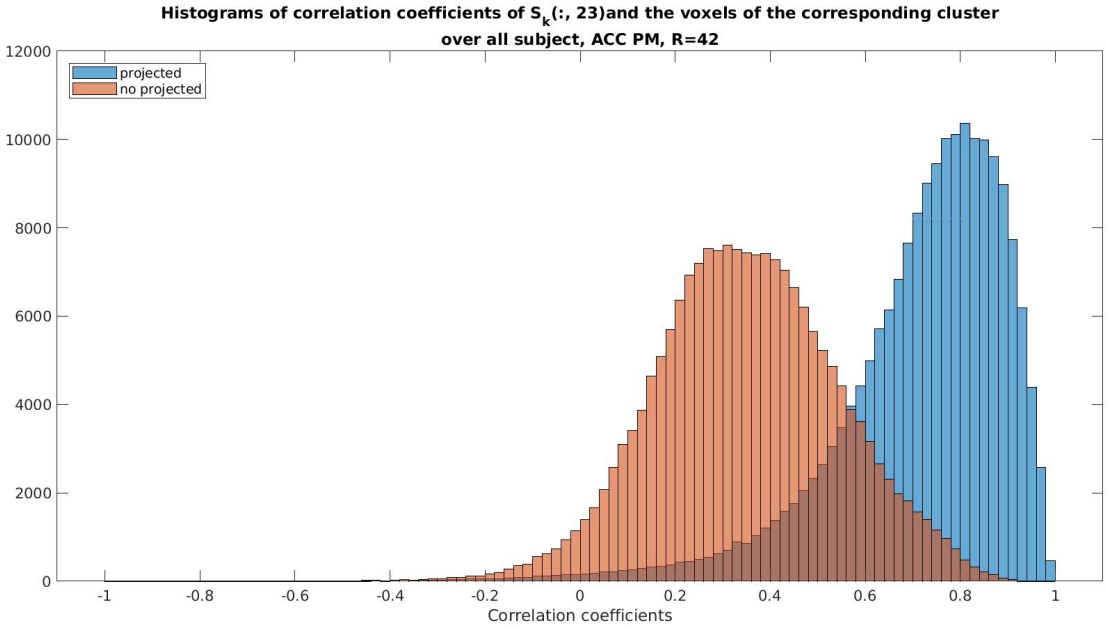


Figure B.394: For the column 23 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 23)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

11. **Cluster 30**

Cluster 30 contains 7008 voxels, which correspond to 3.101599% of the whole brain (background, soft tissue, bone are excluded).

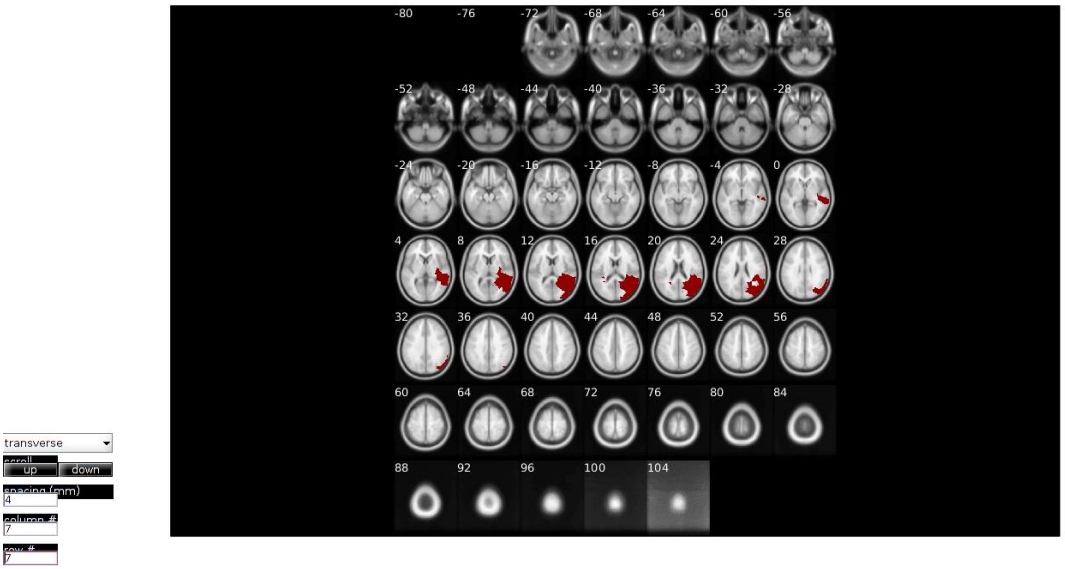


Figure B.395: Cluster 30. Red areas indicate the voxels which belong to cluster 30.

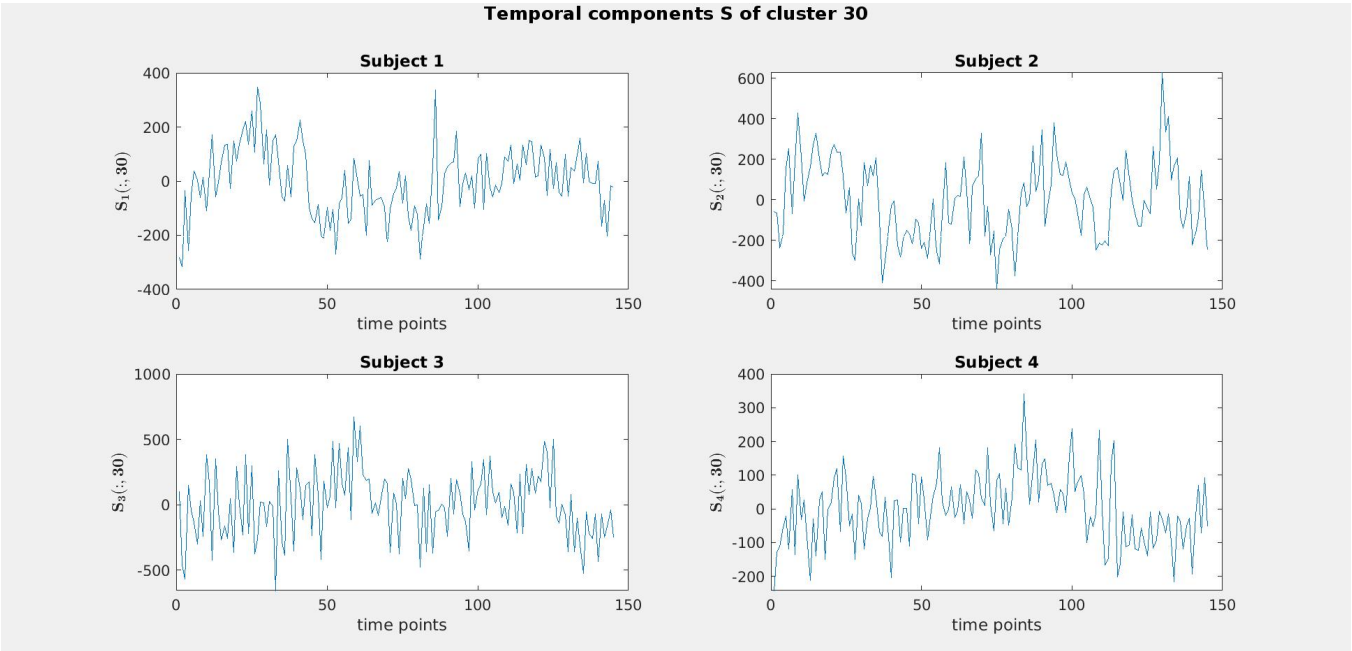


Figure B.396: Waveform of temporal components $\mathbf{S}_k(:, 30)$ for $k = 1, \ldots, 4$ and R=42.

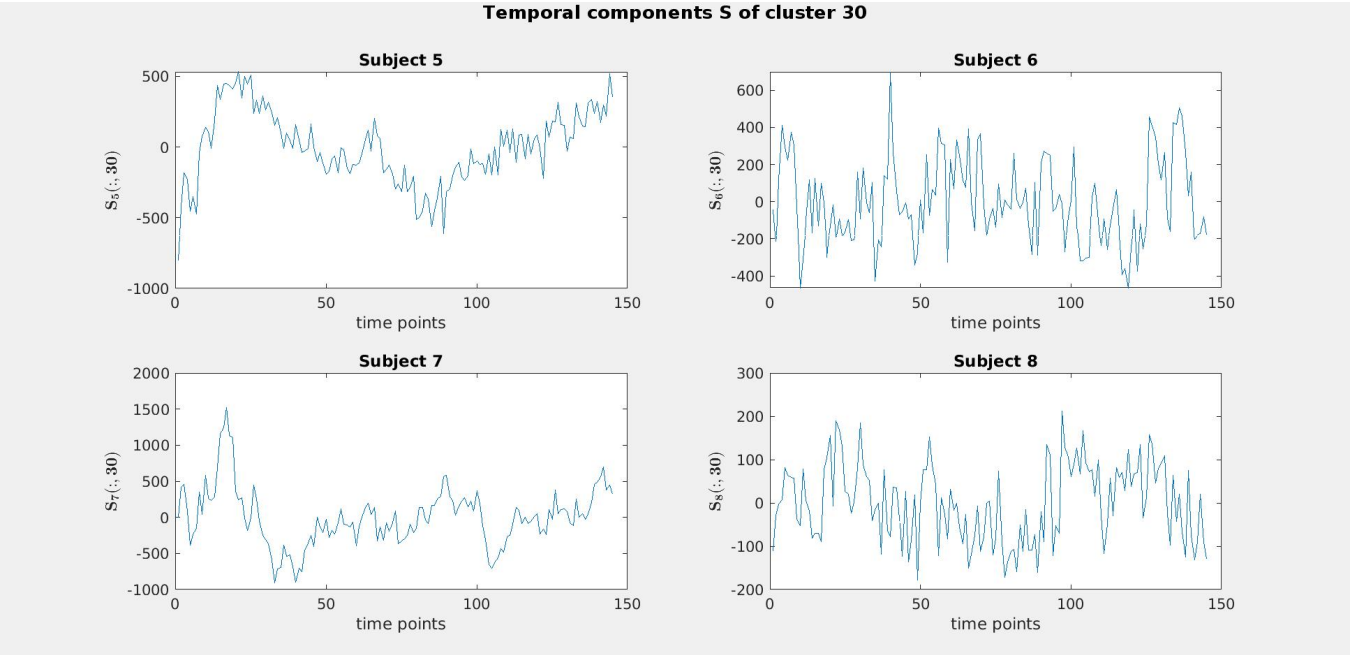


Figure B.397: Waveform of temporal components $\mathbf{S}_k(:, 30)$ for $k = 5, \ldots, 8$ and R=42.

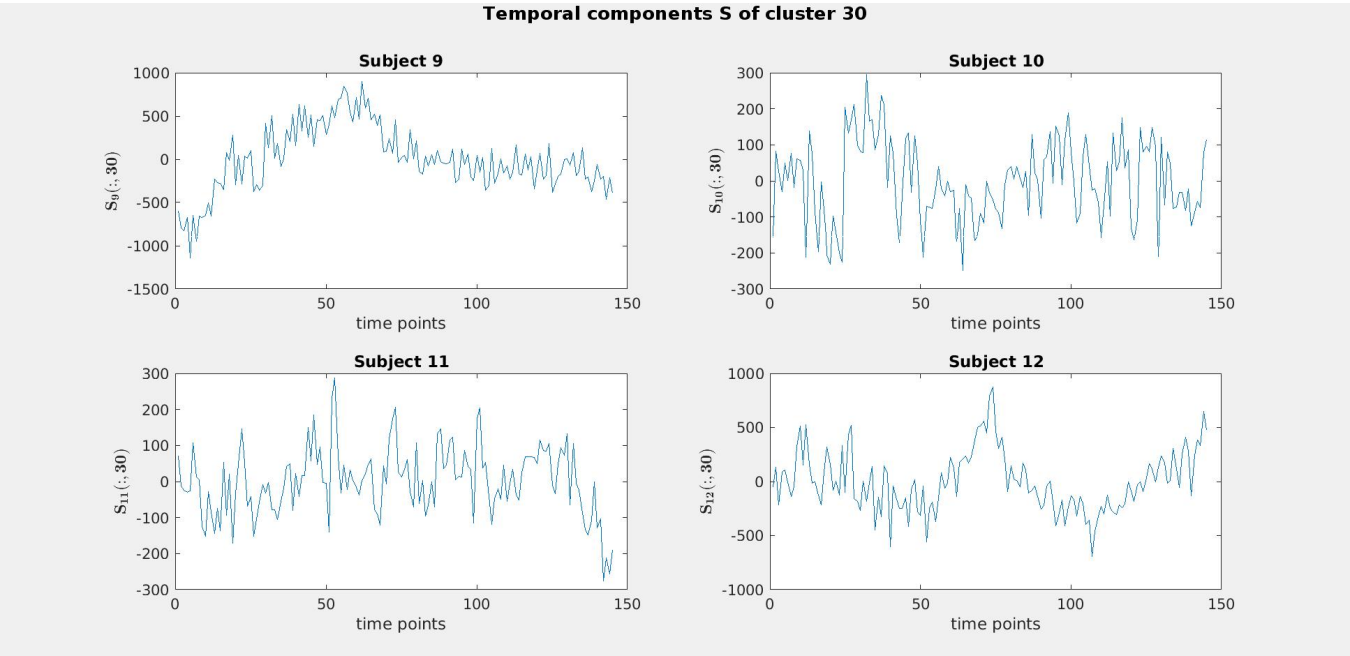


Figure B.398: Waveform of temporal components $\mathbf{S}_k(:, 30)$ for $k = 9, \ldots, 12$ and R=42.

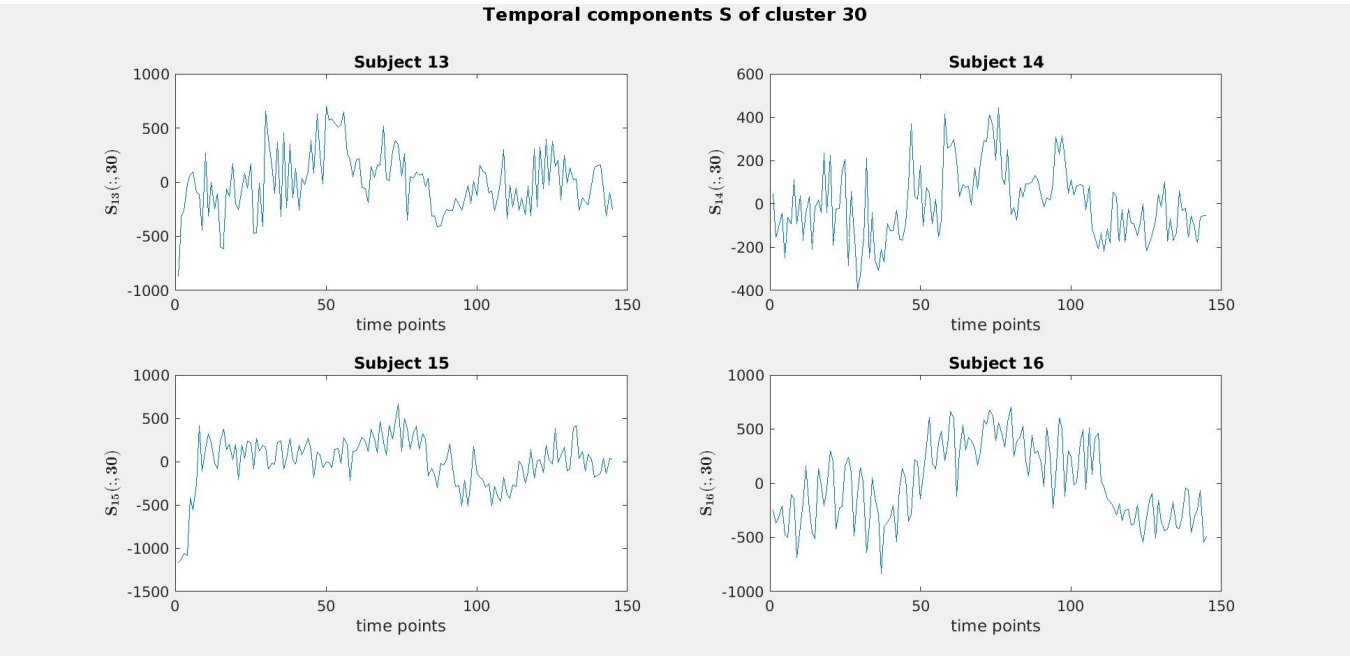


Figure B.399: Waveform of temporal components $\mathbf{S}_k(:, 30)$ for $k = 13, \ldots, 16$ and R=42.

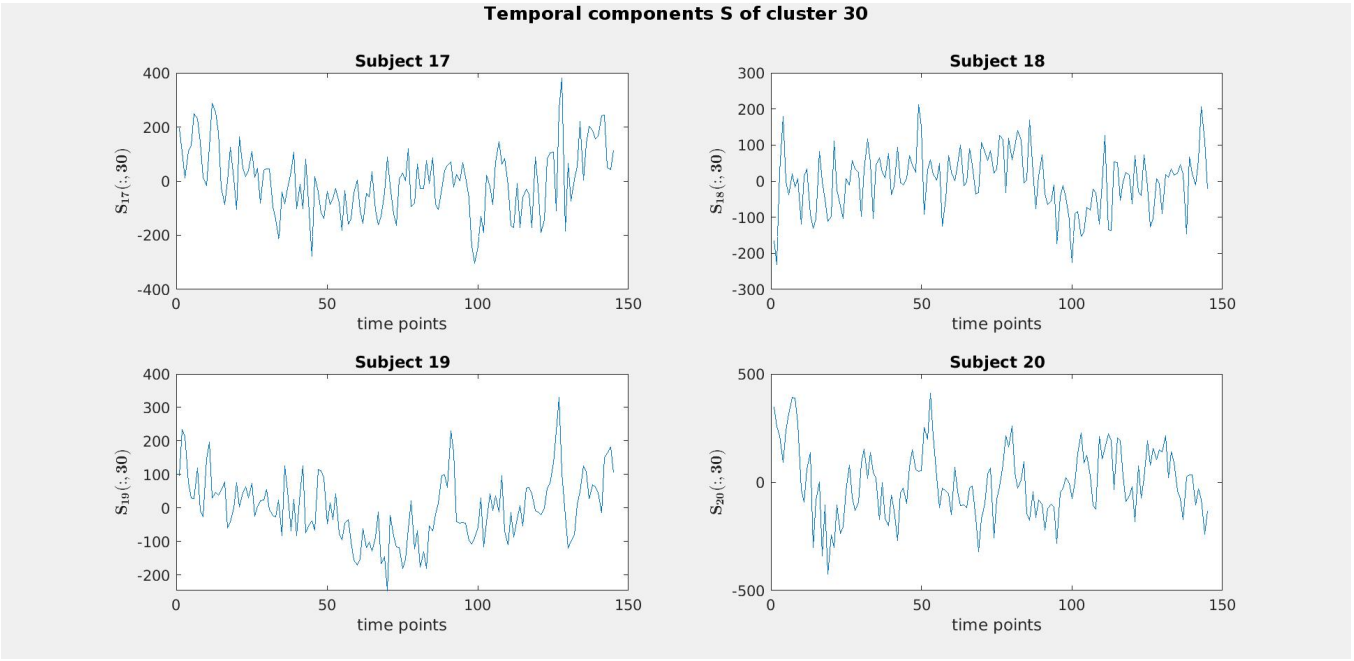


Figure B.400: Waveform of temporal components $\mathbf{S}_k(:, 30)$ for $k = 17, \ldots, 20$ and R=42.

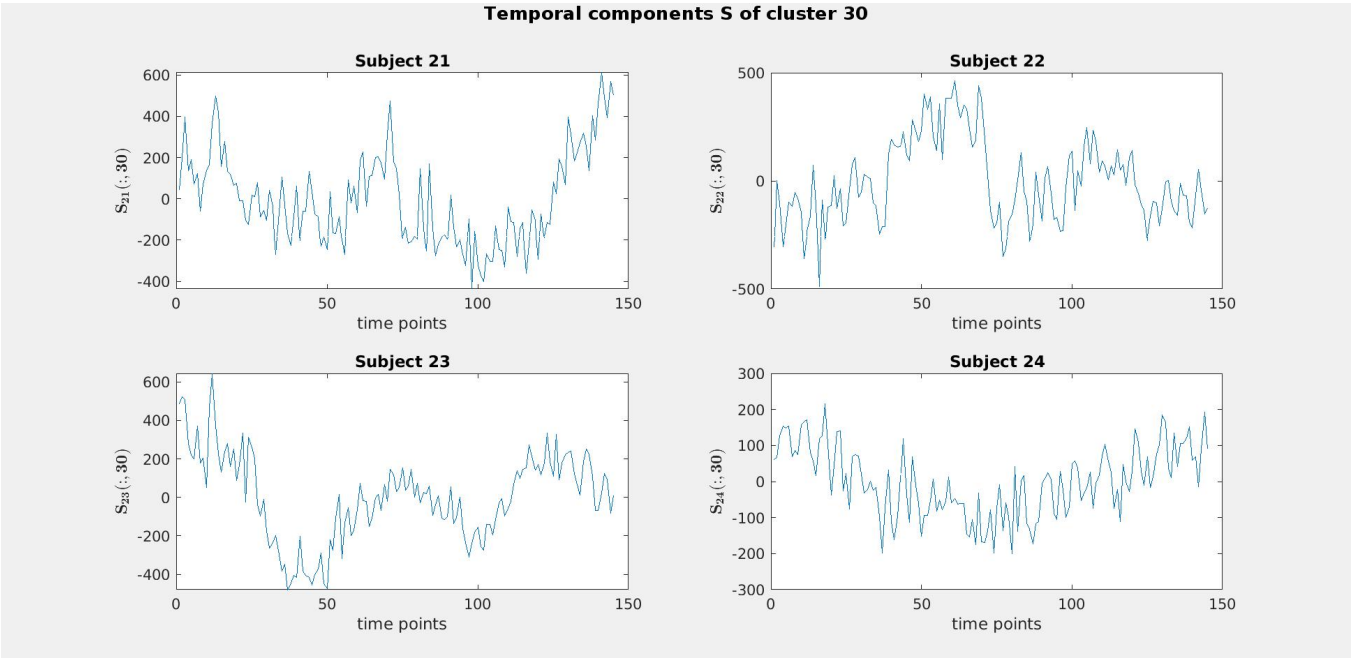


Figure B.401: Waveform of temporal components $\mathbf{S}_k(:, 130)$ for $k = 21, \ldots, 24$ and R=42.

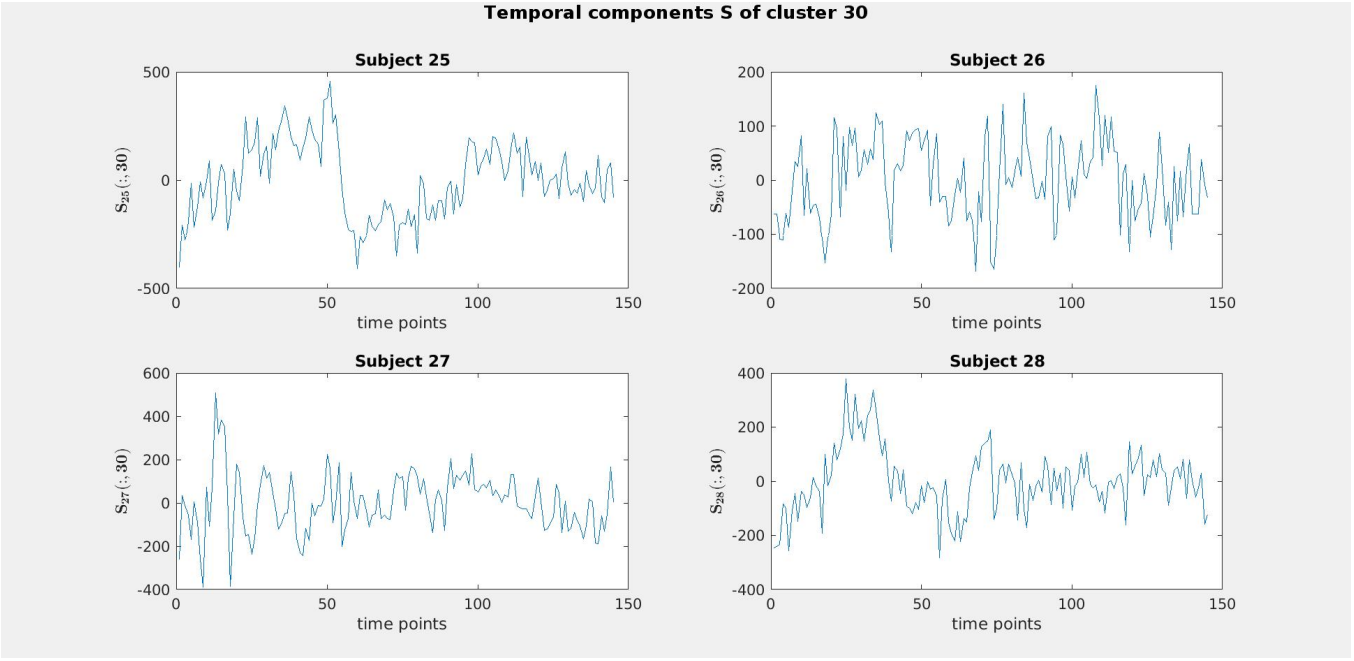


Figure B.402: Waveform of temporal components $\mathbf{S}_k(:, 30)$ for $k = 25, \ldots, 28$ and R=42.

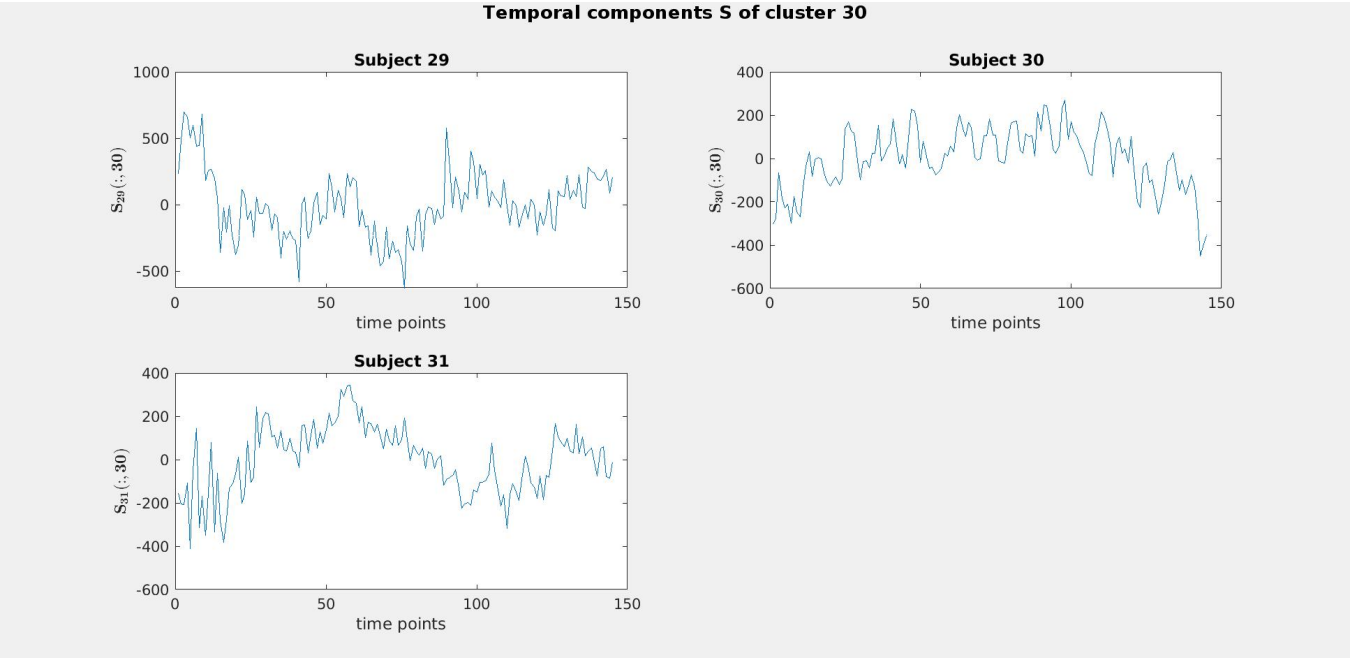


Figure B.403: Waveform of temporal components $\mathbf{S}_k(:, 30)$ for $k = 29, \ldots, 31$ and R=42.

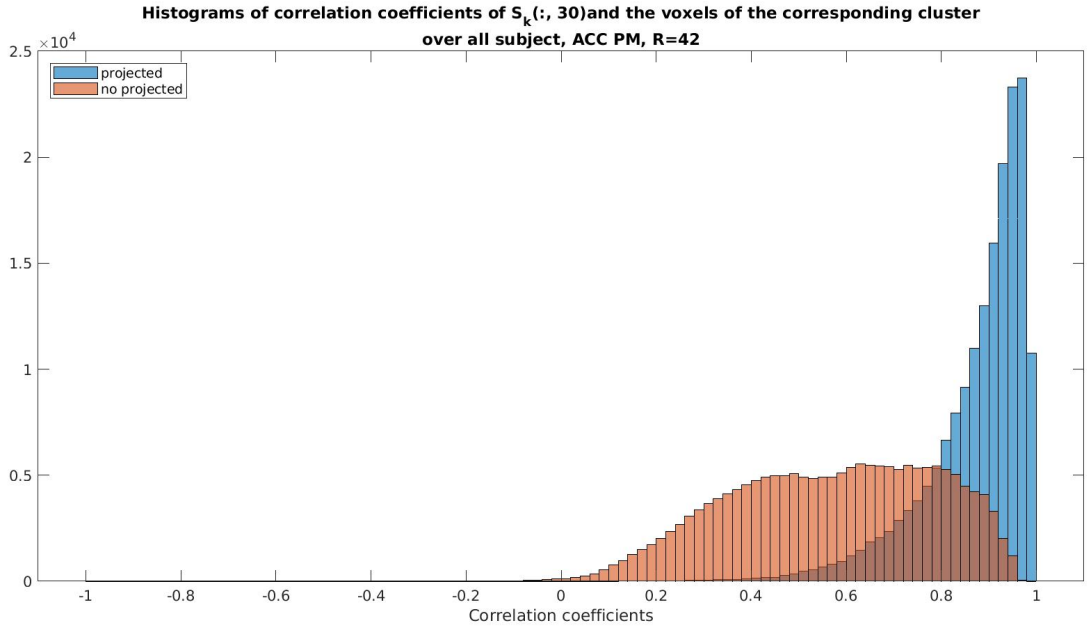


Figure B.404: For the column 30 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 30)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

12. **Cluster 22**

    Cluster 22 contains 6892 voxels, which correspond to 3.050259% of the whole brain (background, soft tissue, bone are excluded).

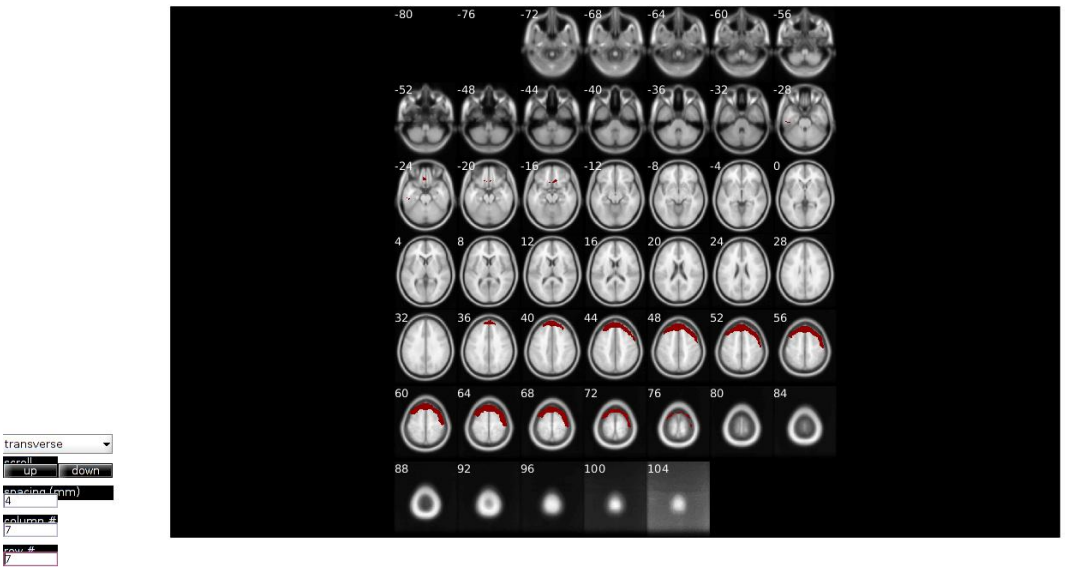


Figure B.405: Cluster 22. Red areas indicate the voxels which belong to cluster 22.

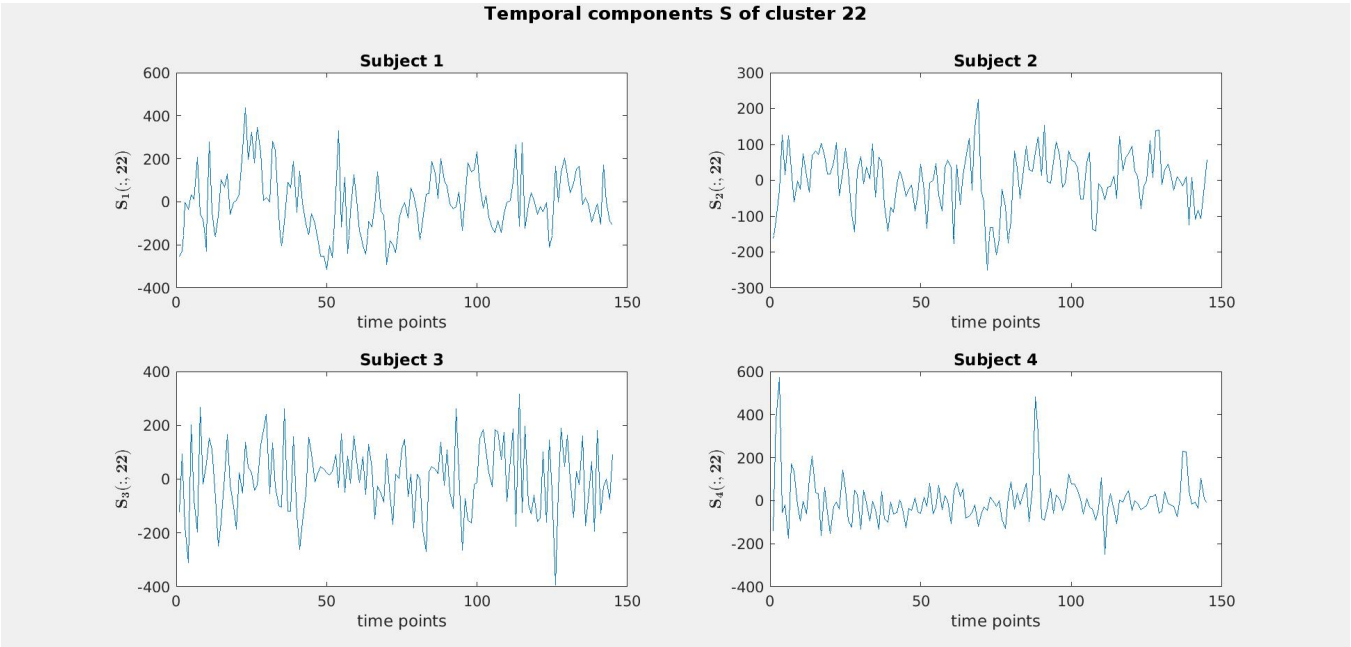


Figure B.406: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 1, \ldots, 4$ and R=42.

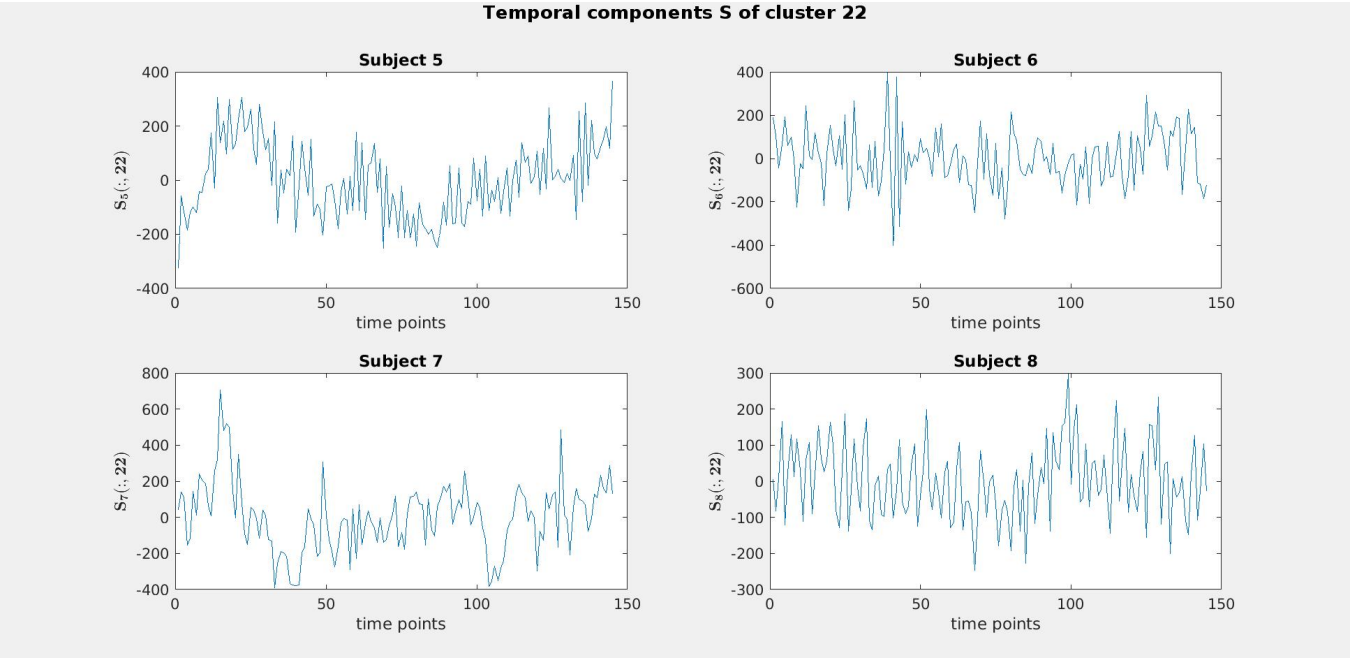


Figure B.407: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 5, \ldots, 8$ and R=42.

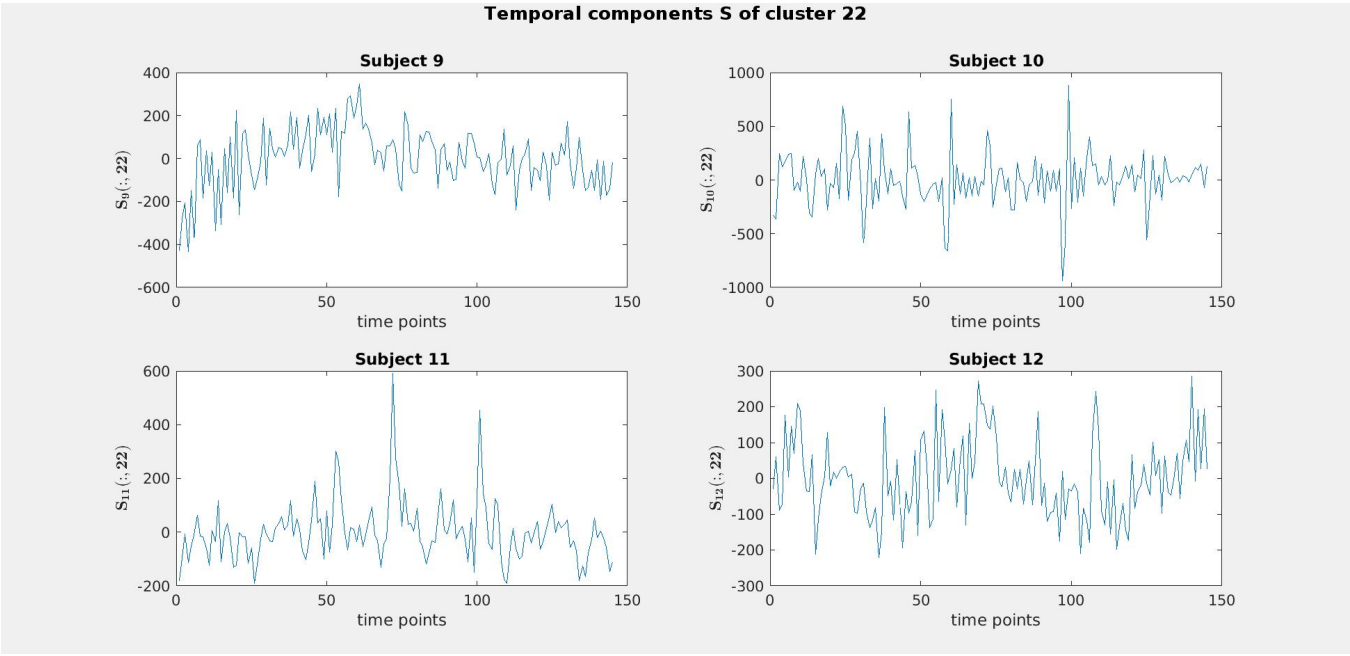


Figure B.408: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 9, \ldots, 12$ and R=42.

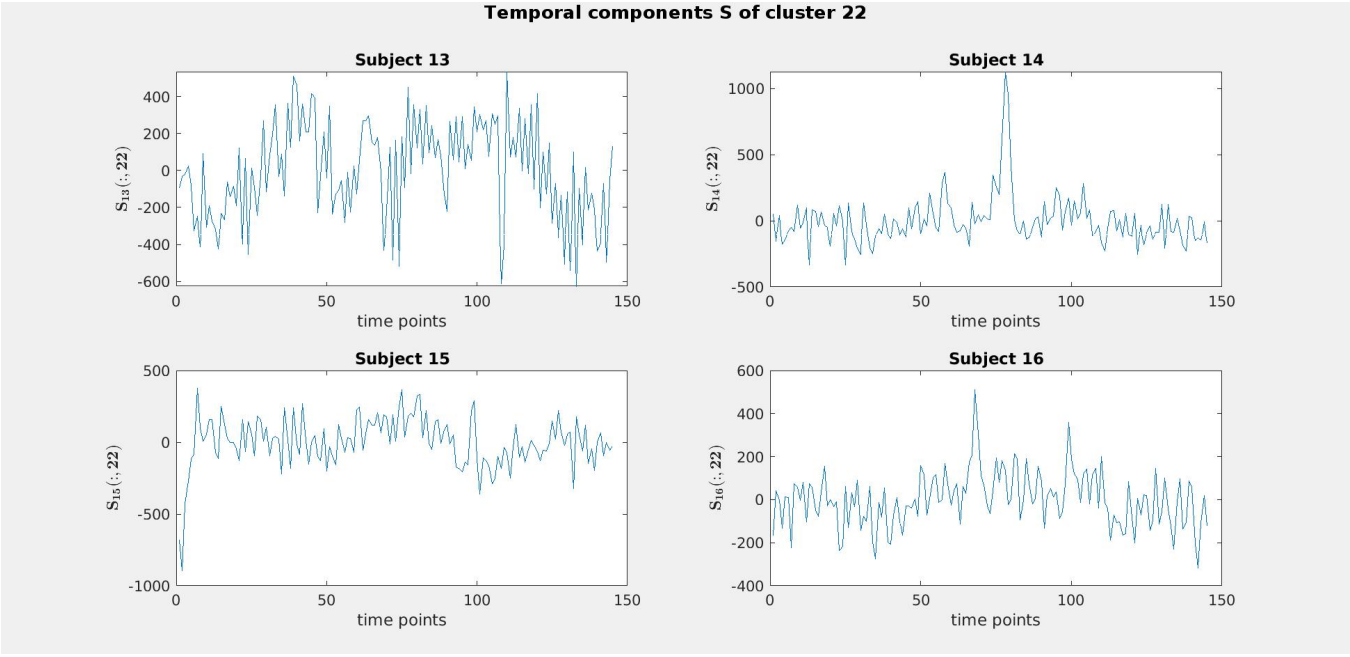


Figure B.409: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 13, \ldots, 16$ and R=42.

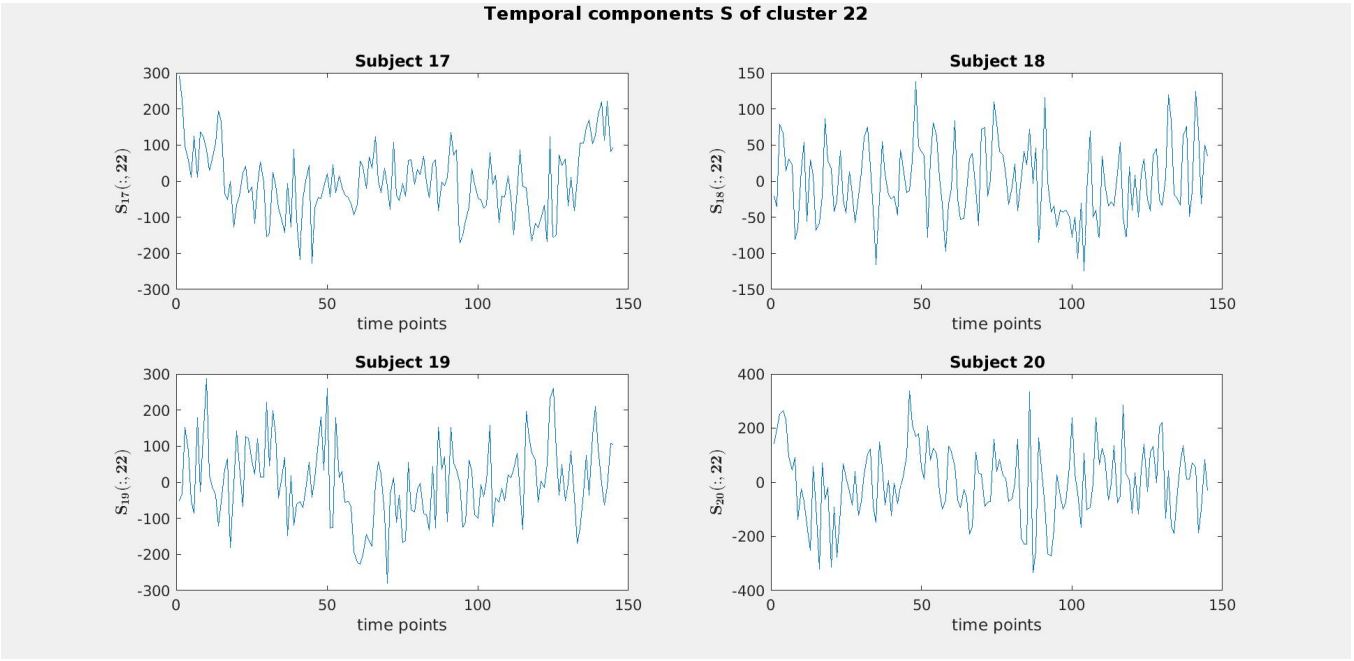


Figure B.410: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 17, \ldots, 20$ and R=42.

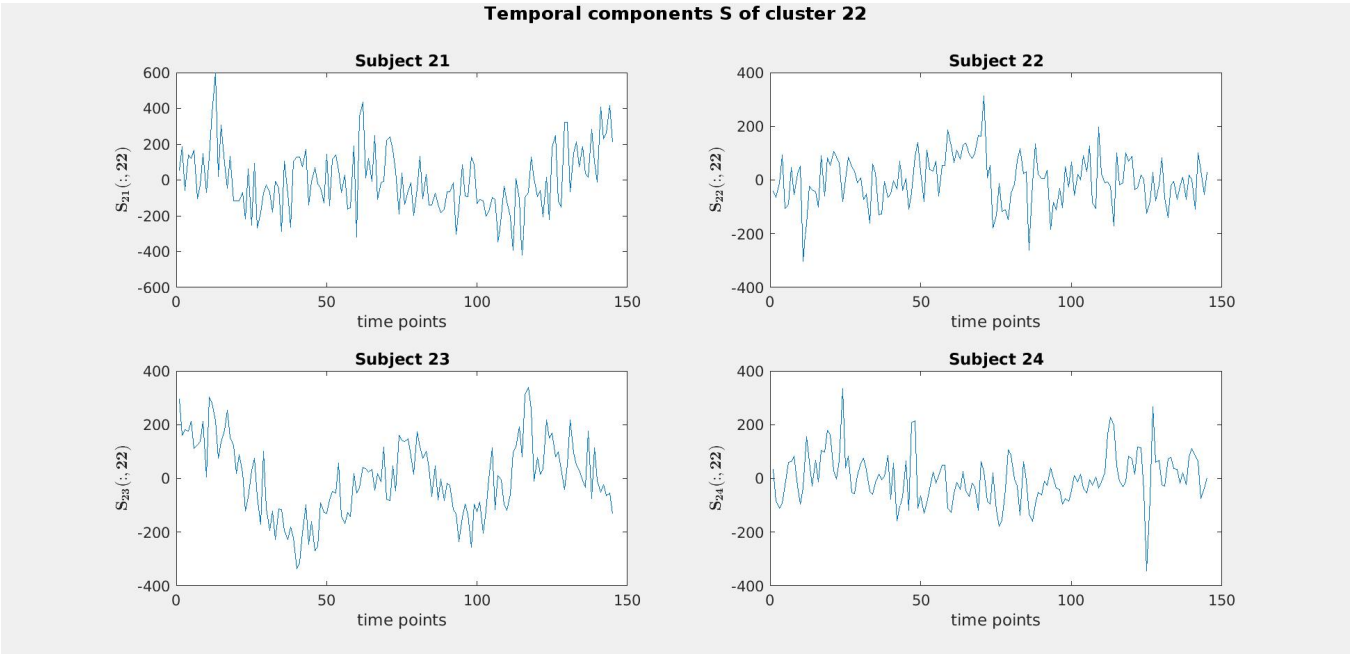


Figure B.411: Waveform of temporal components $\mathbf{S}_k(:, 122)$ for $k = 21, \ldots, 24$ and R=42.

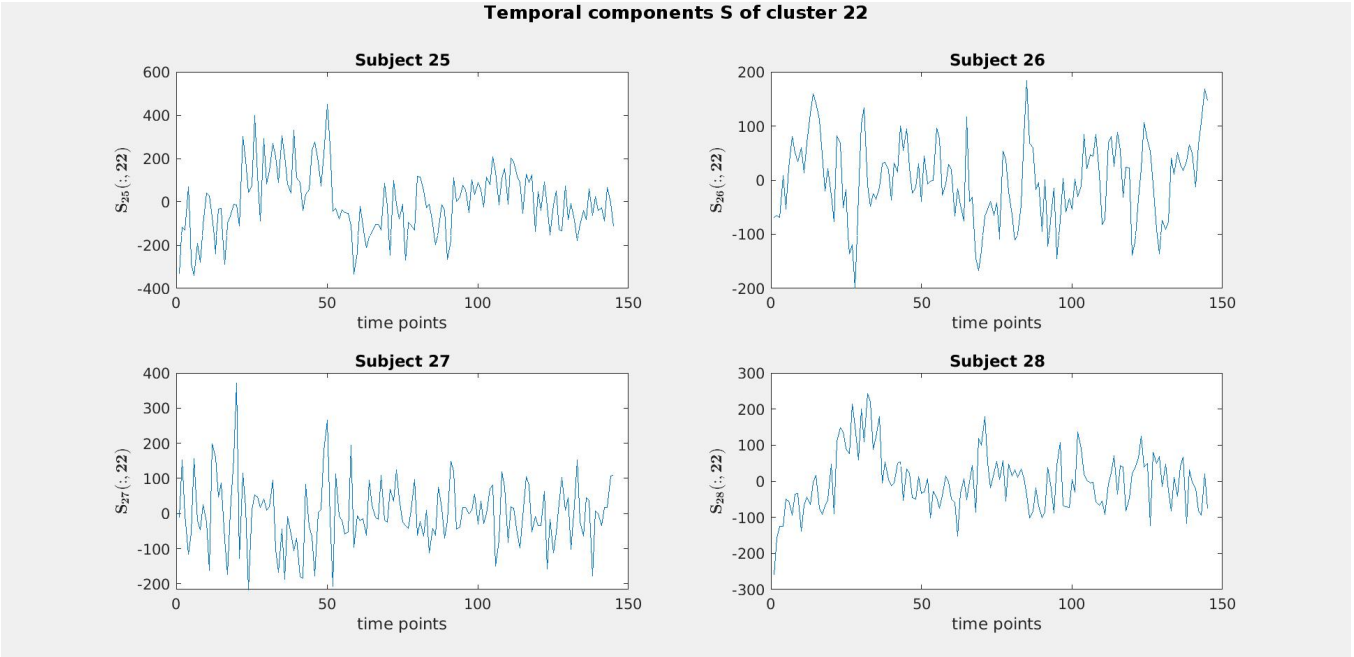


Figure B.412: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 25, \ldots, 28$ and R=42.

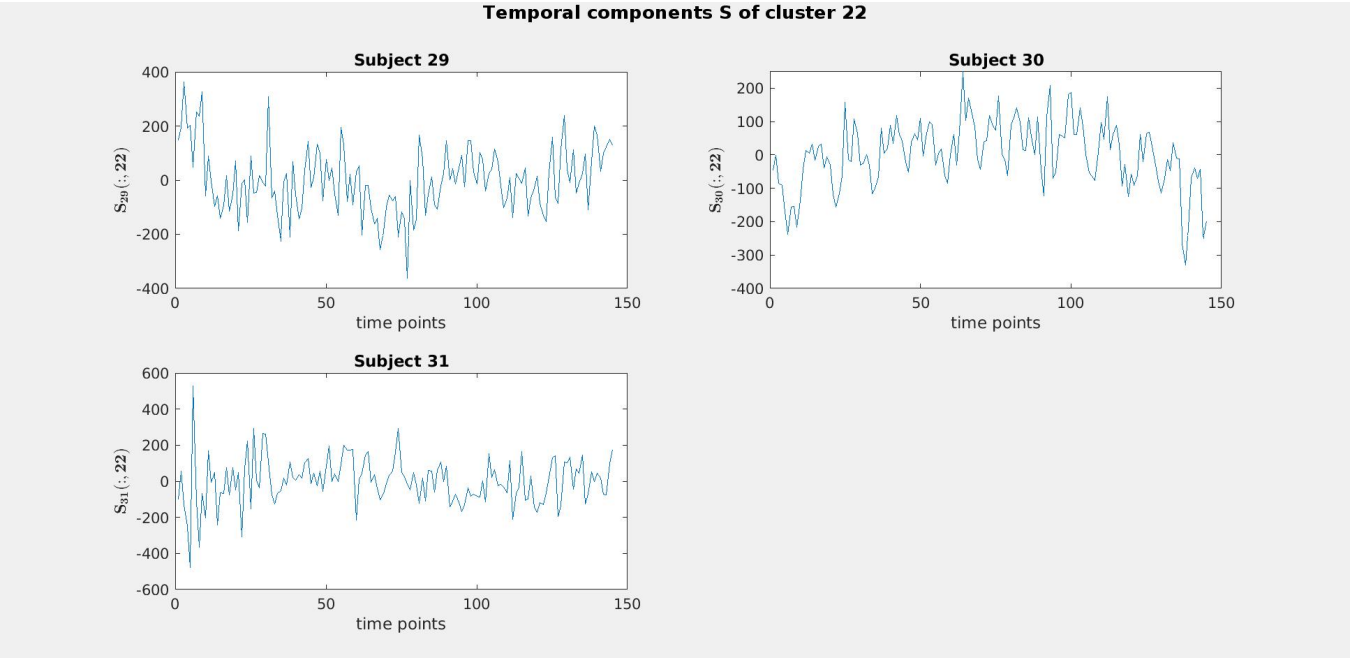


Figure B.413: Waveform of temporal components $\mathbf{S}_k(:, 22)$ for $k = 29, \ldots, 31$ and R=42.

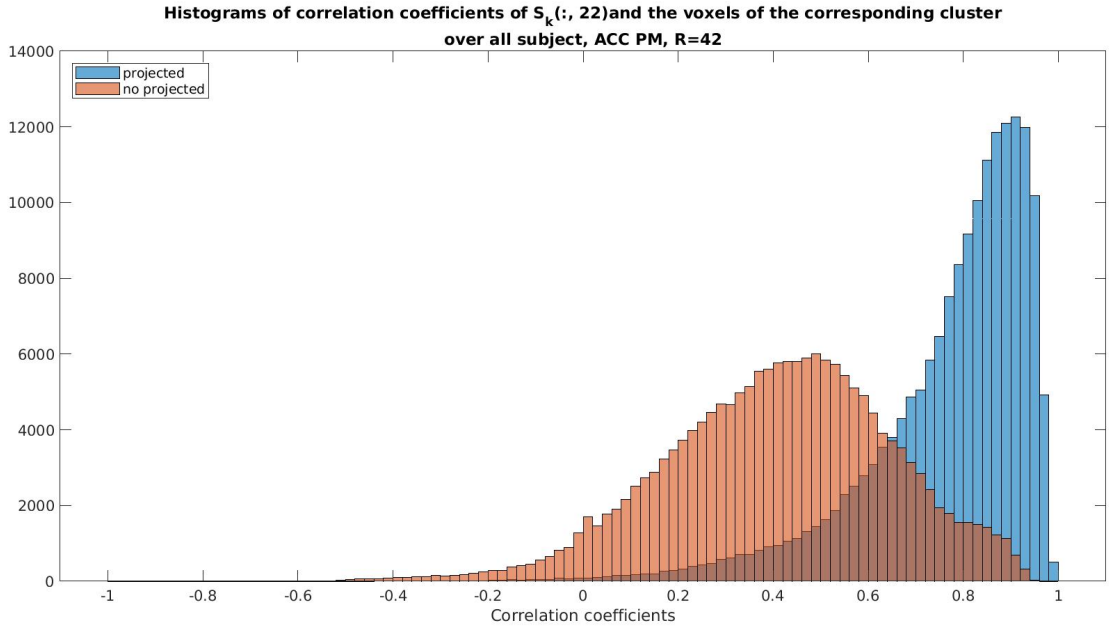


Figure B.414: For the column 22 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 22)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

13. **Cluster 1**

    Cluster 1 contains 6438 voxels, which correspond to 2.849328% of the whole brain (background, soft tissue, bone are excluded).

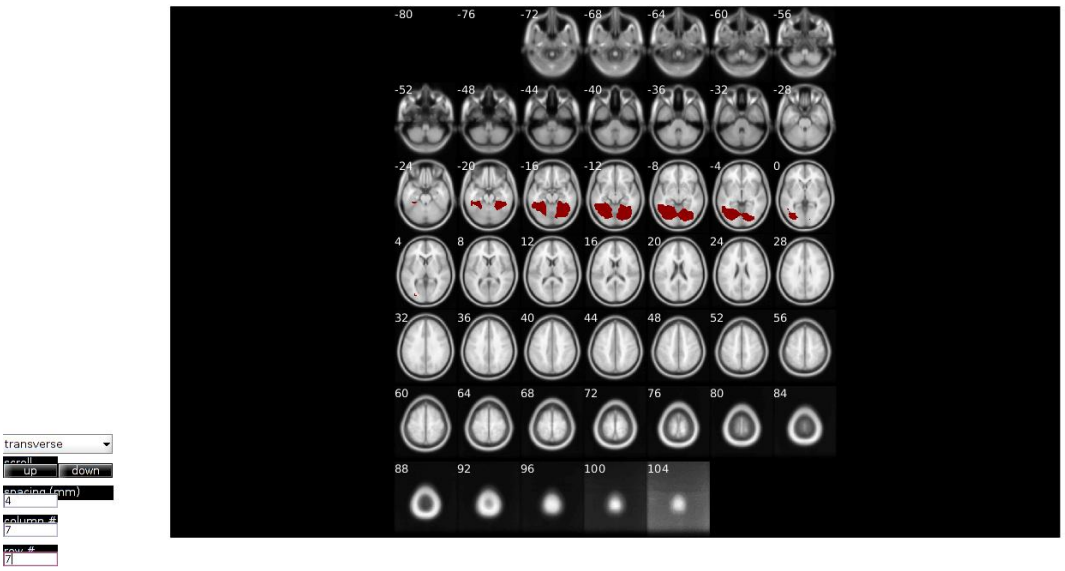


Figure B.415: Cluster 1. Red areas indicate the voxels which belong to cluster 1.

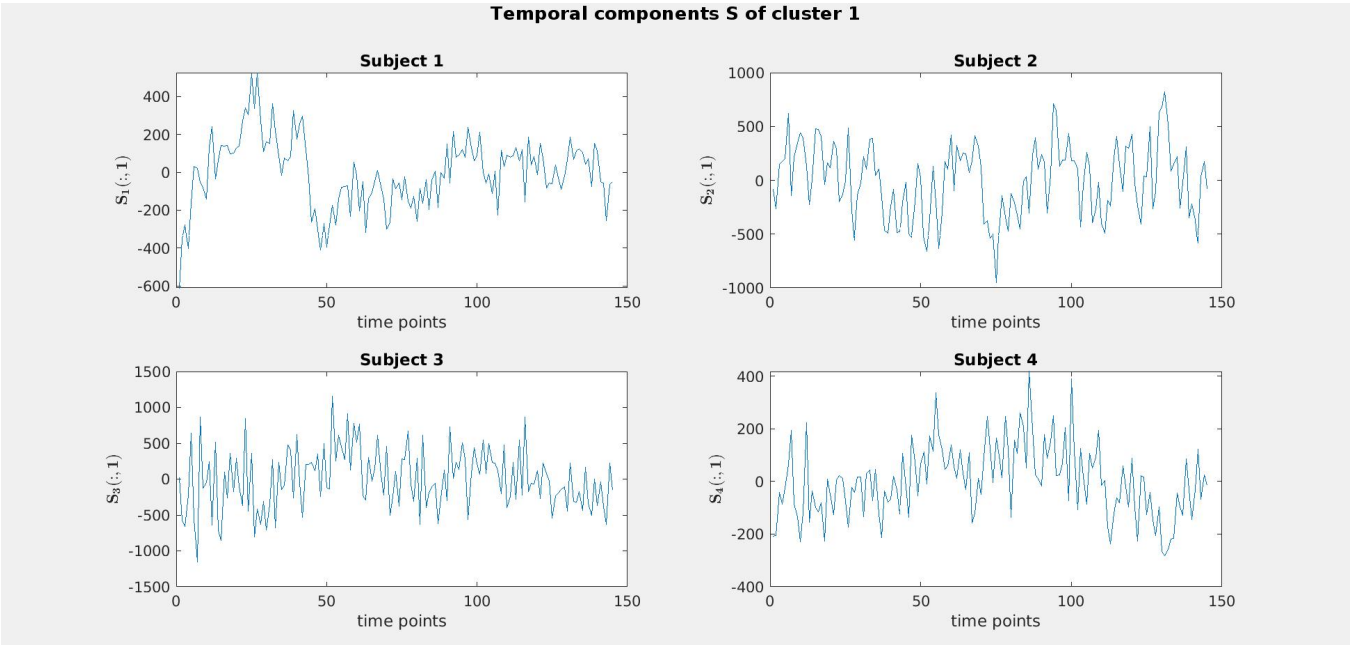


Figure B.416: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 1,\ldots,4$ and R=42.

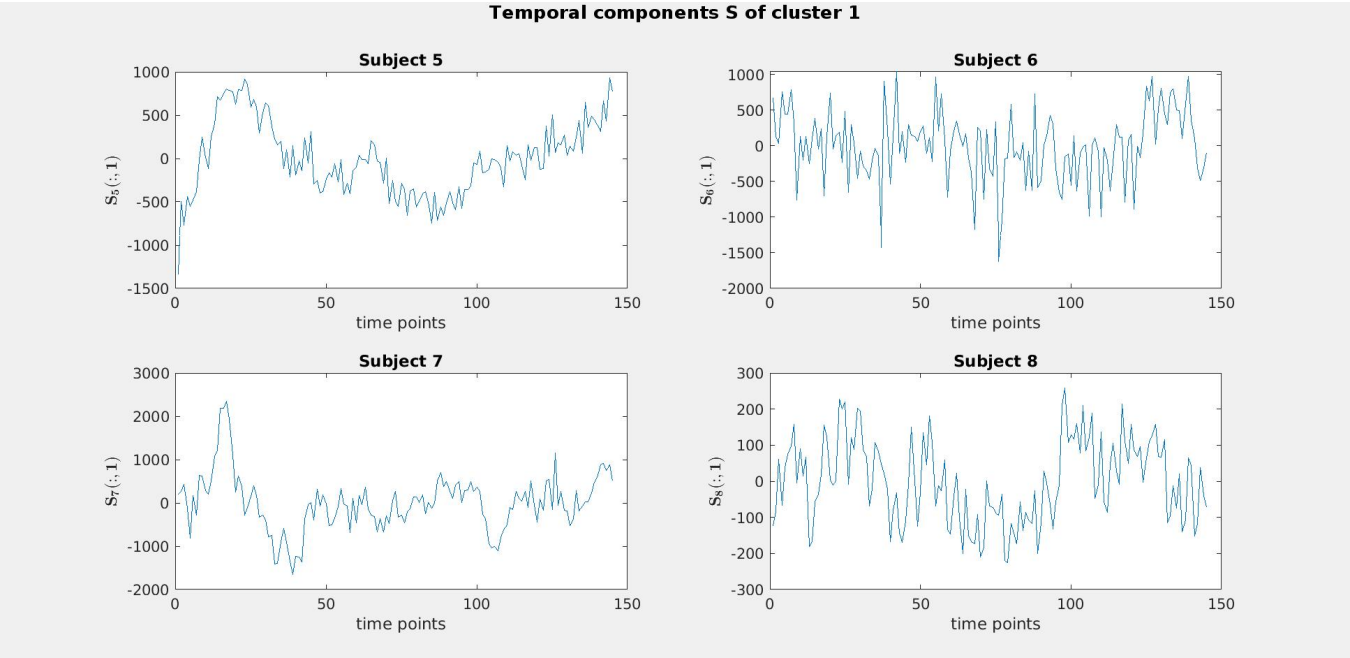


Figure B.417: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 5,\ldots,8$ and R=42.

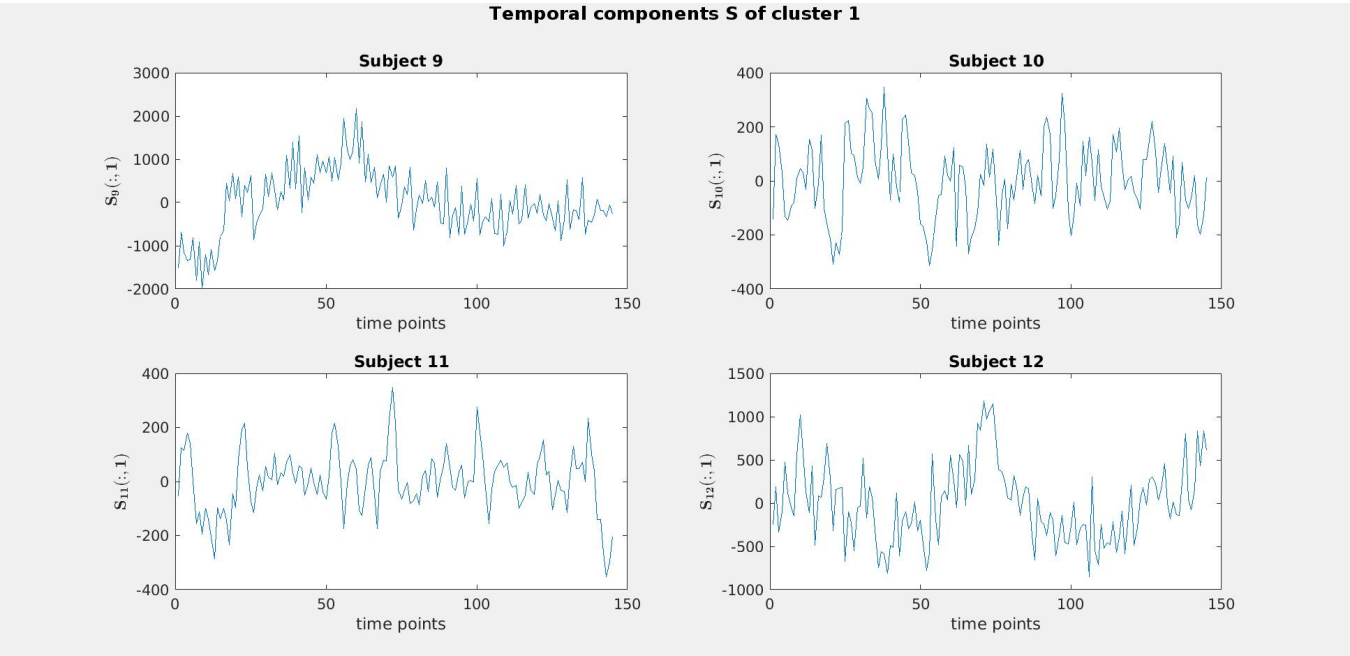


Figure B.418: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 9,\ldots,12$ and R=42.

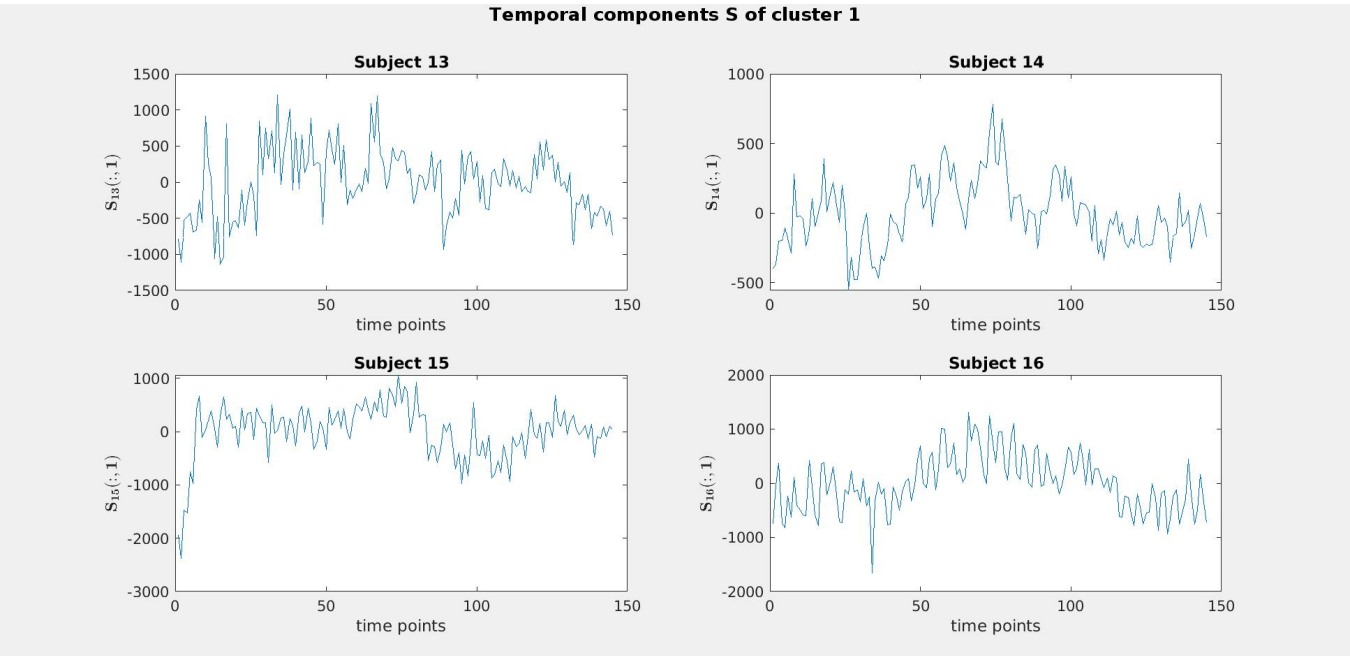


Figure B.419: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 13,\ldots,16$ and R=42.

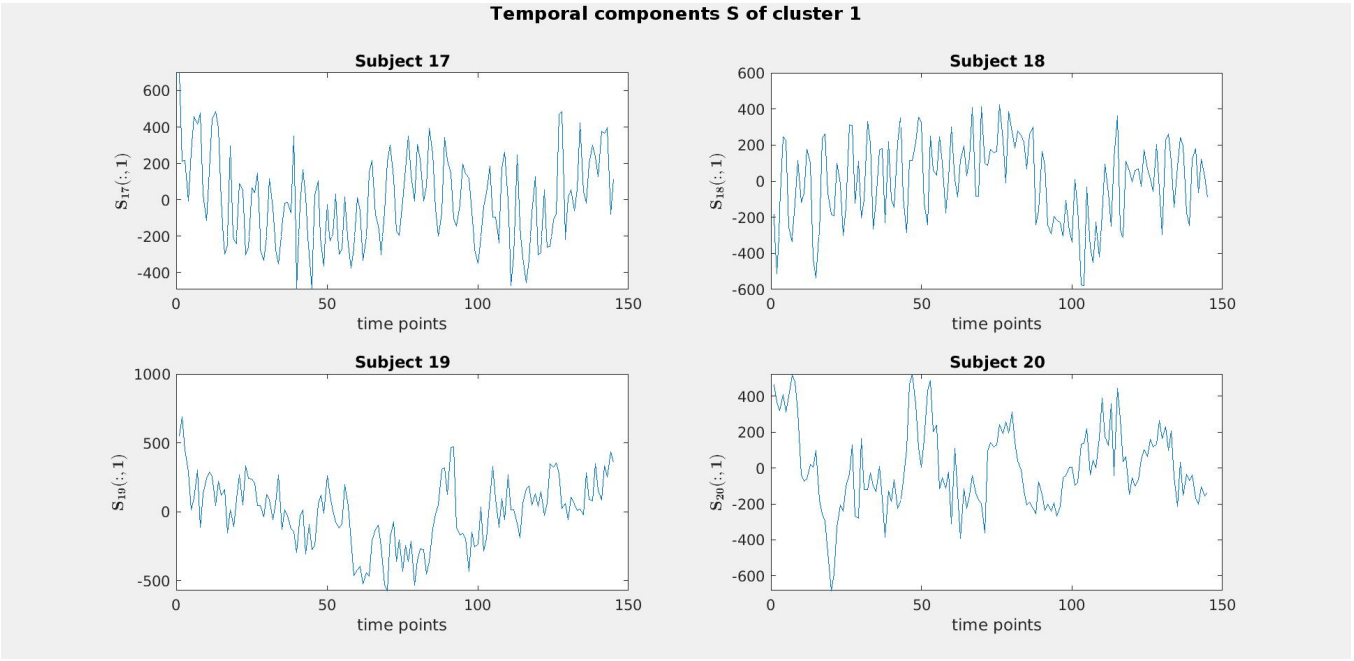


Figure B.420: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 17,\ldots,20$ and R=42.

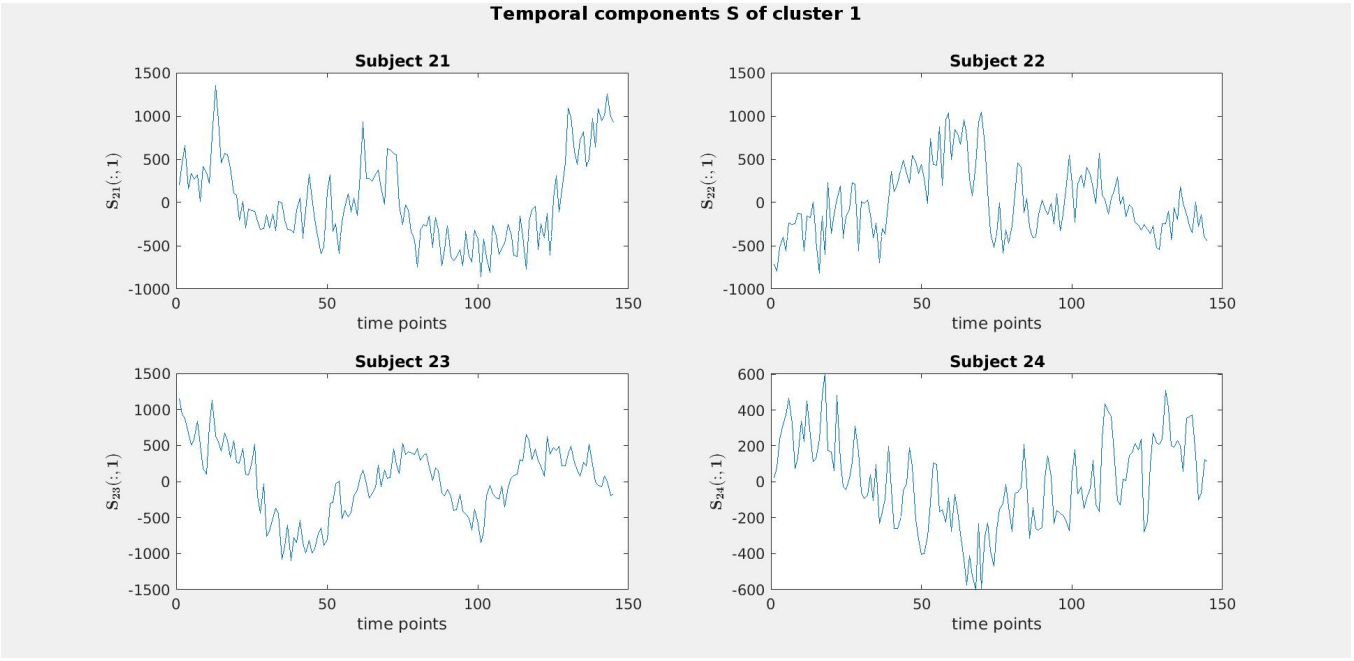


Figure B.421: Waveform of temporal components $\mathbf{S}_k(:,11)$ for $k = 21,\ldots,24$ and R=42.

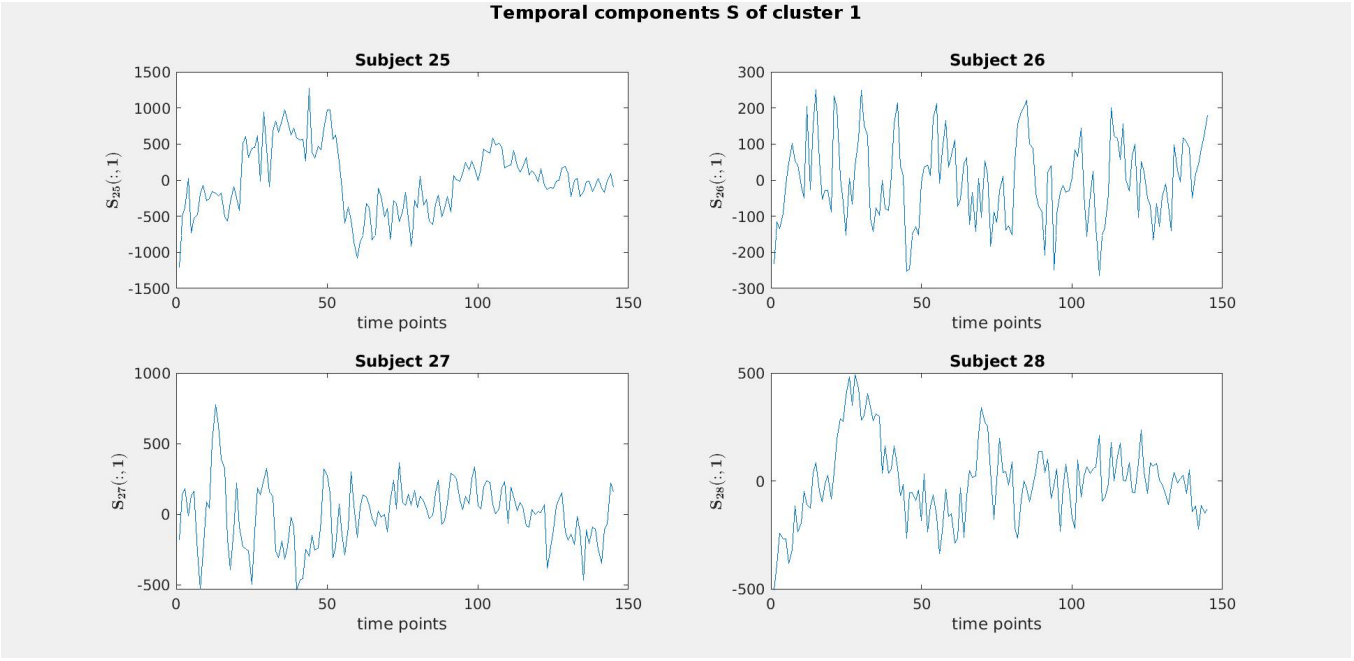


Figure B.422: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 25,\ldots,28$ and R=42.

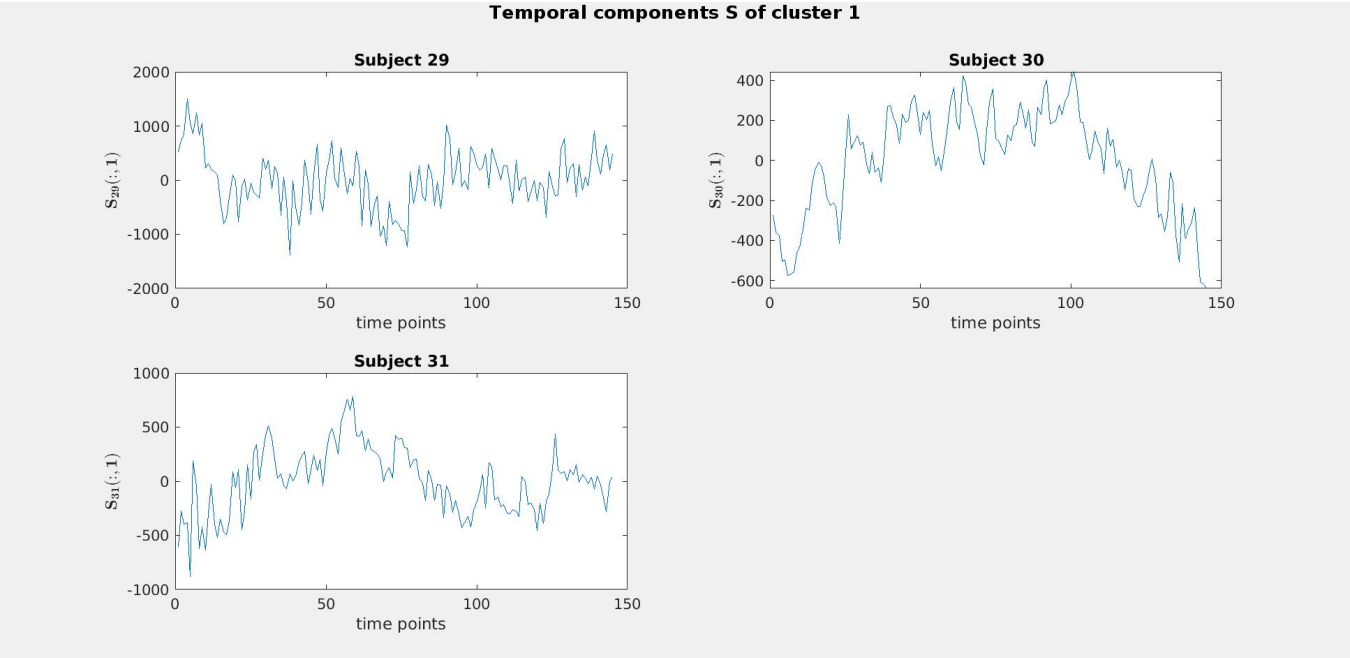


Figure B.423: Waveform of temporal components $\mathbf{S}_k(:,1)$ for $k = 29, \ldots, 31$ and R=42.

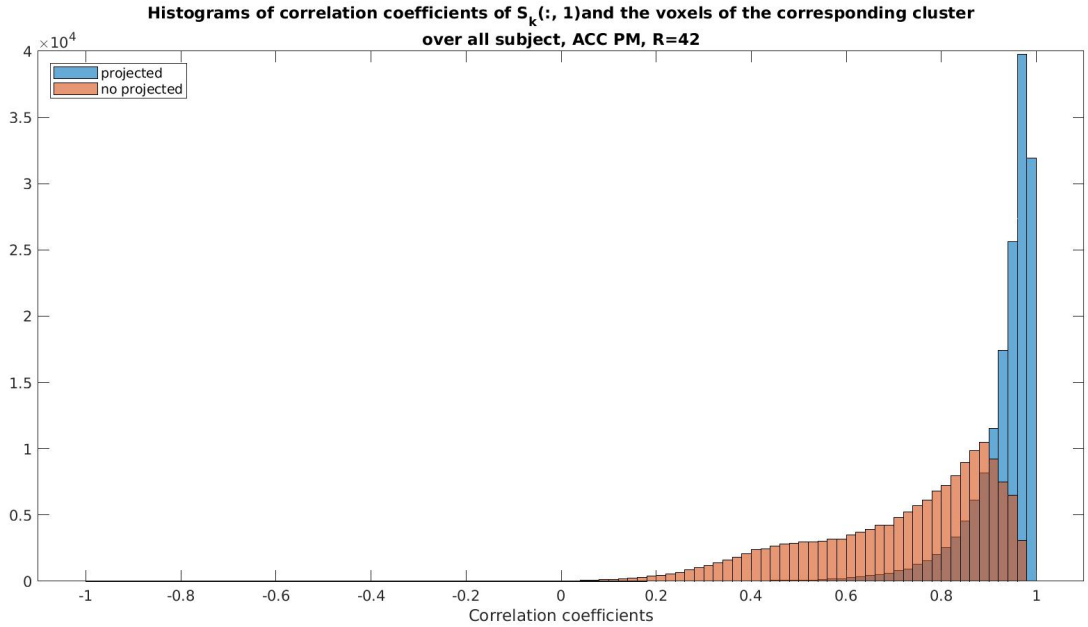


Figure B.424: For the column 1 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,1)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

14. **Cluster 17**

Cluster 17 contains 6387 voxels, which correspond to 2.826757% of the whole brain (background, soft tissue, bone are excluded).

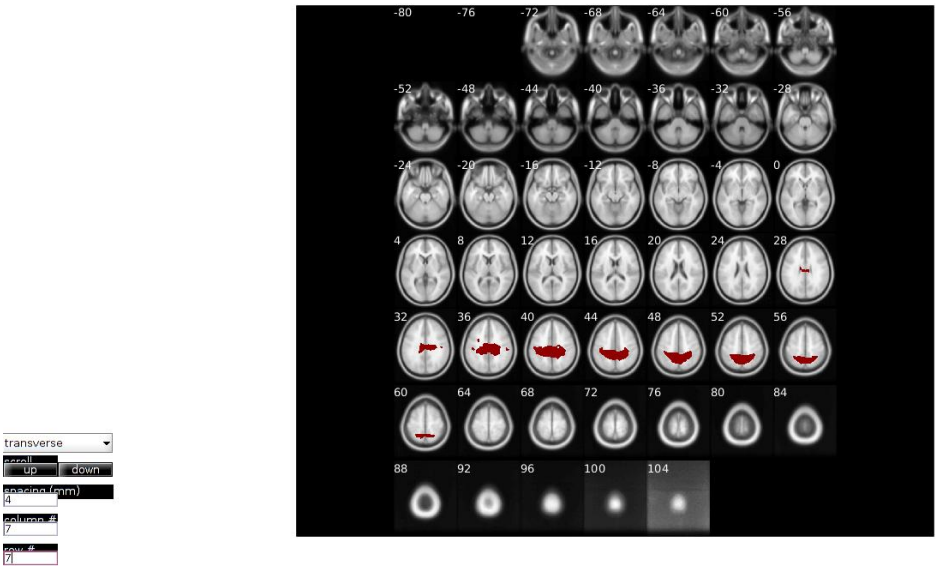


Figure B.425: Cluster 17. Red areas indicate the voxels which belong to cluster 17.

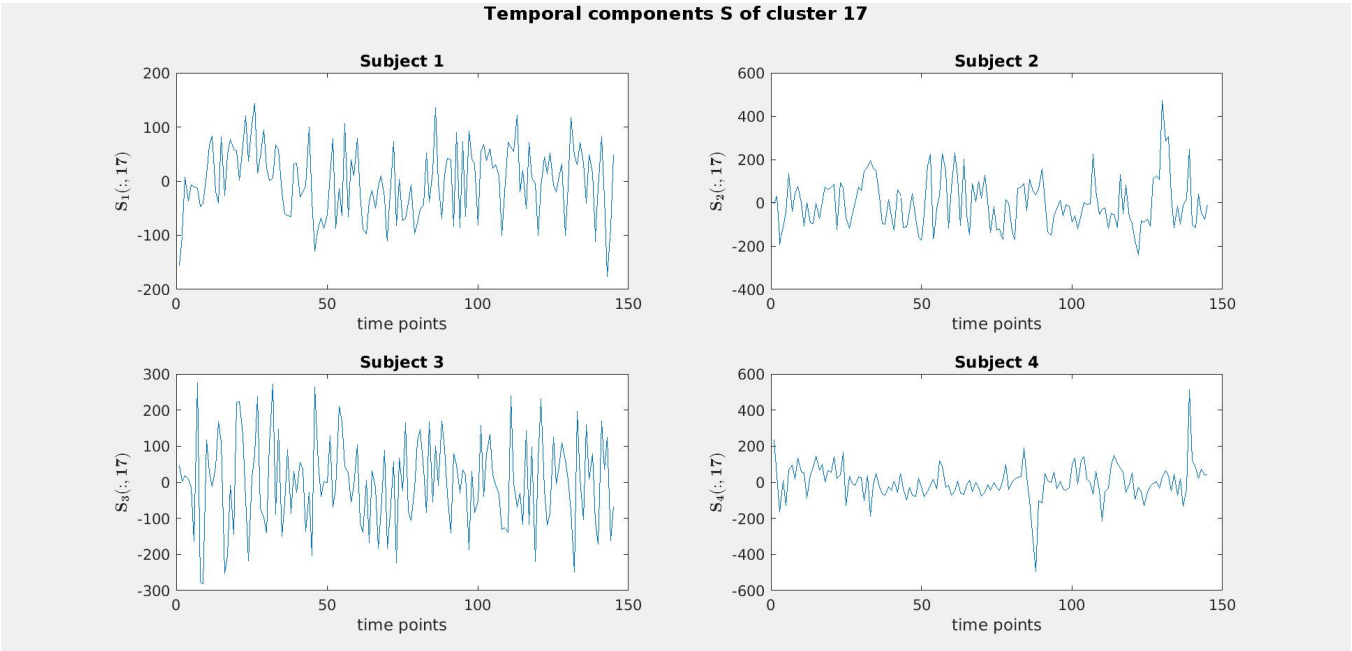


Figure B.426: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 1, \ldots, 4$ and R=42.

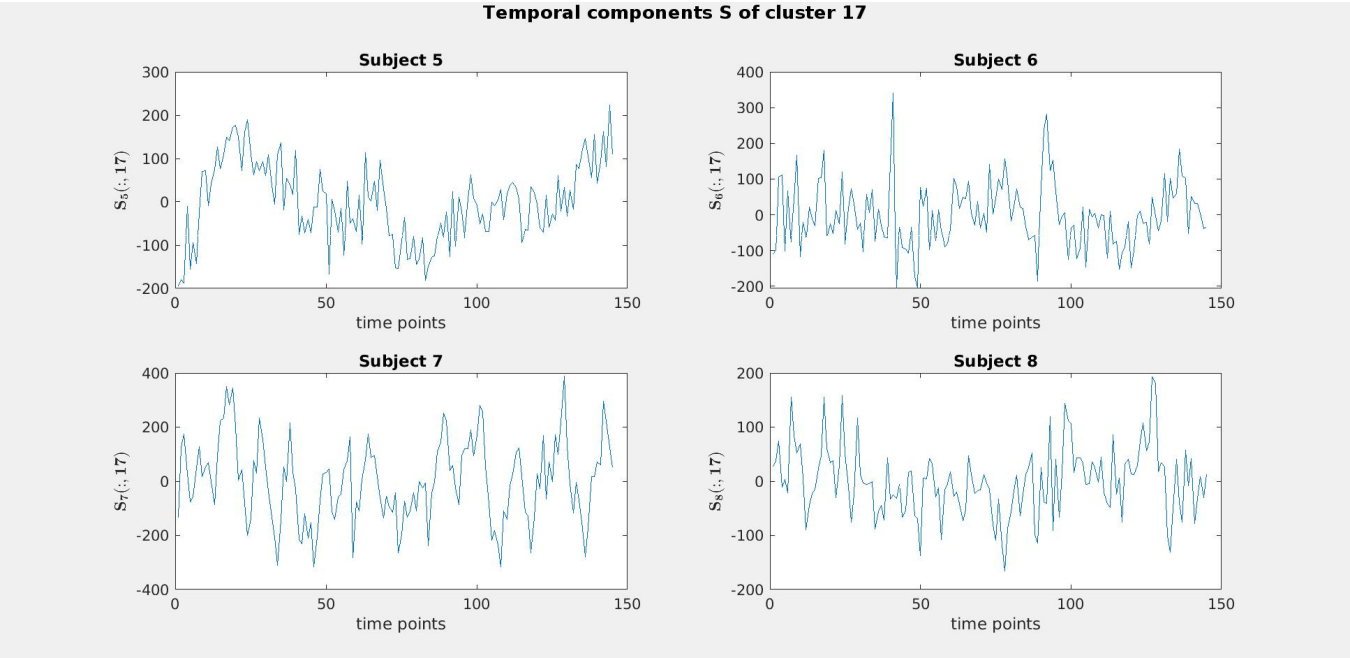


Figure B.427: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 5, \ldots, 8$ and R=42.

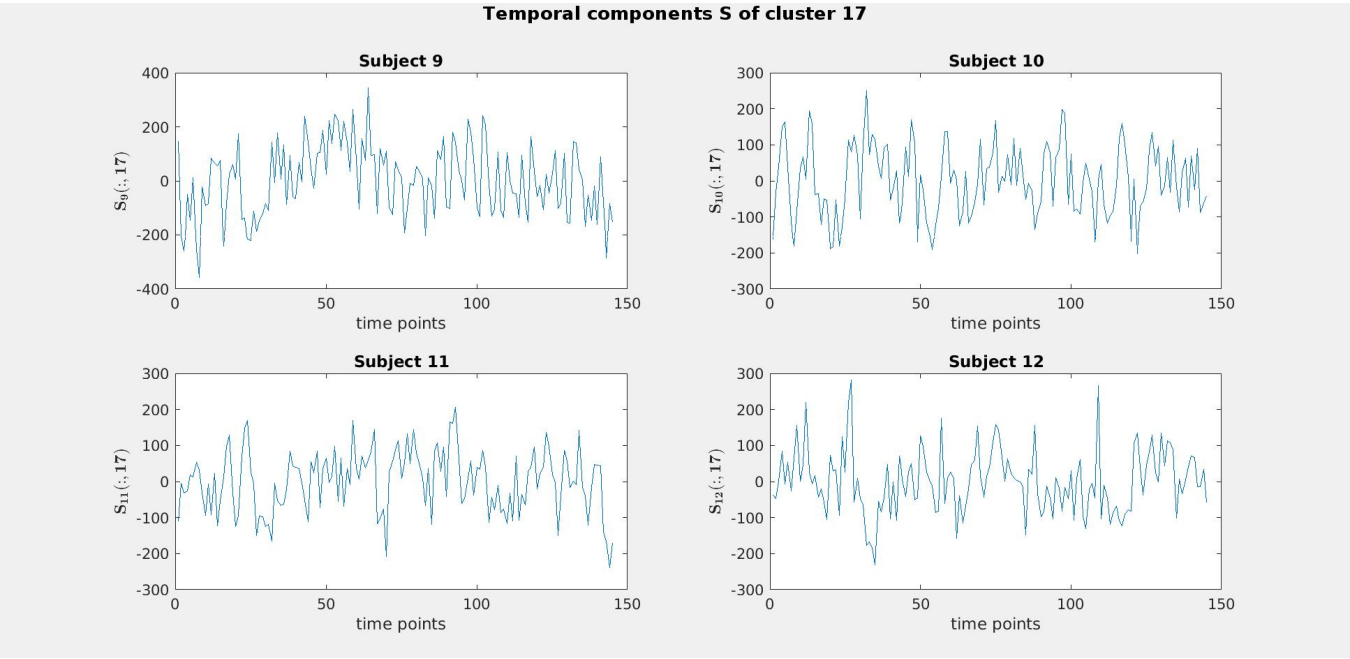


Figure B.428: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 9, \ldots, 12$ and R=42.

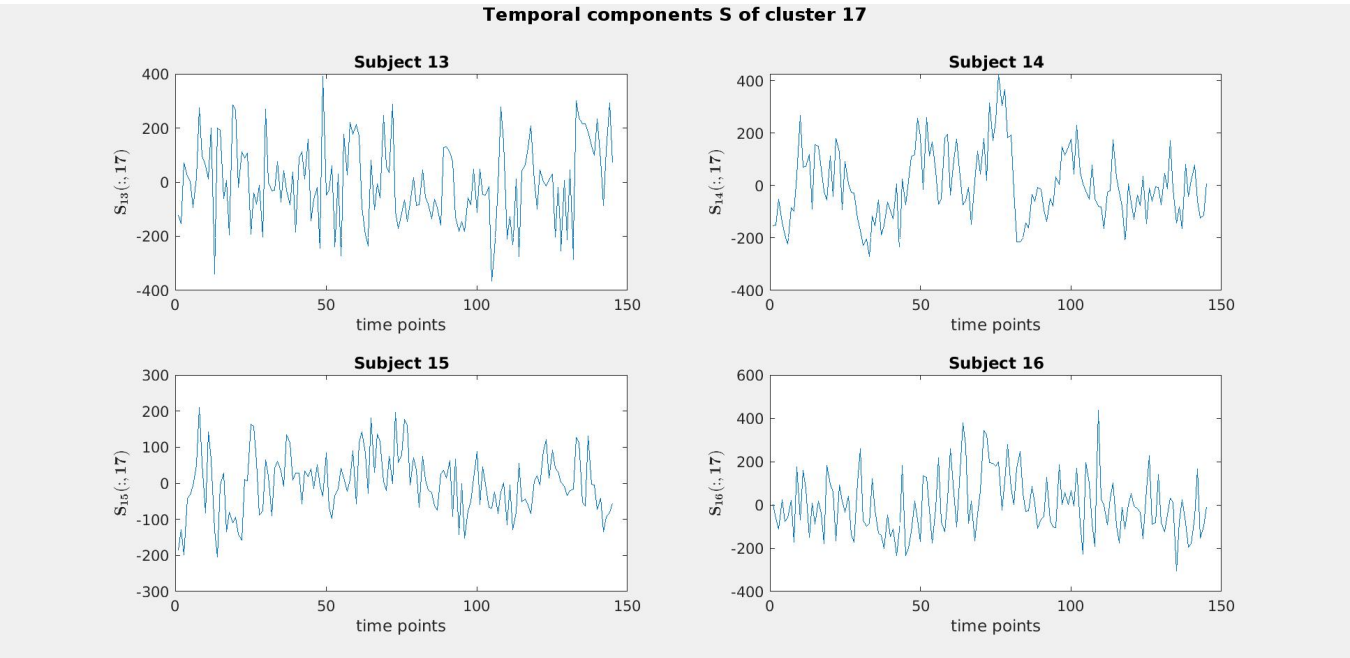


Figure B.429: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 13, \ldots, 16$ and R=42.

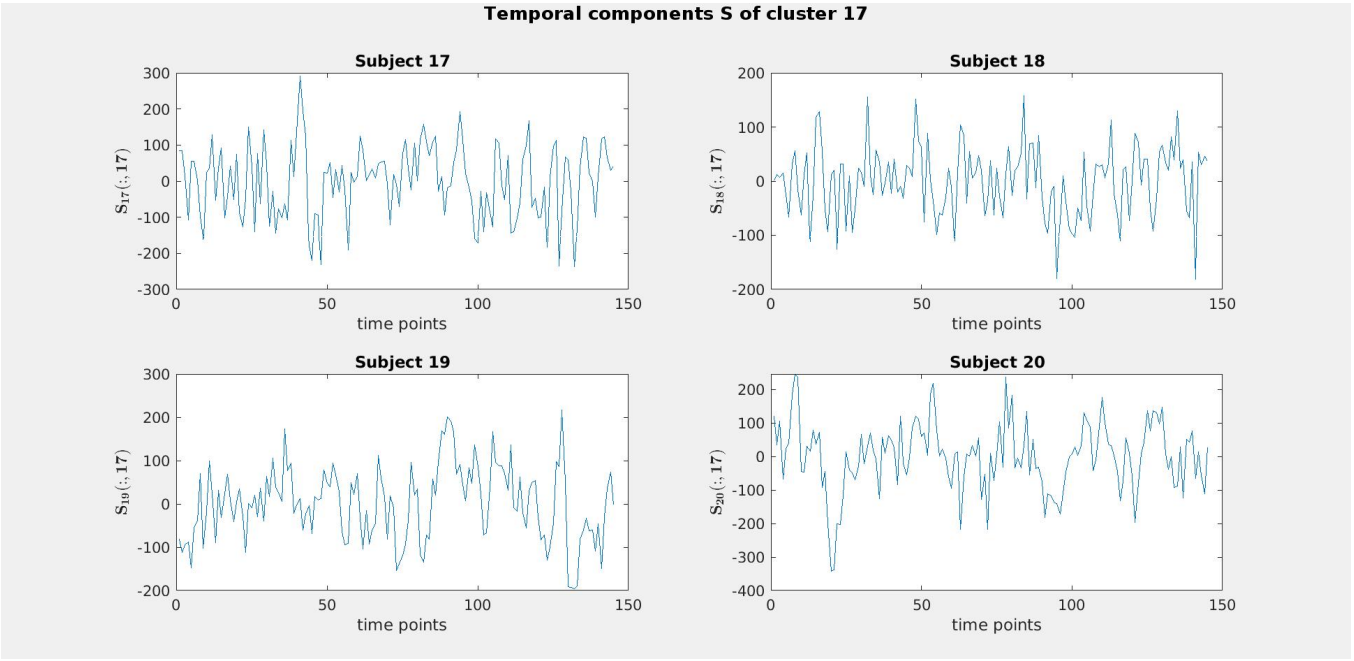


Figure B.430: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 17, \ldots, 20$ and R=42.

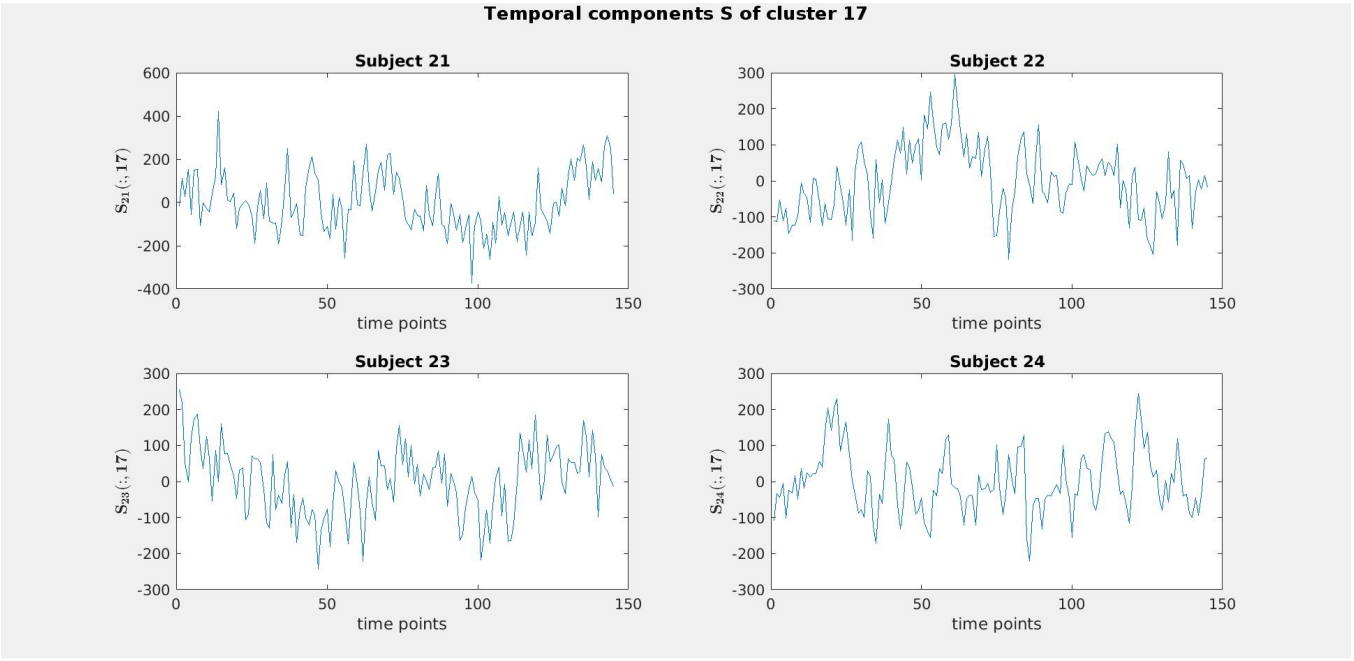


Figure B.431: Waveform of temporal components $\mathbf{S}_k(:, 117)$ for $k = 21, \ldots, 24$ and R=42.

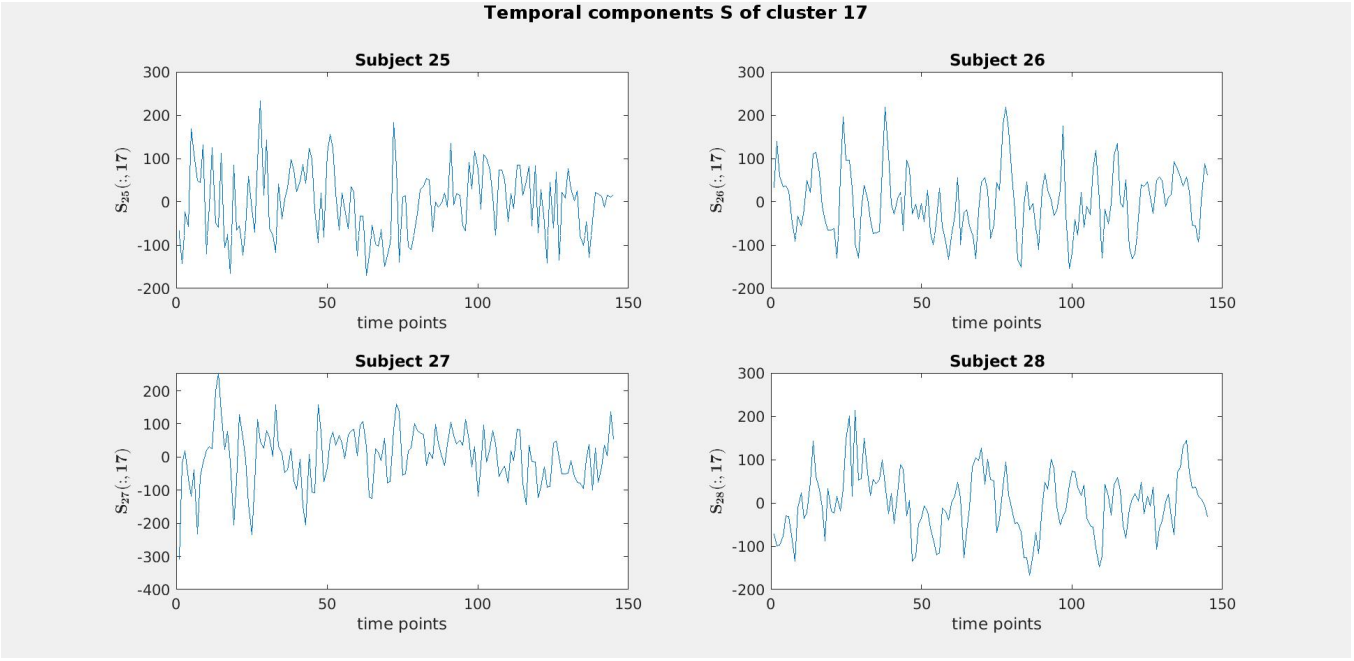


Figure B.432: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 25, \ldots, 28$ and R=42.

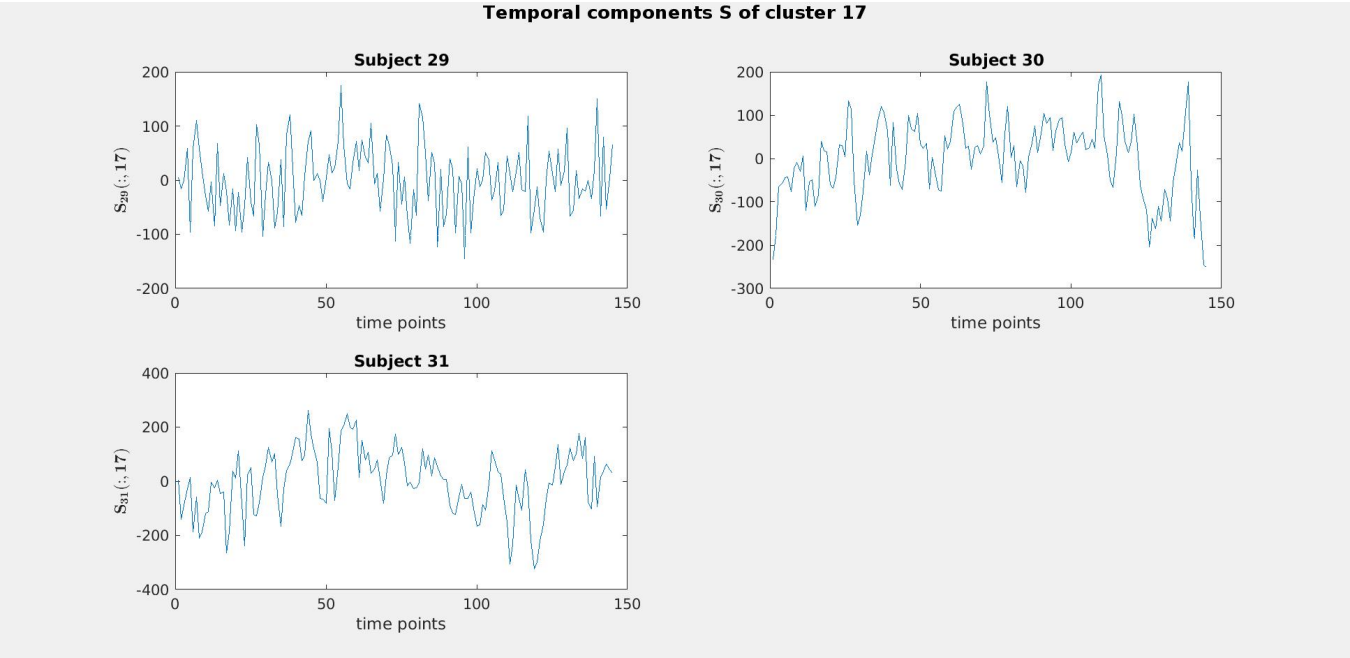


Figure B.433: Waveform of temporal components $\mathbf{S}_k(:, 17)$ for $k = 29, \ldots, 31$ and R=42.

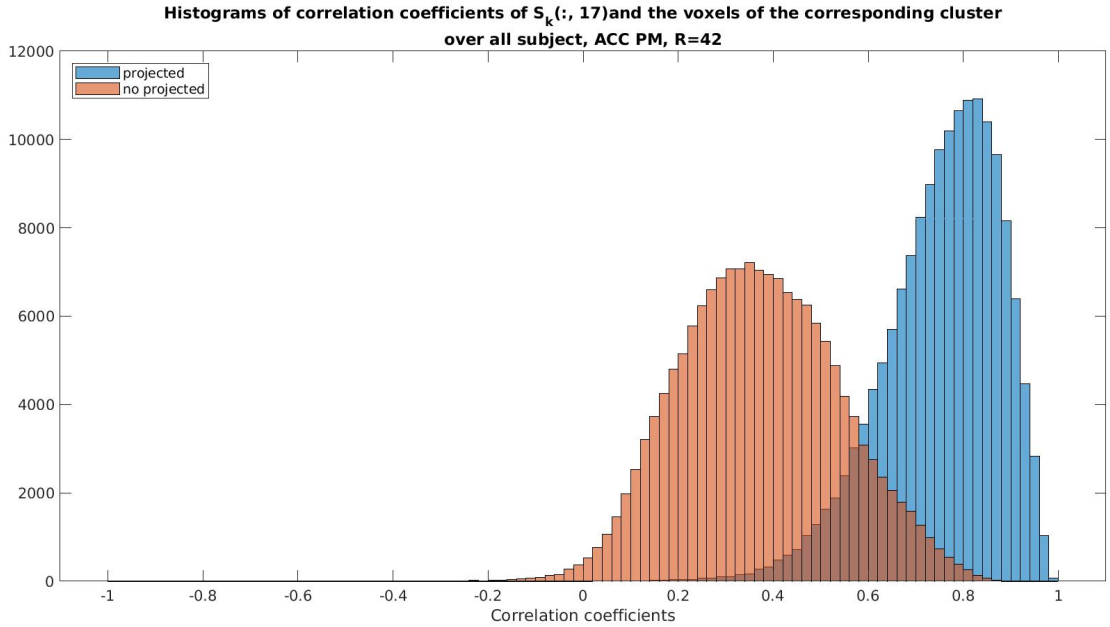


Figure B.434: For the column 17 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 17)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

15. **Cluster 16**

Cluster 16 contains 6268 voxels, which correspond to 2.774090% of the whole brain (background, soft tissue, bone are excluded).

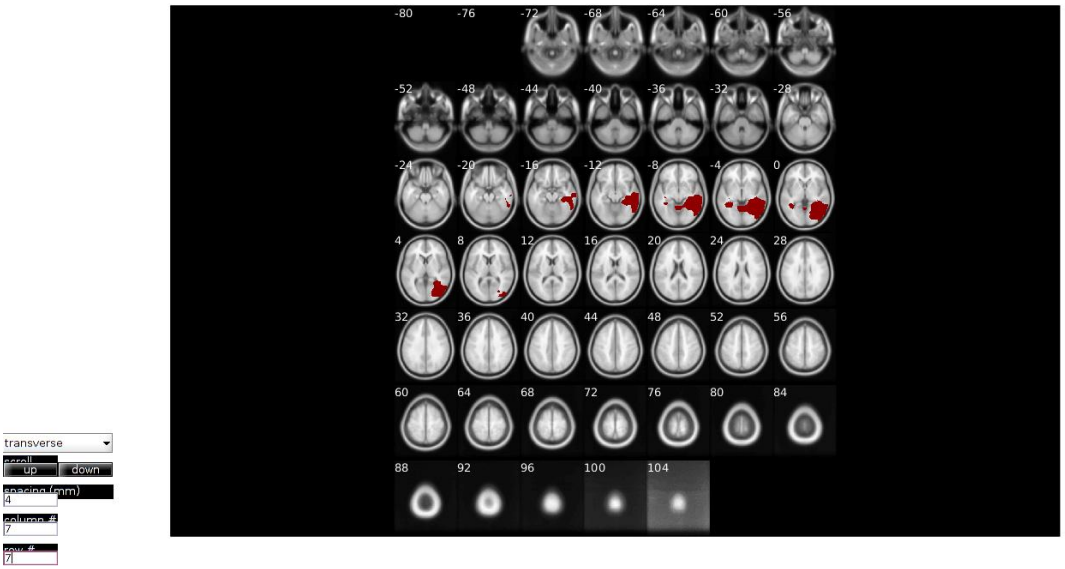


Figure B.435: Cluster 16. Red areas indicate the voxels which belong to cluster 16.

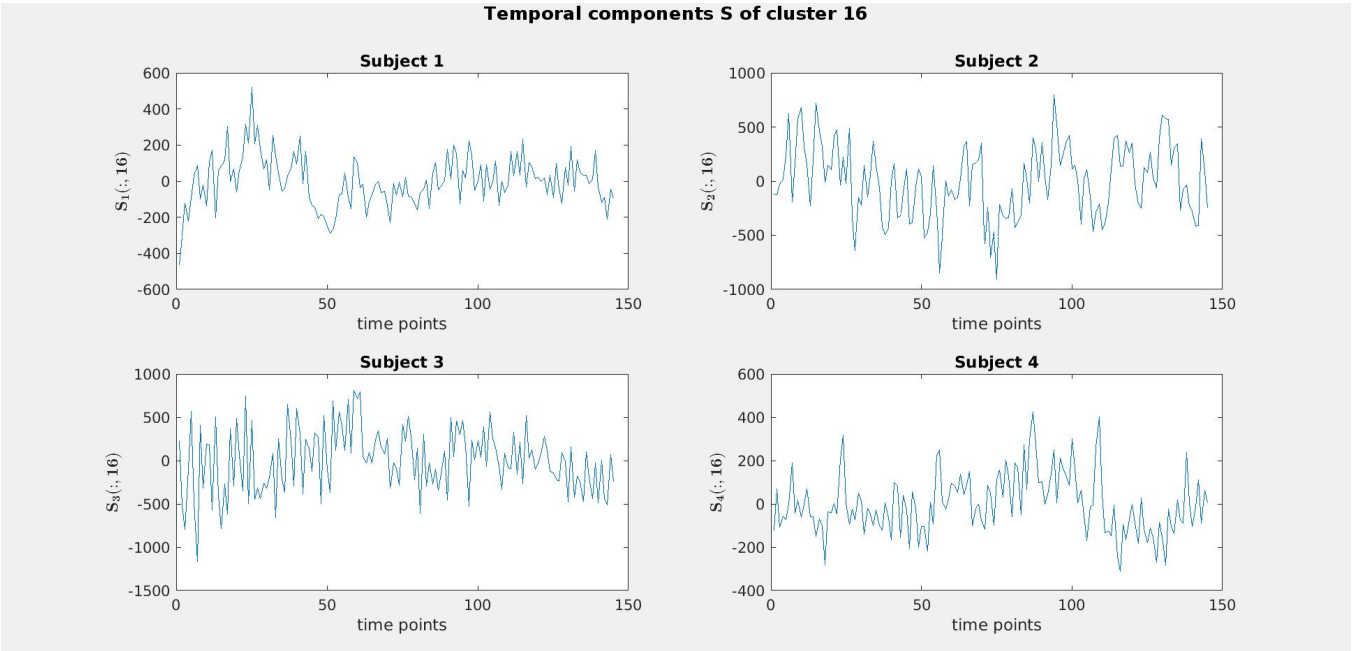


Figure B.436: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 1, \ldots, 4$ and R=42.

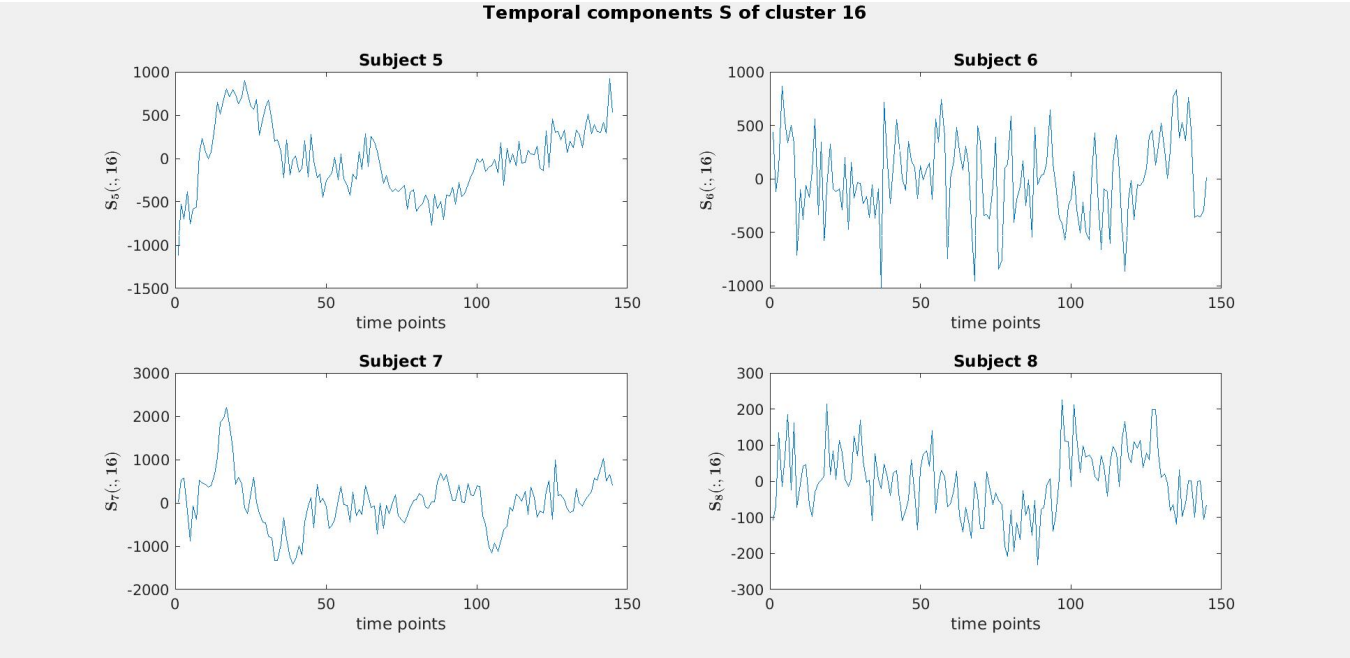


Figure B.437: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 5, \ldots, 8$ and R=42.

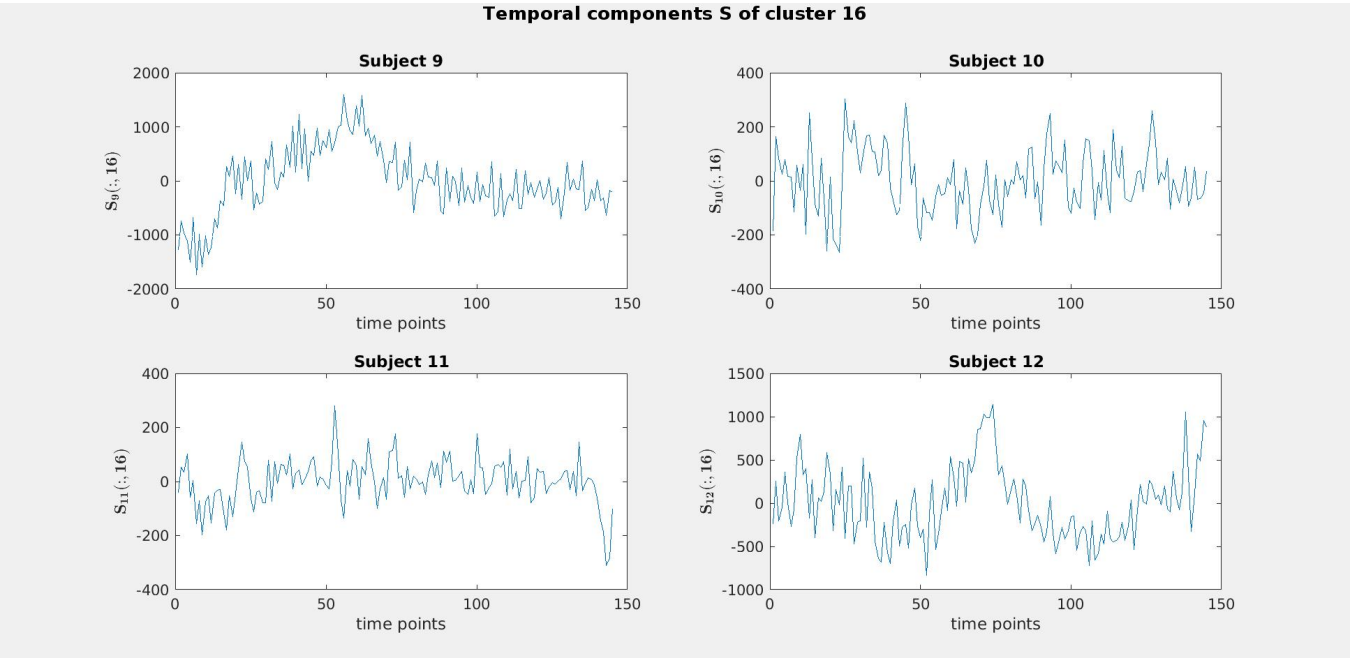


Figure B.438: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 9, \ldots, 12$ and R=42.

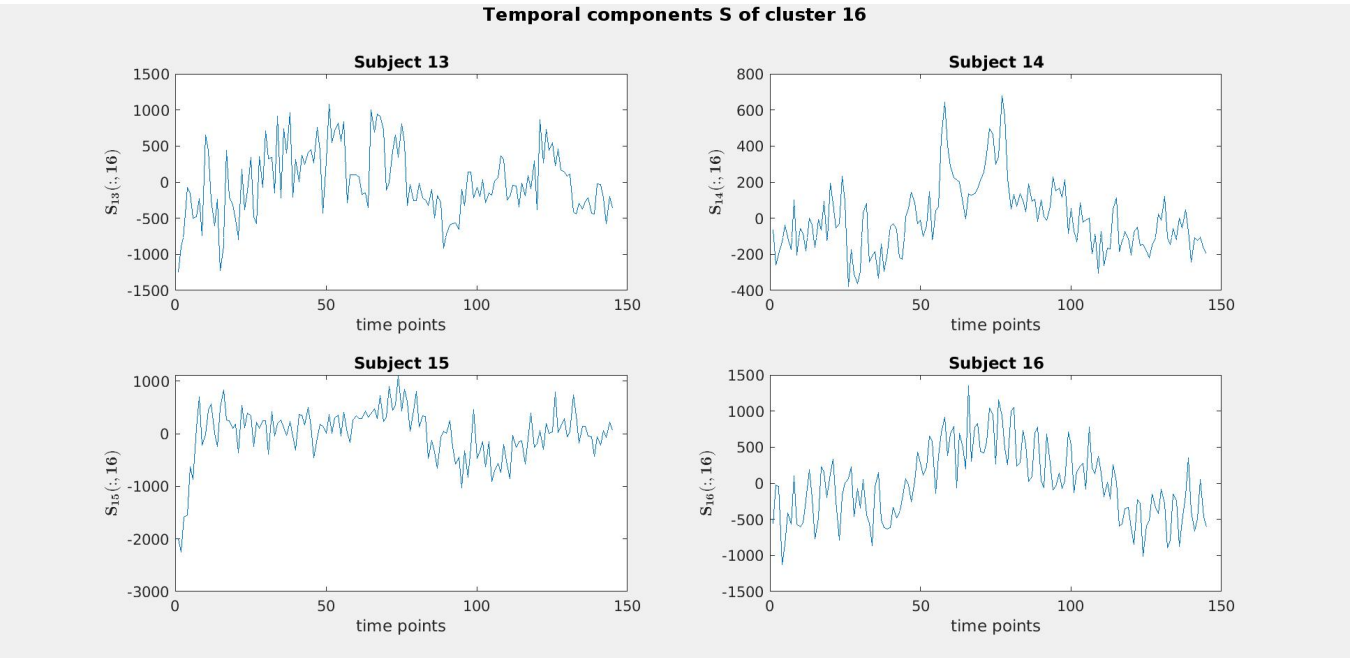


Figure B.439: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 13, \ldots, 16$ and R=42.

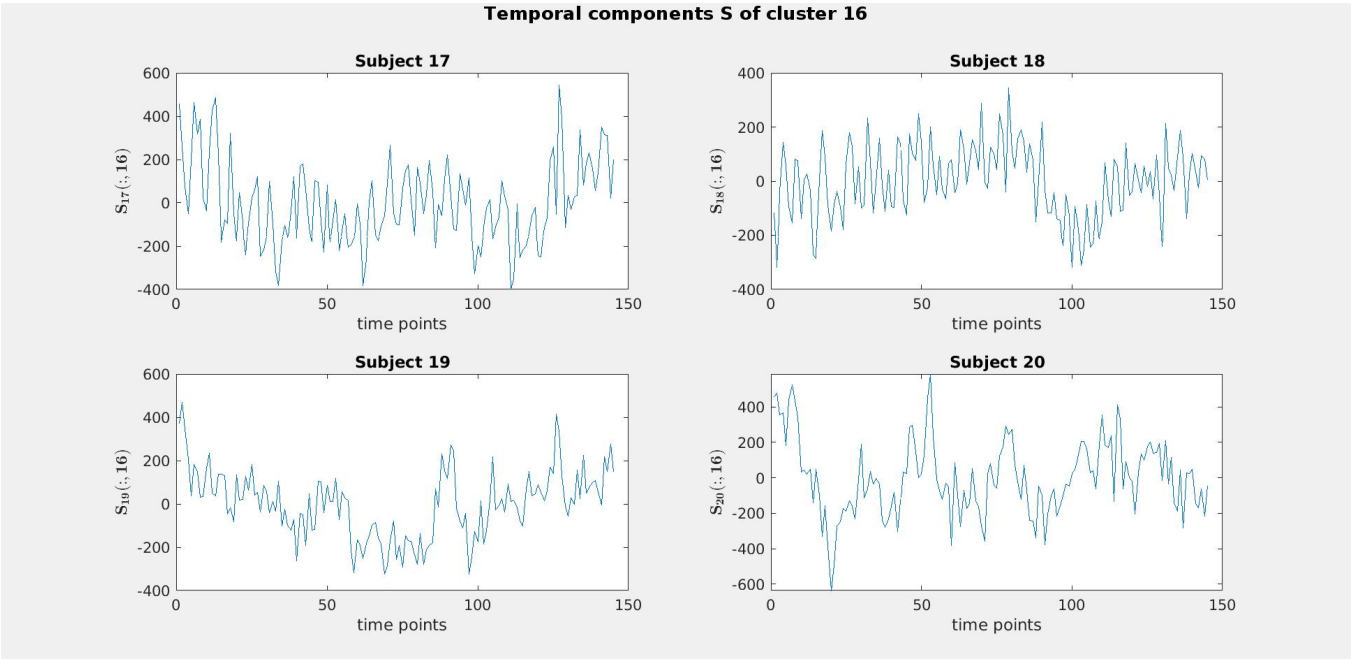


Figure B.440: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 17, \ldots, 20$ and R=42.

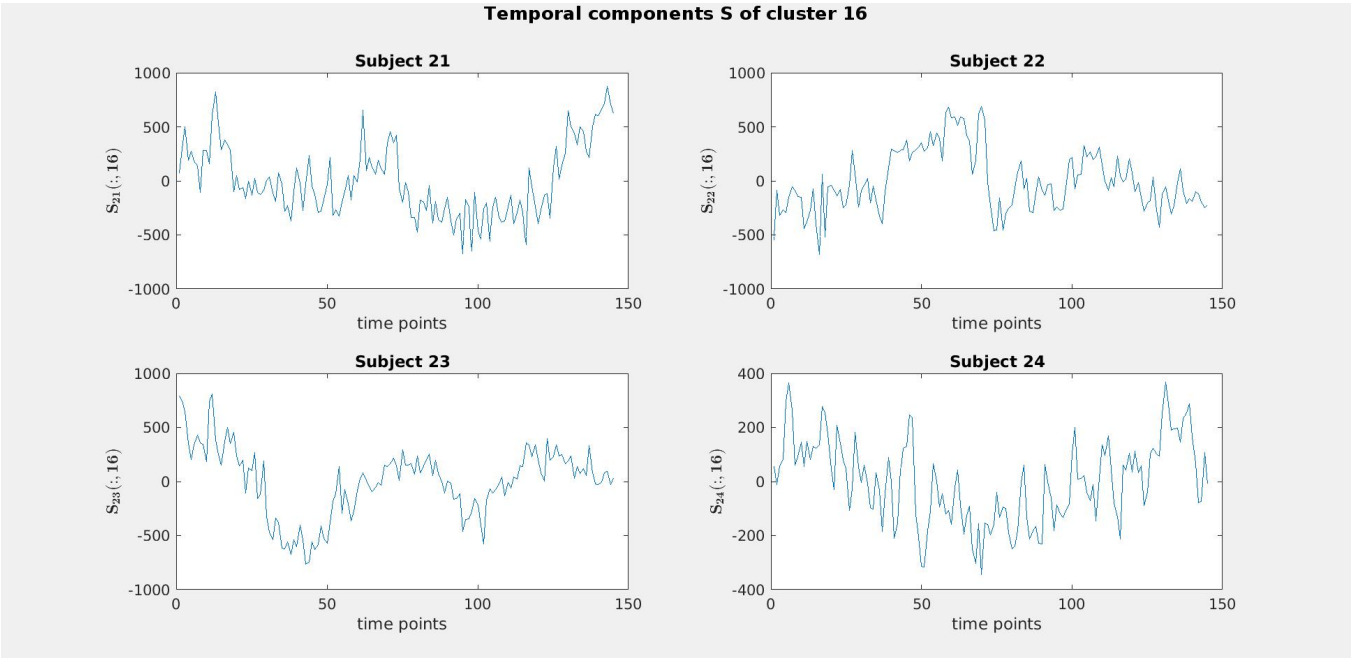


Figure B.441: Waveform of temporal components $\mathbf{S}_k(:, 116)$ for $k = 21, \ldots, 24$ and R=42.

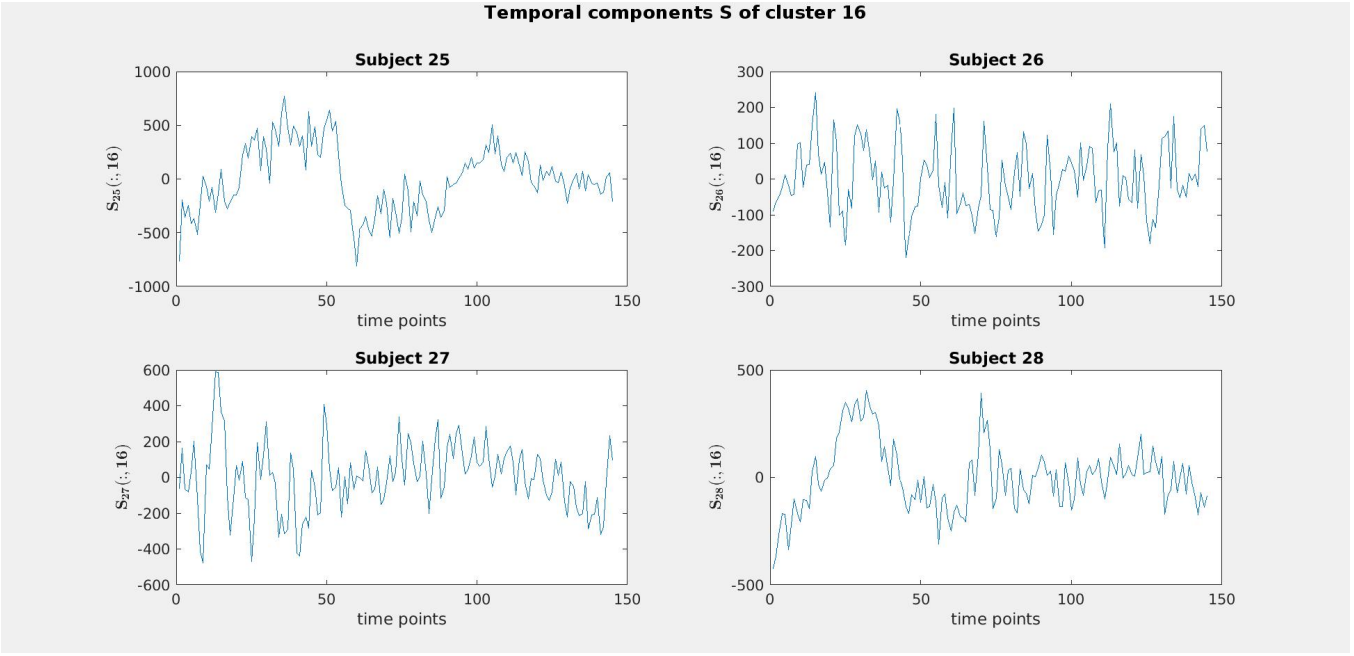


Figure B.442: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 25, \ldots, 28$ and R=42.

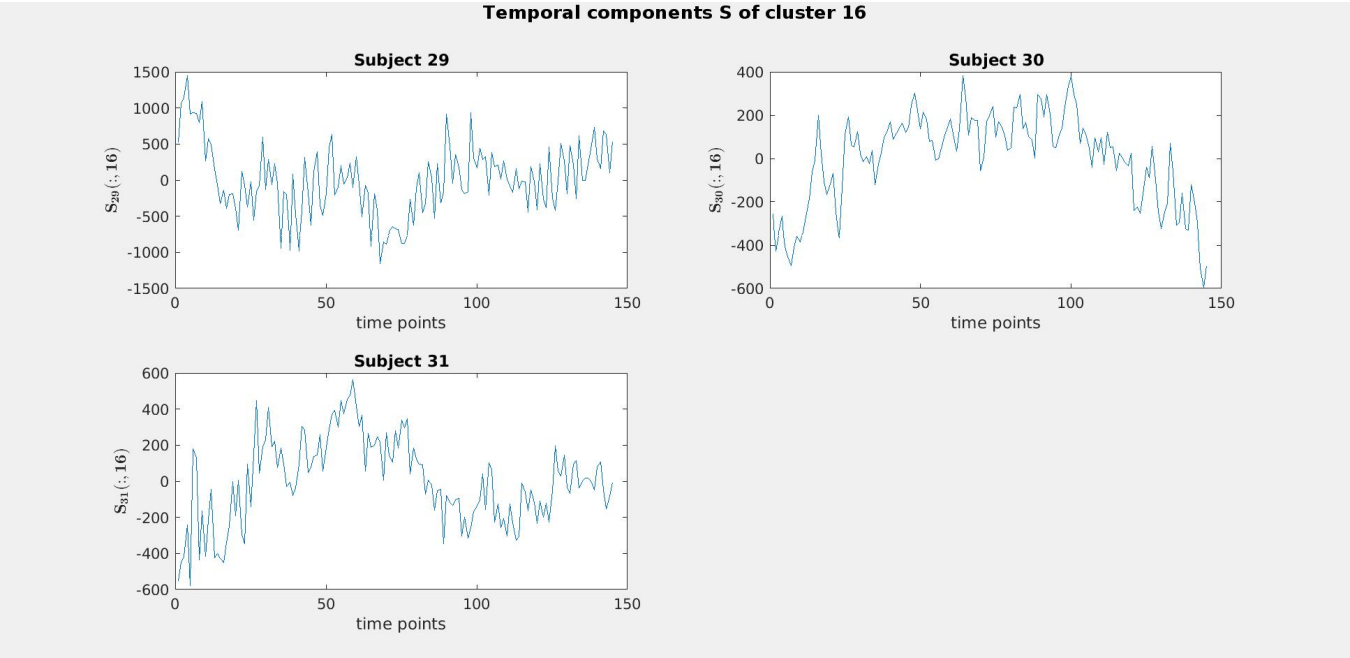


Figure B.443: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 29, \ldots, 31$ and R=42.

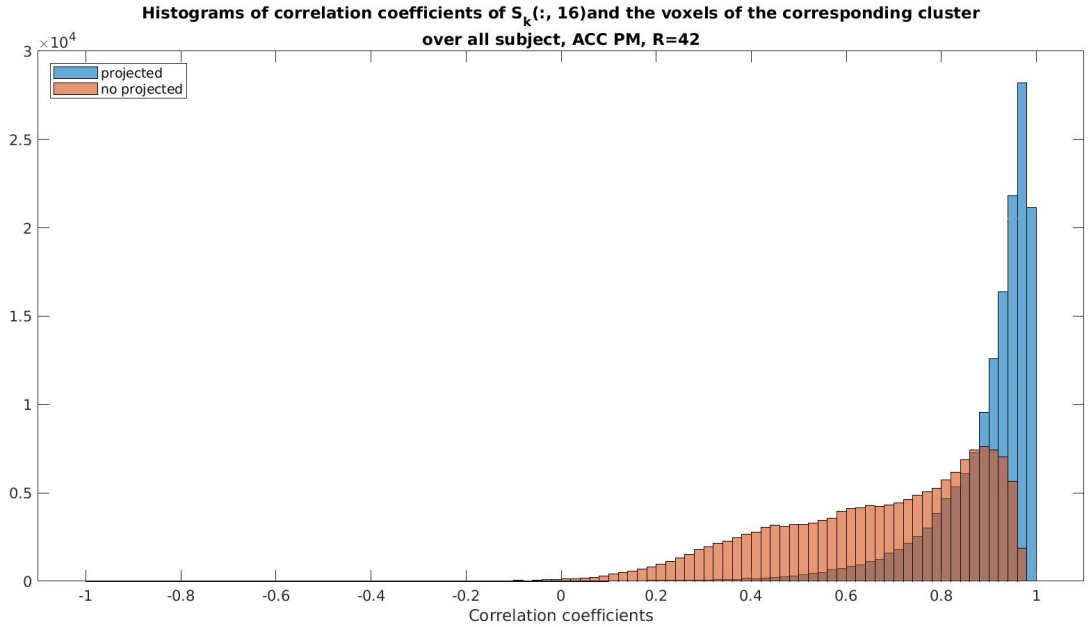


Figure B.444: For the column 16 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 16)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

16. **Cluster 37**

    Cluster 37 contains 6266 voxels, which correspond to 2.773204% of the whole brain (background, soft tissue, bone are excluded).

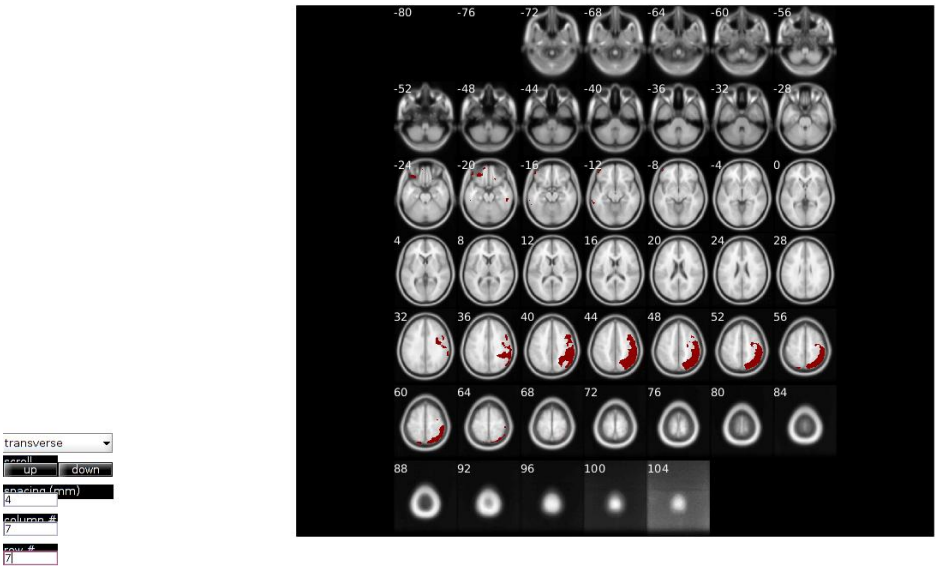


Figure B.445: Cluster 37. Red areas indicate the voxels which belong to cluster 37.

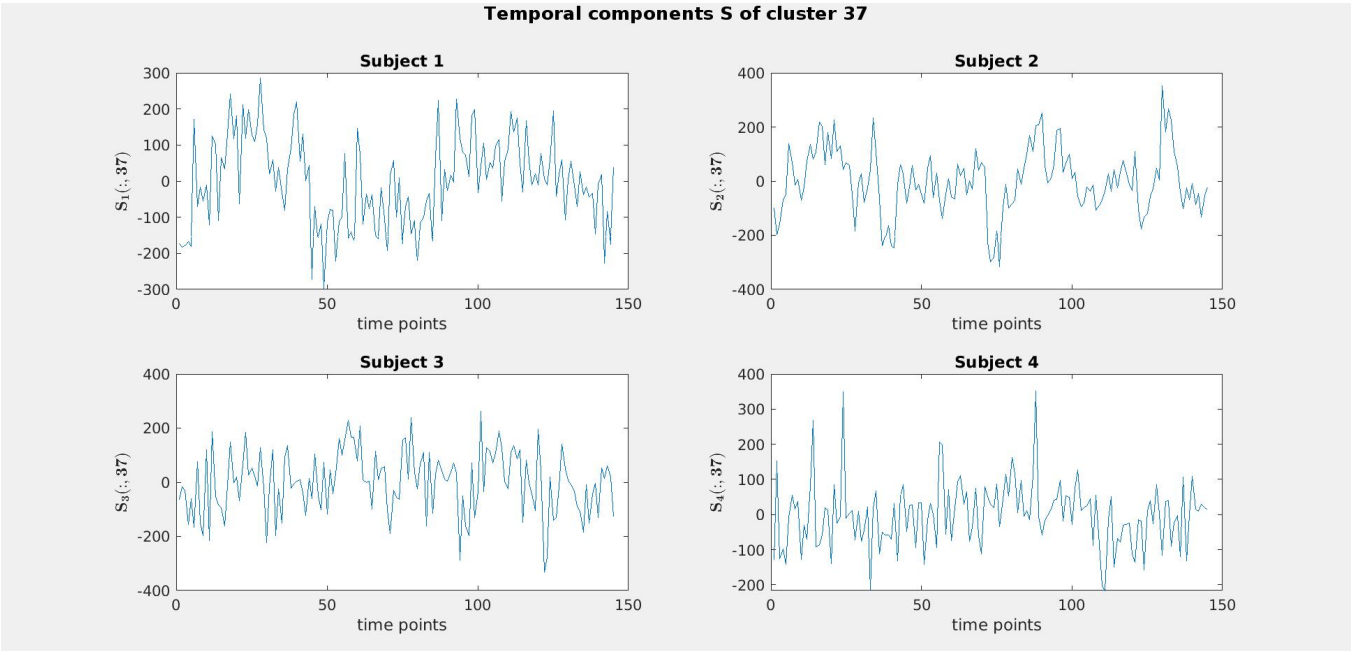


Figure B.446: Waveform of temporal components $\mathbf{S}_k(:, 37)$ for $k = 1, \ldots, 4$ and R=42.

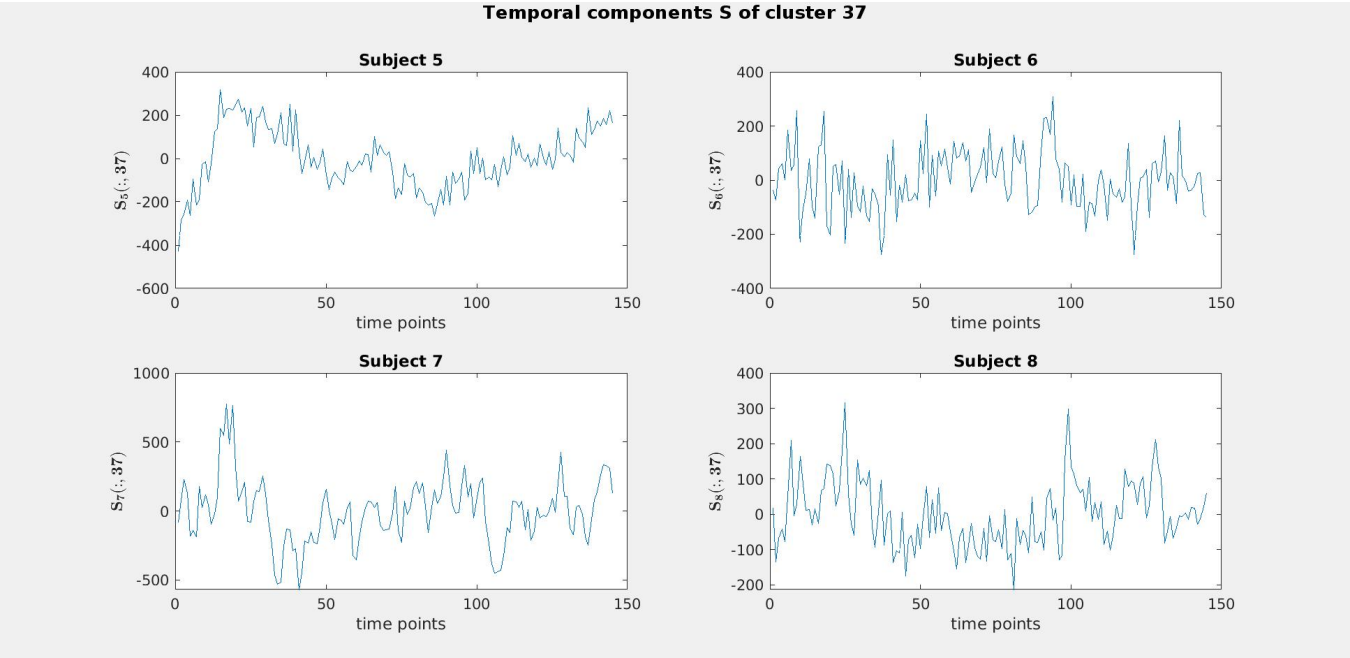


Figure B.447: Waveform of temporal components $\mathbf{S}_k(:, 37)$ for $k = 5, \ldots, 8$ and R=42.

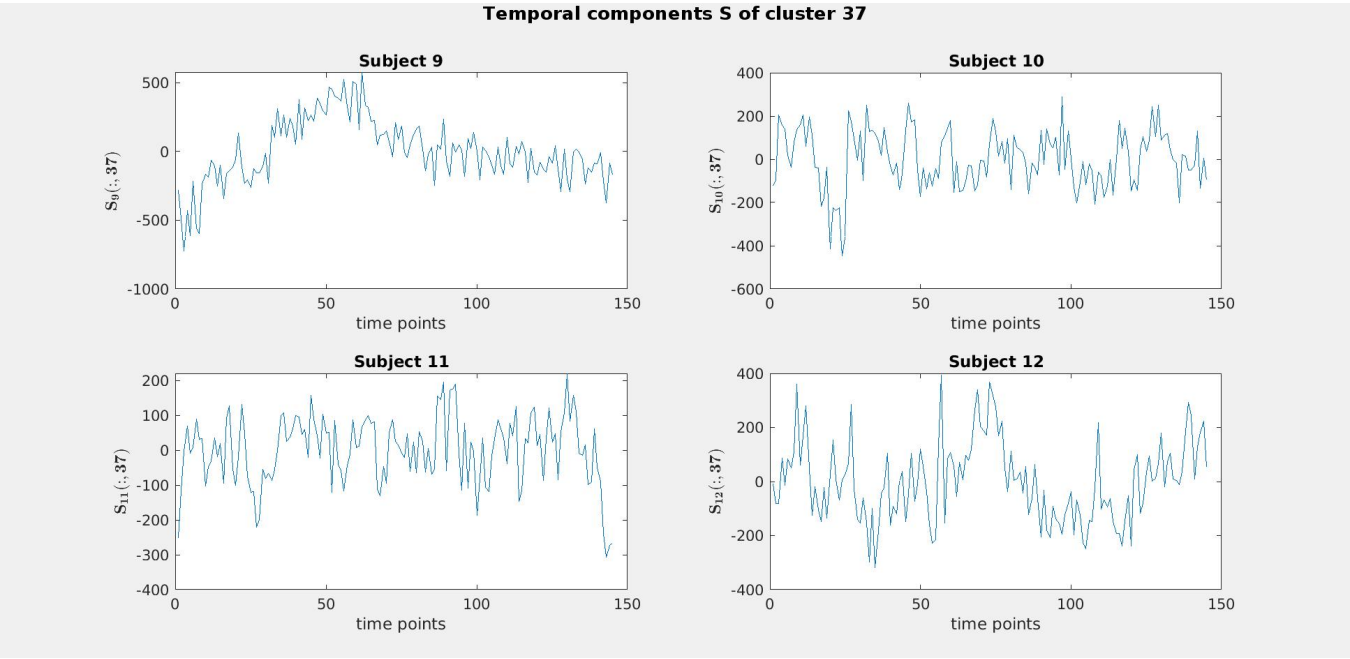


Figure B.448: Waveform of temporal components $\mathbf{S}_k(:, 37)$ for $k = 9, \ldots, 12$ and R=42.

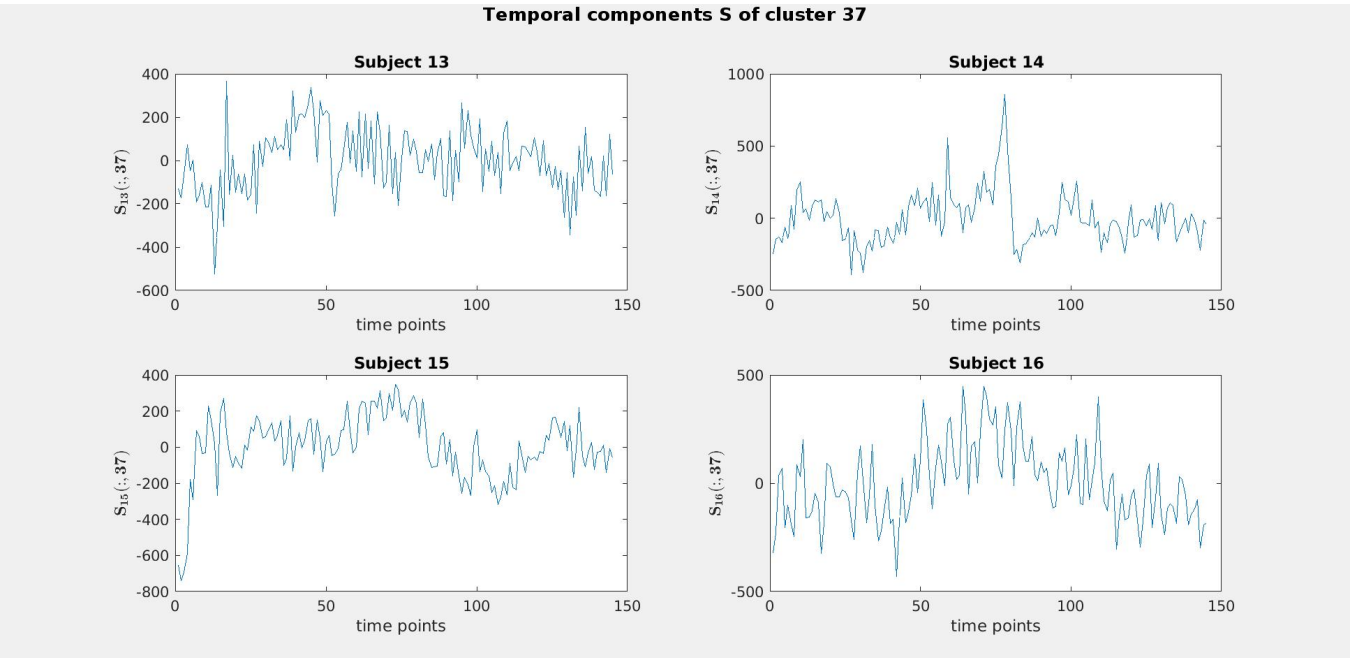


Figure B.449: Waveform of temporal components $\mathbf{S}_k(:, 37)$ for $k = 13, \ldots, 16$ and R=42.

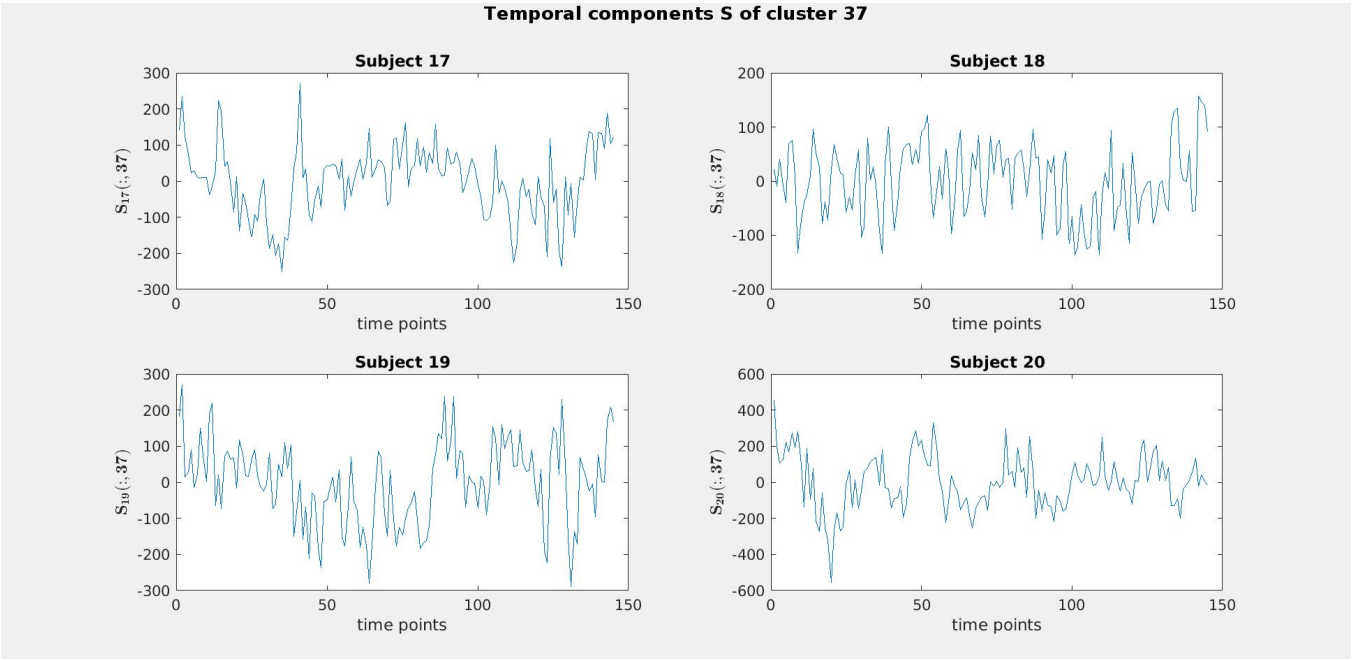


Figure B.450: Waveform of temporal components $\mathbf{S}_k(:, 37)$ for $k = 17, \ldots, 20$ and R=42.

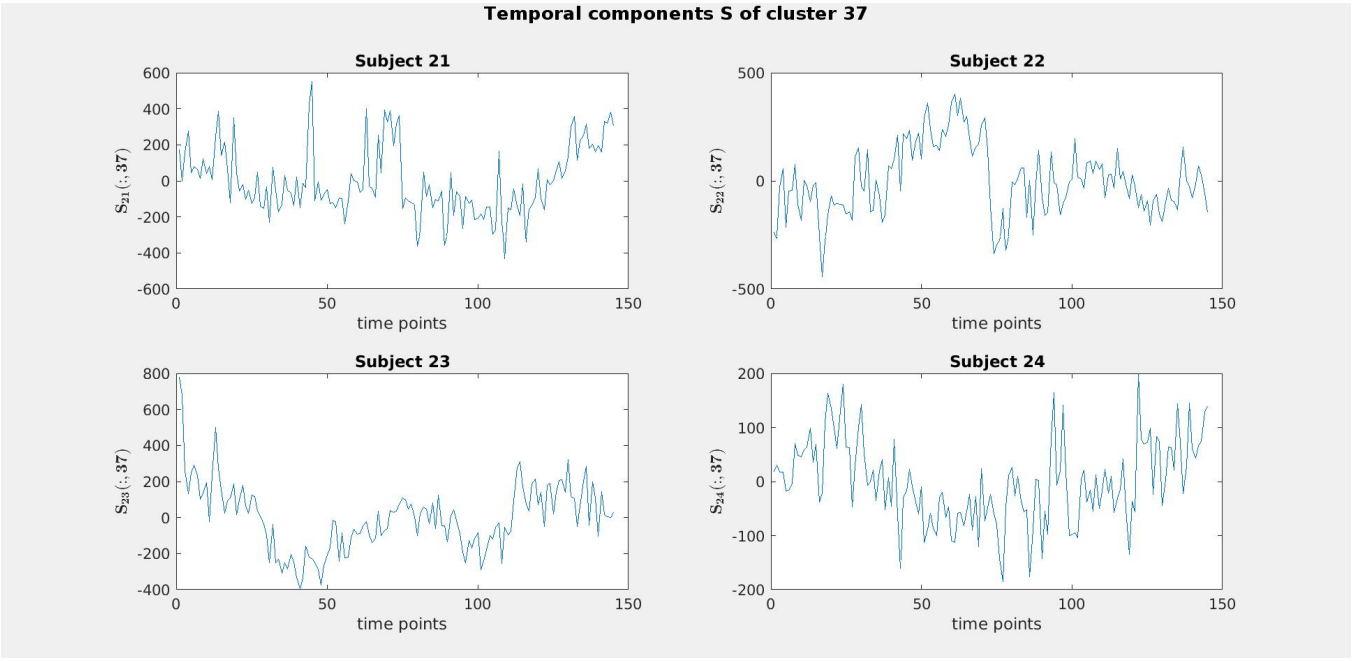


Figure B.451: Waveform of temporal components $\mathbf{S}_k(:, 137)$ for $k = 21, \ldots, 24$ and R=42.

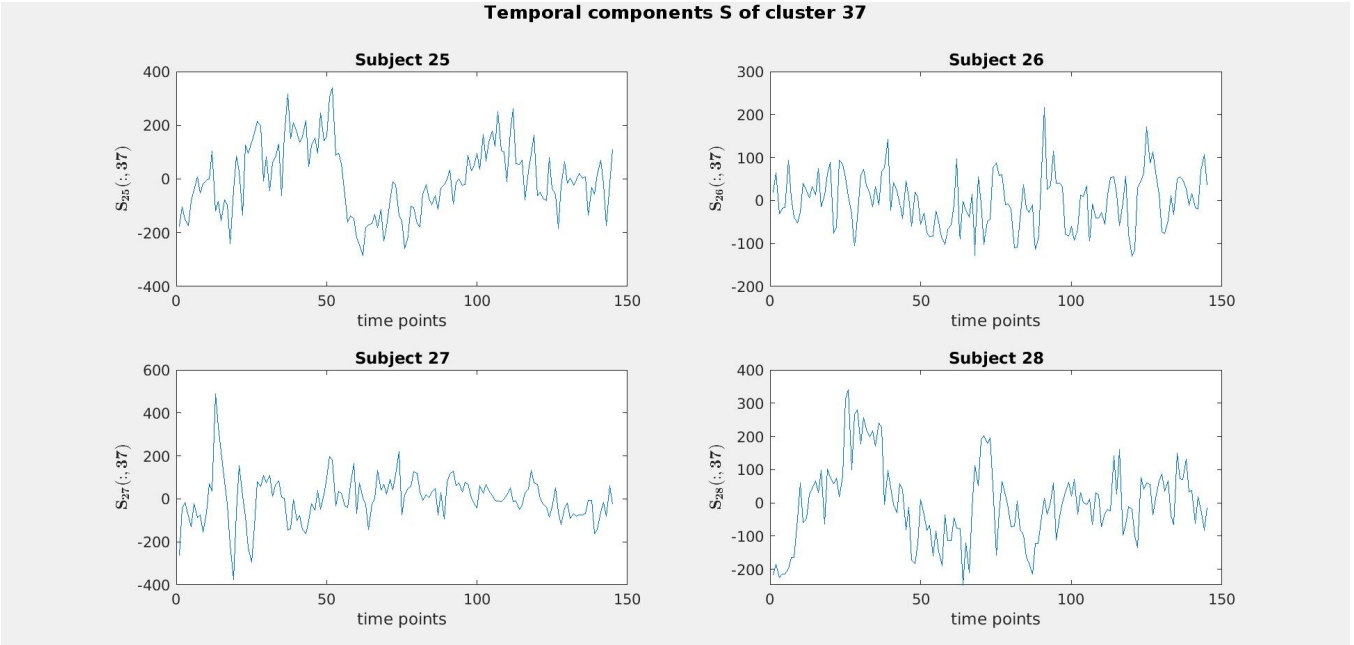


Figure B.452: Waveform of temporal components $\mathbf{S}_k(:, 37)$ for $k = 25, \ldots, 28$ and R=42.

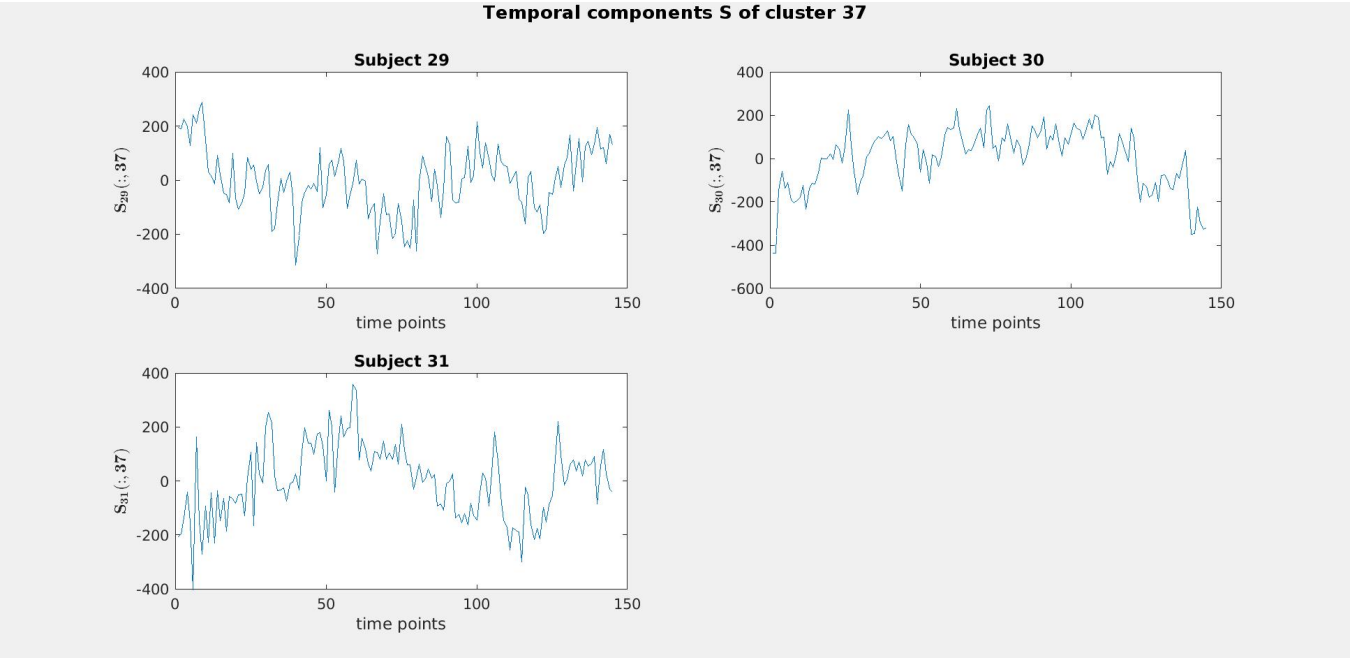


Figure B.453: Waveform of temporal components $\mathbf{S}_k(:, 37)$ for $k = 29, \ldots, 31$ and R=42.

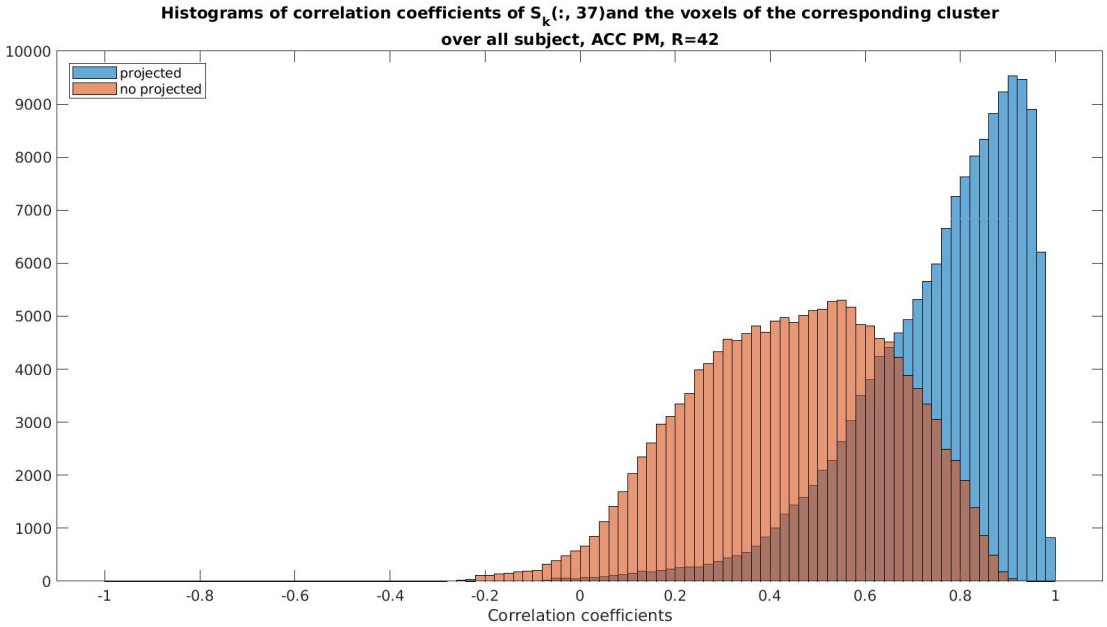


Figure B.454: For the column 37 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 37)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

17. **Cluster 7**

    Cluster 7 contains 5973 voxels, which correspond to 2.643529% of the whole brain (background, soft tissue, bone are excluded).

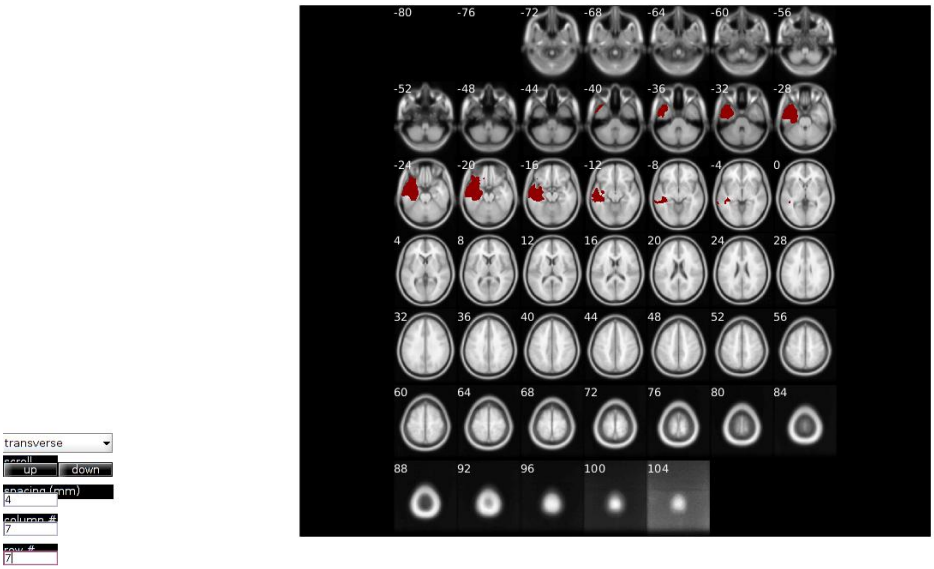


Figure B.455: Cluster 7. Red areas indicate the voxels which belong to cluster 7.

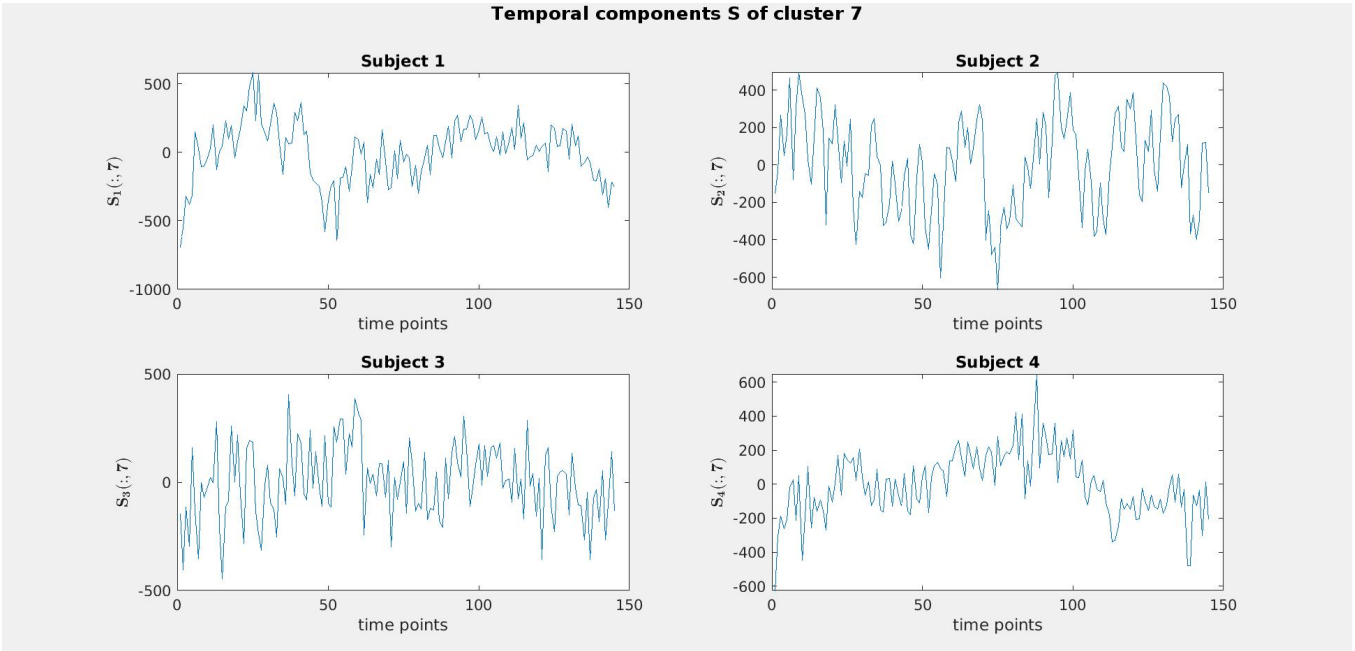


Figure B.456: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 1,\ldots,4$ and R=42.

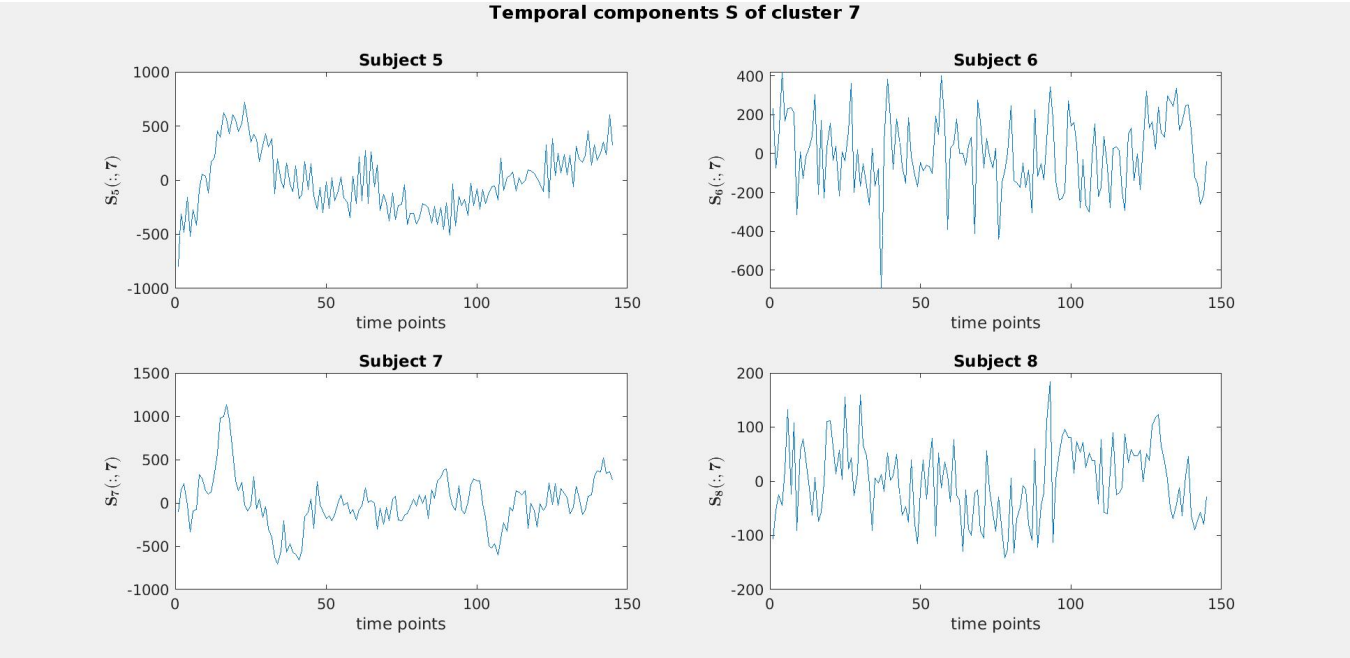


Figure B.457: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 5,\ldots,8$ and R=42.

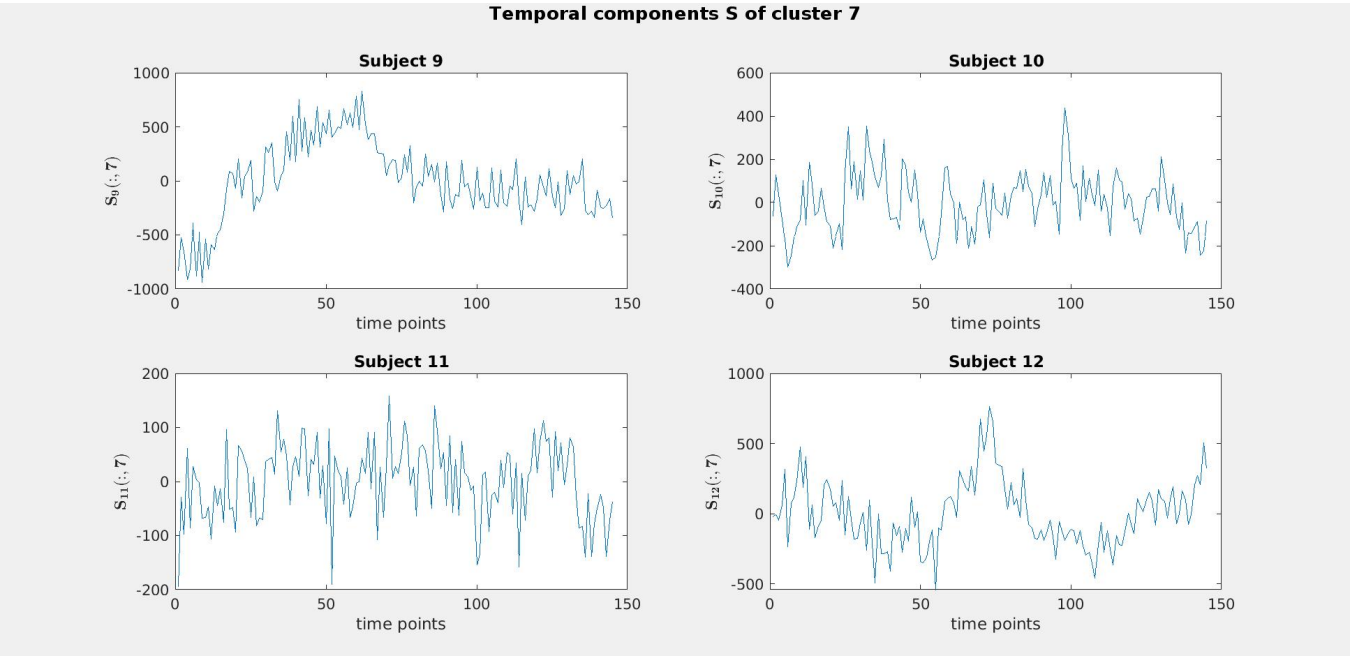


Figure B.458: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 9,\ldots,12$ and R=42.

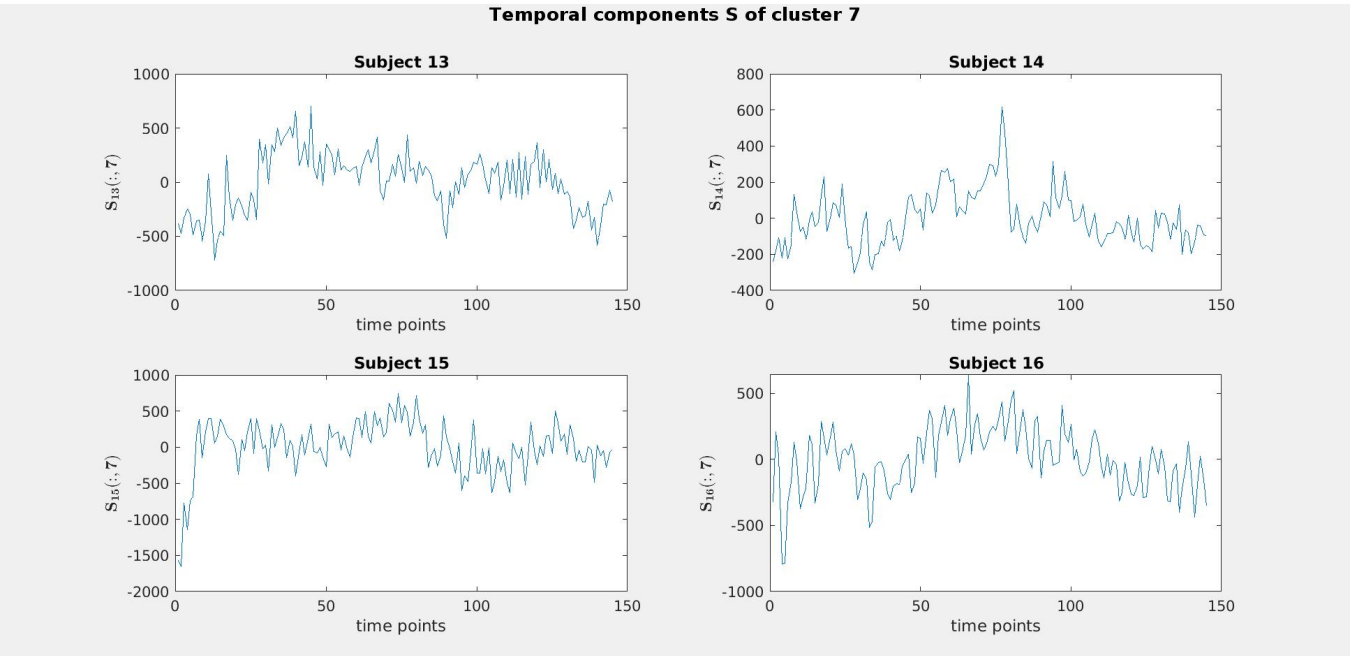


Figure B.459: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 13,\ldots,16$ and R=42.

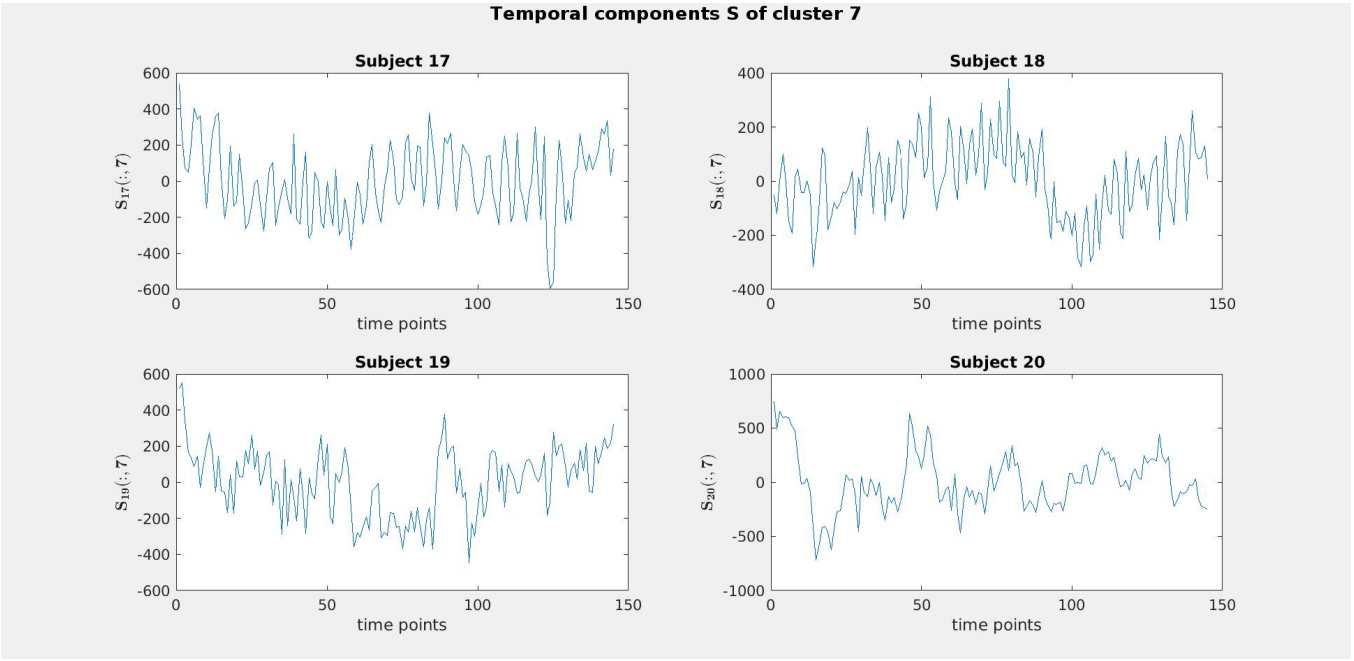


Figure B.460: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 17,\ldots,20$ and R=42.

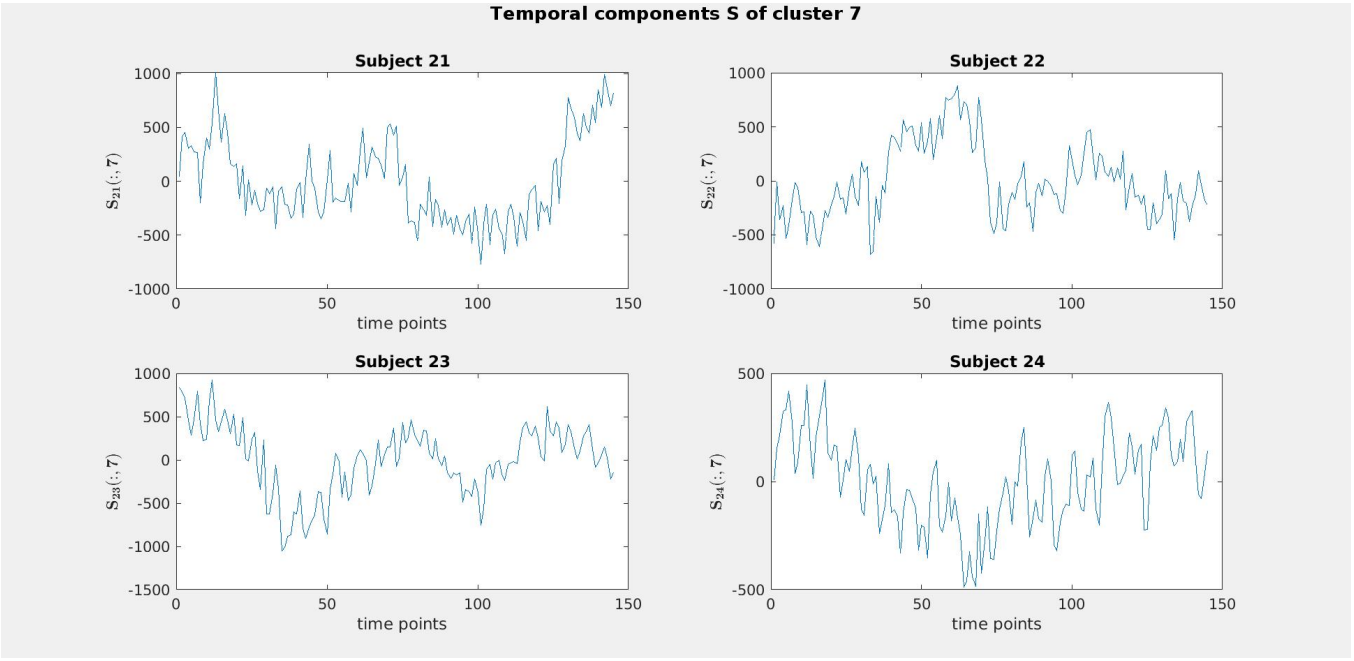


Figure B.461: Waveform of temporal components $\mathbf{S}_k(:,17)$ for $k = 21,\ldots,24$ and R=42.

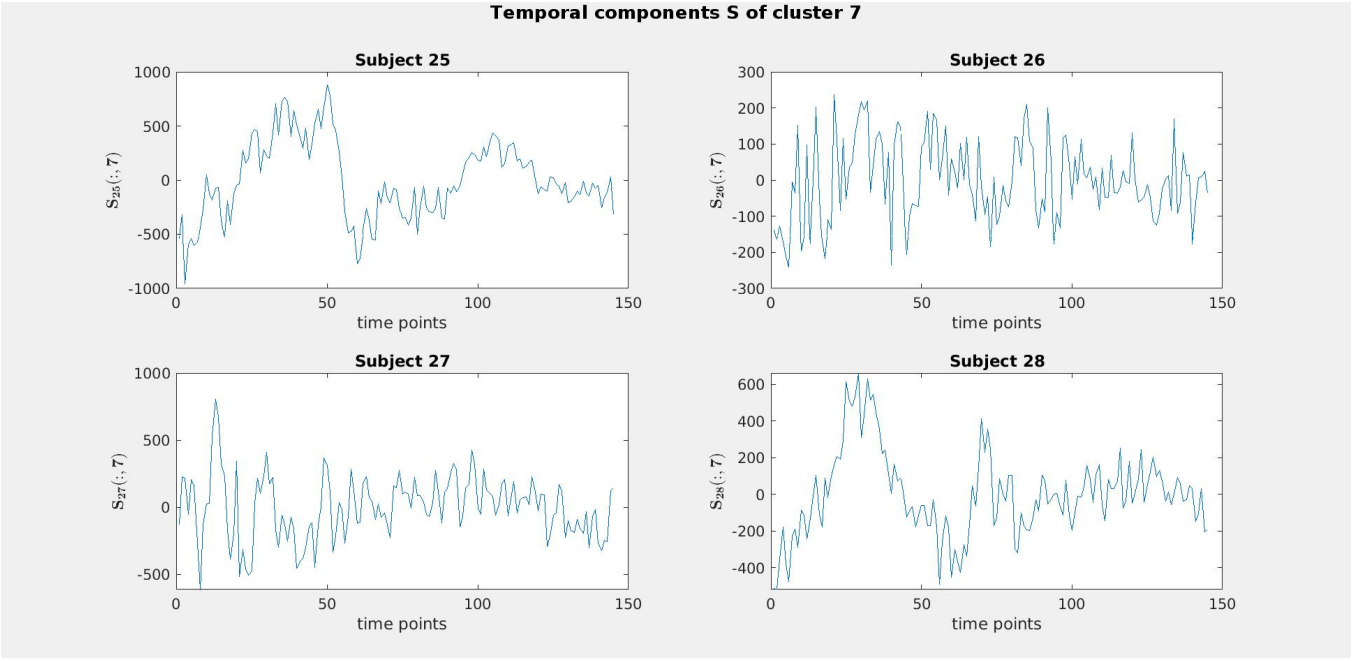


Figure B.462: Waveform of temporal components $\mathbf{S}_k(:,7)$ for $k = 25,\ldots,28$ and R=42.

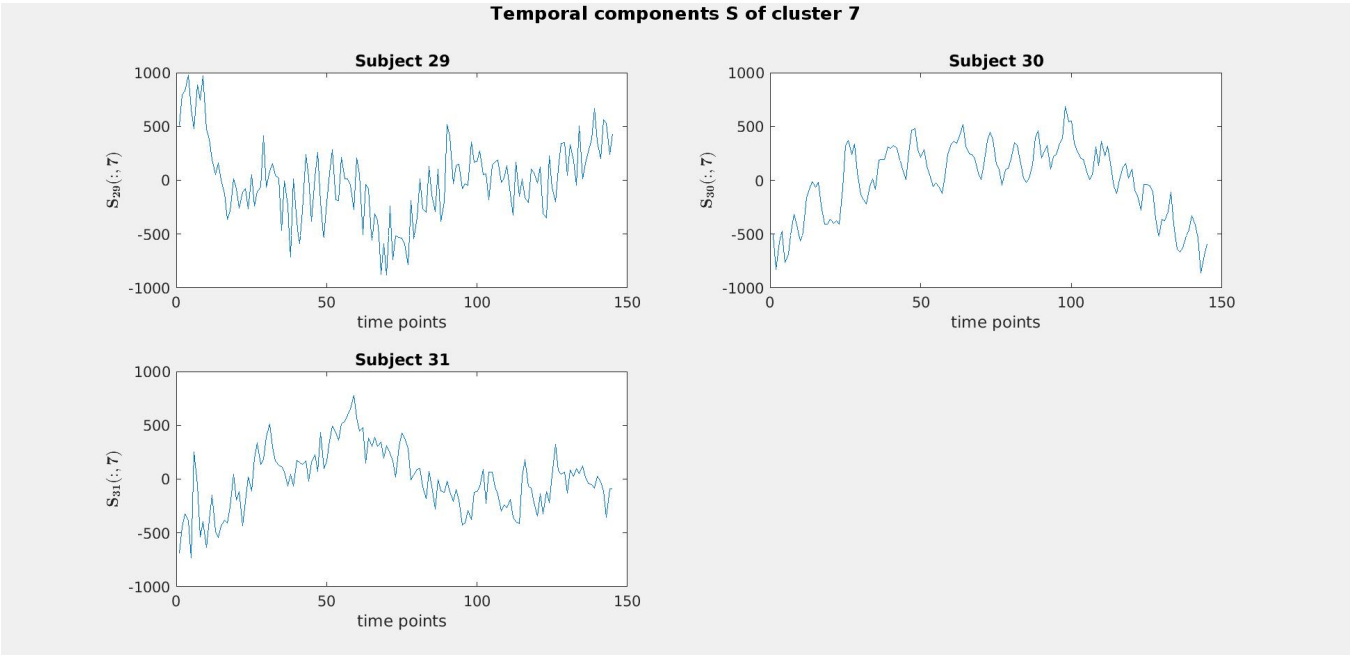


Figure B.463: Waveform of temporal components $\mathbf{S}_k(:, 7)$ for $k = 29, \ldots, 31$ and R=42.

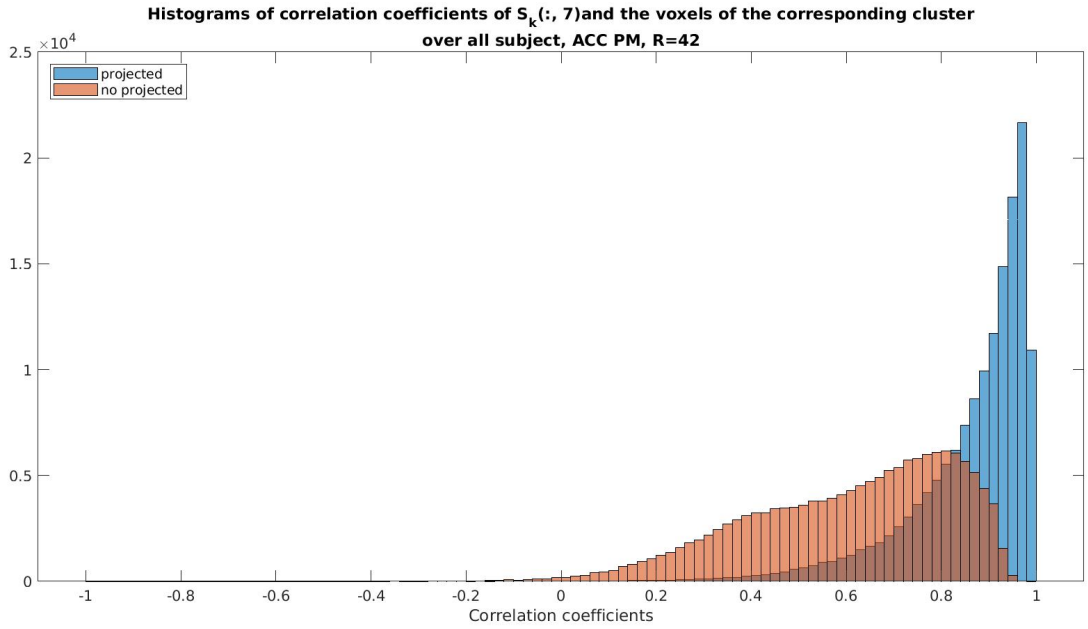


Figure B.464: For the column 7 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 7)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

18. **Cluster 35**

Cluster 35 contains 5851 voxels, which correspond to 2.589534% of the whole brain (background, soft tissue, bone are excluded).

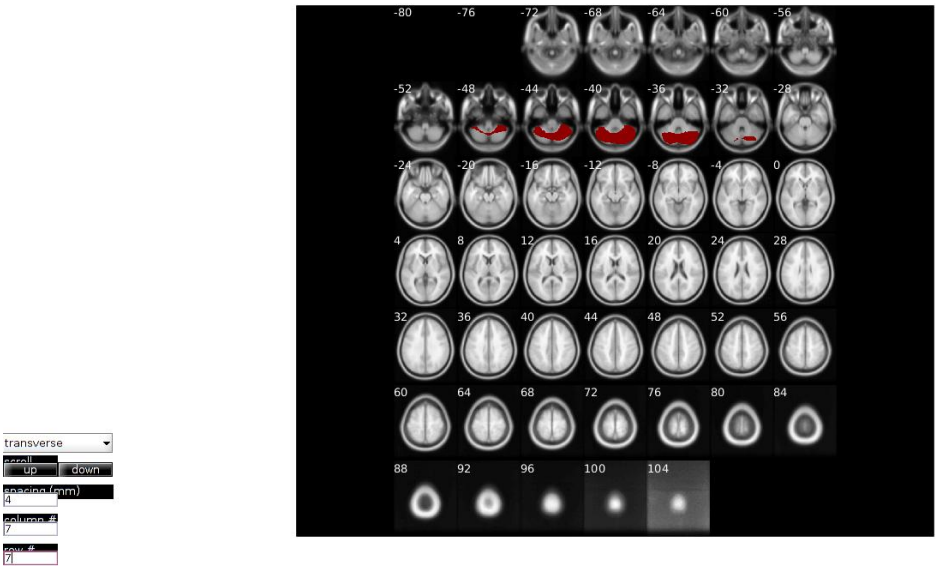


Figure B.465: Cluster 35. Red areas indicate the voxels which belong to cluster 35.

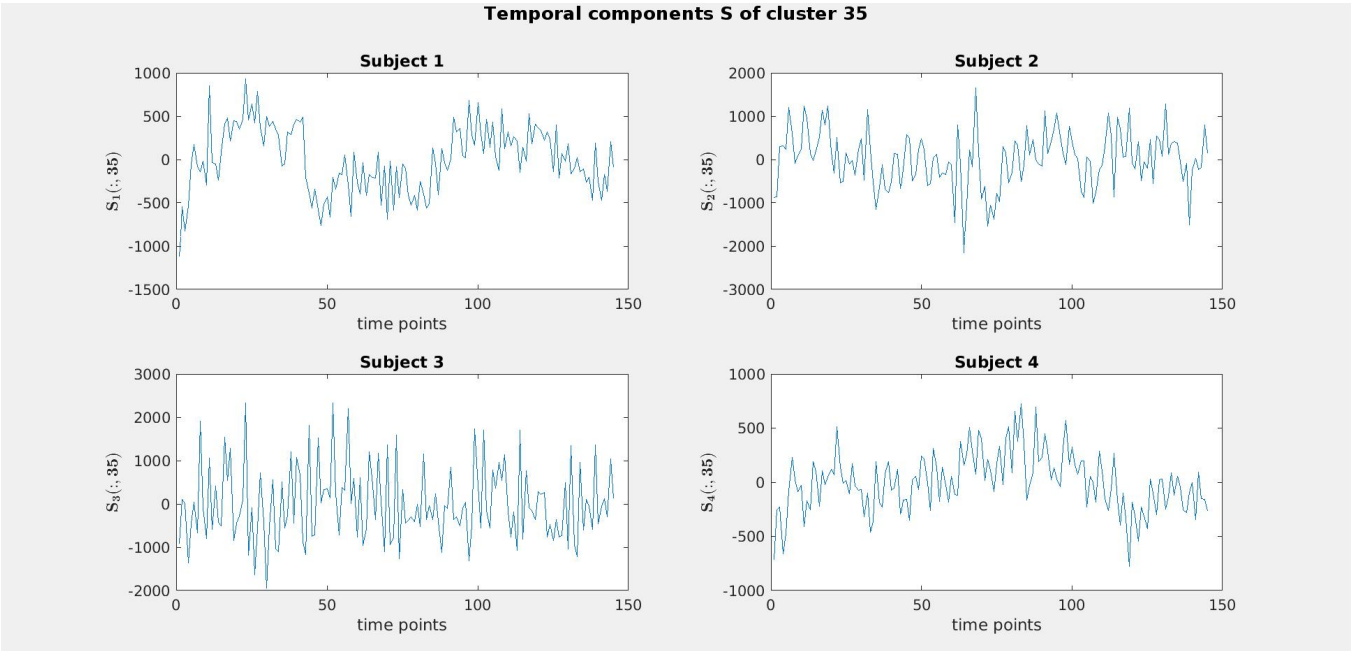


Figure B.466: Waveform of temporal components $\mathbf{S}_k(:, 35)$ for $k = 1, \ldots, 4$ and R=42.

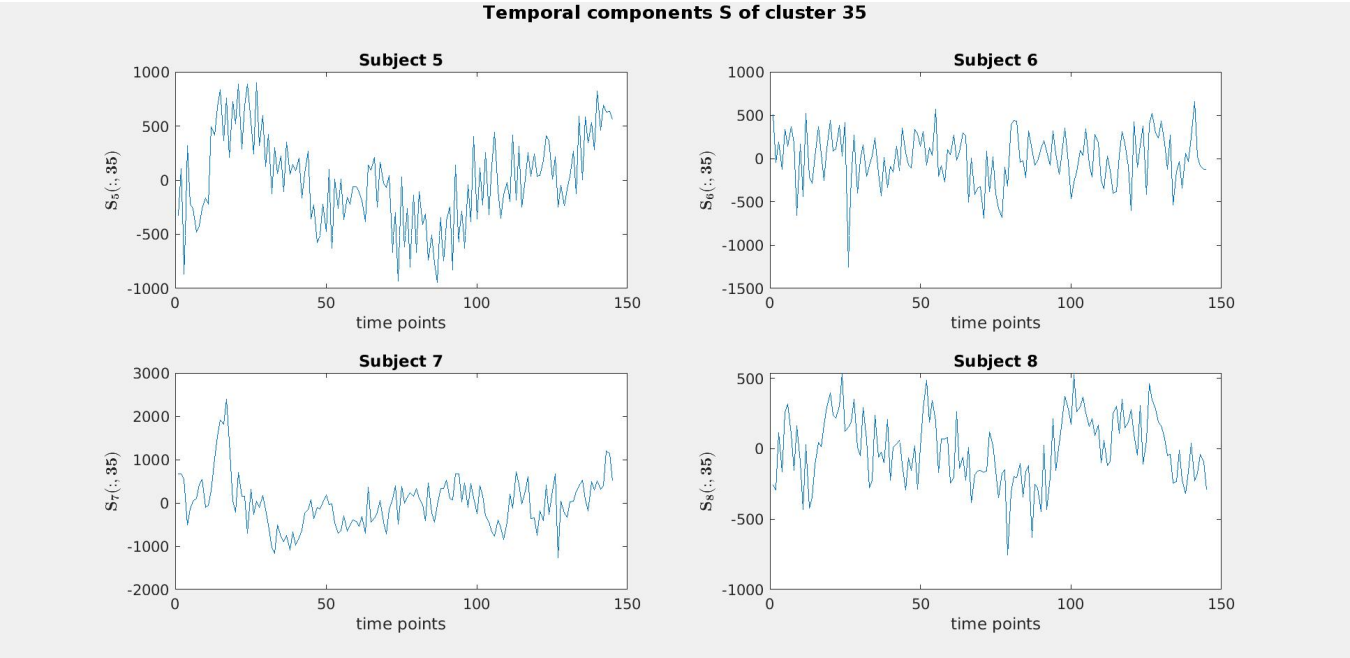


Figure B.467: Waveform of temporal components $\mathbf{S}_k(:, 35)$ for $k = 5, \ldots, 8$ and R=42.

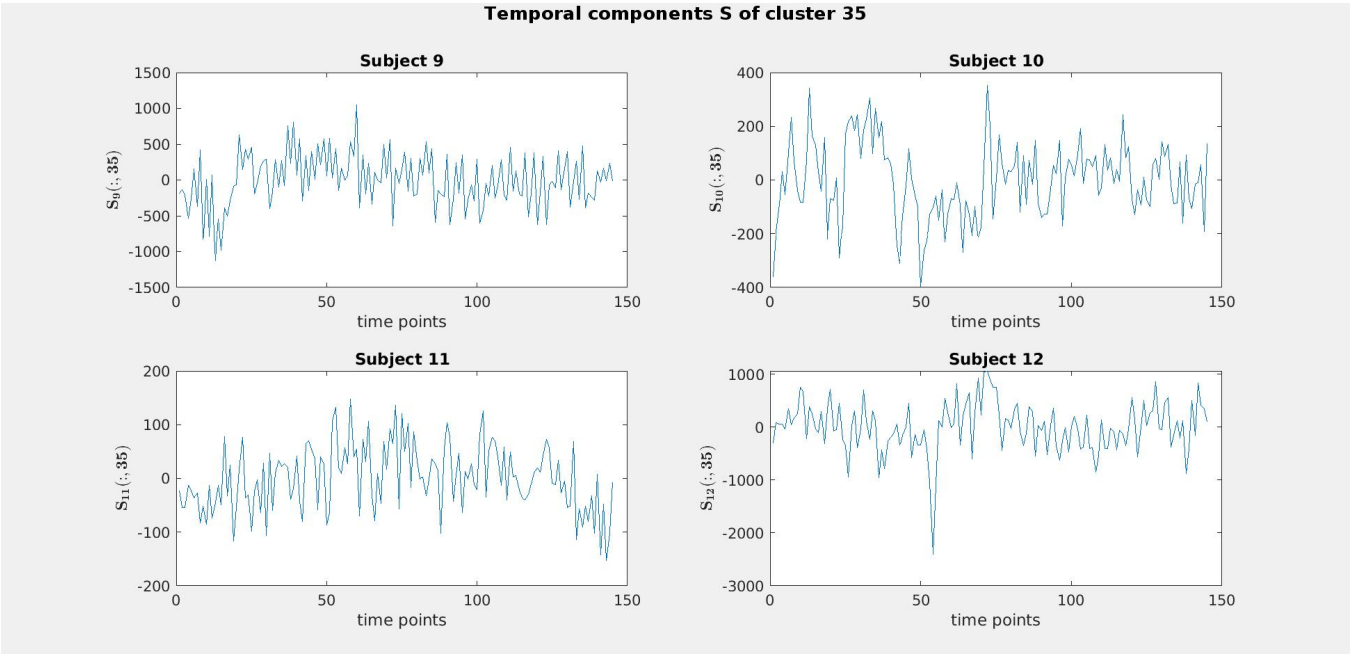


Figure B.468: Waveform of temporal components $\mathbf{S}_k(:, 35)$ for $k = 9, \ldots, 12$ and R=42.

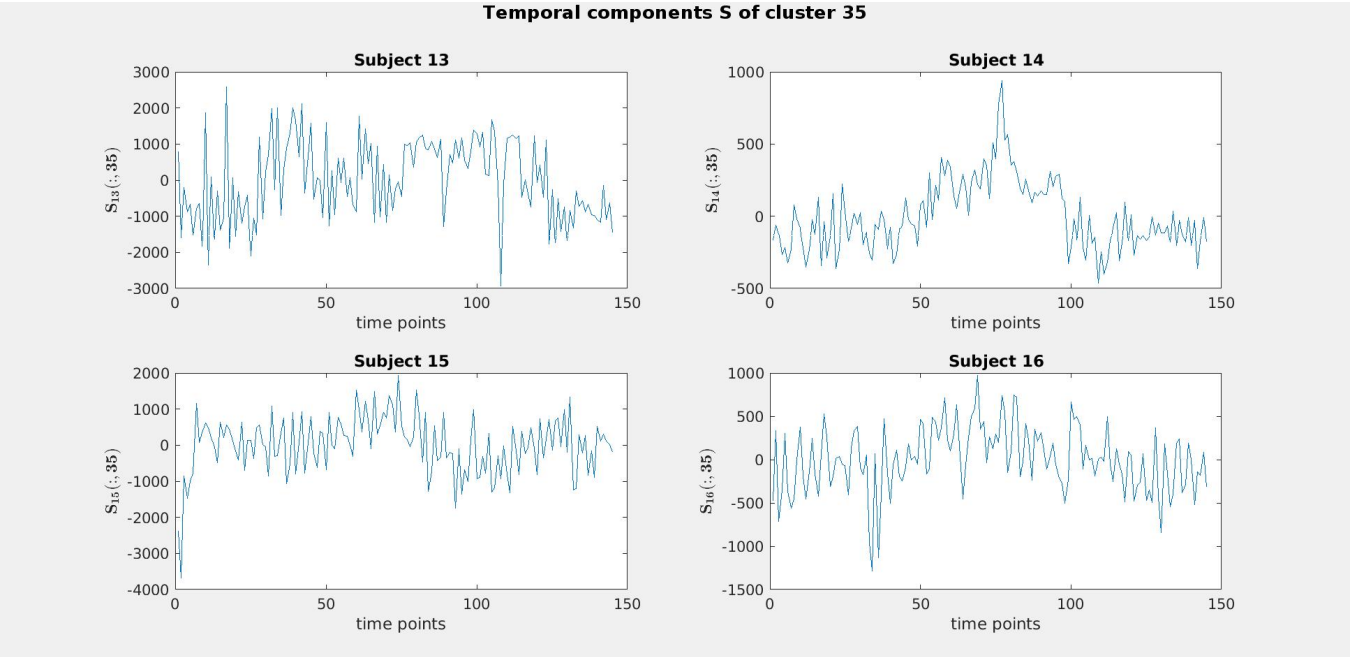


Figure B.469: Waveform of temporal components $\mathbf{S}_k(:, 35)$ for $k = 13, \ldots, 16$ and R=42.

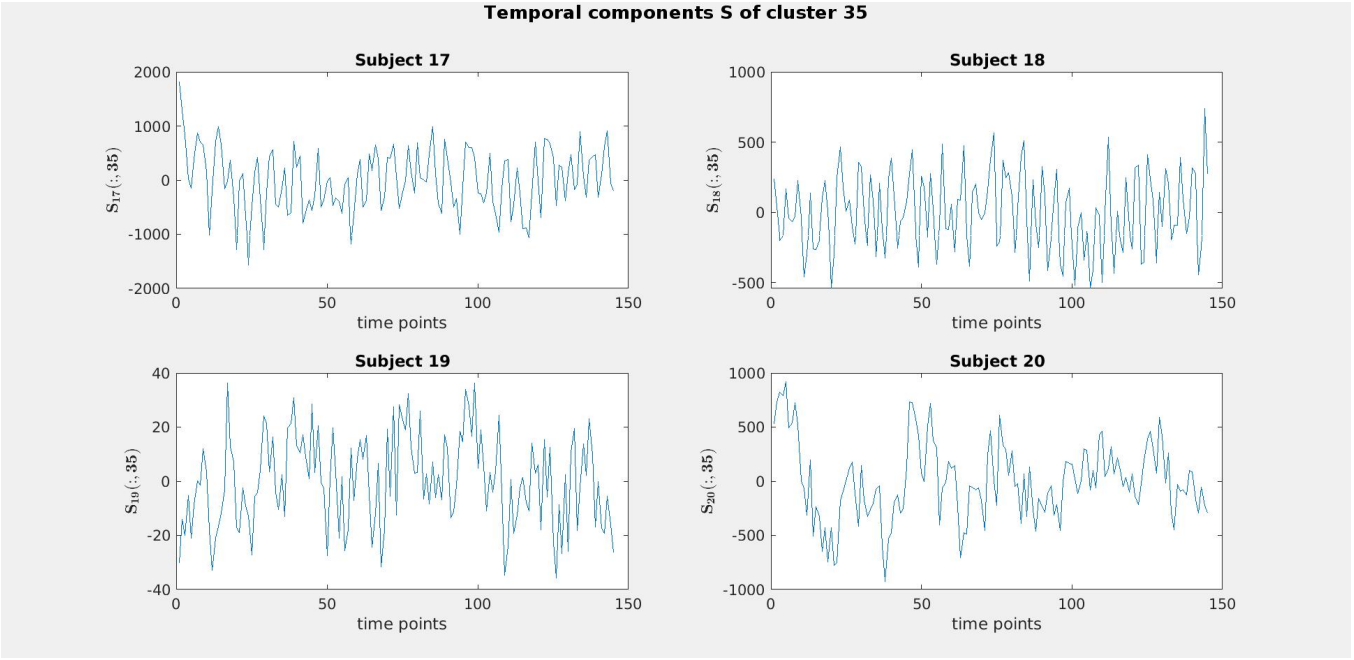


Figure B.470: Waveform of temporal components $\mathbf{S}_k(:, 35)$ for $k = 17, \ldots, 20$ and R=42.

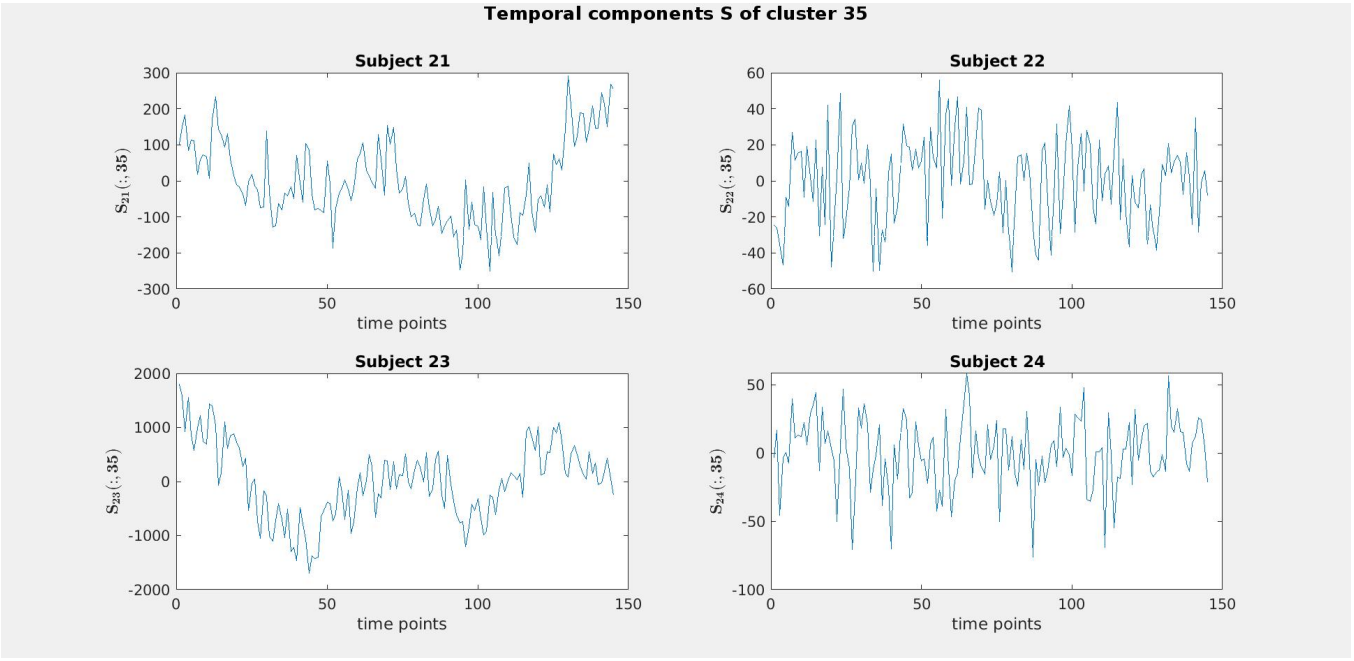


Figure B.471: Waveform of temporal components $\mathbf{S}_k(:, 135)$ for $k = 21, \ldots, 24$ and R=42.

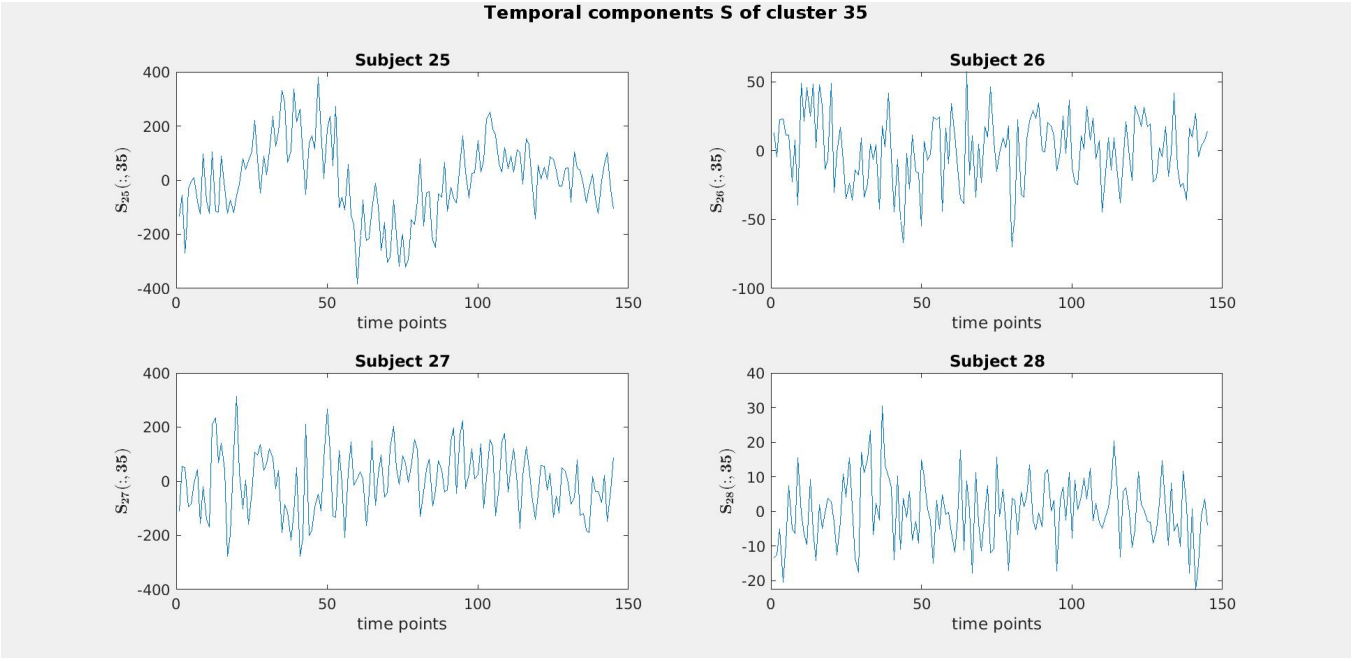


Figure B.472: Waveform of temporal components $\mathbf{S}_k(:, 35)$ for $k = 25, \ldots, 28$ and R=42.

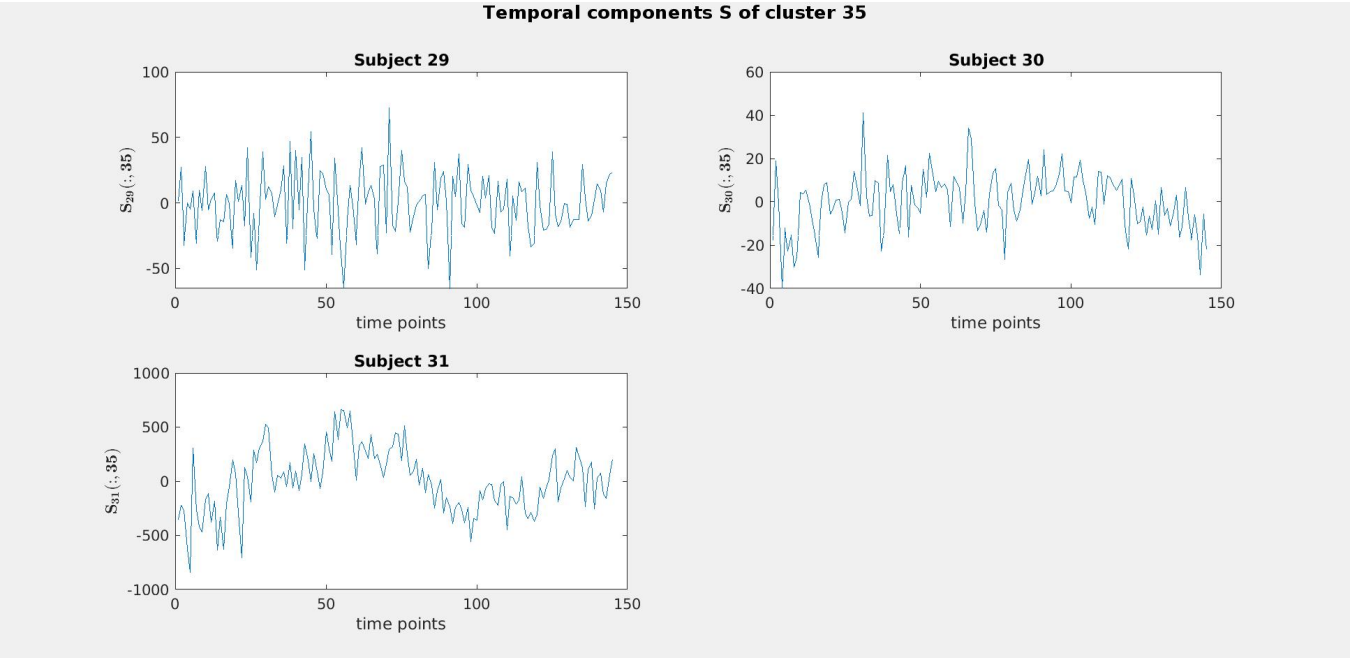


Figure B.473: Waveform of temporal components $\mathbf{S}_k(:, 35)$ for $k = 29, \ldots, 31$ and R=42.

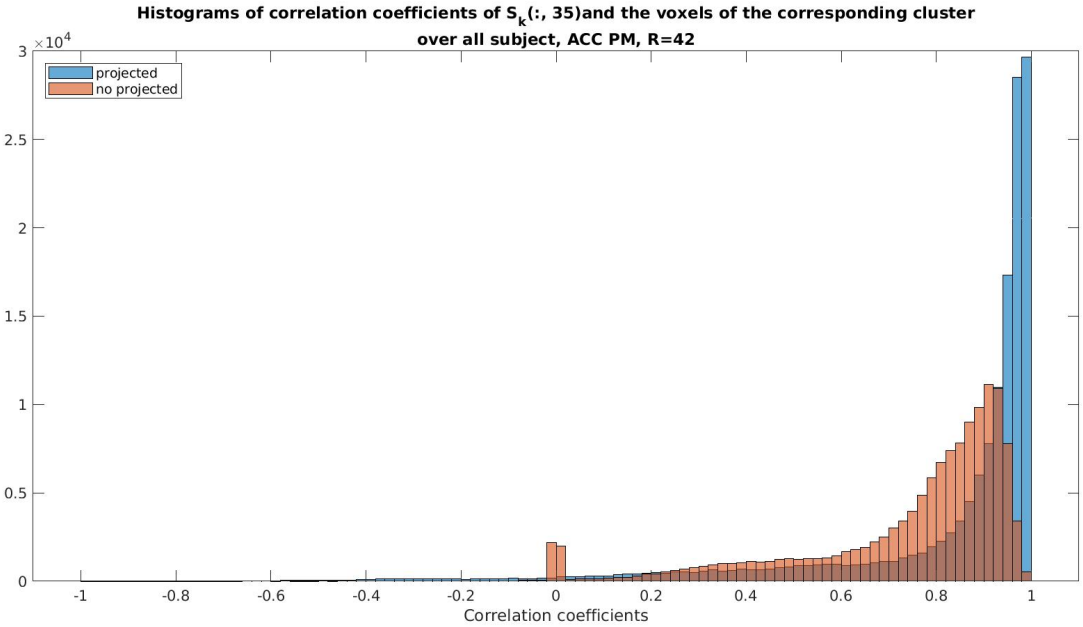


Figure B.474: For the column 35 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 35)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

19. **Cluster 19**

Cluster 19 contains 5538 voxels, which correspond to 2.451006% of the whole brain (background, soft tissue, bone are excluded).

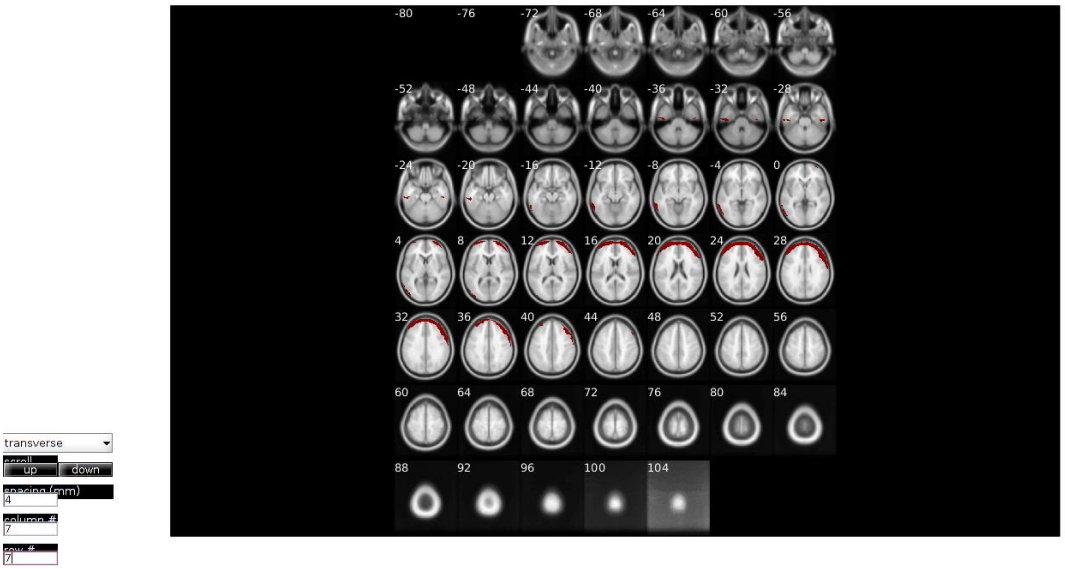


Figure B.475: Cluster 19. Red areas indicate the voxels which belong to cluster 19.

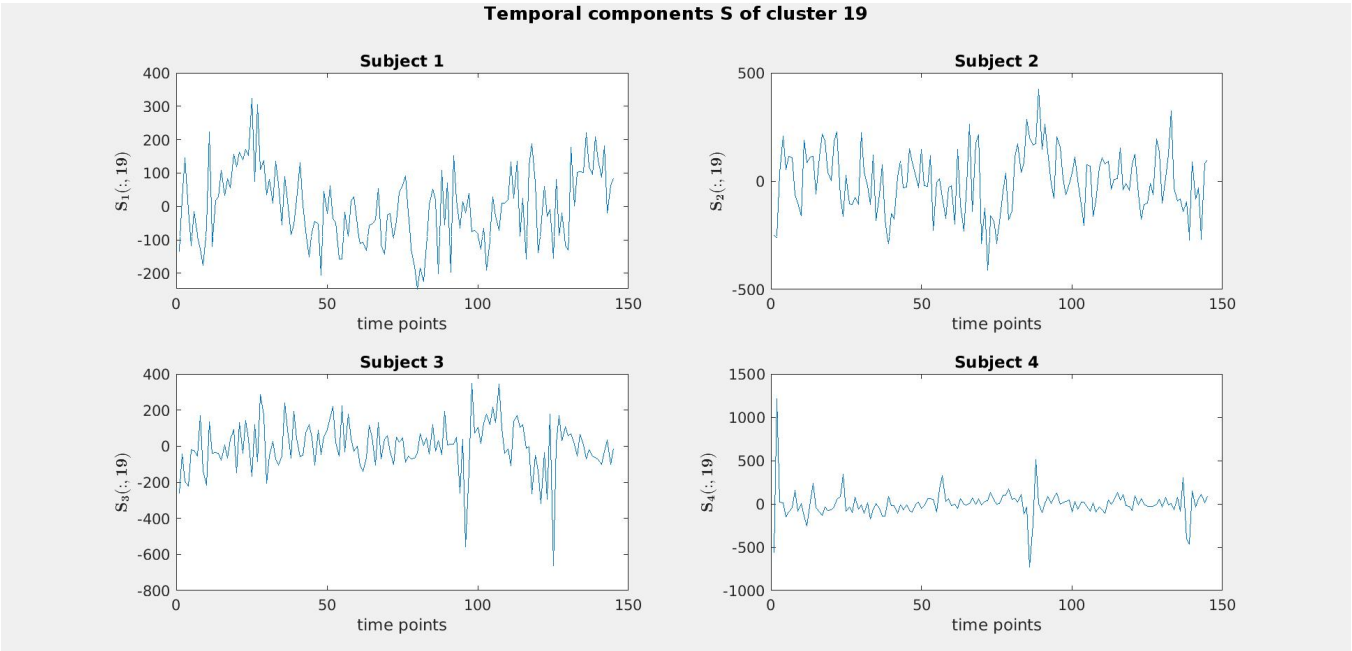


Figure B.476: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 1, \ldots, 4$ and R=42.

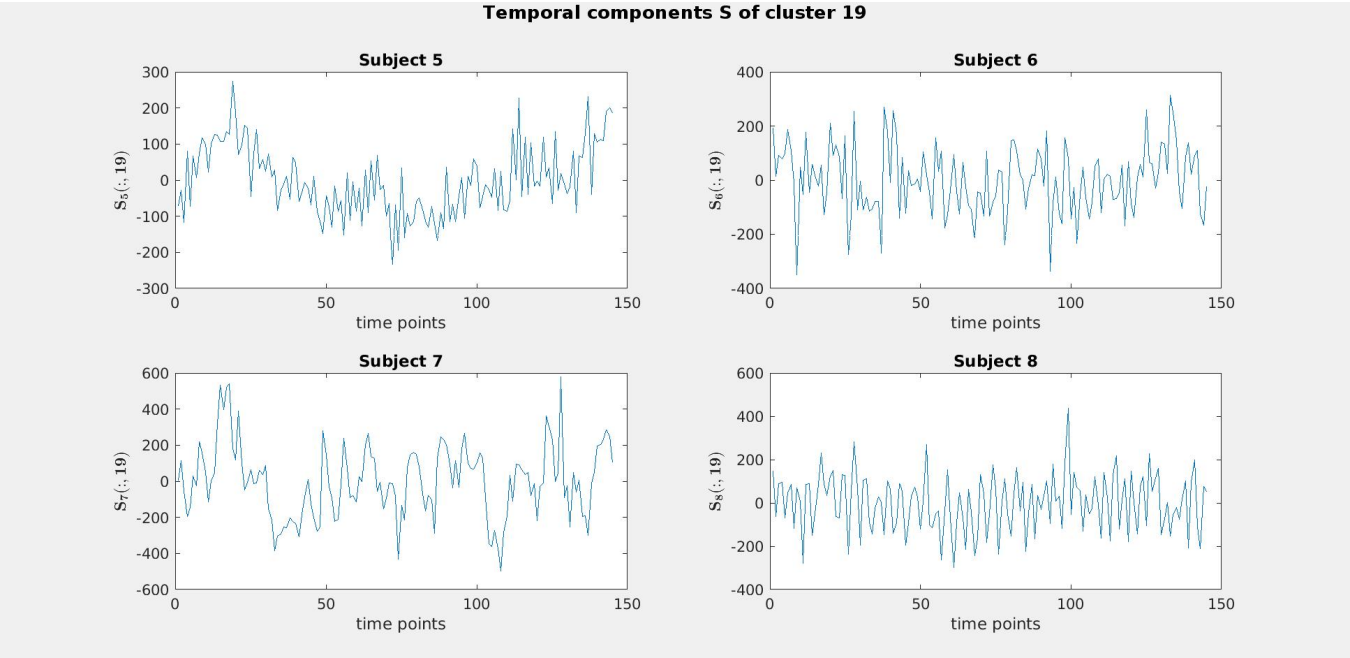


Figure B.477: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 5, \ldots, 8$ and R=42.

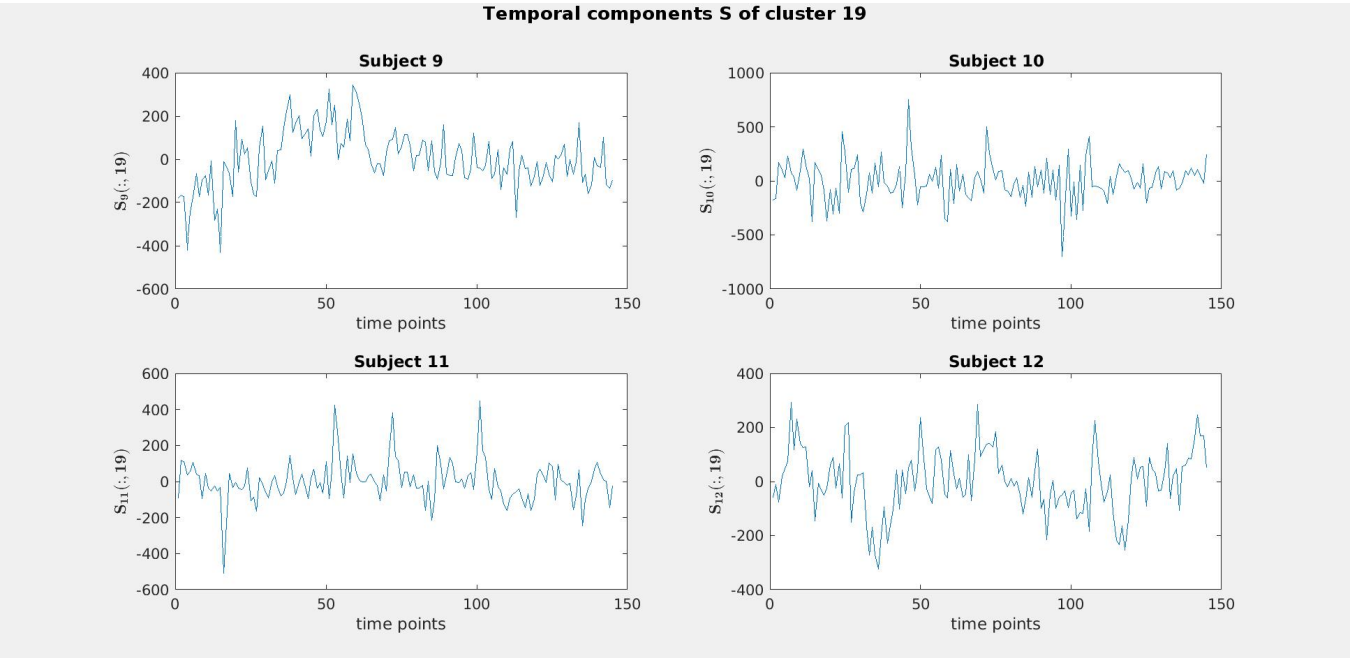


Figure B.478: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 9, \ldots, 12$ and R=42.

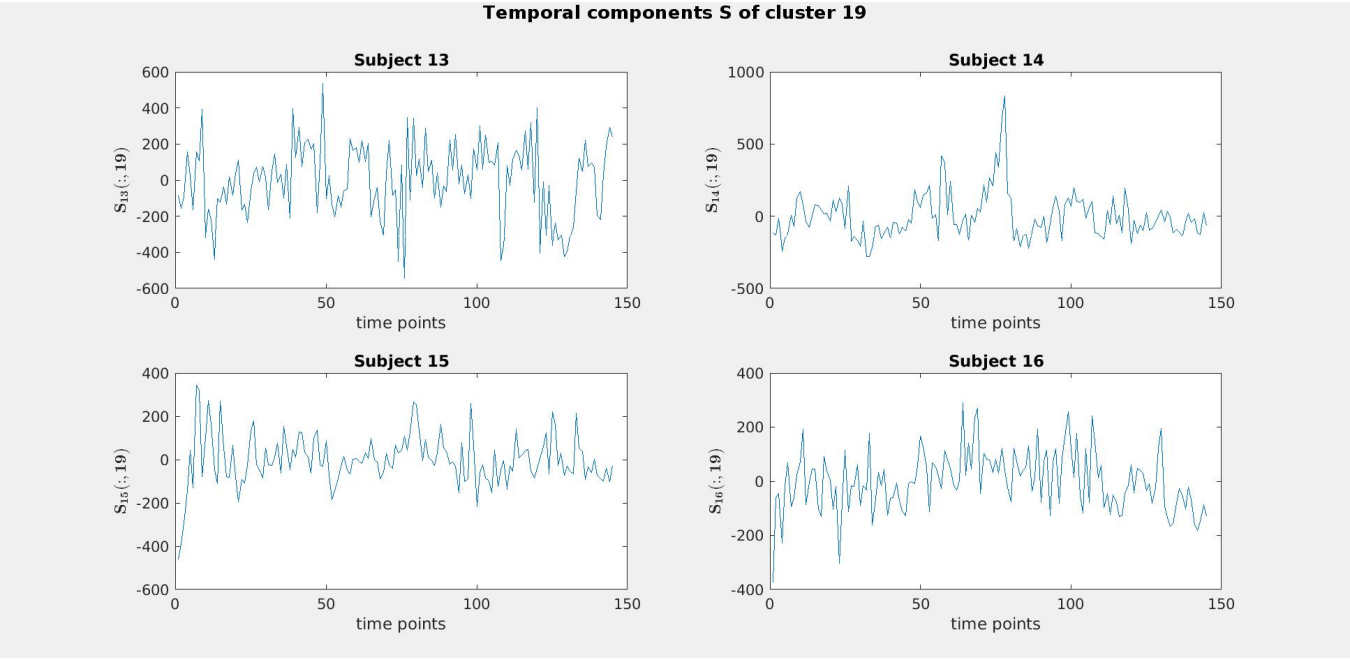


Figure B.479: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 13, \ldots, 16$ and R=42.

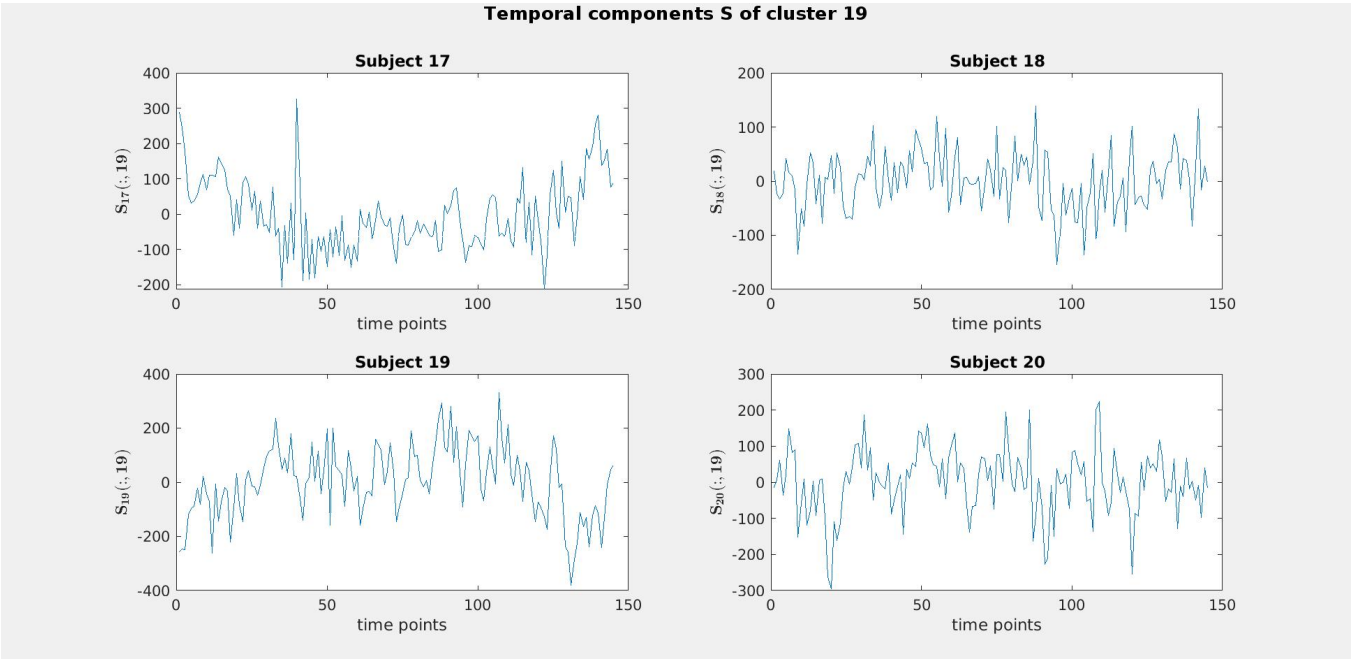


Figure B.480: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 17, \ldots, 20$ and R=42.

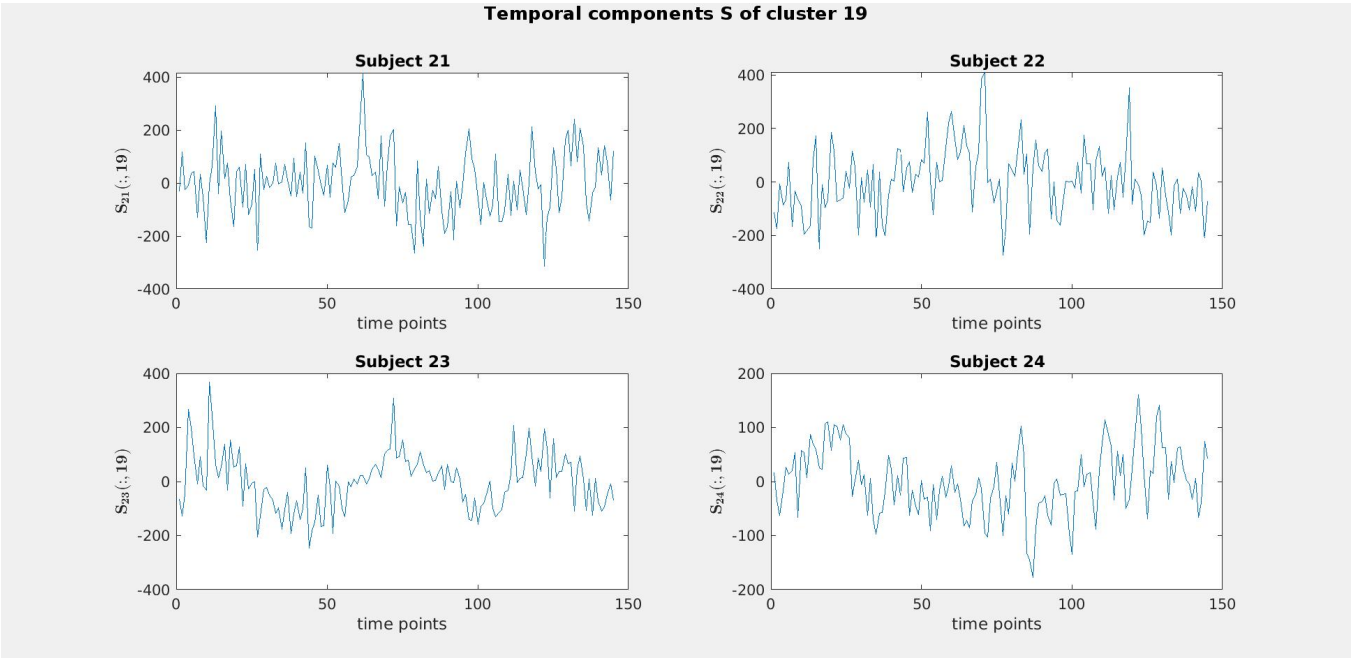


Figure B.481: Waveform of temporal components $\mathbf{S}_k(:, 119)$ for $k = 21, \ldots, 24$ and R=42.

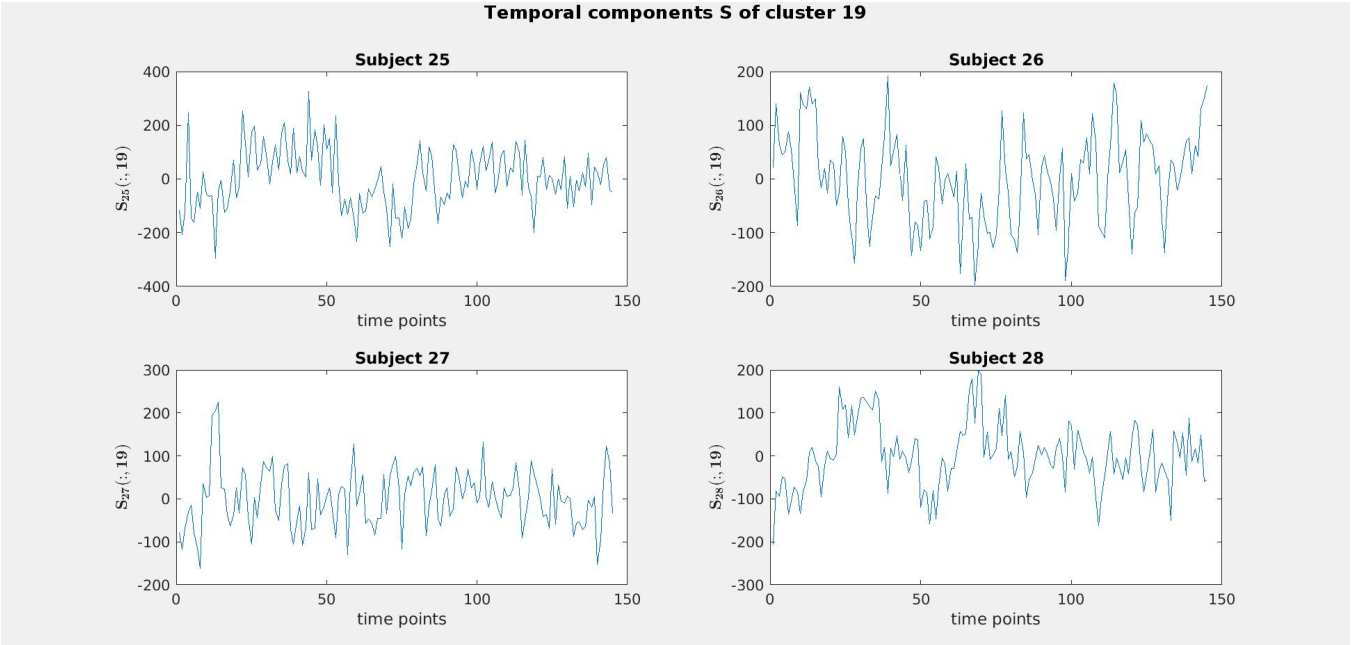


Figure B.482: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 25, \ldots, 28$ and R=42.

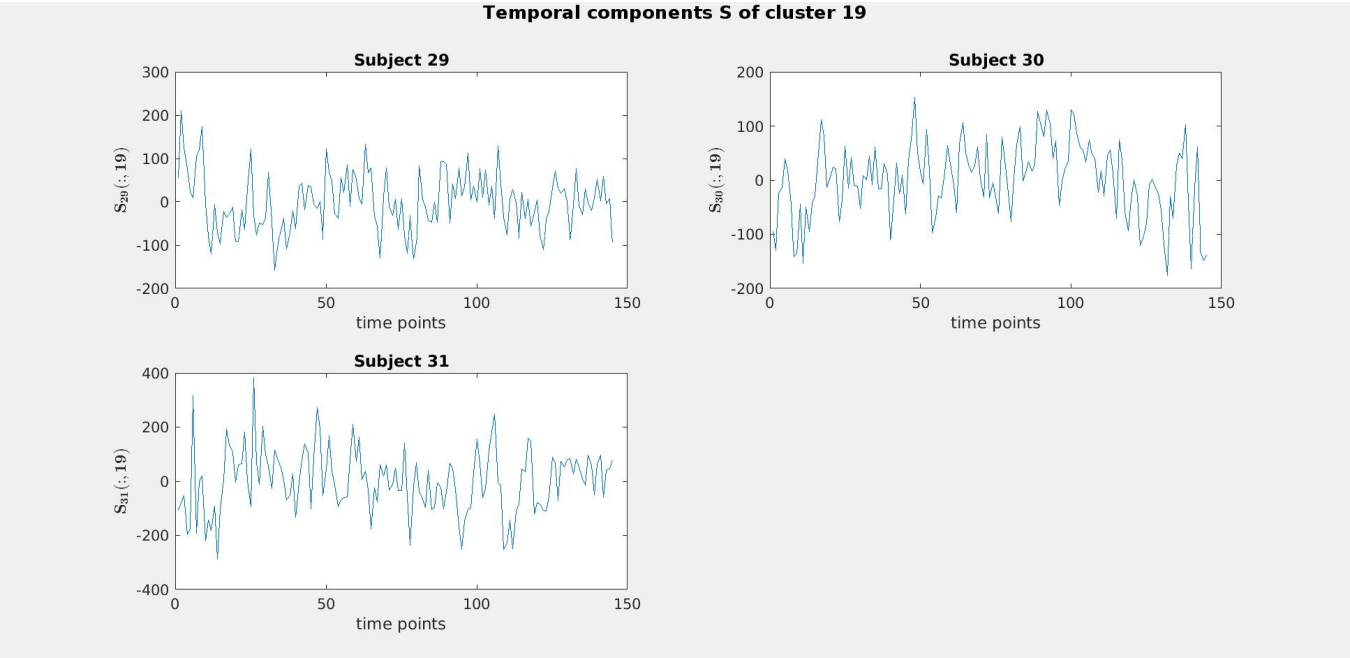


Figure B.483: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 29, \ldots, 31$ and R=42.

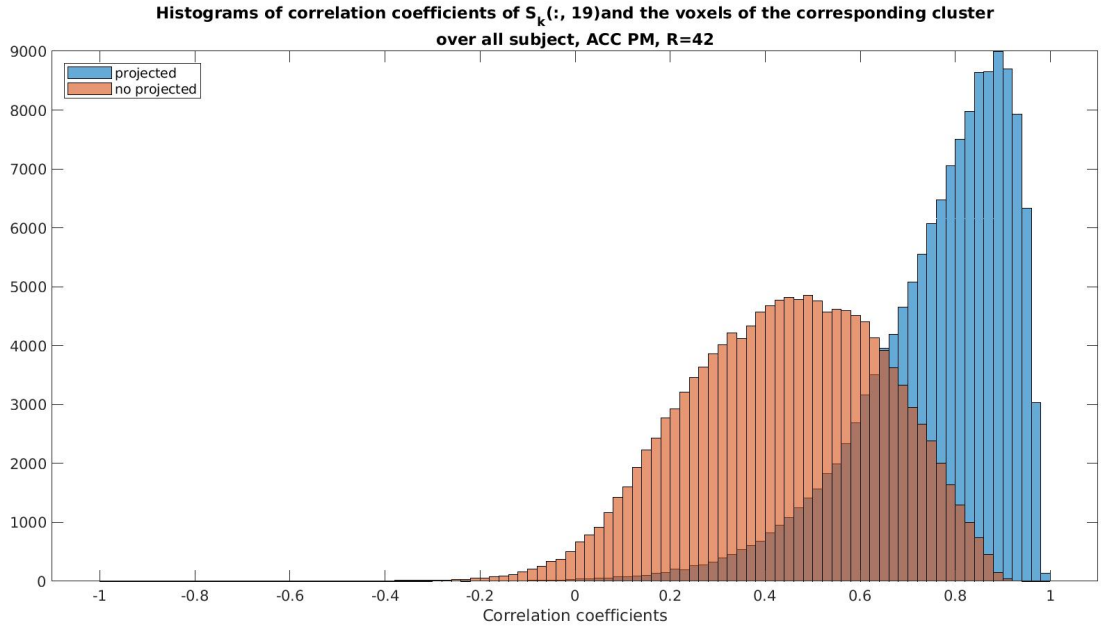


Figure B.484: For the column 19 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 19)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

20. **Cluster 14**

    Cluster 14 contains 5484 voxels, which correspond to 2.427107% of the whole brain (background, soft tissue, bone are excluded).

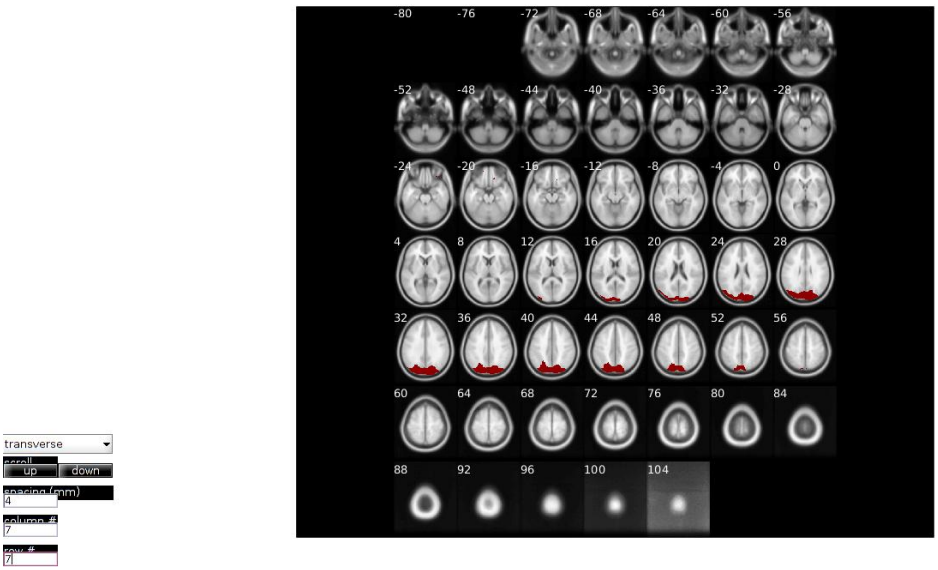


Figure B.485: Cluster 14. Red areas indicate the voxels which belong to cluster 14.

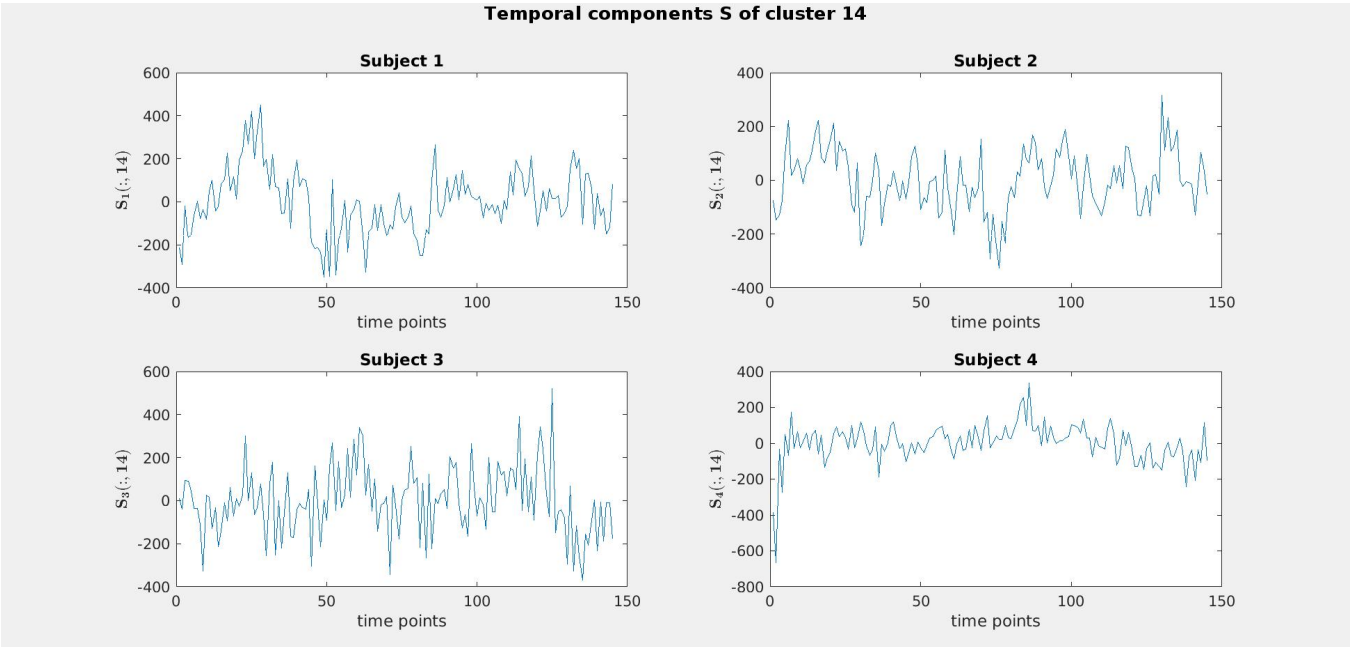


Figure B.486: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 1, \ldots, 4$ and R=42.

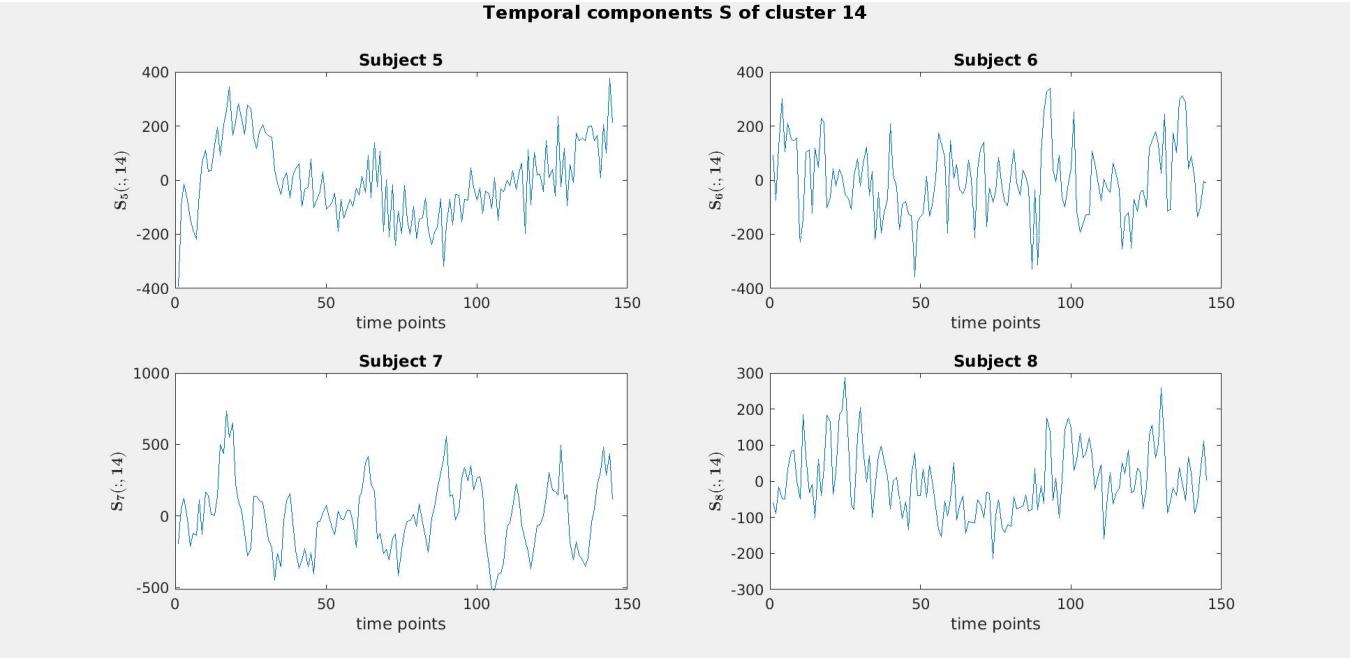


Figure B.487: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 5, \ldots, 8$ and R=42.

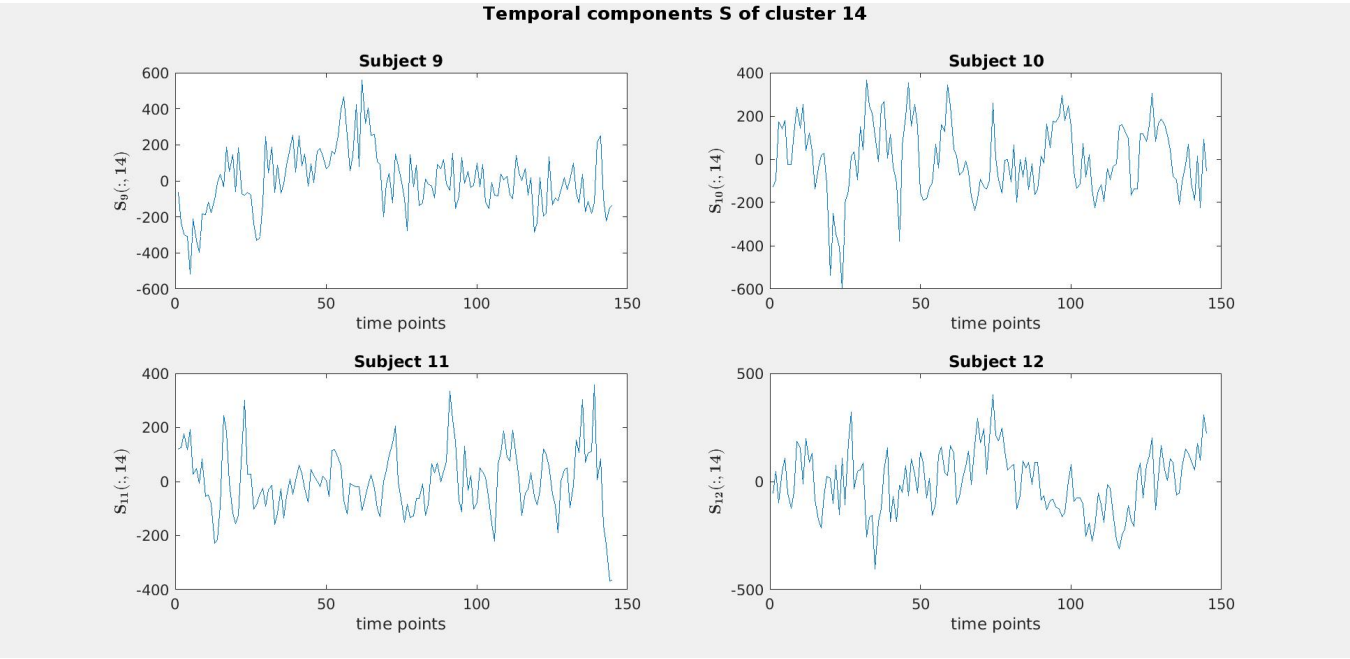


Figure B.488: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 9, \ldots, 12$ and R=42.

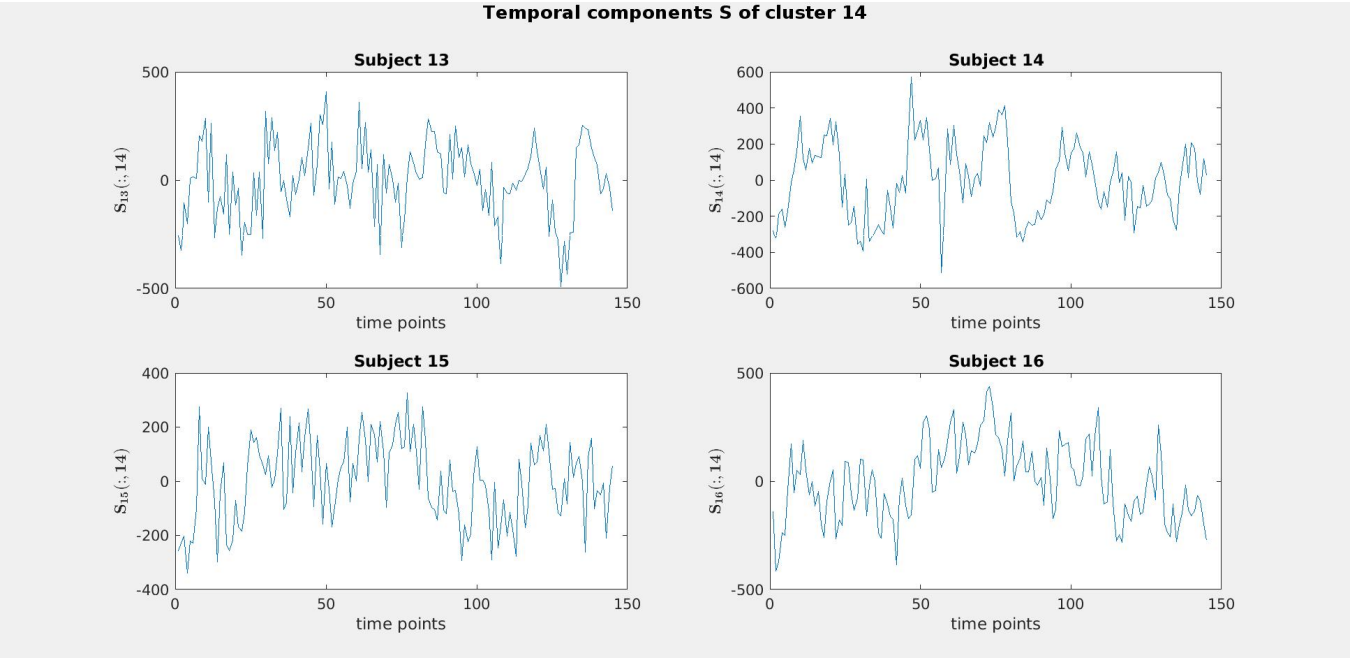


Figure B.489: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 13, \ldots, 16$ and R=42.

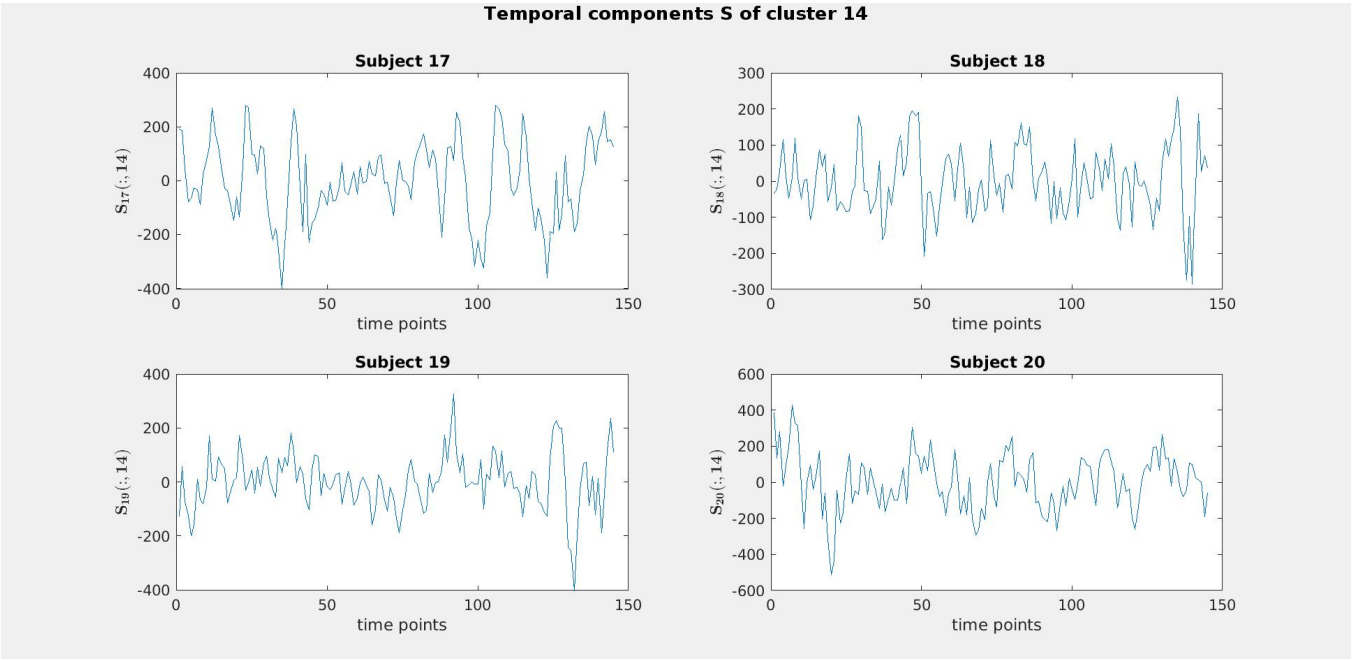


Figure B.490: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 17, \ldots, 20$ and R=42.

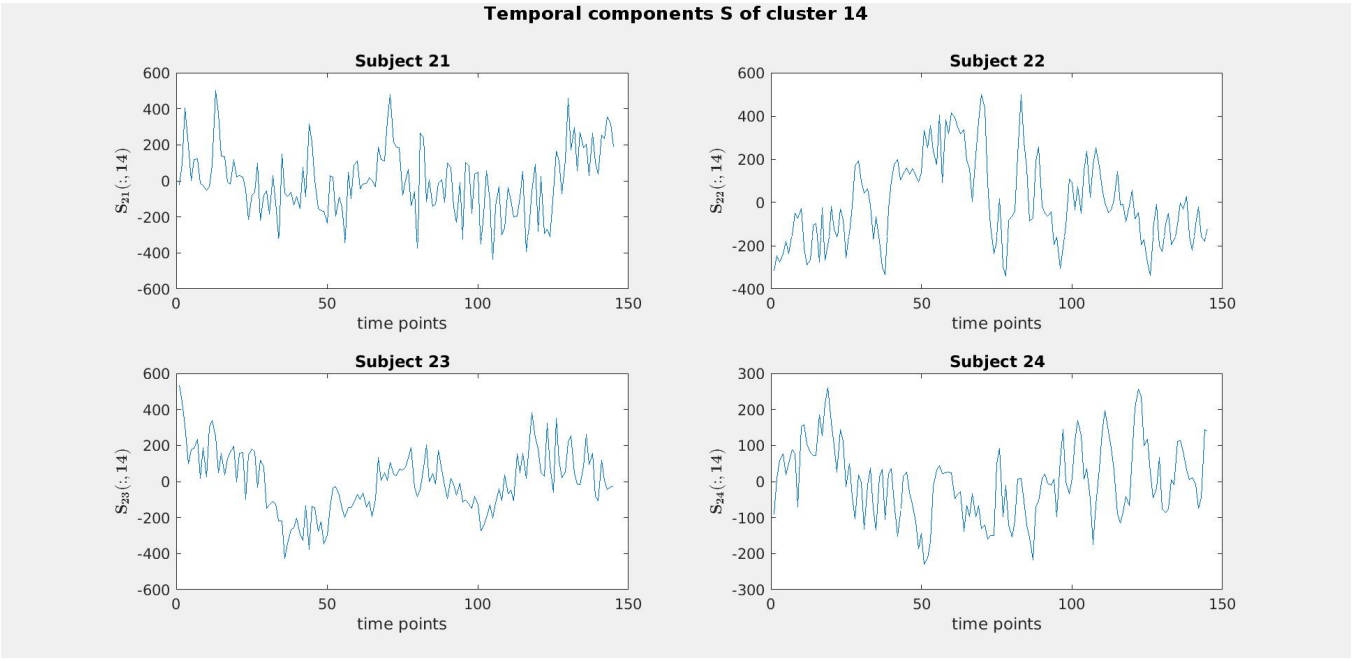


Figure B.491: Waveform of temporal components $\mathbf{S}_k(:, 114)$ for $k = 21, \ldots, 24$ and R=42.

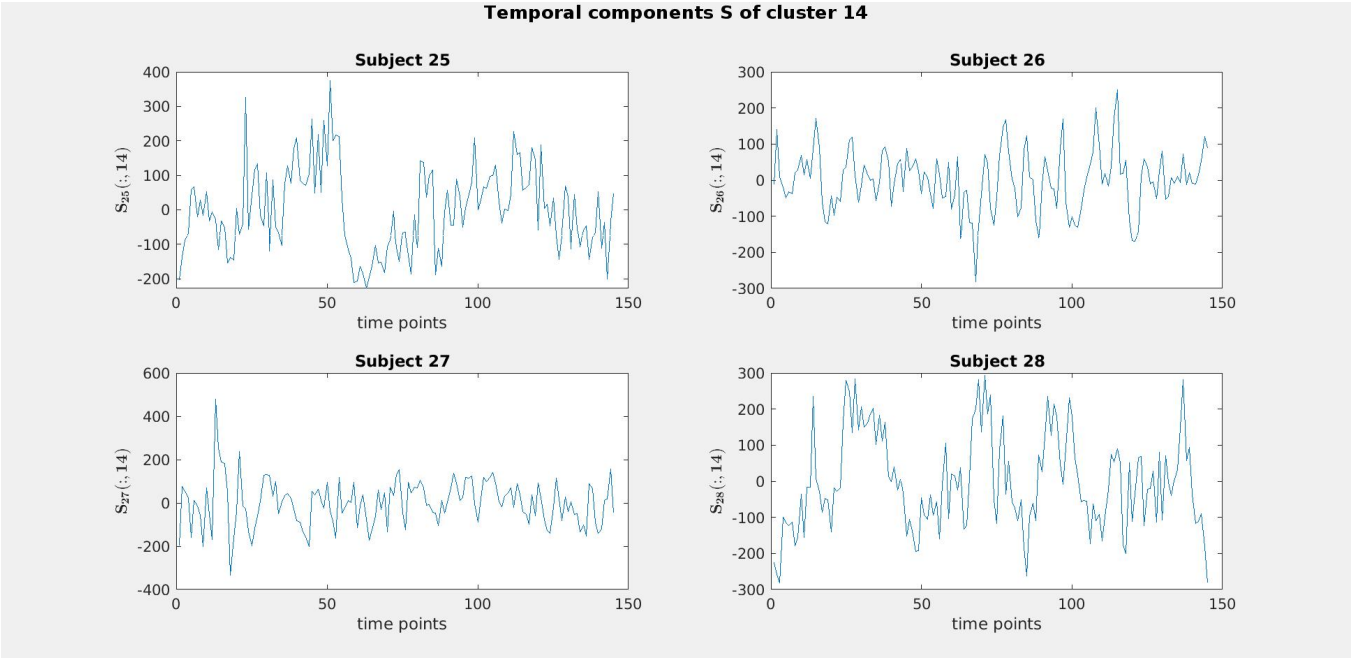


Figure B.492: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 25, \ldots, 28$ and R=42.

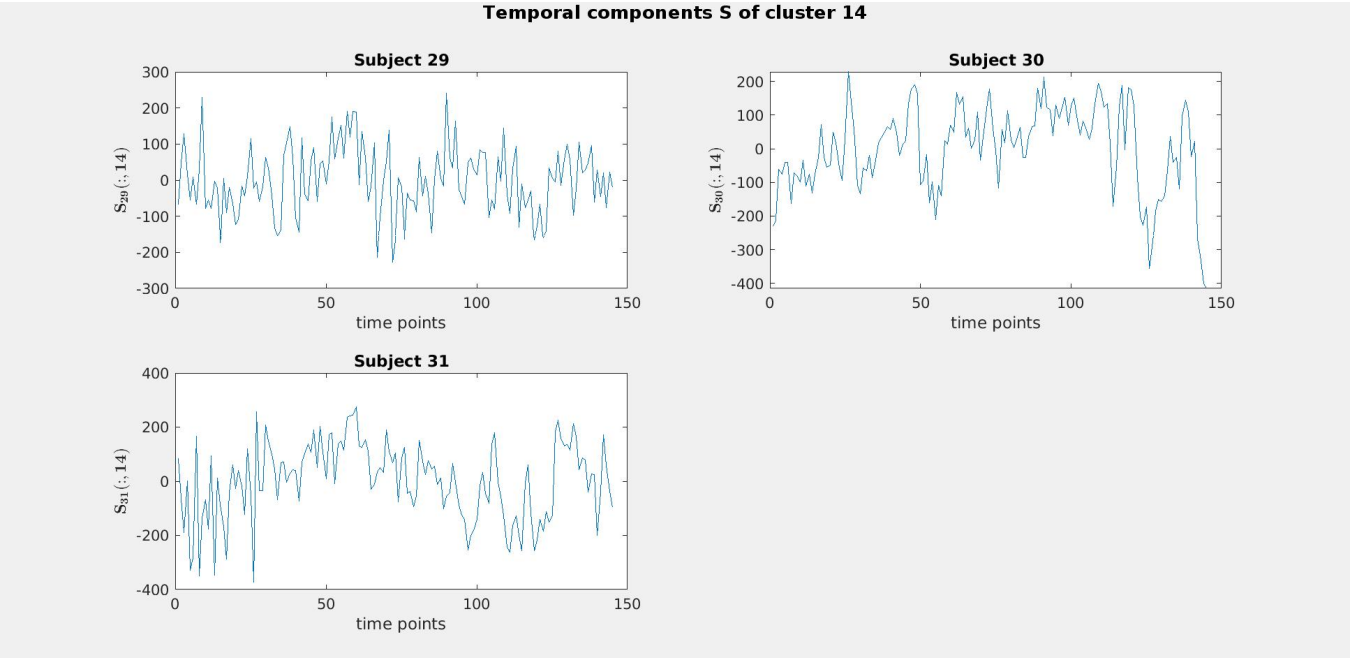


Figure B.493: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 29, \ldots, 31$ and R=42.

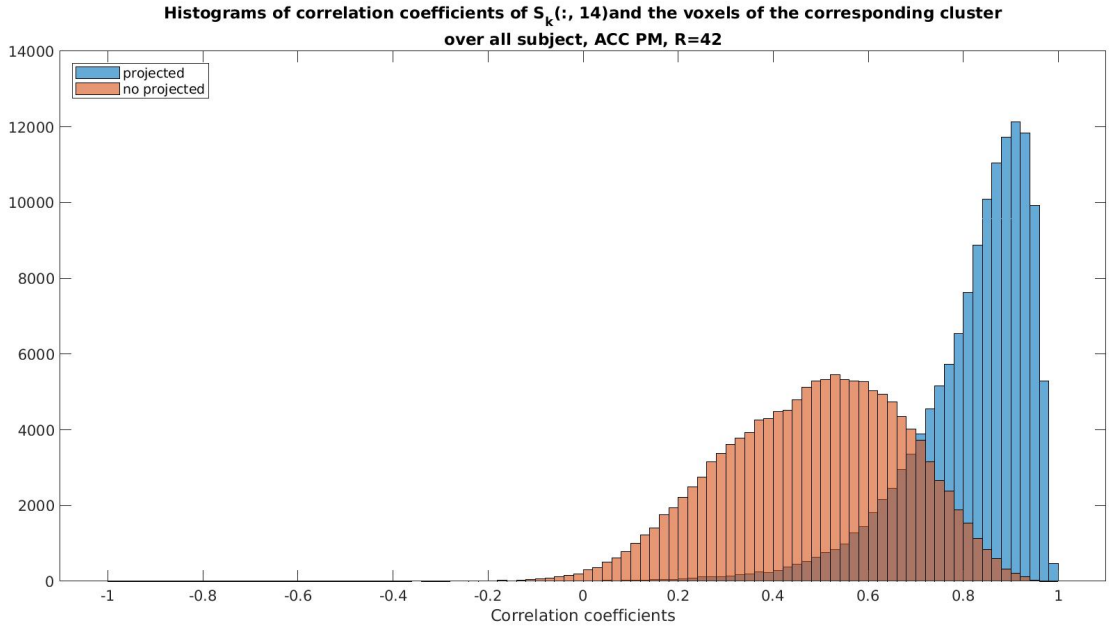


Figure B.494: For the column 14 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 14)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

21. **Cluster 2**

    Cluster 2 contains 4746 voxels, which correspond to 2.100483% of the whole brain (background, soft tissue, bone are excluded).

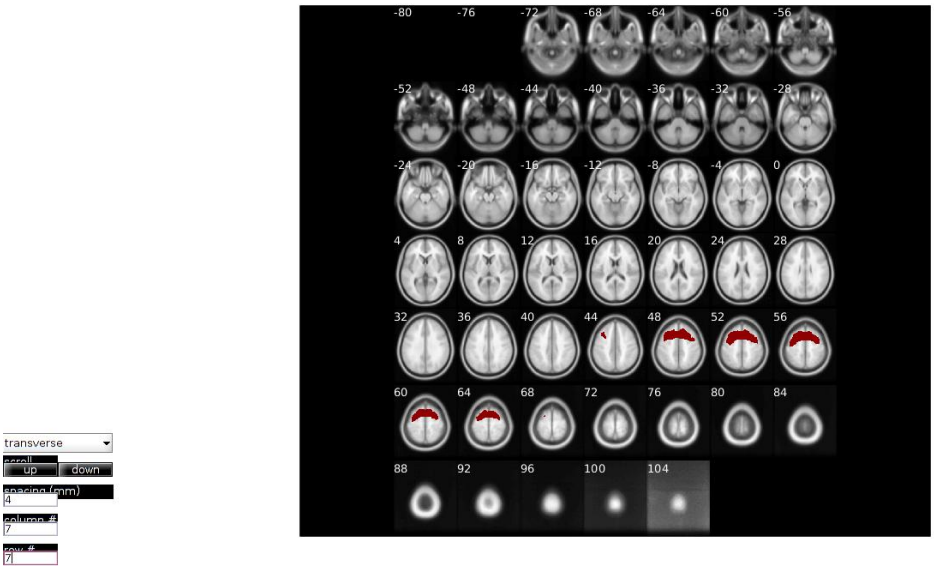


Figure B.495: Cluster 2. Red areas indicate the voxels which belong to cluster 2.

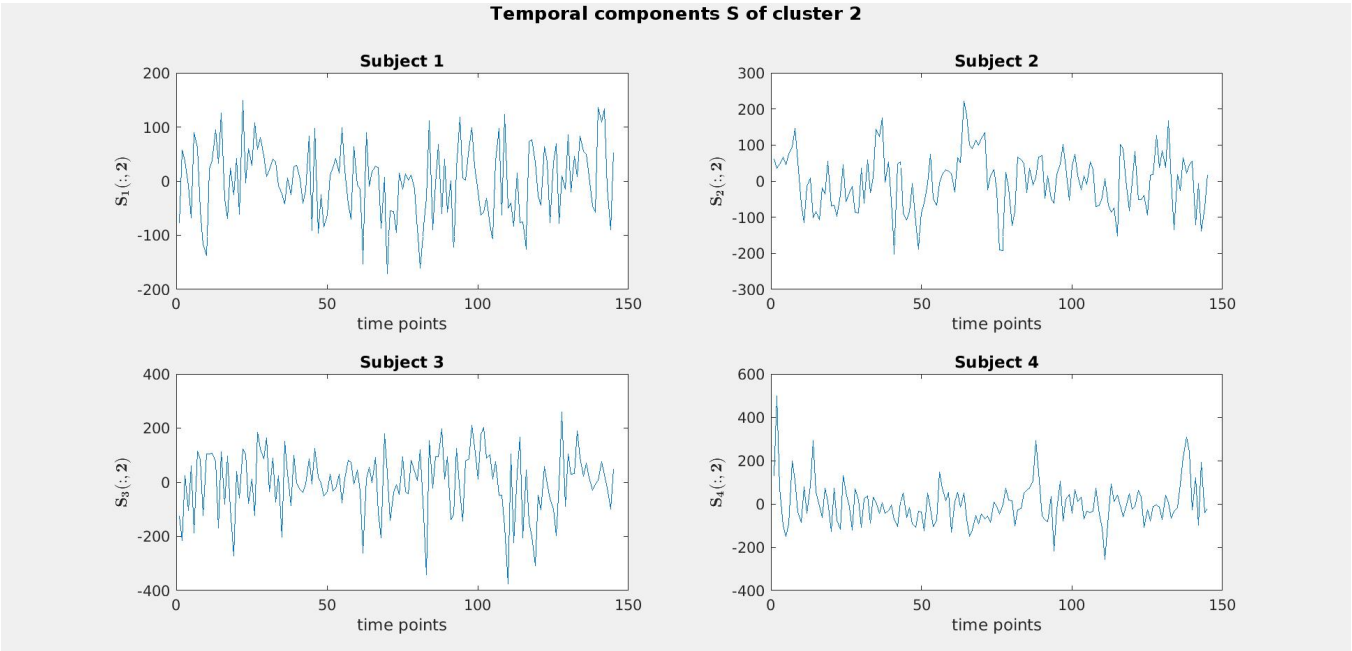


Figure B.496: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 1,\ldots,4$ and R=42.

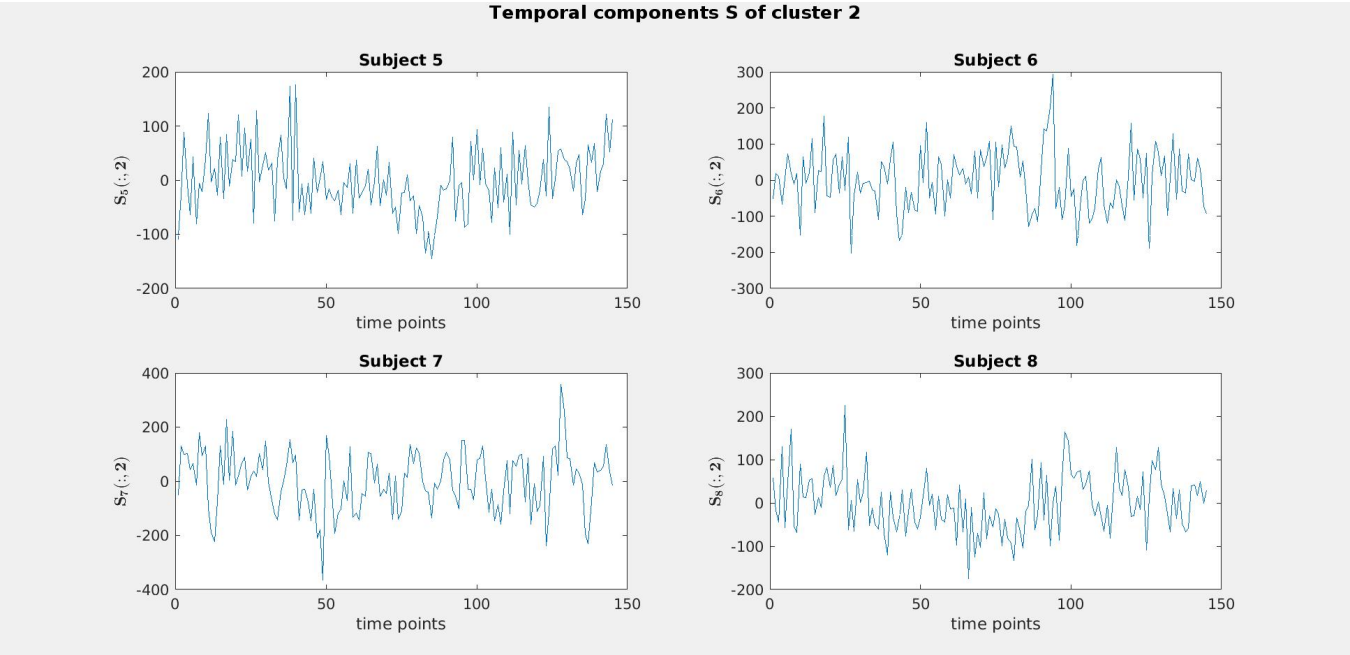


Figure B.497: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 5,\ldots,8$ and R=42.

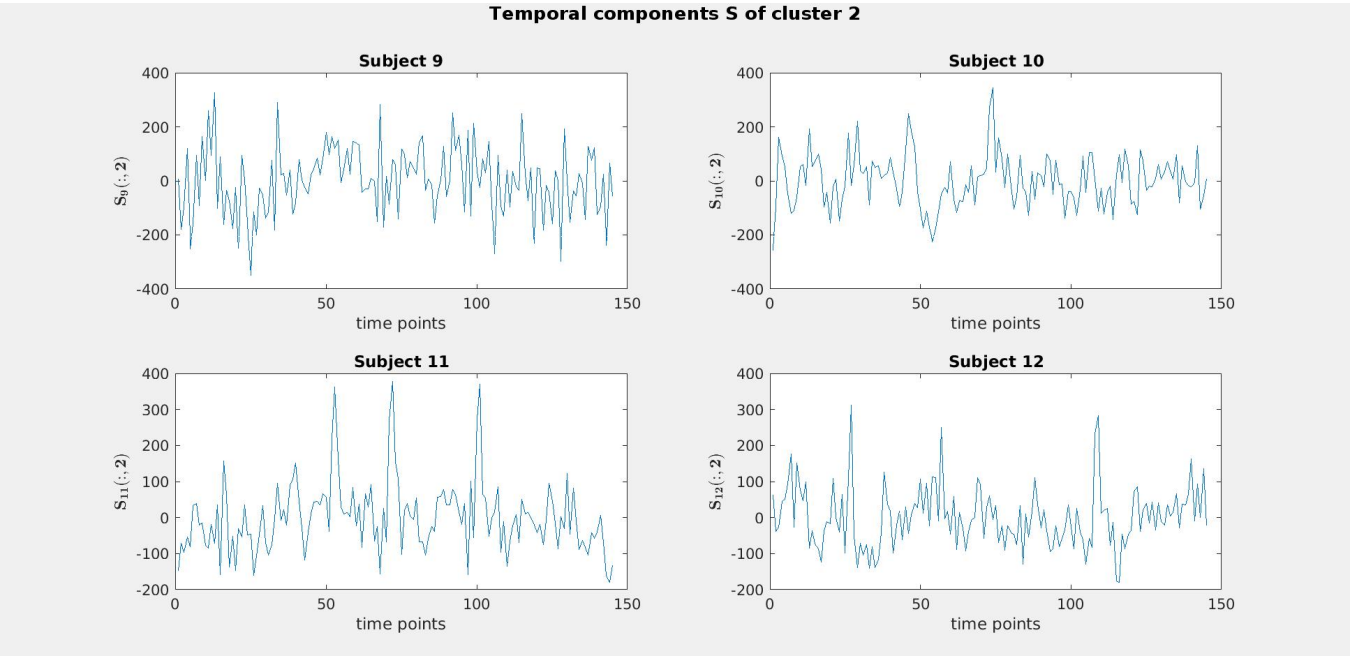


Figure B.498: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 9,\ldots,12$ and R=42.

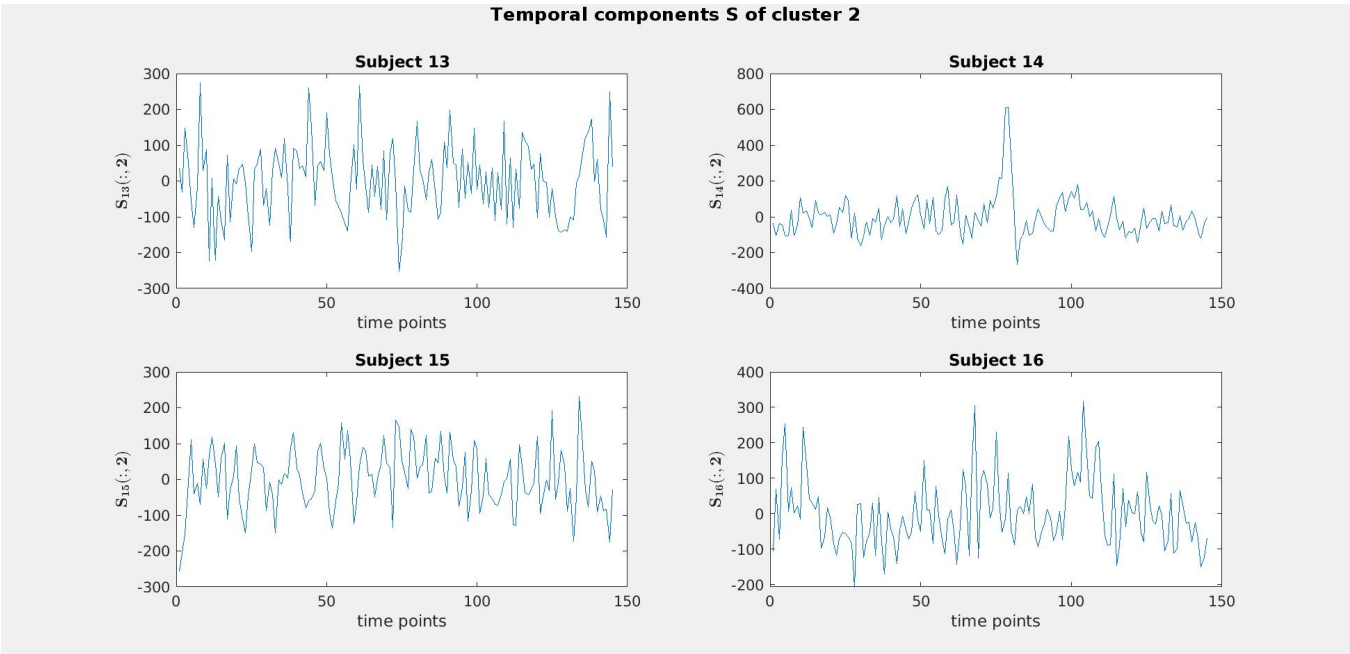


Figure B.499: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 13,\ldots,16$ and R=42.

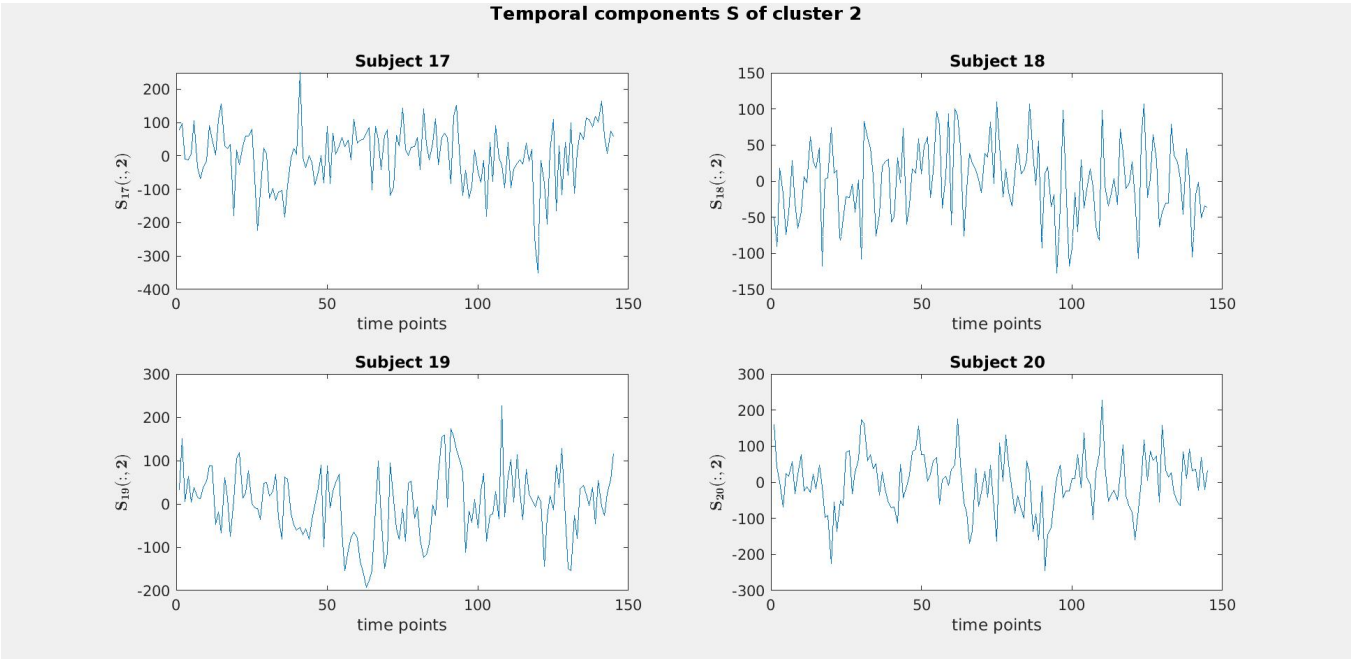


Figure B.500: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 17,\ldots,20$ and R=42.

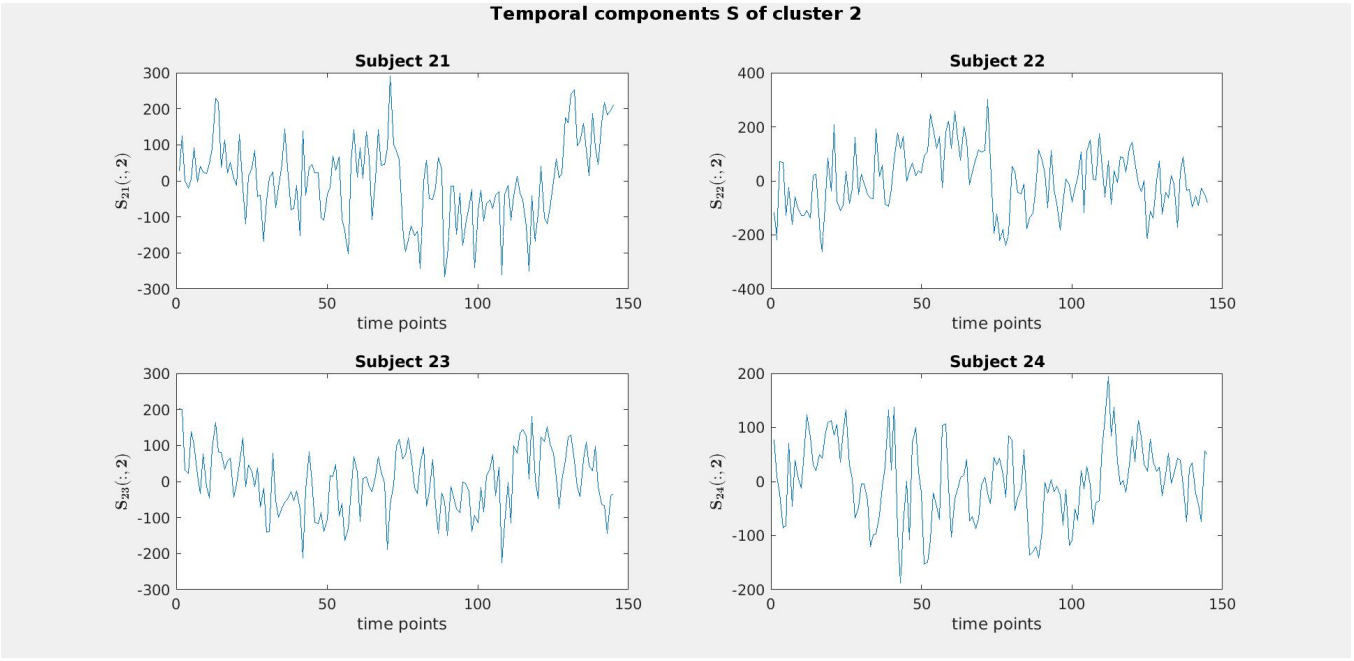


Figure B.501: Waveform of temporal components $\mathbf{S}_k(:,12)$ for $k = 21,\ldots,24$ and R=42.

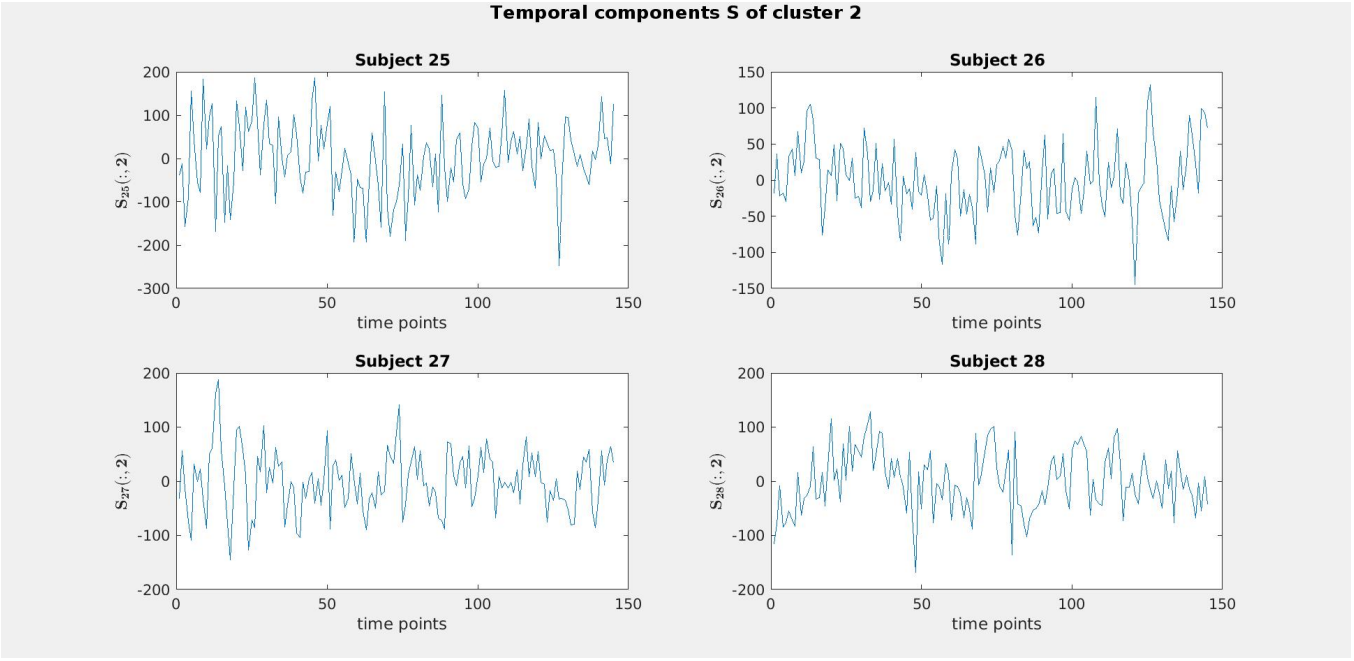


Figure B.502: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 25,\ldots,28$ and R=42.

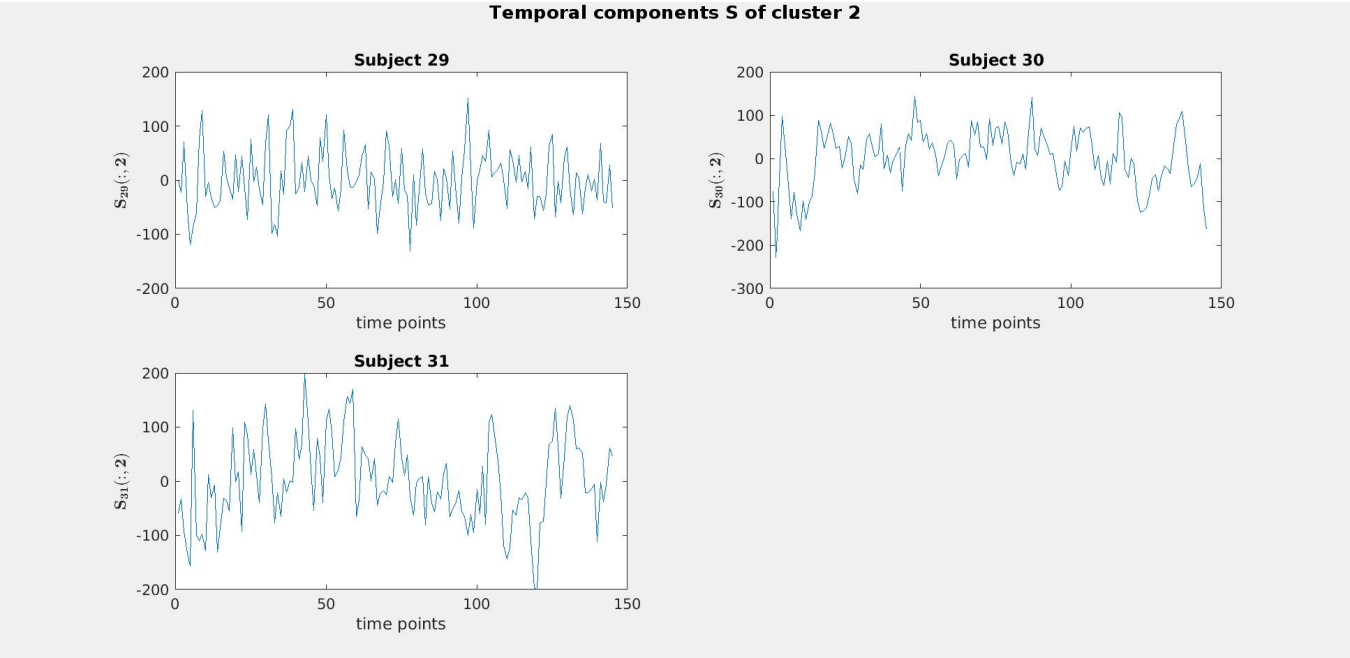


Figure B.503: Waveform of temporal components $\mathbf{S}_k(:,2)$ for $k = 29, \ldots, 31$ and R=42.

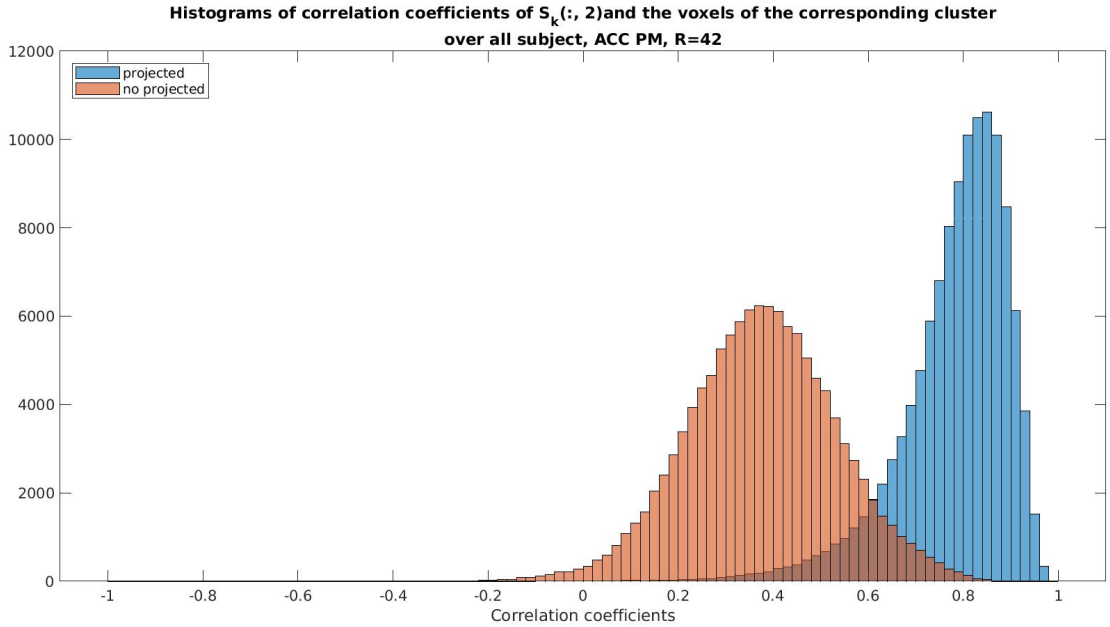


Figure B.504: For the column 2 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,2)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

22. **Cluster 40**

Cluster 40 contains 4478 voxels, which correspond to 1.981872% of the whole brain (background, soft tissue, bone are excluded).

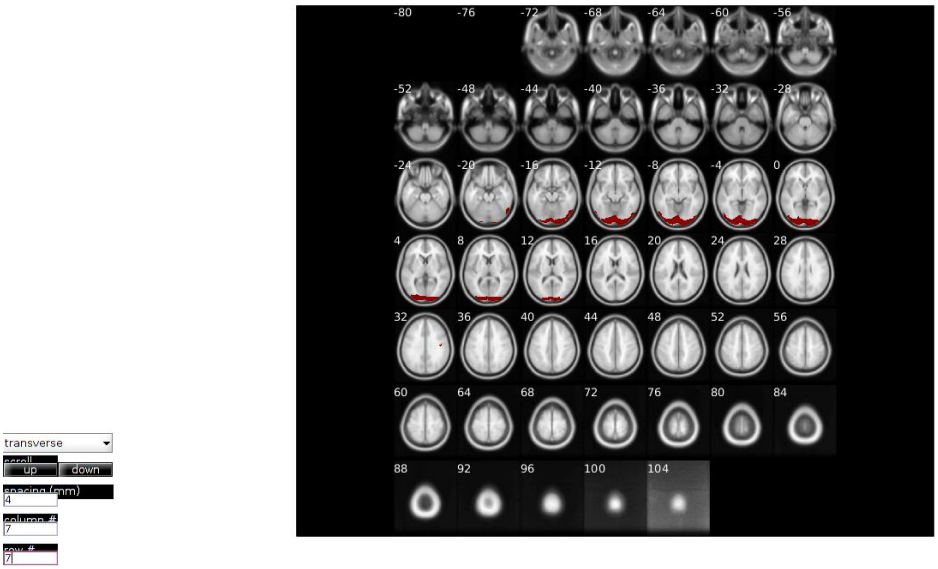


Figure B.505: Cluster 40. Red areas indicate the voxels which belong to cluster 40.

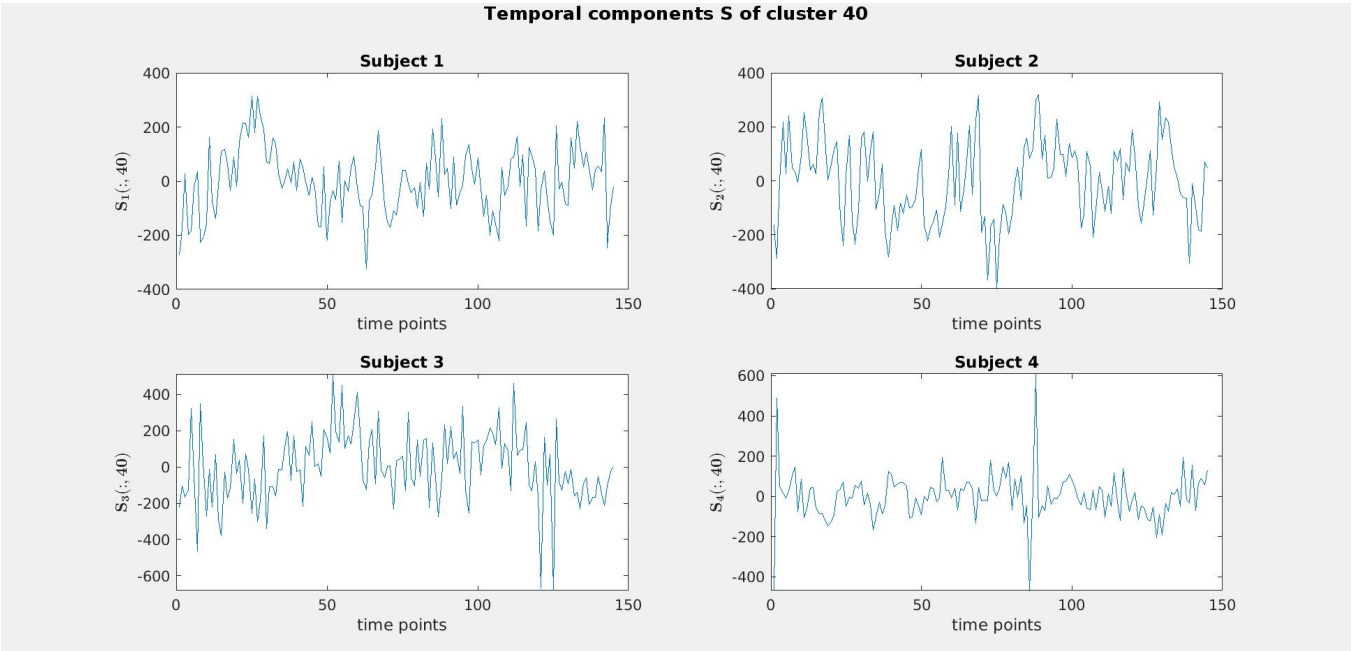


Figure B.506: Waveform of temporal components $\mathbf{S}_k(:, 40)$ for $k = 1, \ldots, 4$ and R=42.

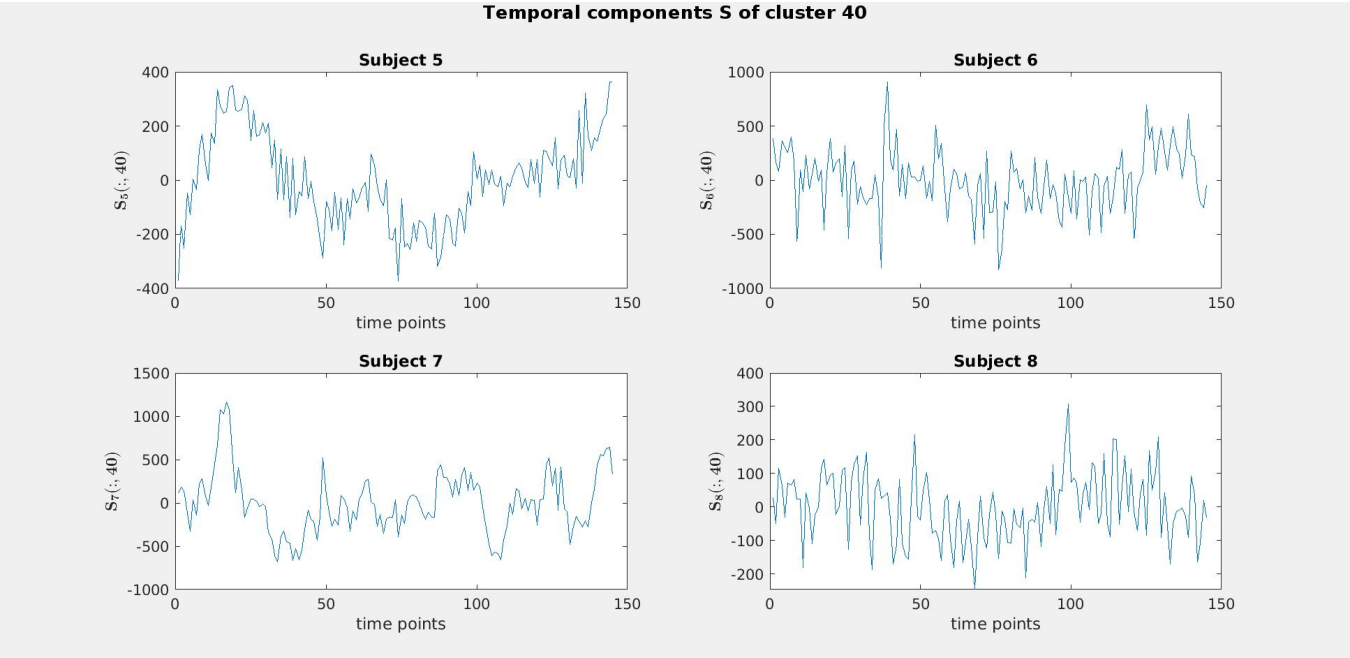


Figure B.507: Waveform of temporal components $\mathbf{S}_k(:, 40)$ for $k = 5, \ldots, 8$ and R=42.

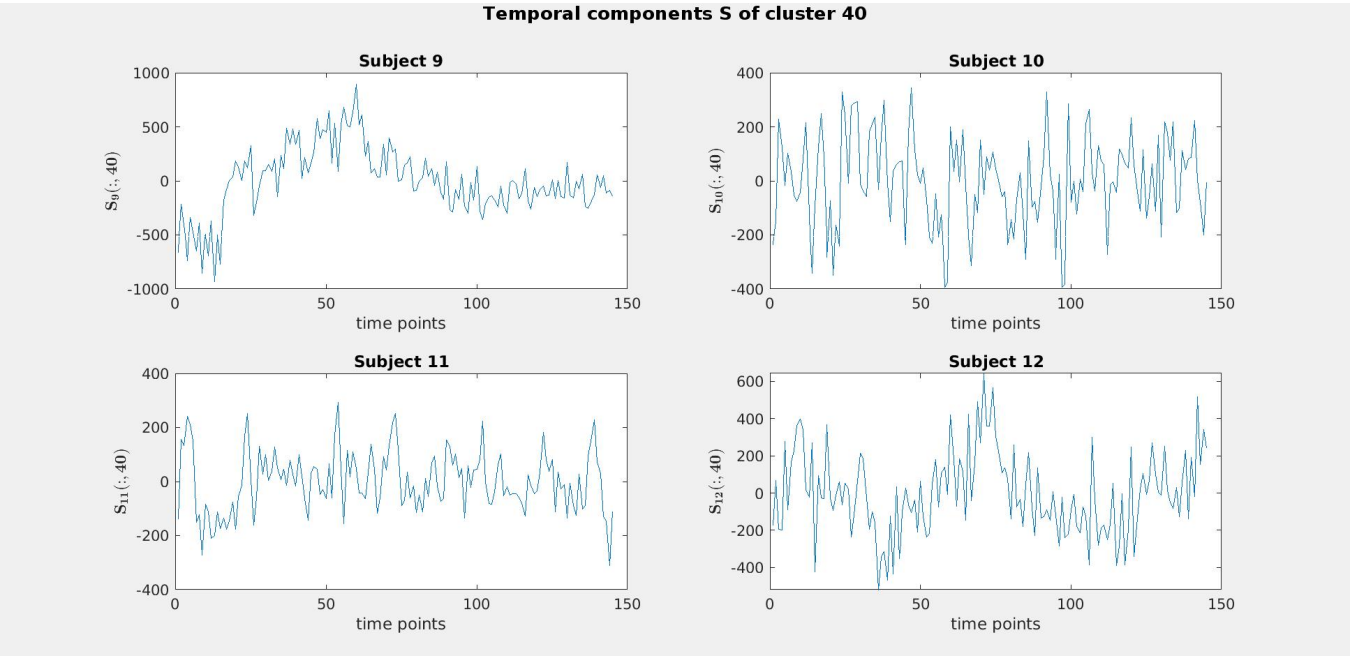


Figure B.508: Waveform of temporal components $\mathbf{S}_k(:, 40)$ for $k = 9, \ldots, 12$ and R=42.

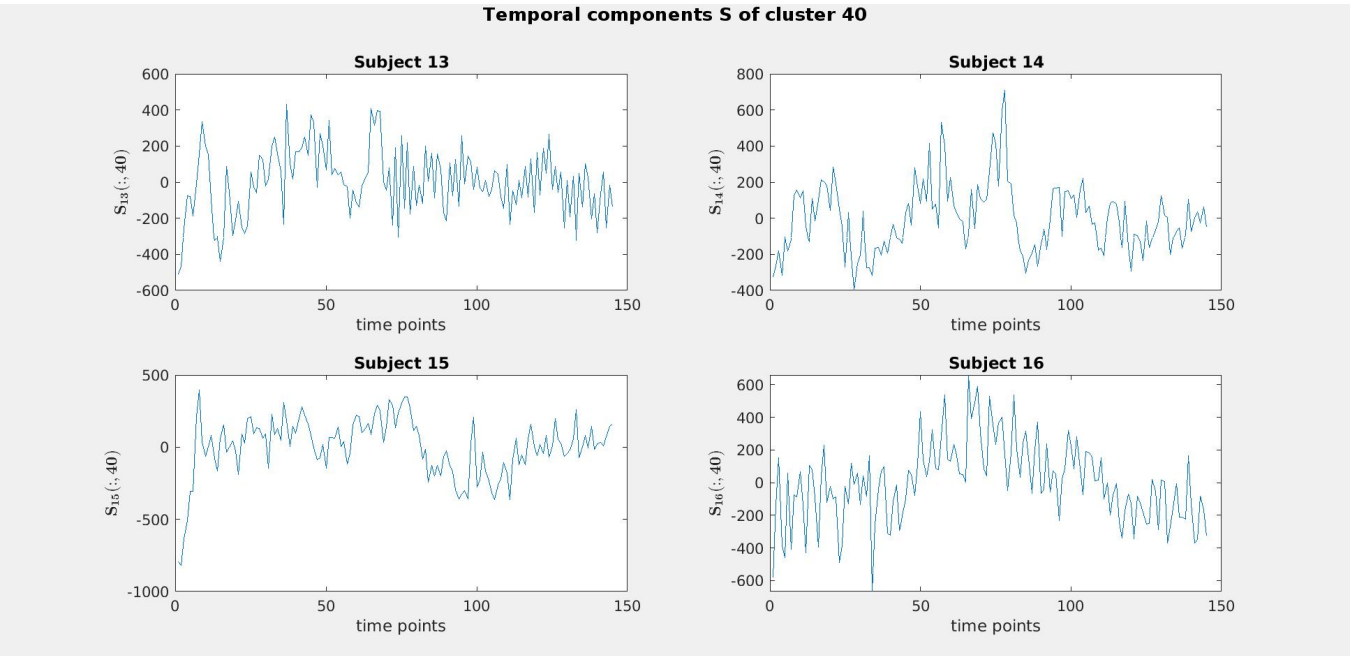


Figure B.509: Waveform of temporal components $\mathbf{S}_k(:, 40)$ for $k = 13, \ldots, 16$ and R=42.

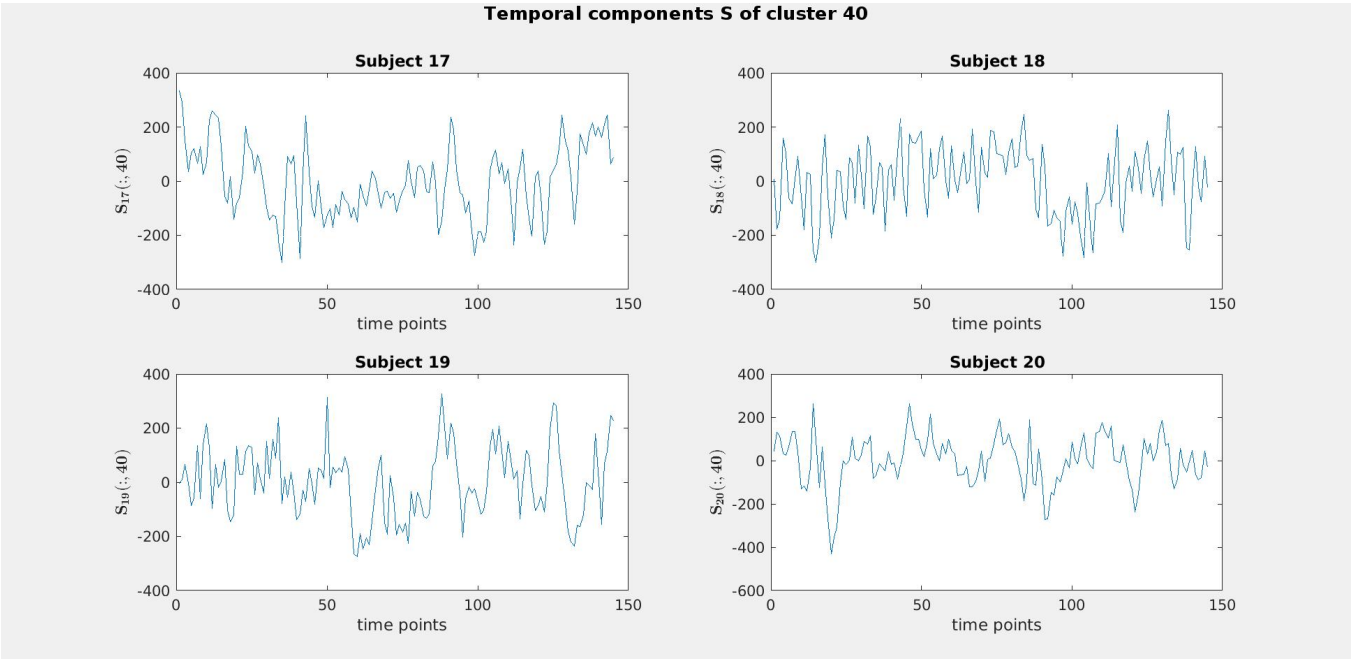


Figure B.510: Waveform of temporal components $\mathbf{S}_k(:, 40)$ for $k = 17, \ldots, 20$ and R=42.

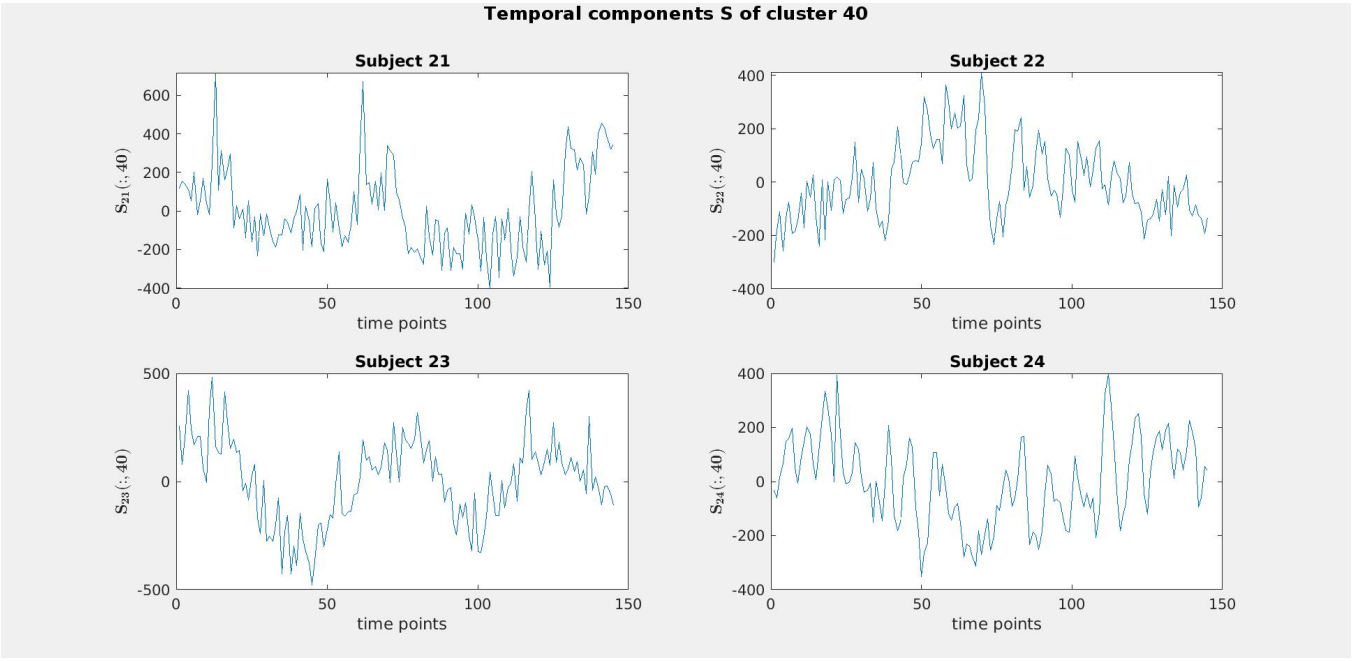


Figure B.511: Waveform of temporal components $\mathbf{S}_k(:, 140)$ for $k = 21, \ldots, 24$ and R=42.

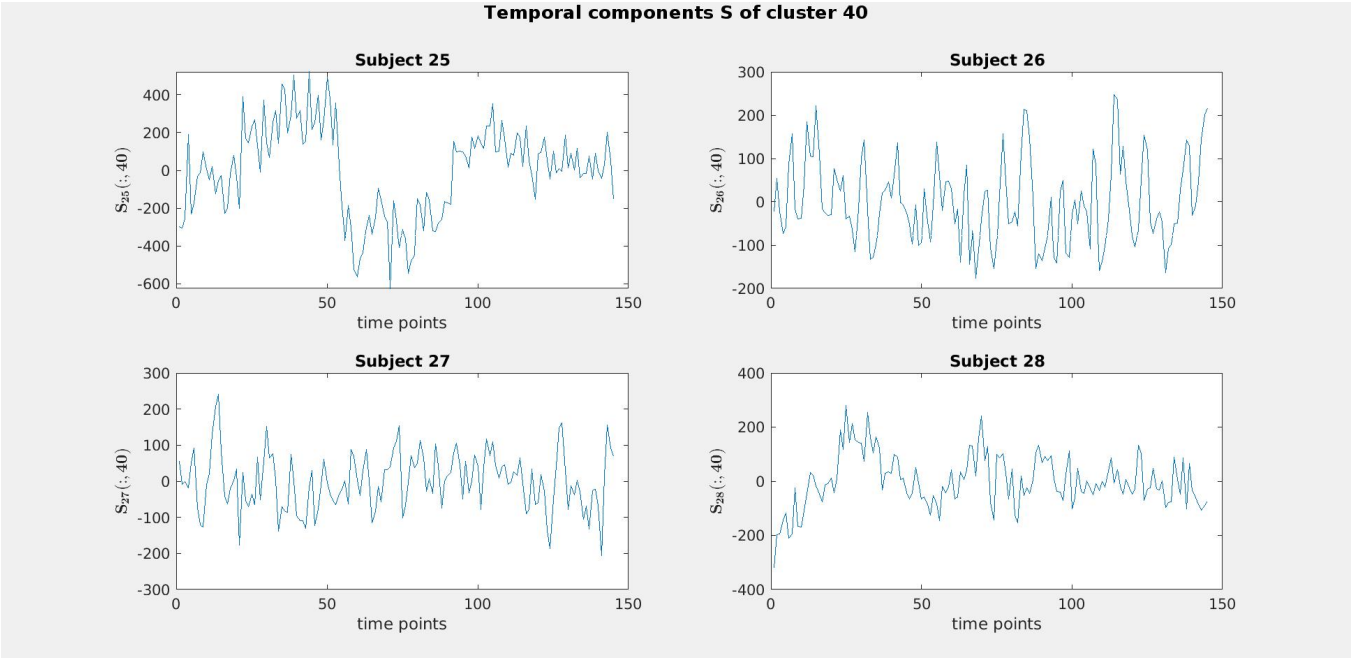


Figure B.512: Waveform of temporal components $\mathbf{S}_k(:, 40)$ for $k = 25, \ldots, 28$ and R=42.

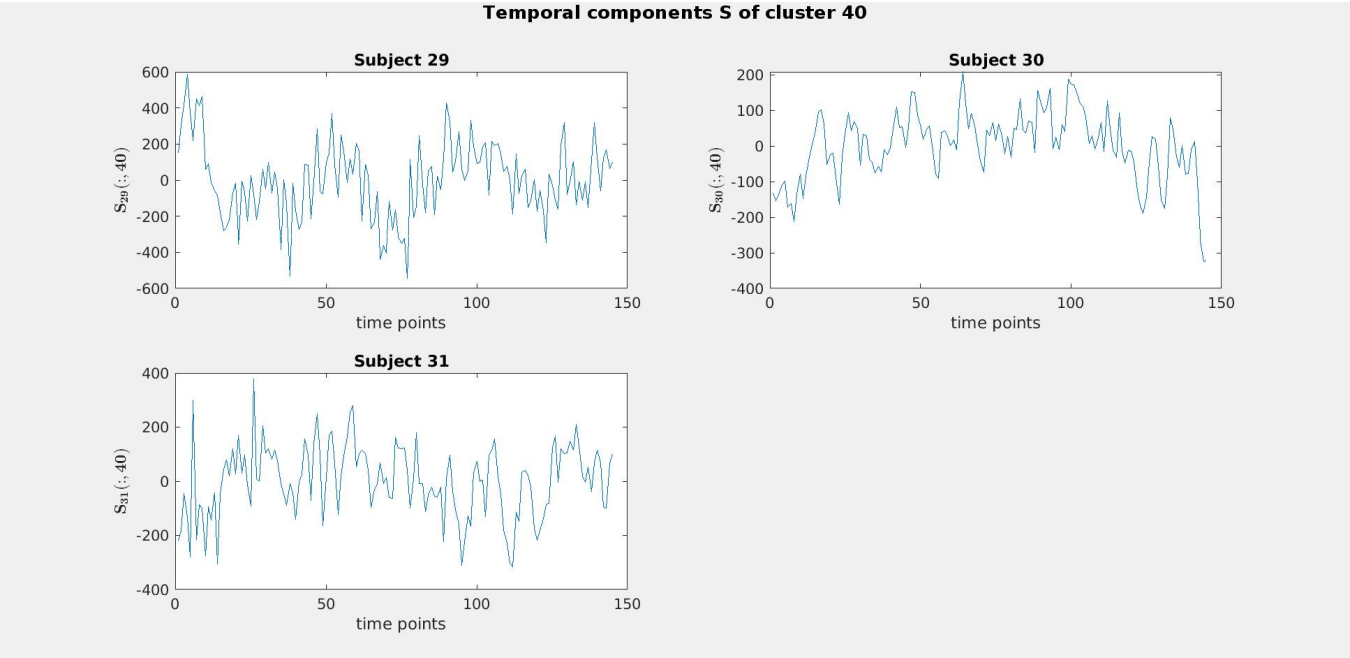


Figure B.513: Waveform of temporal components $\mathbf{S}_k(:, 40)$ for $k = 29, \ldots, 31$ and R=42.

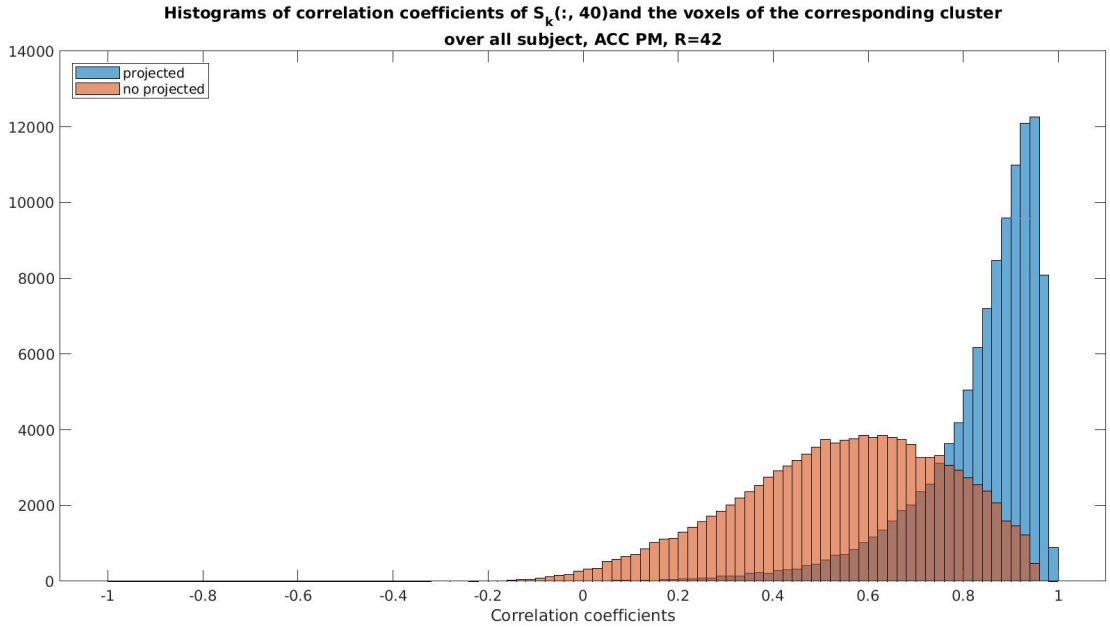


Figure B.514: For the column 40 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 40)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

23. **Cluster 21**

Cluster 21 contains 4357 voxels, which correspond to 1.928320% of the whole brain (background, soft tissue, bone are excluded).

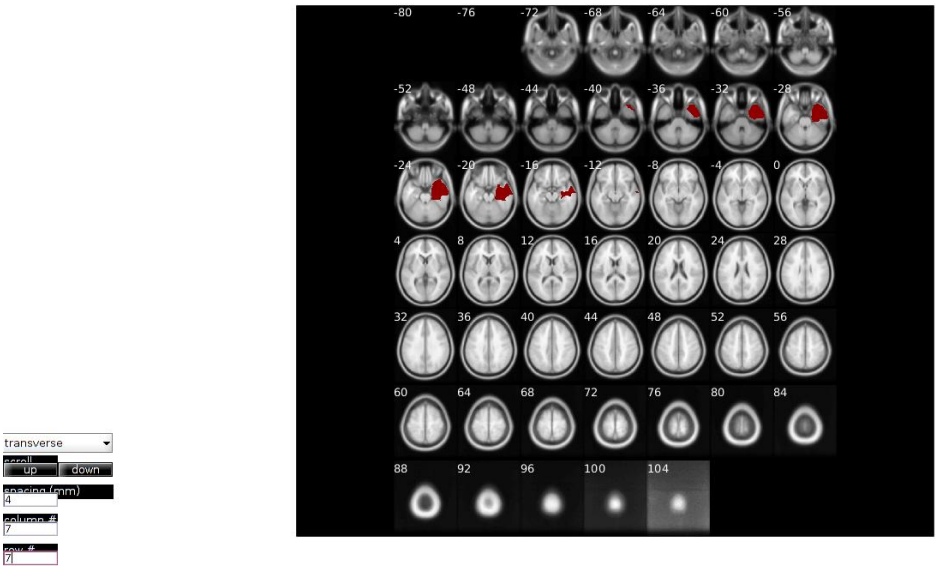


Figure B.515: Cluster 21. Red areas indicate the voxels which belong to cluster 21.

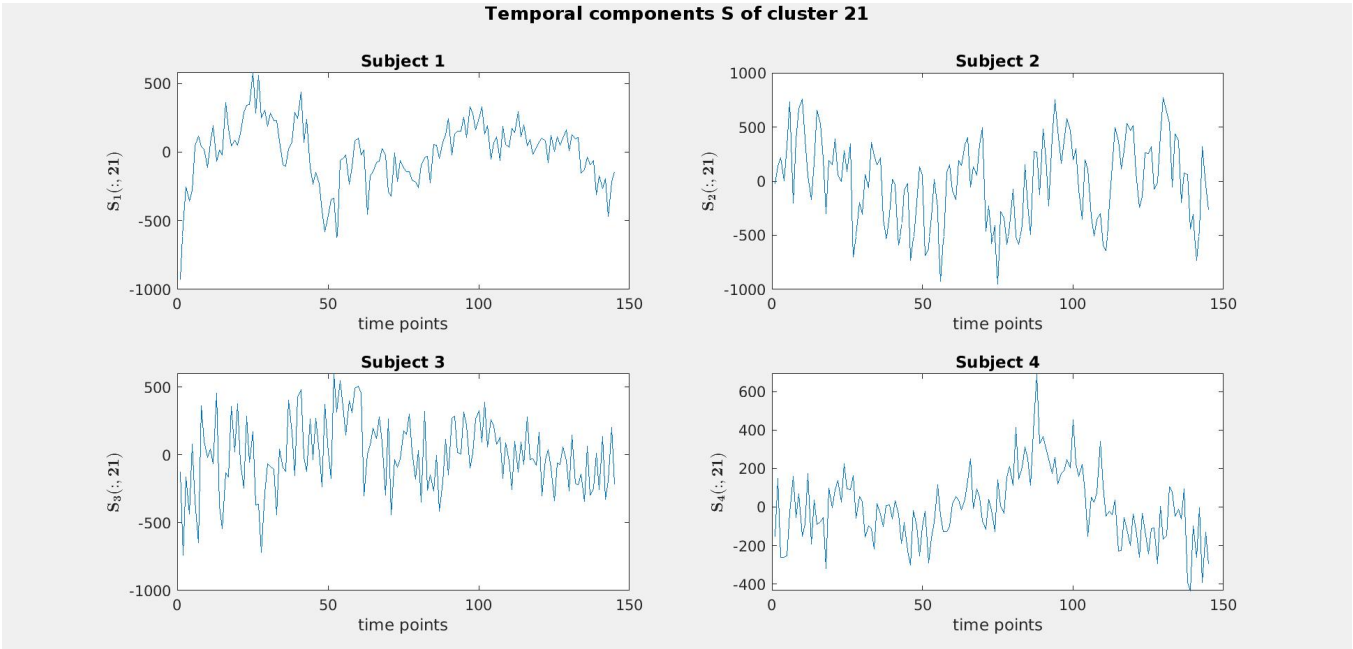


Figure B.516: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 1, \ldots, 4$ and R=42.

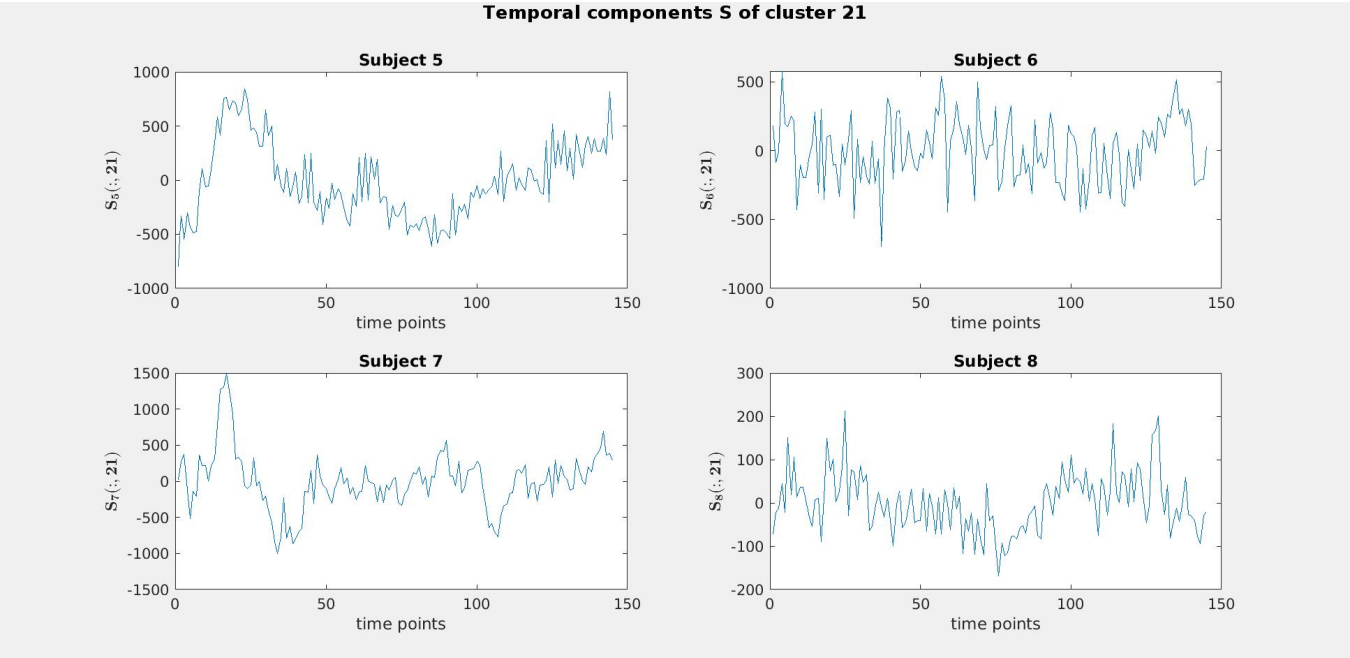


Figure B.517: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 5, \ldots, 8$ and R=42.

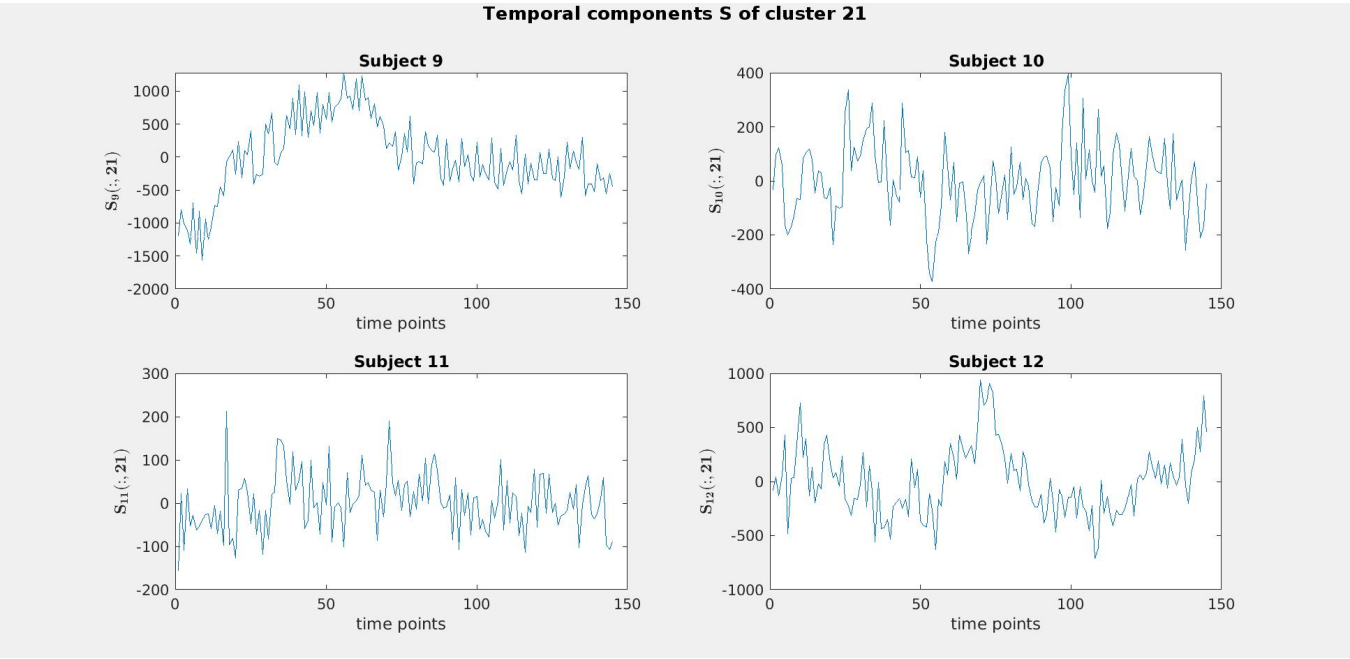


Figure B.518: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 9, \ldots, 12$ and R=42.

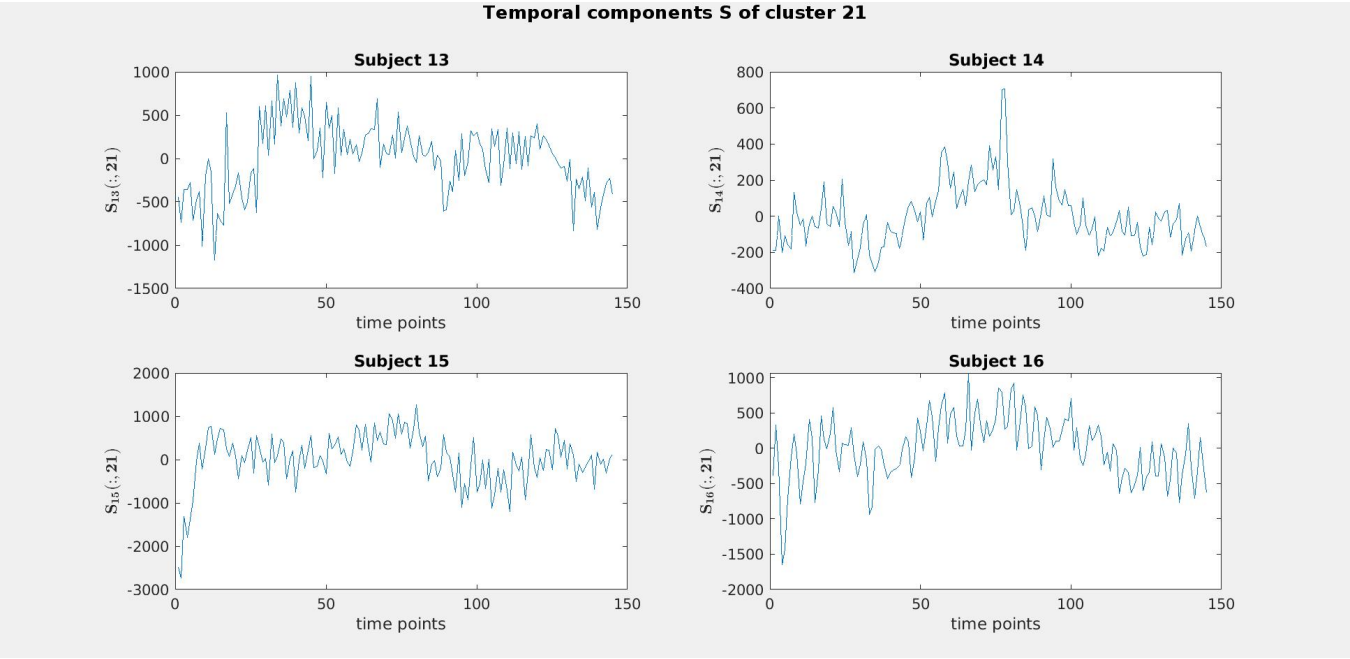


Figure B.519: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 13, \ldots, 16$ and R=42.

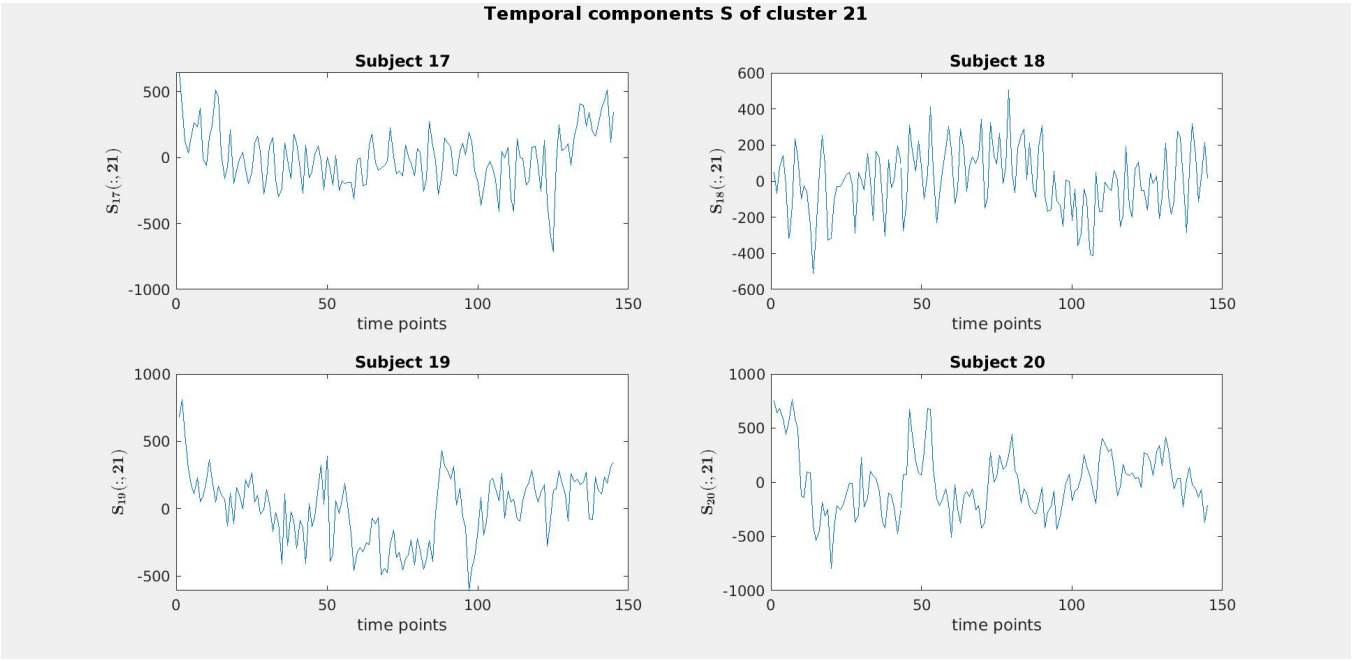


Figure B.520: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 17, \ldots, 20$ and R=42.

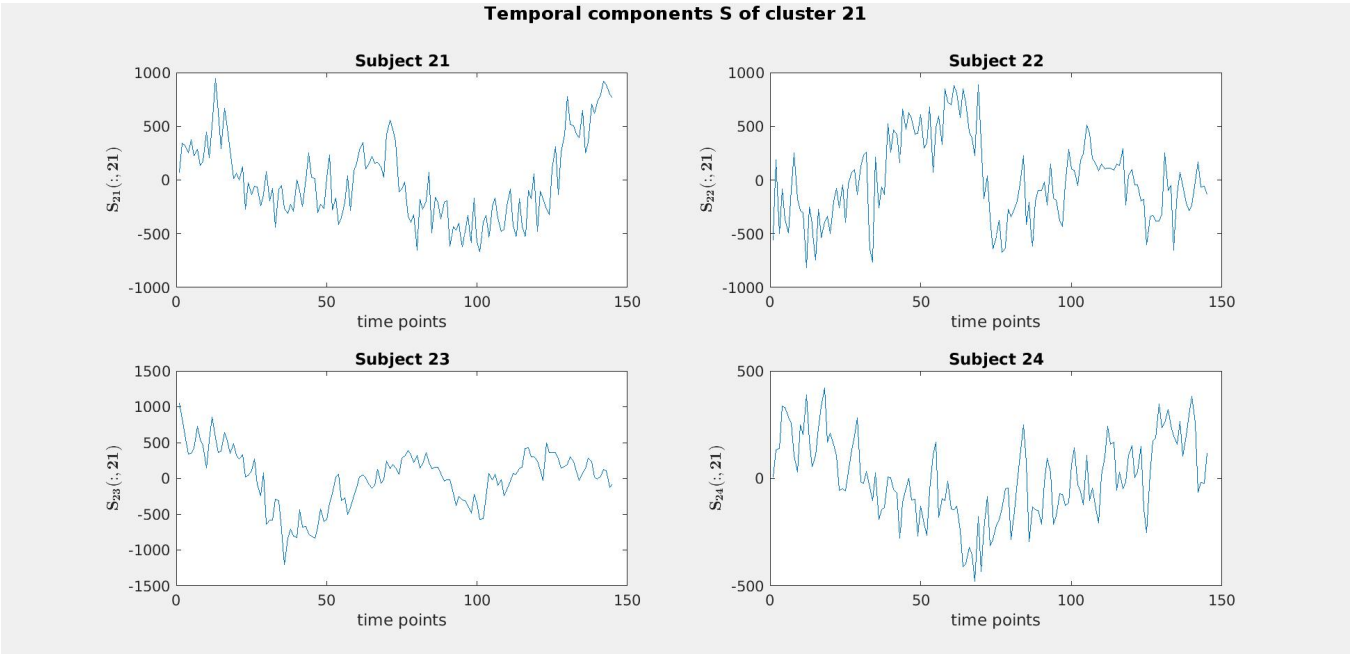


Figure B.521: Waveform of temporal components $\mathbf{S}_k(:, 121)$ for $k = 21, \ldots, 24$ and R=42.

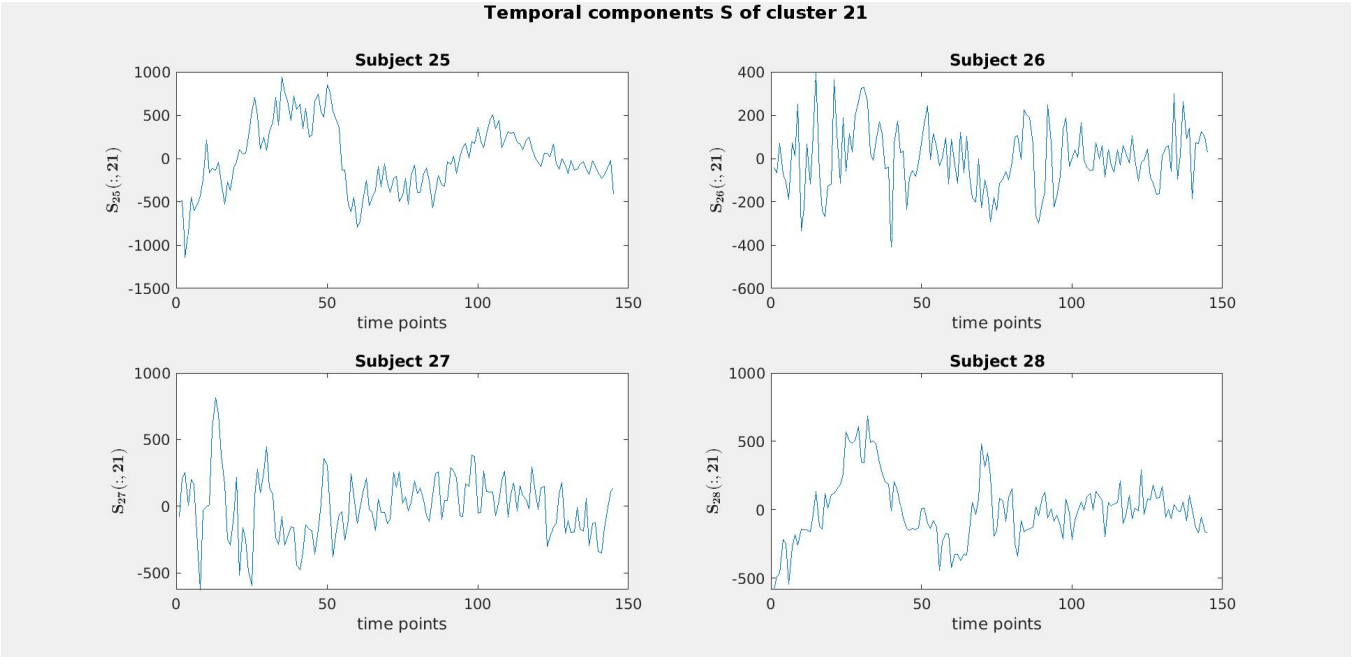


Figure B.522: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 25, \ldots, 28$ and R=42.

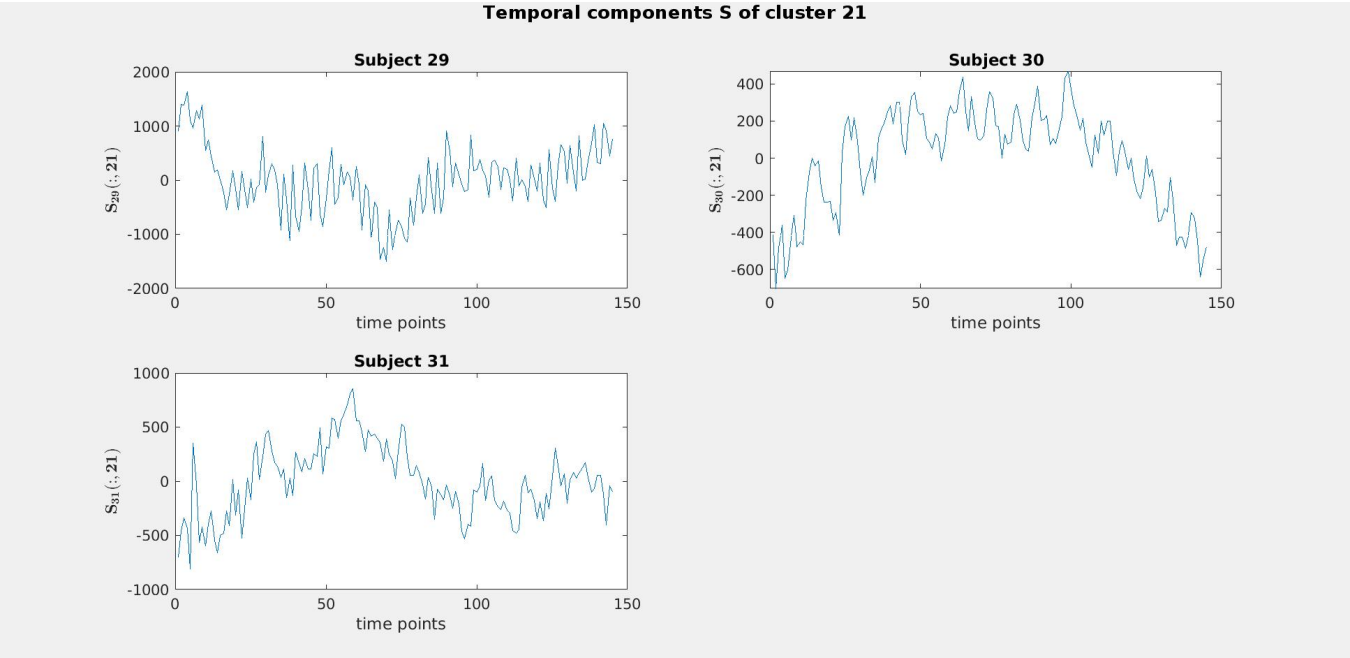


Figure B.523: Waveform of temporal components $\mathbf{S}_k(:, 21)$ for $k = 29, \ldots, 31$ and R=42.

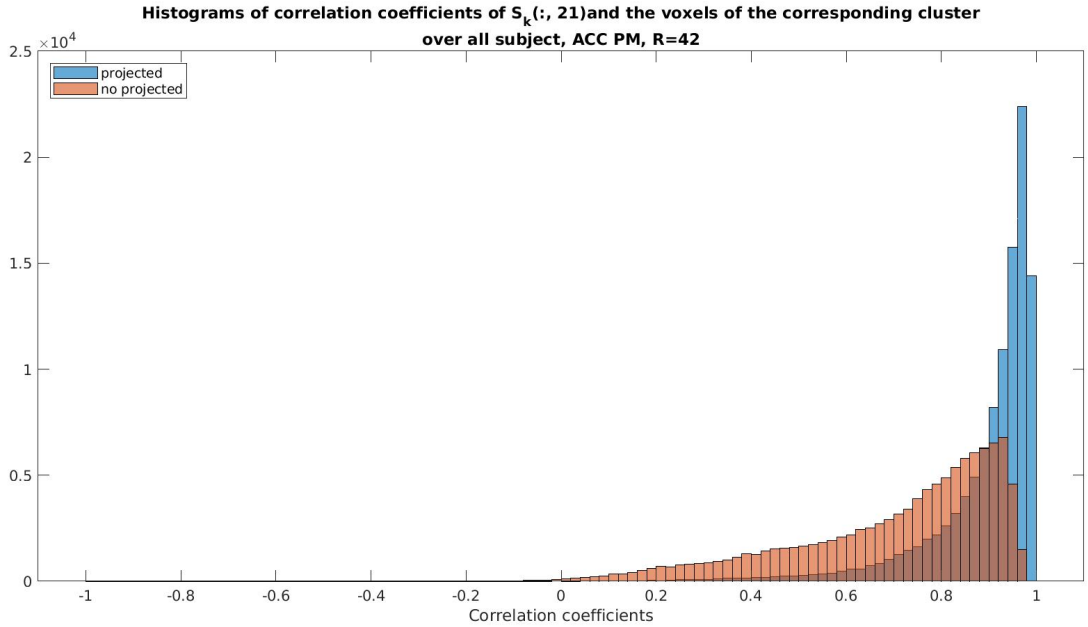


Figure B.524: For the column 21 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 21)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

24. **Cluster 18**

Cluster 18 contains 4009 voxels, which correspond to 1.774302% of the whole brain (background, soft tissue, bone are excluded).

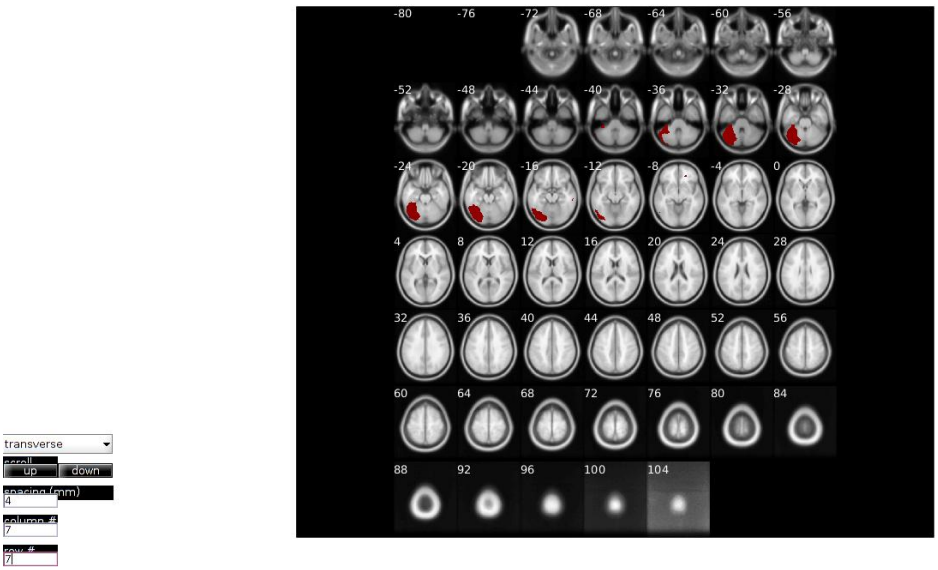


Figure B.525: Cluster 18. Red areas indicate the voxels which belong to cluster 18.

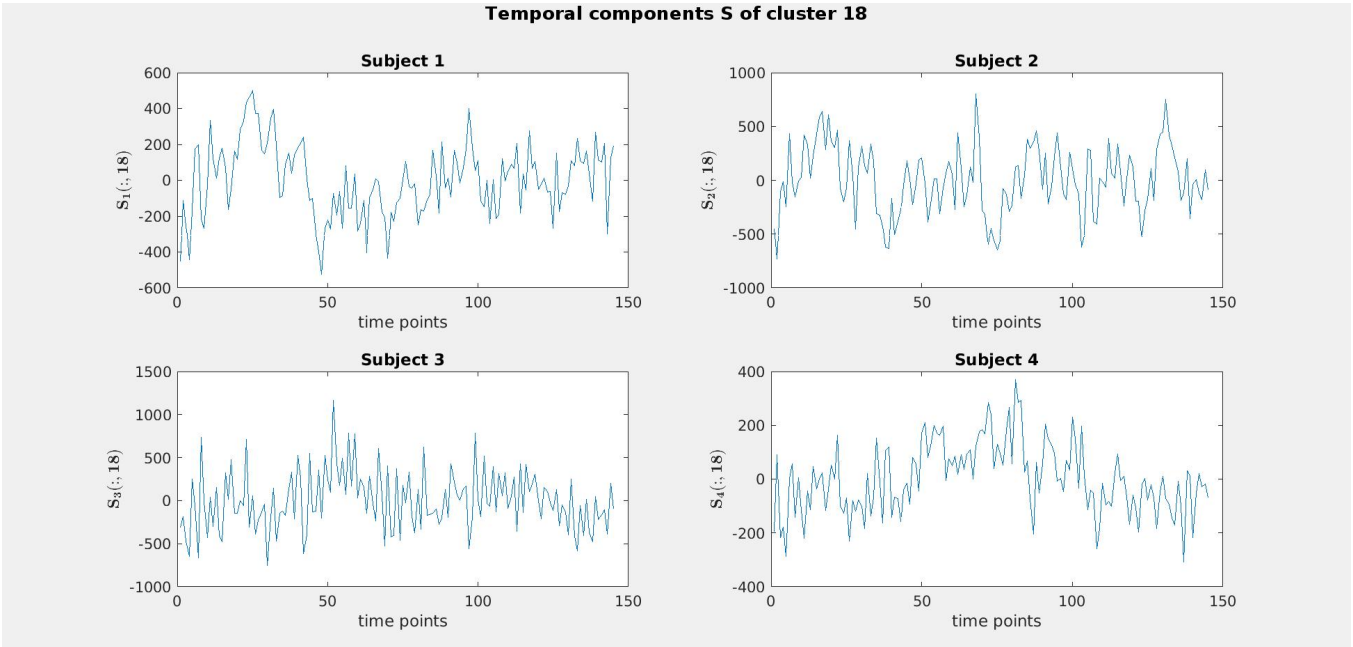


Figure B.526: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 1, \ldots, 4$ and R=42.

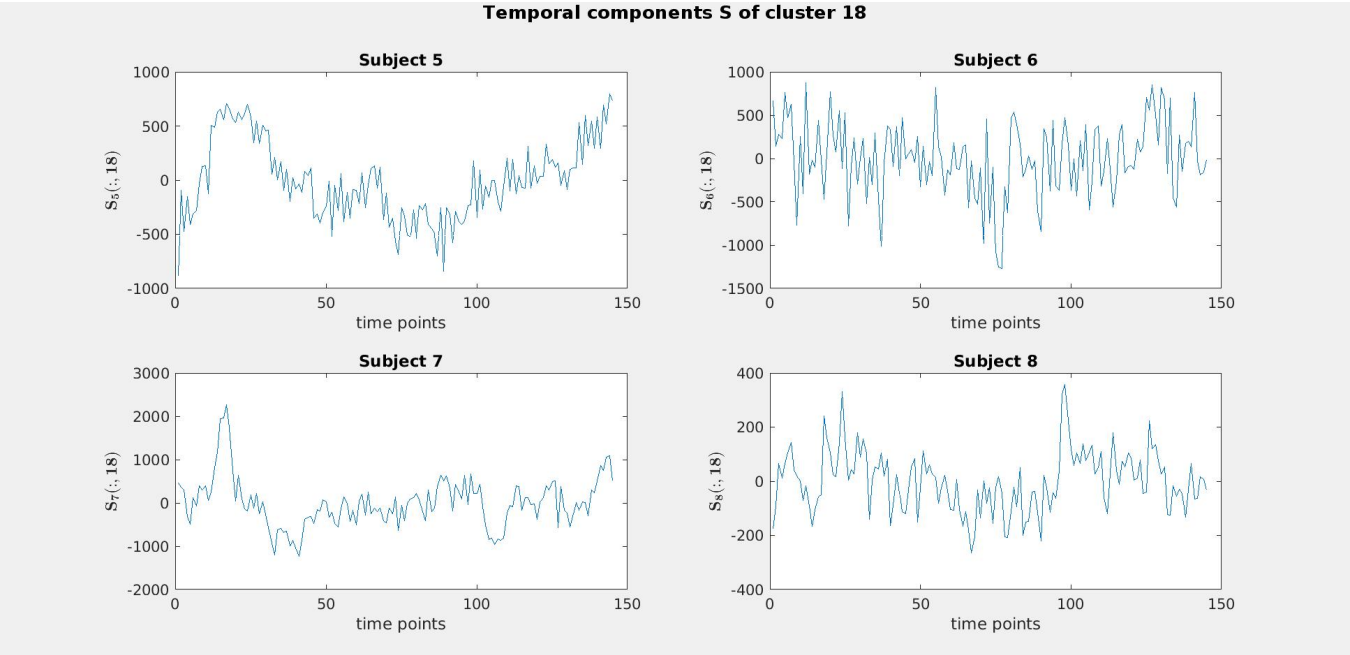


Figure B.527: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 5, \ldots, 8$ and R=42.

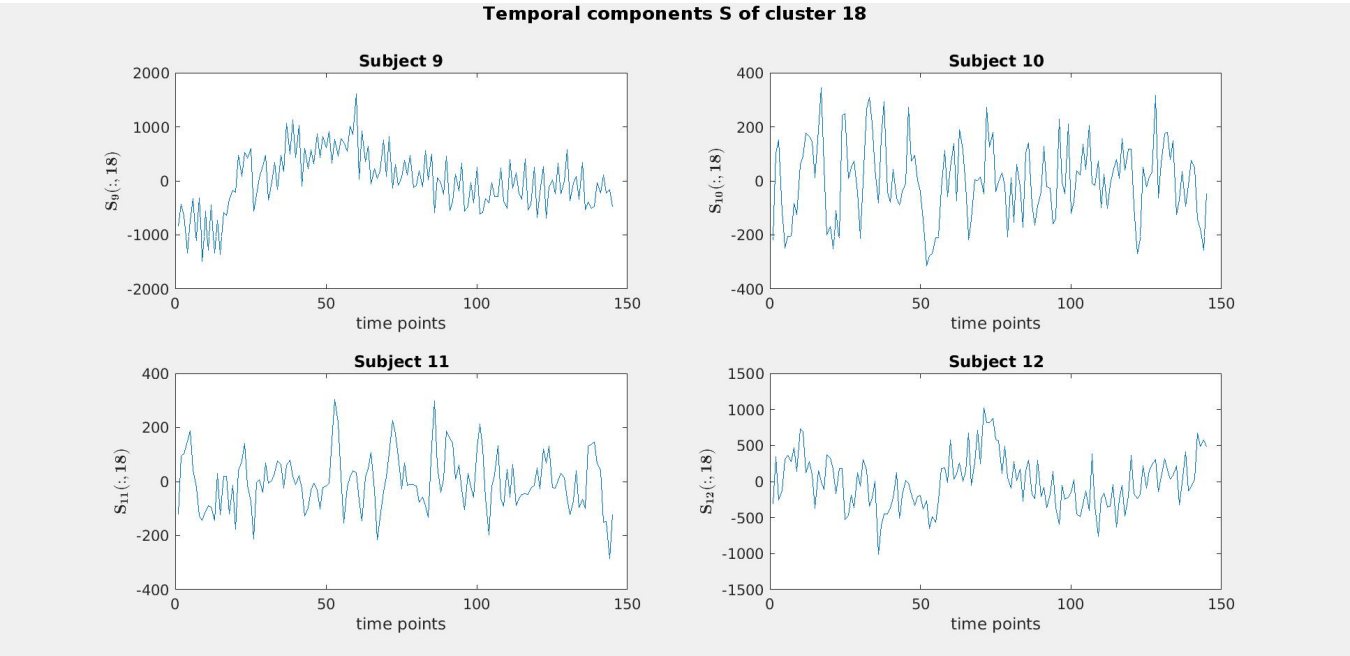


Figure B.528: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 9, \ldots, 12$ and R=42.

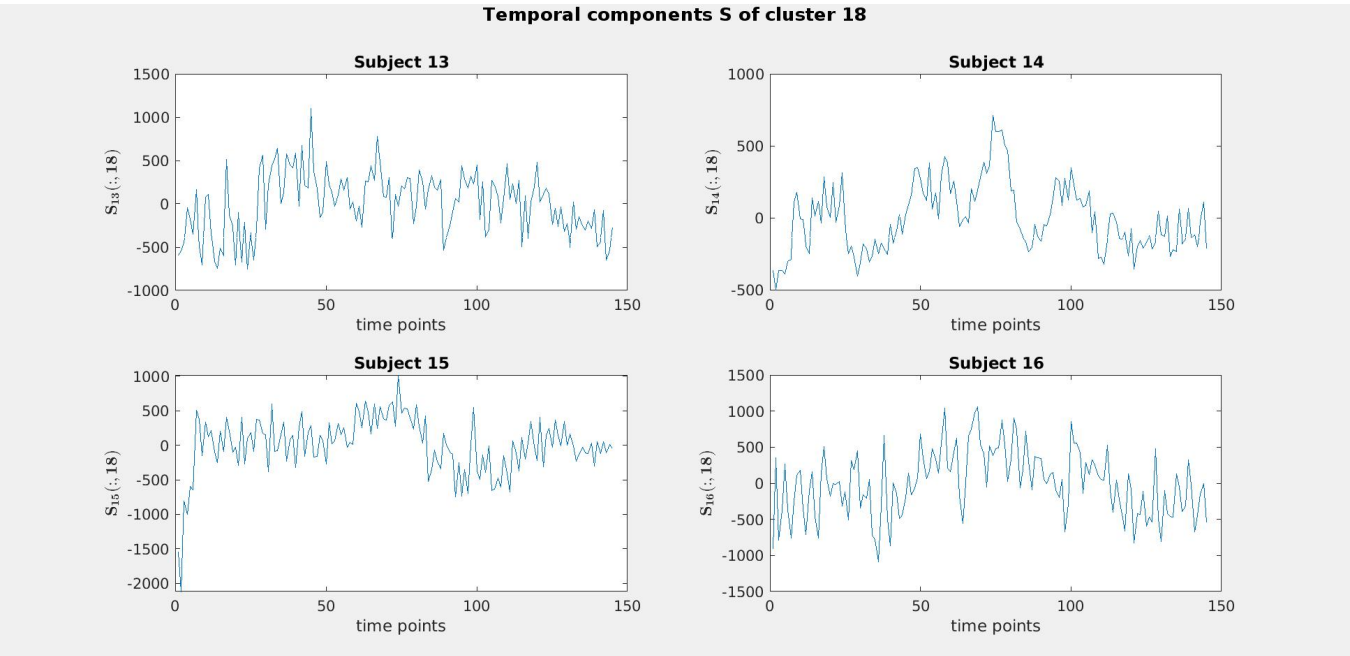


Figure B.529: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 13, \ldots, 16$ and R=42.

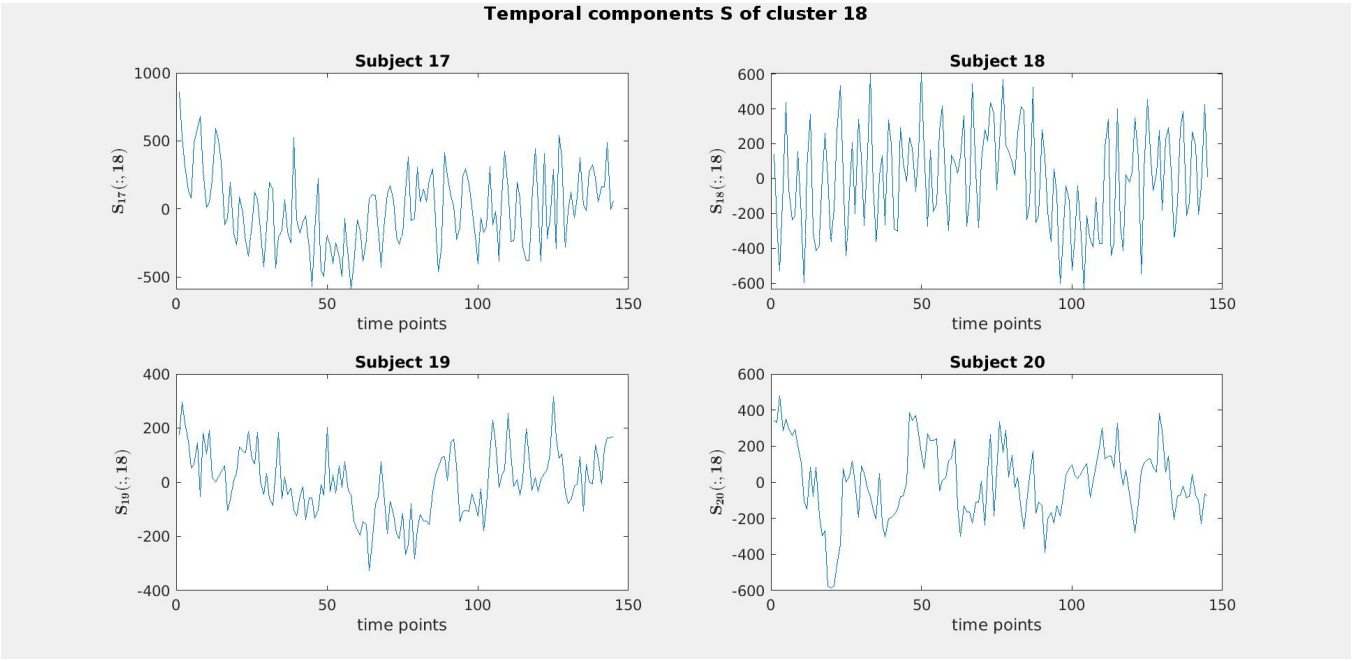


Figure B.530: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 17, \ldots, 20$ and R=42.

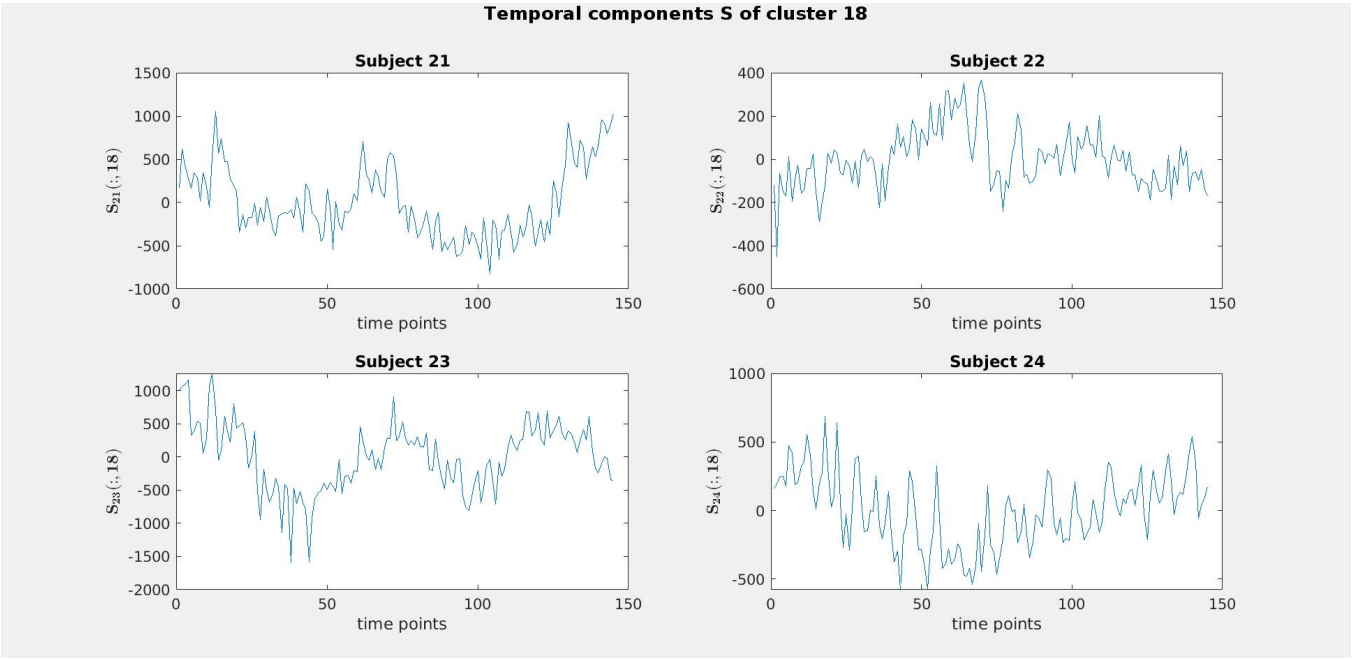


Figure B.531: Waveform of temporal components $\mathbf{S}_k(:, 118)$ for $k = 21, \ldots, 24$ and R=42.

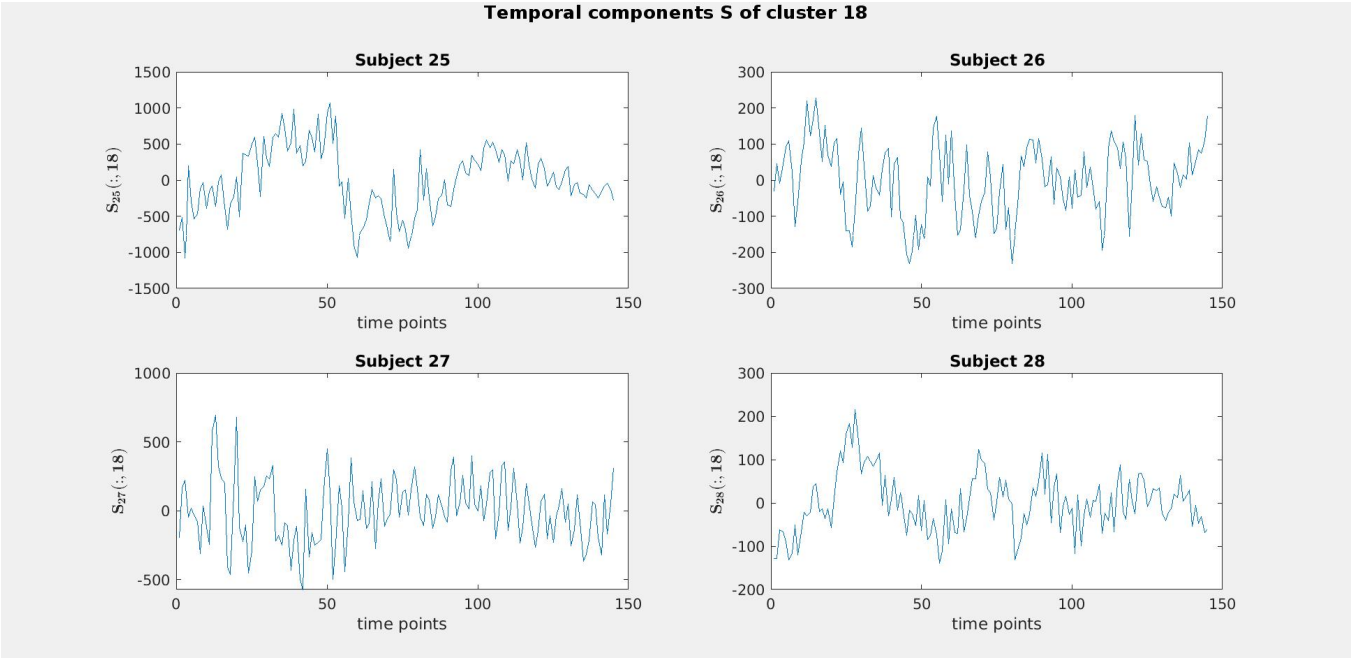


Figure B.532: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 25, \ldots, 28$ and R=42.

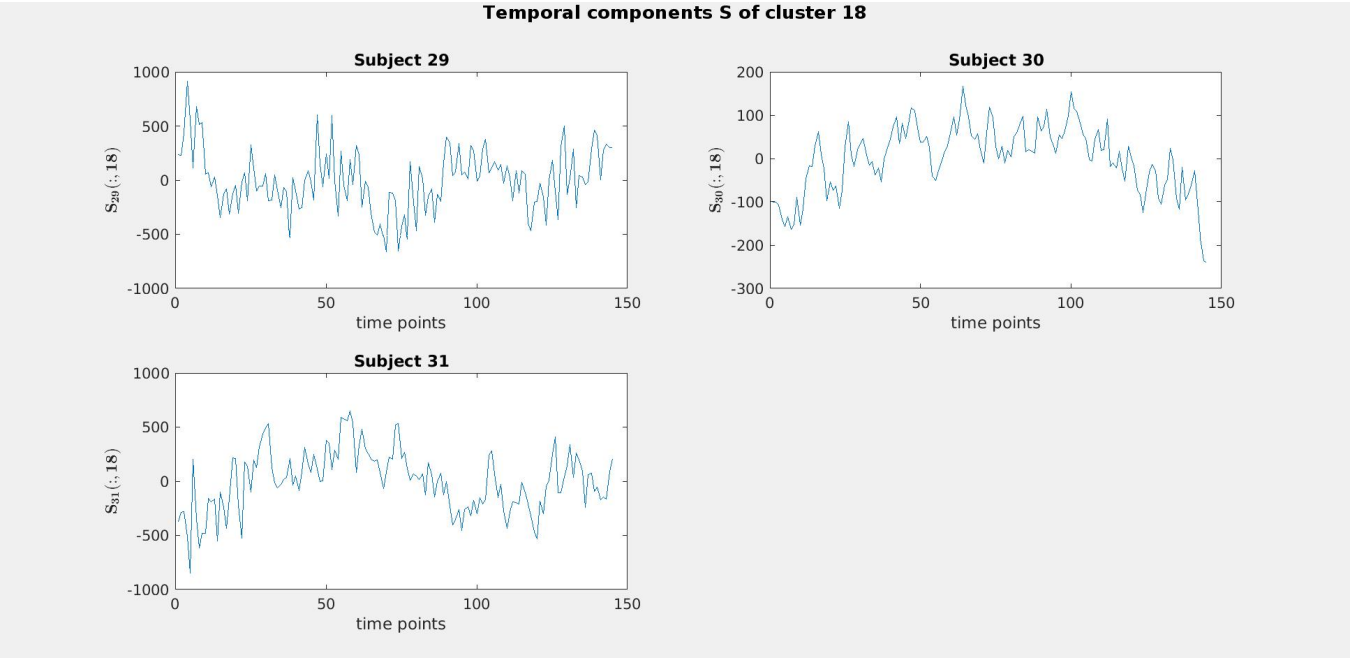


Figure B.533: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 29, \ldots, 31$ and R=42.

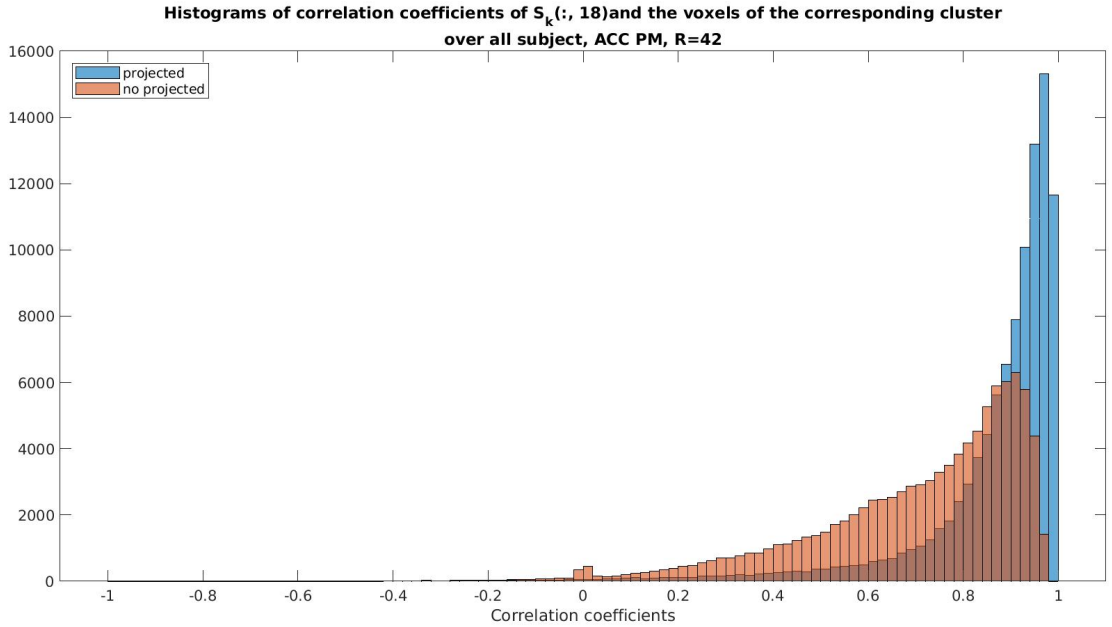


Figure B.534: For the column 18 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 18)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

25. **Cluster 10**

Cluster 10 contains 3435 voxels, which correspond to 1.520261% of the whole brain (background, soft tissue, bone are excluded).

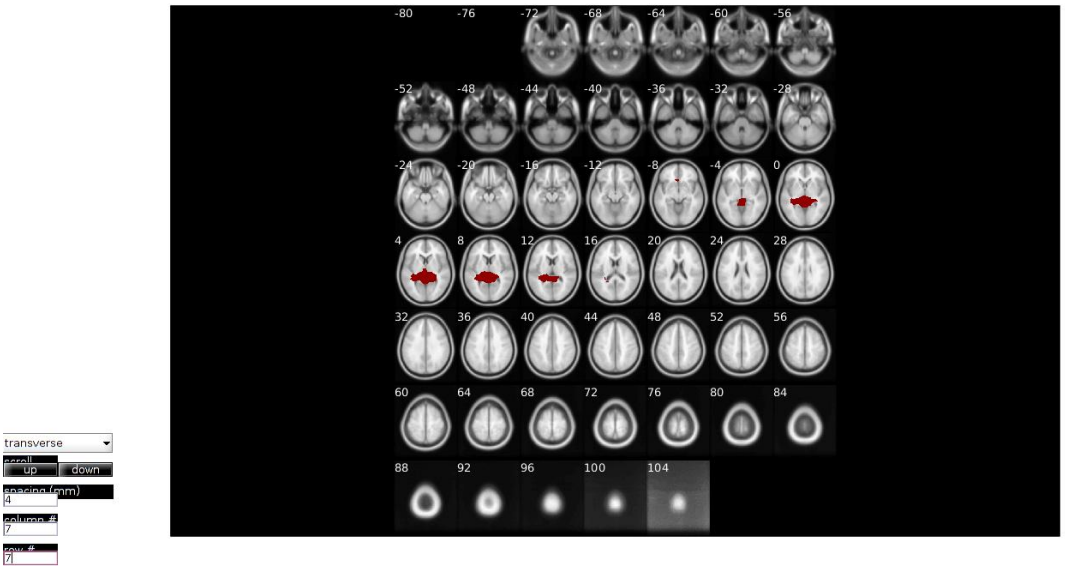


Figure B.535: Cluster 10. Red areas indicate the voxels which belong to cluster 10.

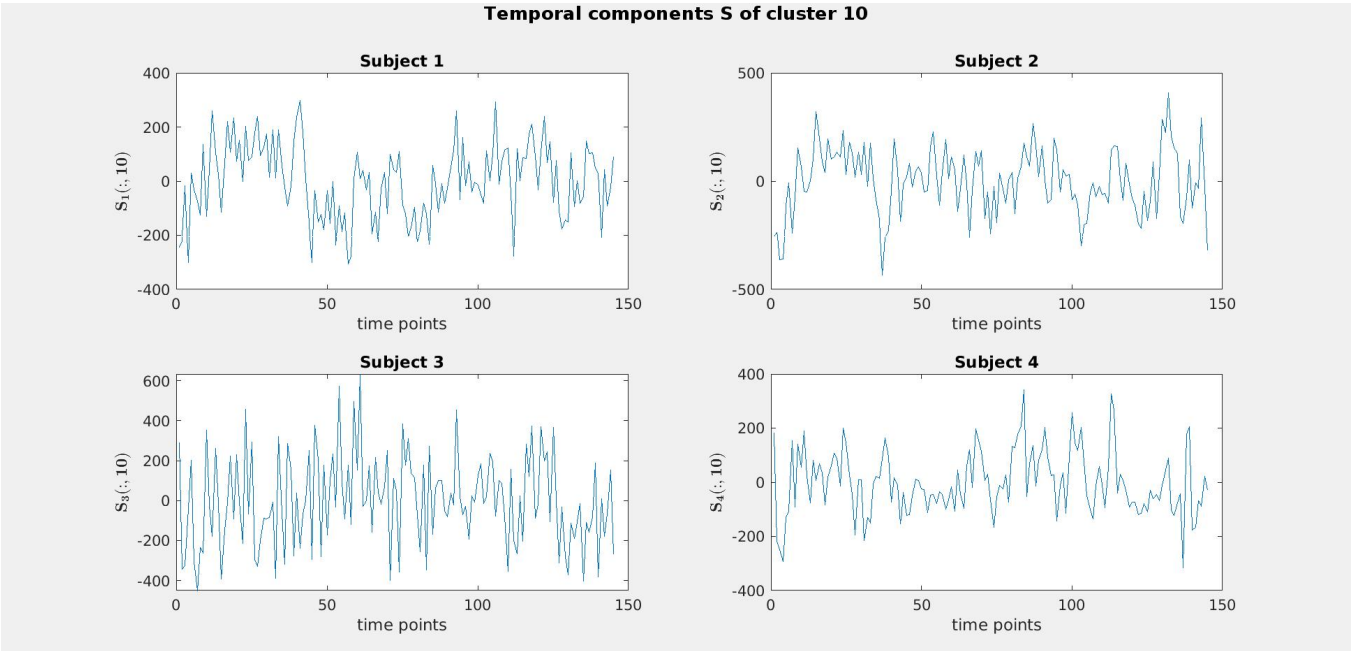


Figure B.536: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 1, \ldots, 4$ and R=42.

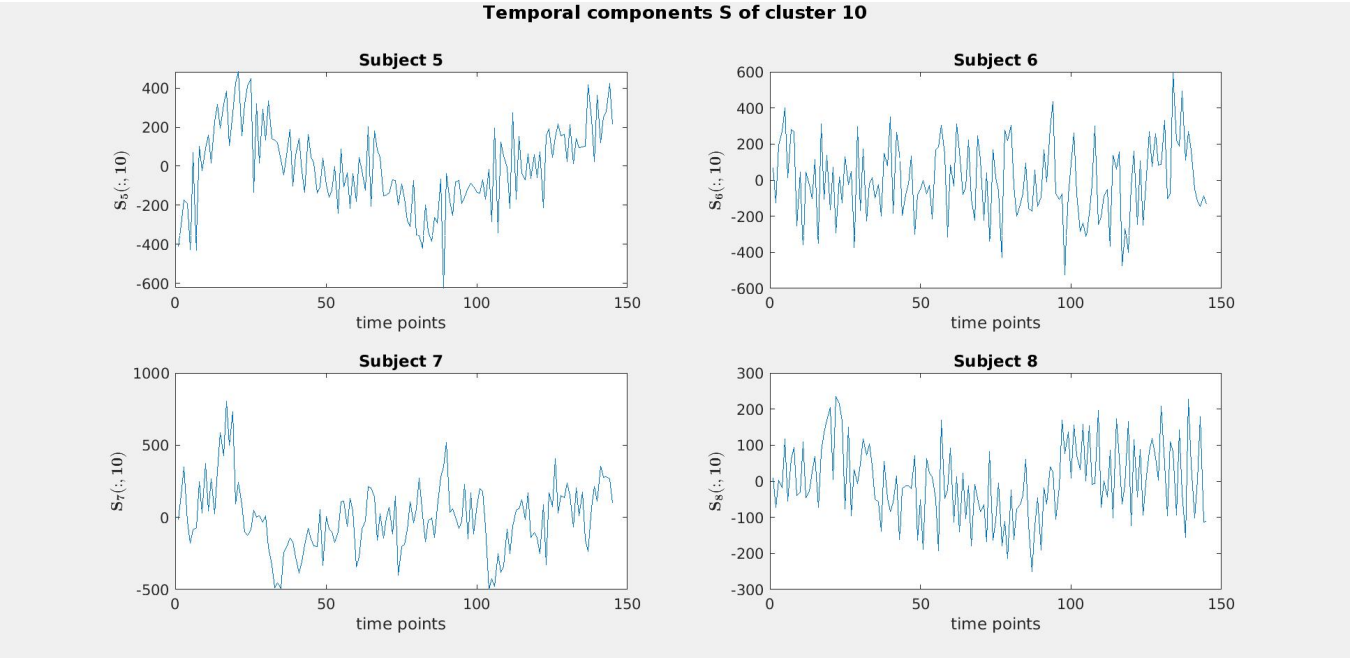


Figure B.537: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 5, \ldots, 8$ and R=42.

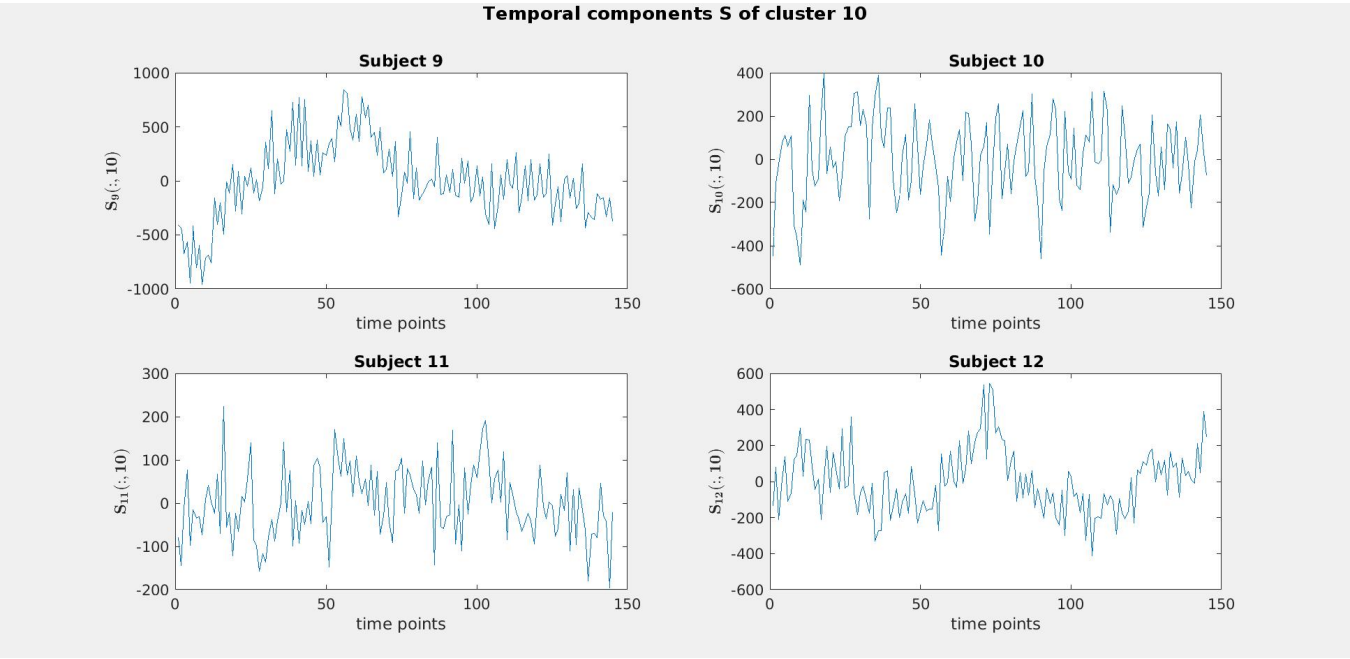


Figure B.538: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 9, \ldots, 12$ and R=42.

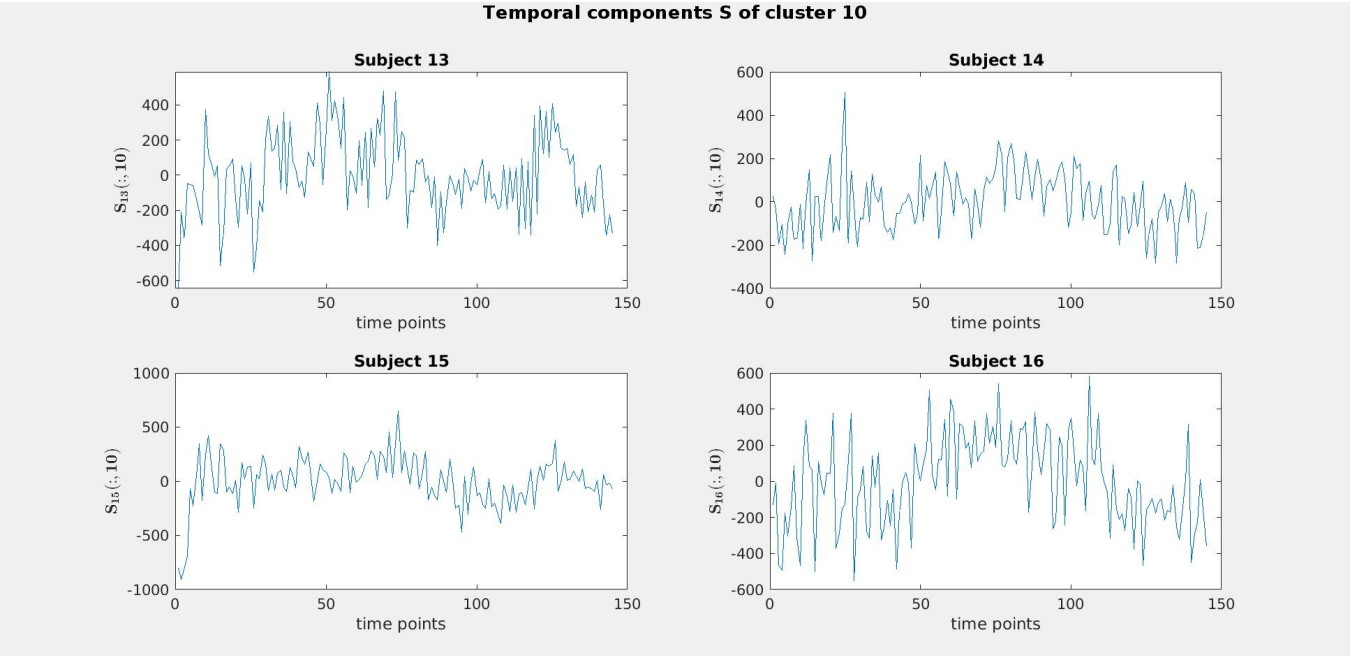


Figure B.539: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 13, \ldots, 16$ and R=42.

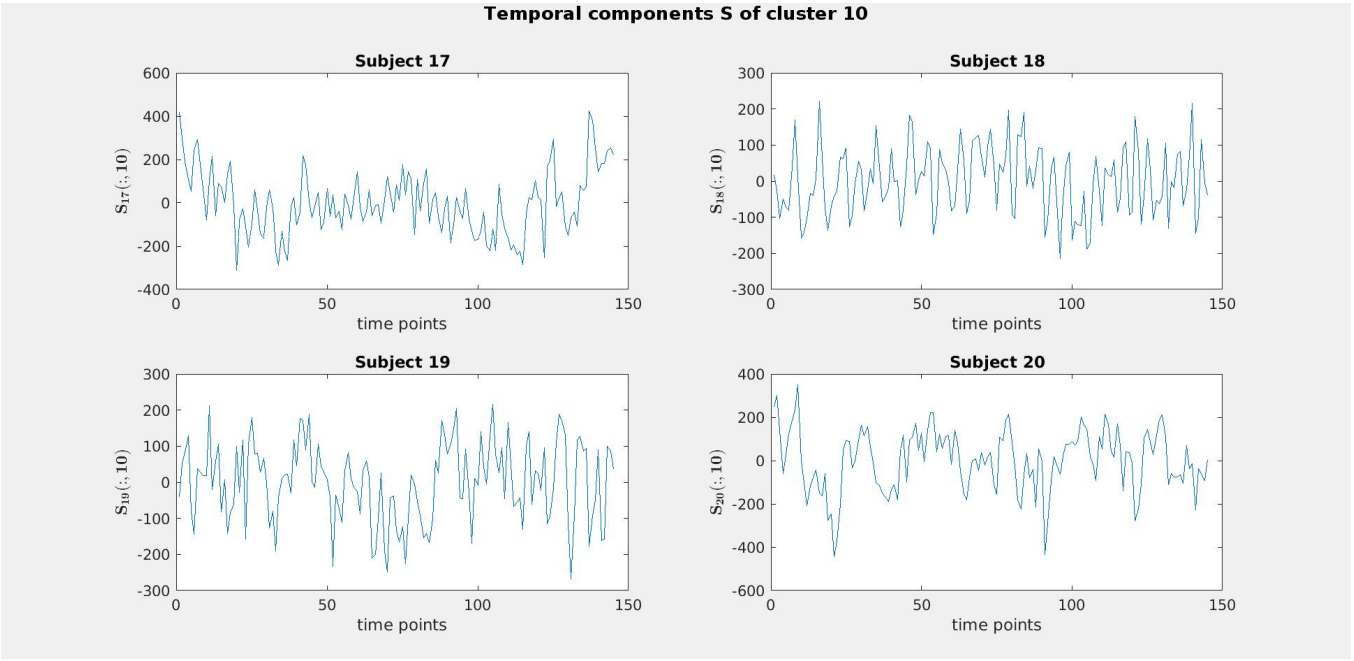


Figure B.540: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 17, \ldots, 20$ and R=42.

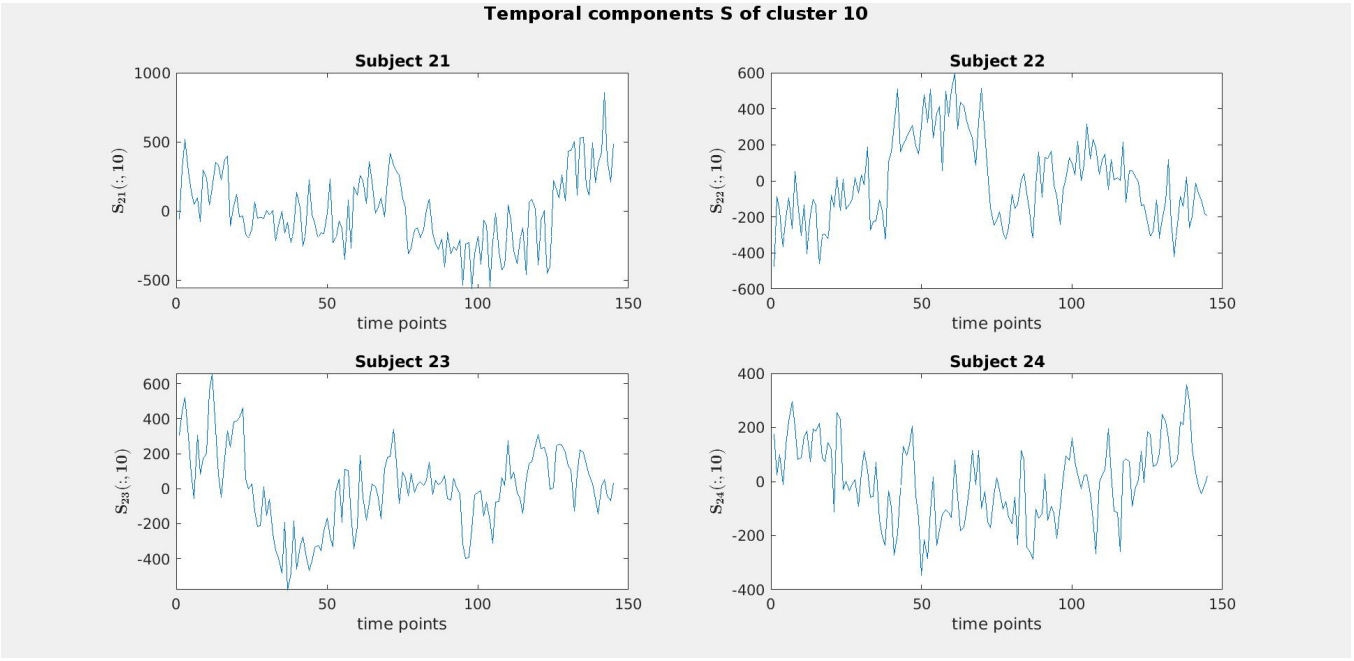


Figure B.541: Waveform of temporal components $\mathbf{S}_k(:, 110)$ for $k = 21, \ldots, 24$ and R=42.

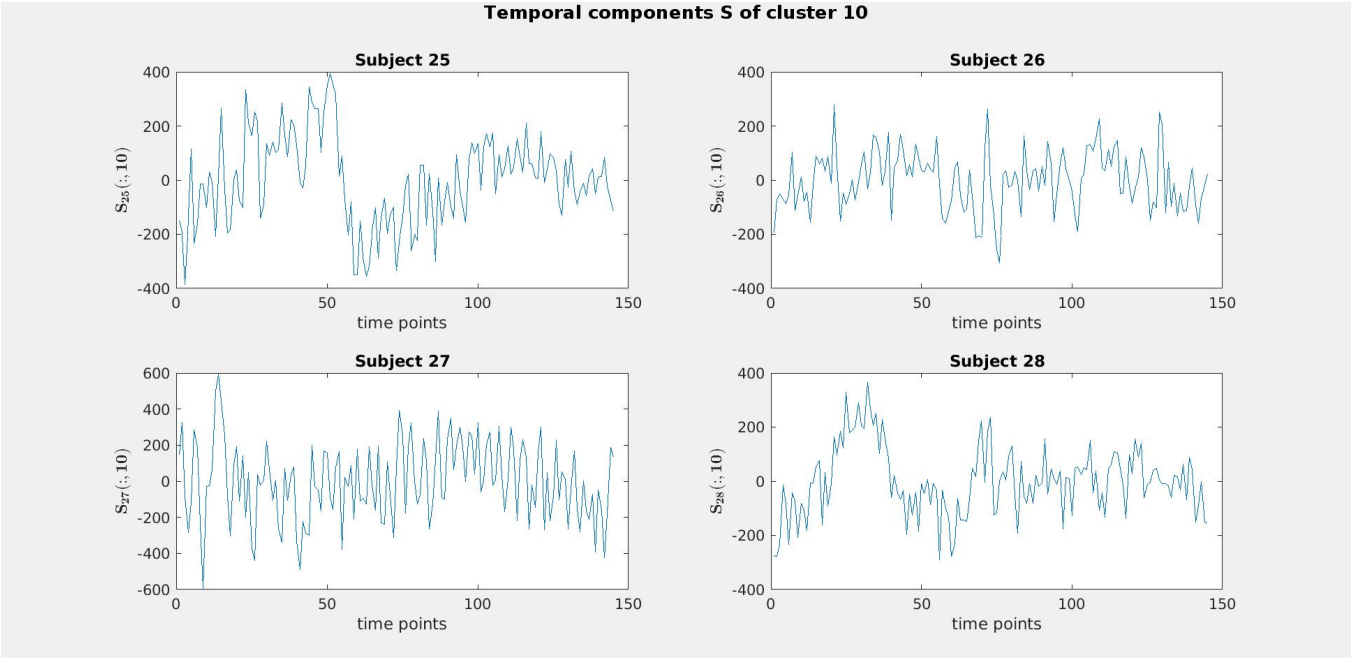


Figure B.542: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 25, \ldots, 28$ and R=42.

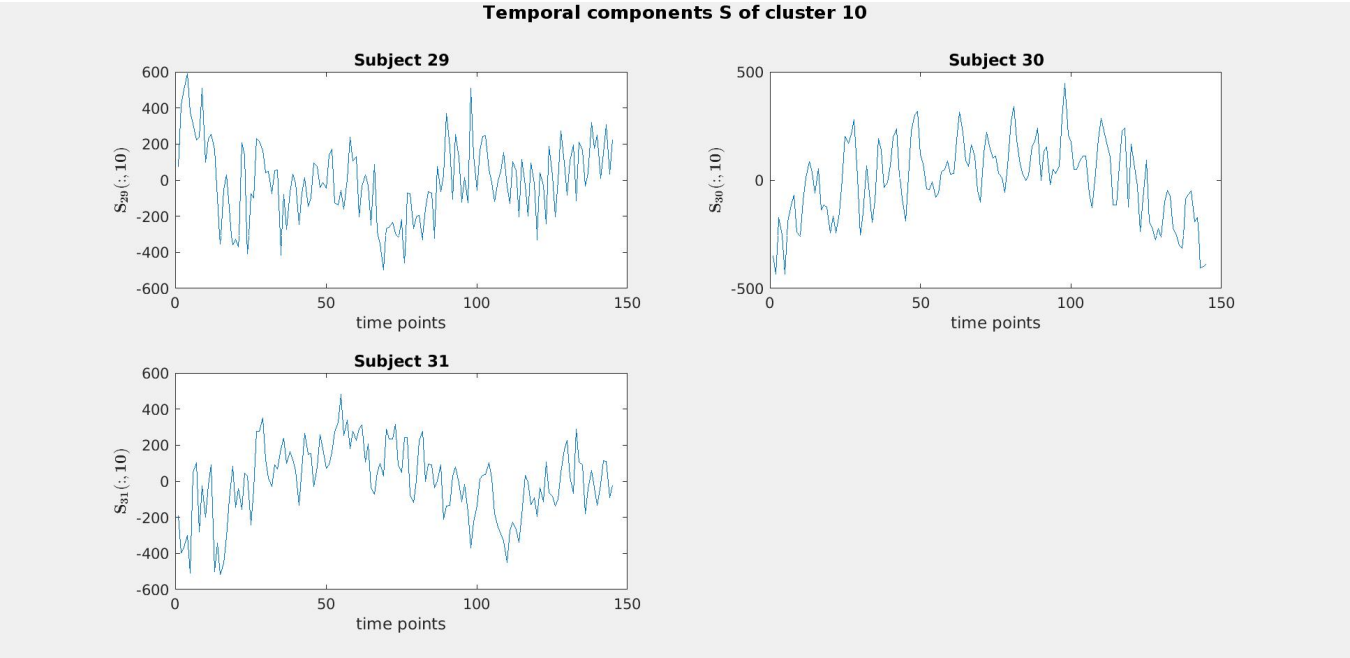


Figure B.543: Waveform of temporal components $\mathbf{S}_k(:, 10)$ for $k = 29, \ldots, 31$ and R=42.

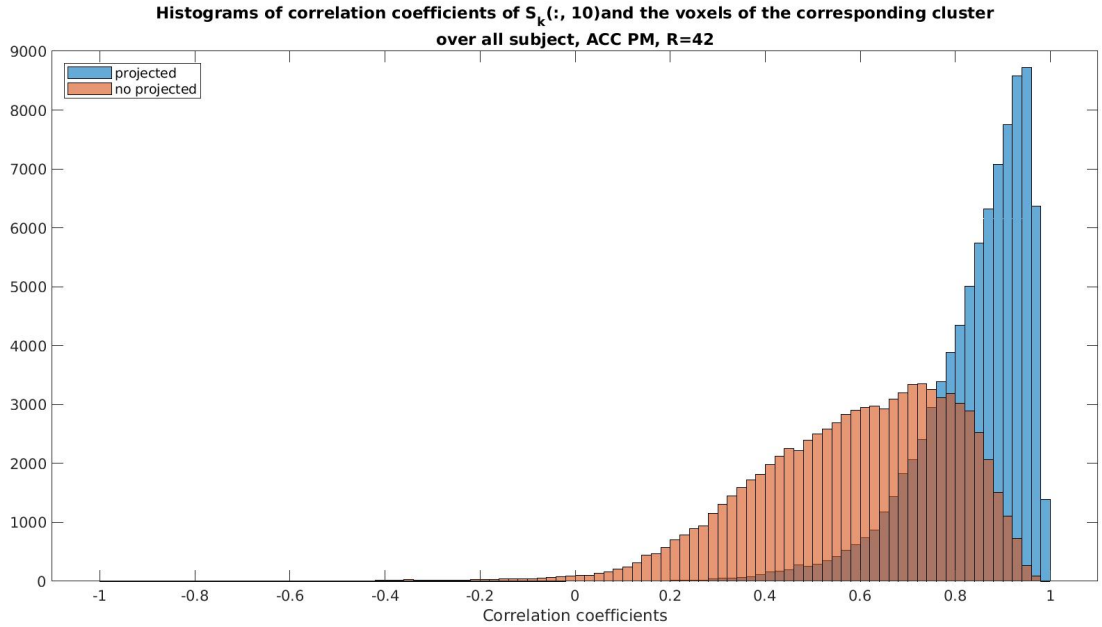


Figure B.544: For the column 10 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 10)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

26. **Cluster 28**

Cluster 28 contains 3421 voxels, which correspond to 1.514065% of the whole brain (background, soft tissue, bone are excluded).

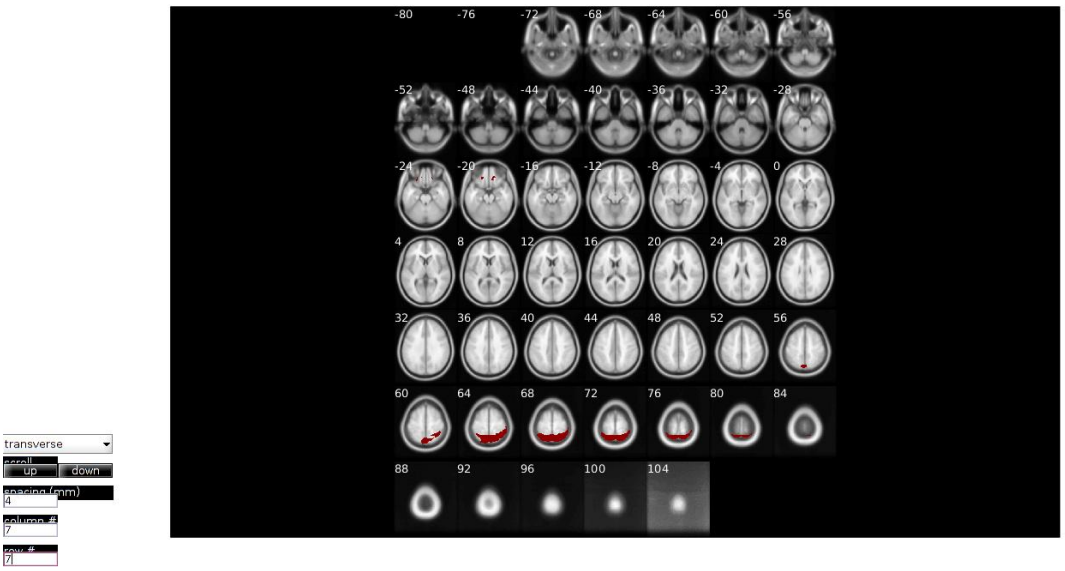


Figure B.545: Cluster 28. Red areas indicate the voxels which belong to cluster 28.

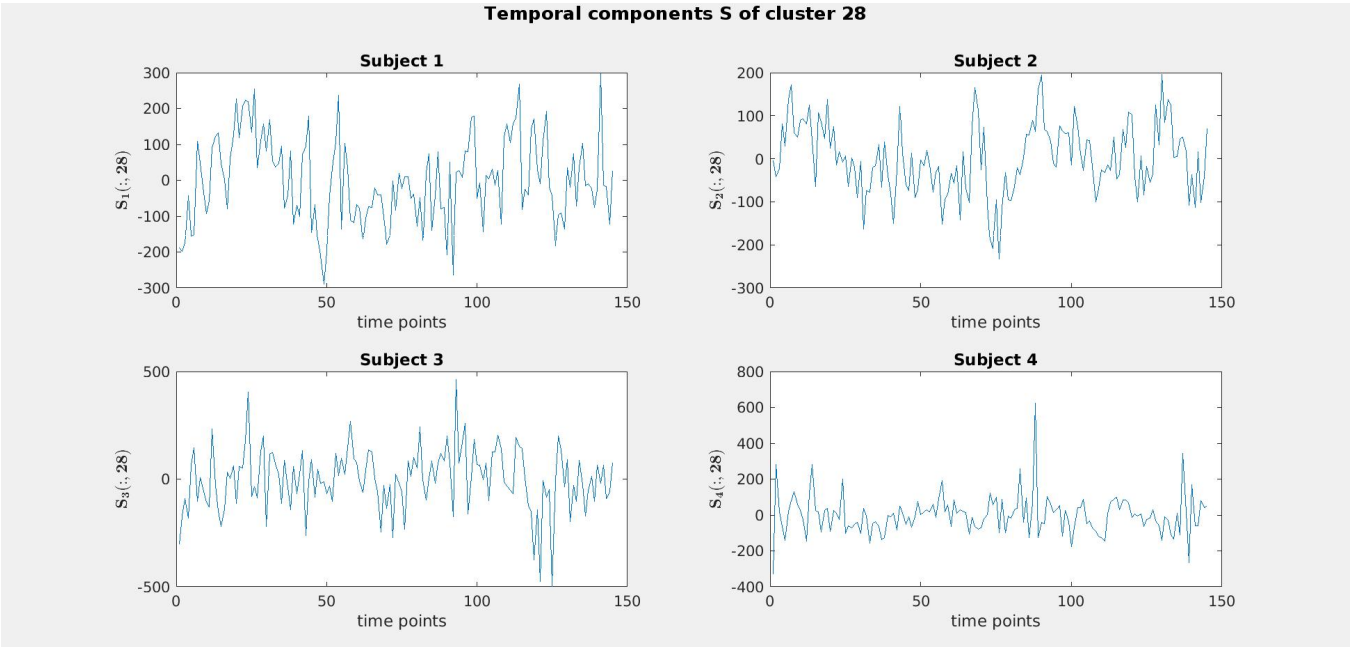


Figure B.546: Waveform of temporal components $\mathbf{S}_k(:, 28)$ for $k = 1, \ldots, 4$ and R=42.

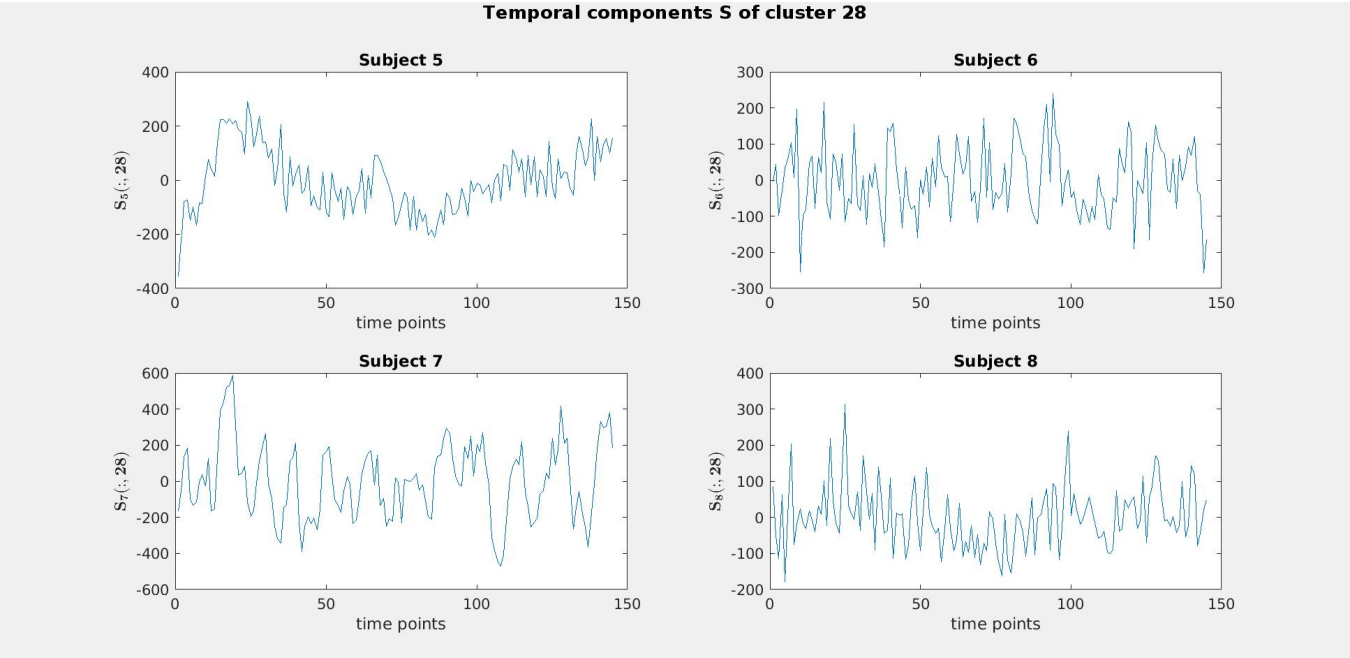


Figure B.547: Waveform of temporal components $\mathbf{S}_k(:, 28)$ for $k = 5, \ldots, 8$ and R=42.

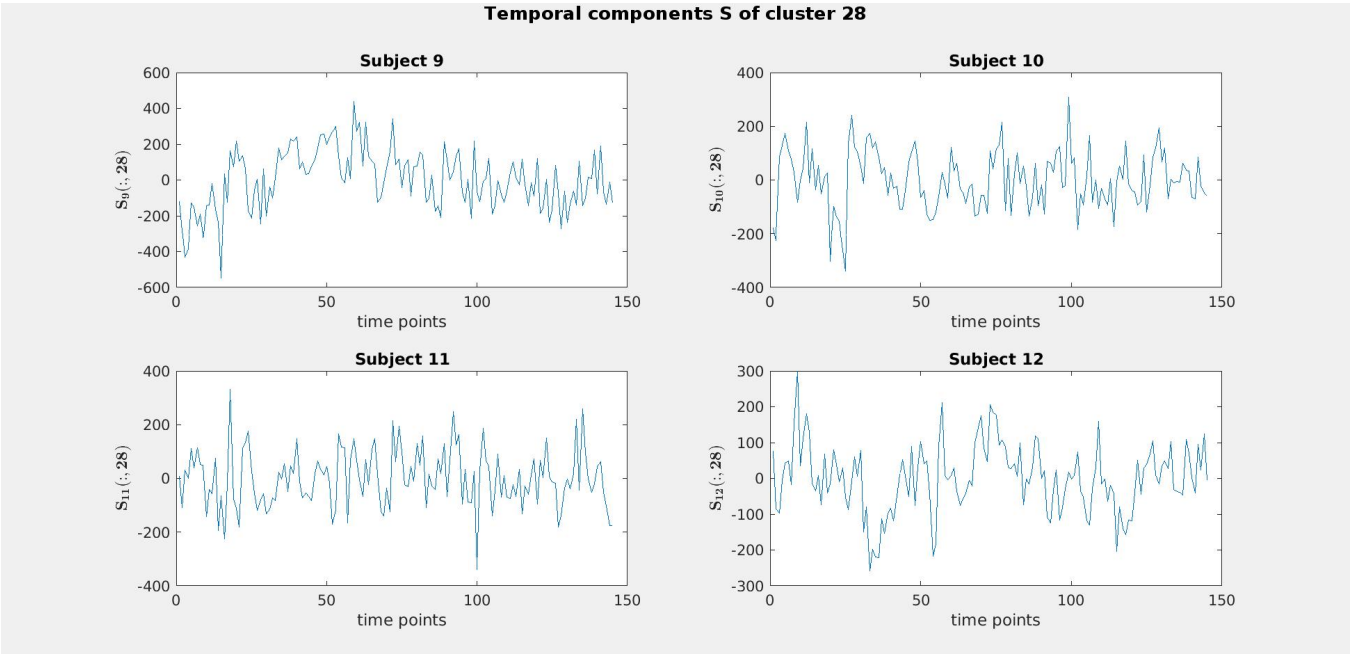


Figure B.548: Waveform of temporal components $\mathbf{S}_k(:, 28)$ for $k = 9, \ldots, 12$ and R=42.

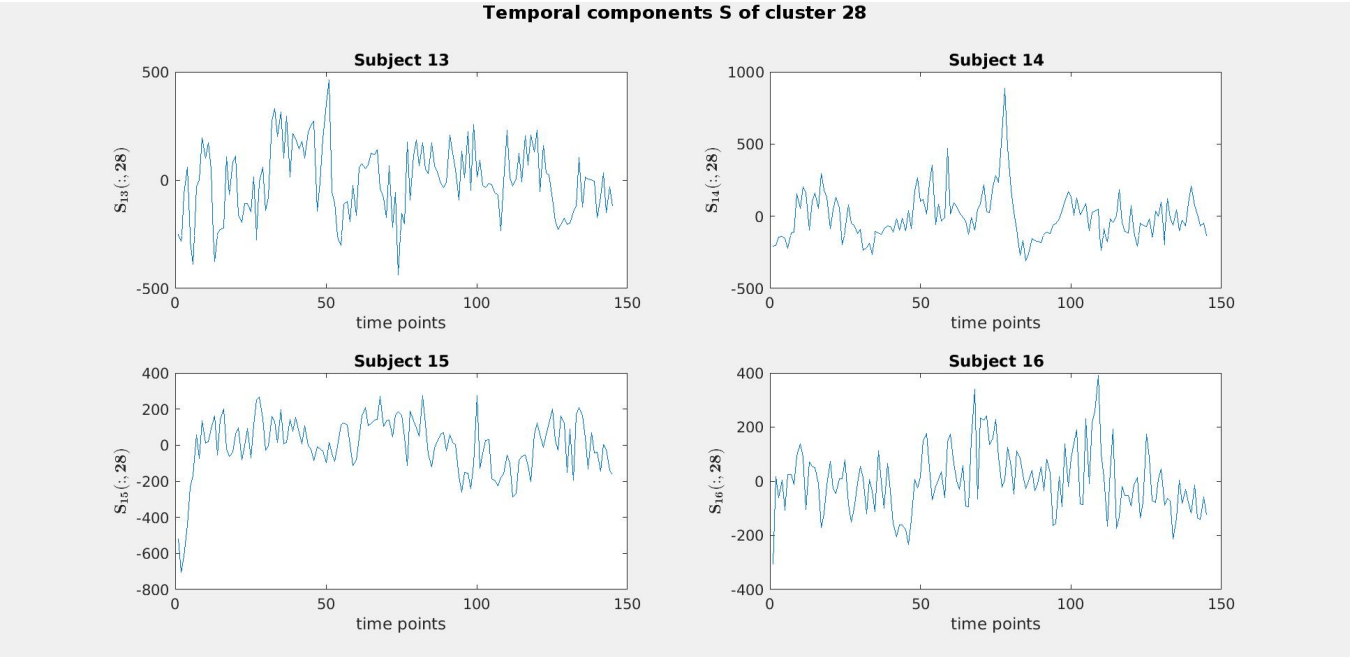


Figure B.549: Waveform of temporal components $\mathbf{S}_k(:, 28)$ for $k = 13, \ldots, 16$ and R=42.

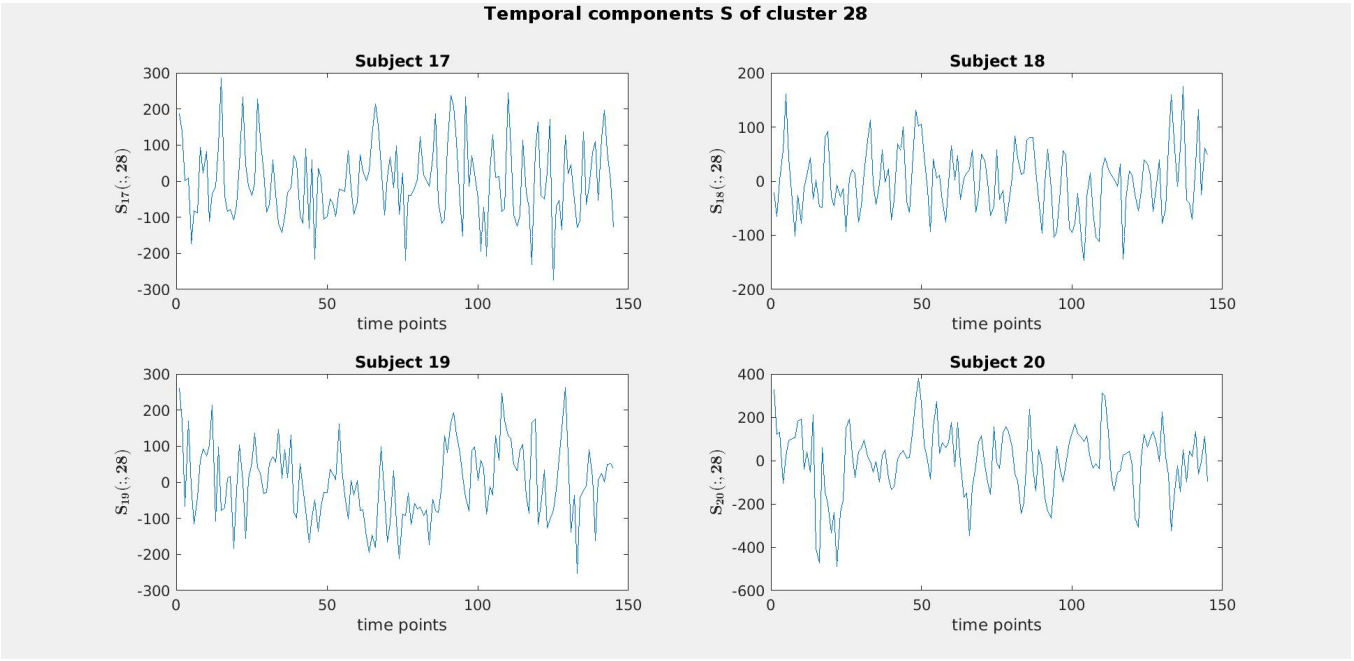


Figure B.550: Waveform of temporal components $\mathbf{S}_k(:, 28)$ for $k = 17, \ldots, 20$ and R=42.

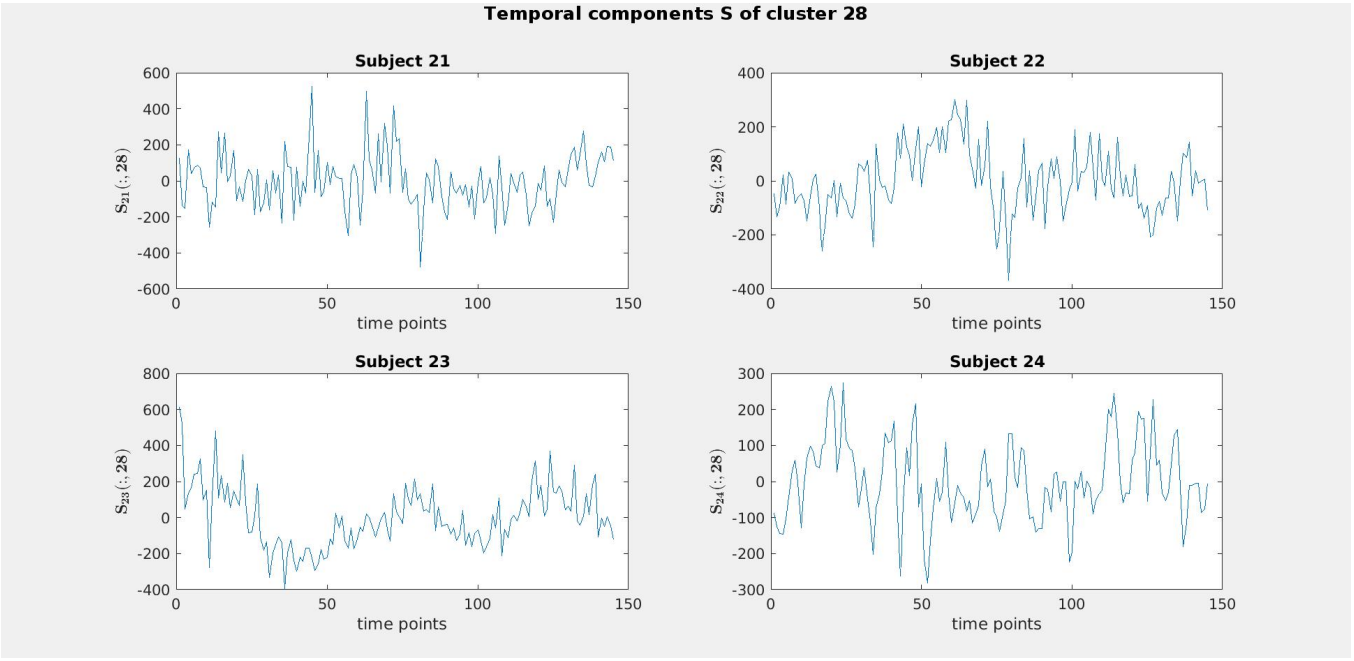


Figure B.551: Waveform of temporal components $\mathbf{S}_k(:, 128)$ for $k = 21, \ldots, 24$ and R=42.

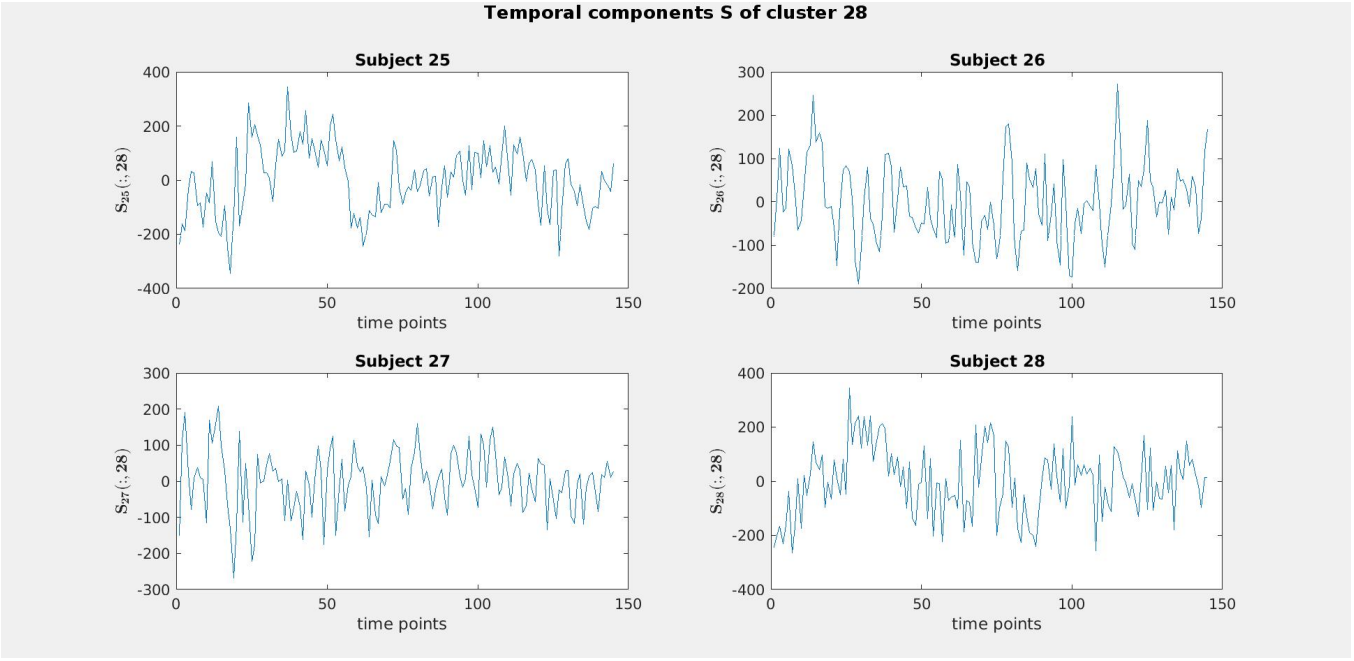


Figure B.552: Waveform of temporal components $\mathbf{S}_k(:, 28)$ for $k = 25, \ldots, 28$ and R=42.

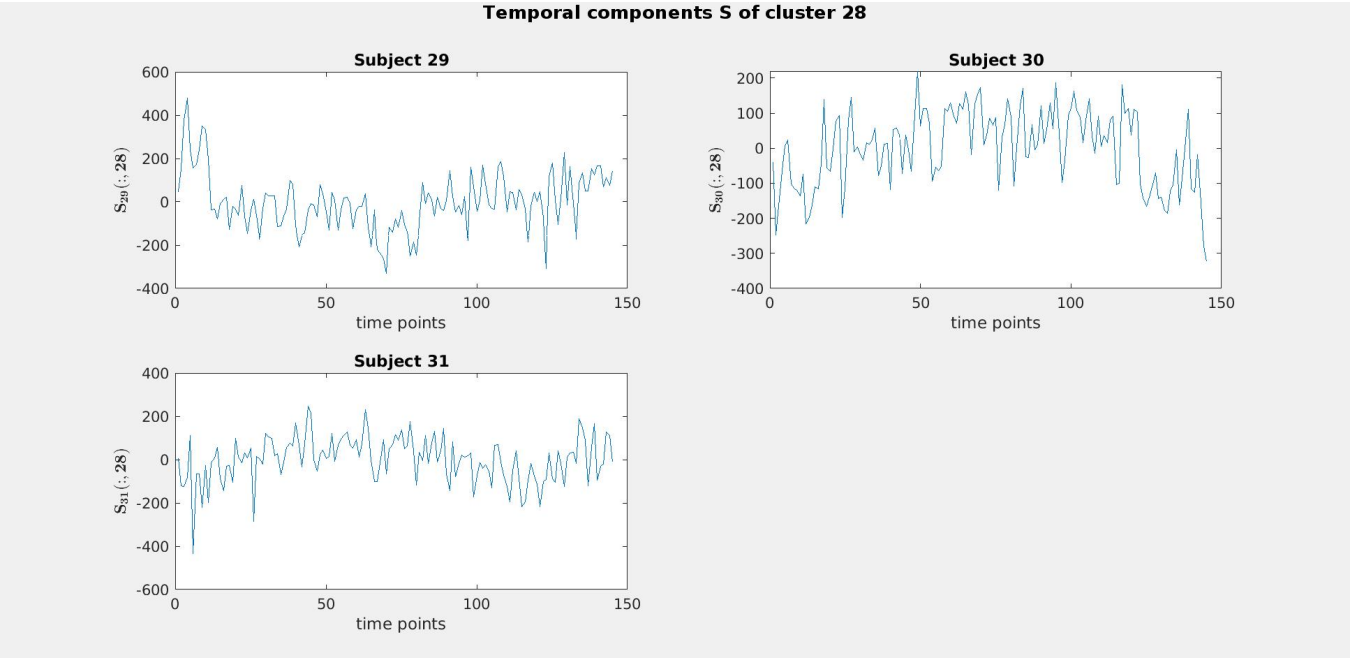


Figure B.553: Waveform of temporal components $\mathbf{S}_k(:, 28)$ for $k = 29, \ldots, 31$ and R=42.

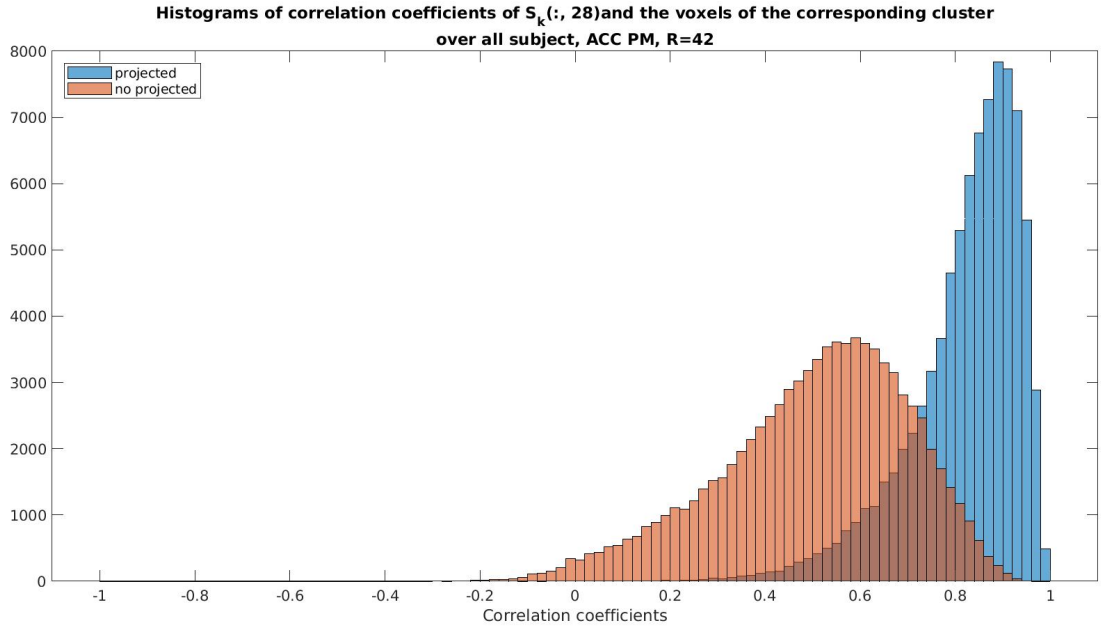


Figure B.554: For the column 28 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 28)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

27. **Cluster 9**

Cluster 9 contains 3374 voxels, which correspond to 1.493264% of the whole brain (background, soft tissue, bone are excluded).

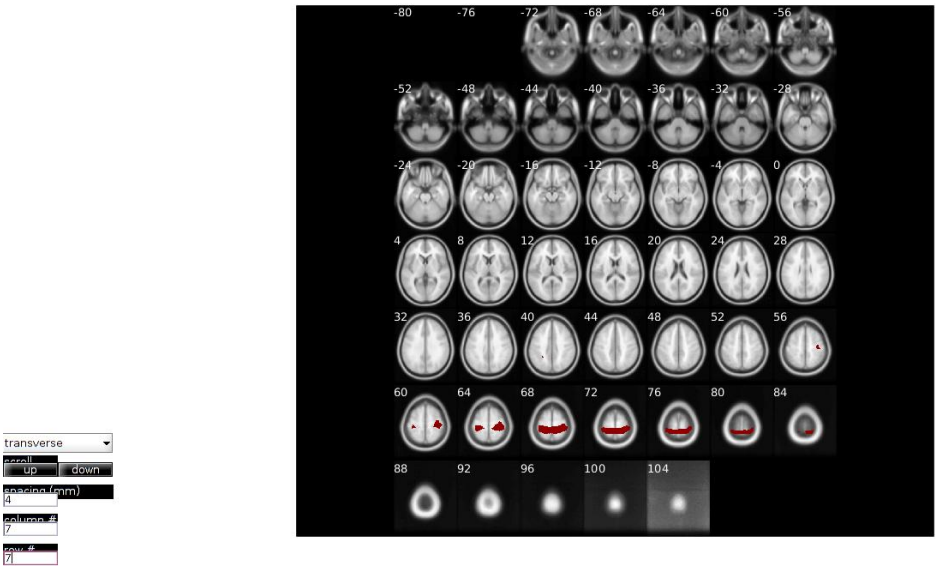


Figure B.555: Cluster 9. Red areas indicate the voxels which belong to cluster 9.

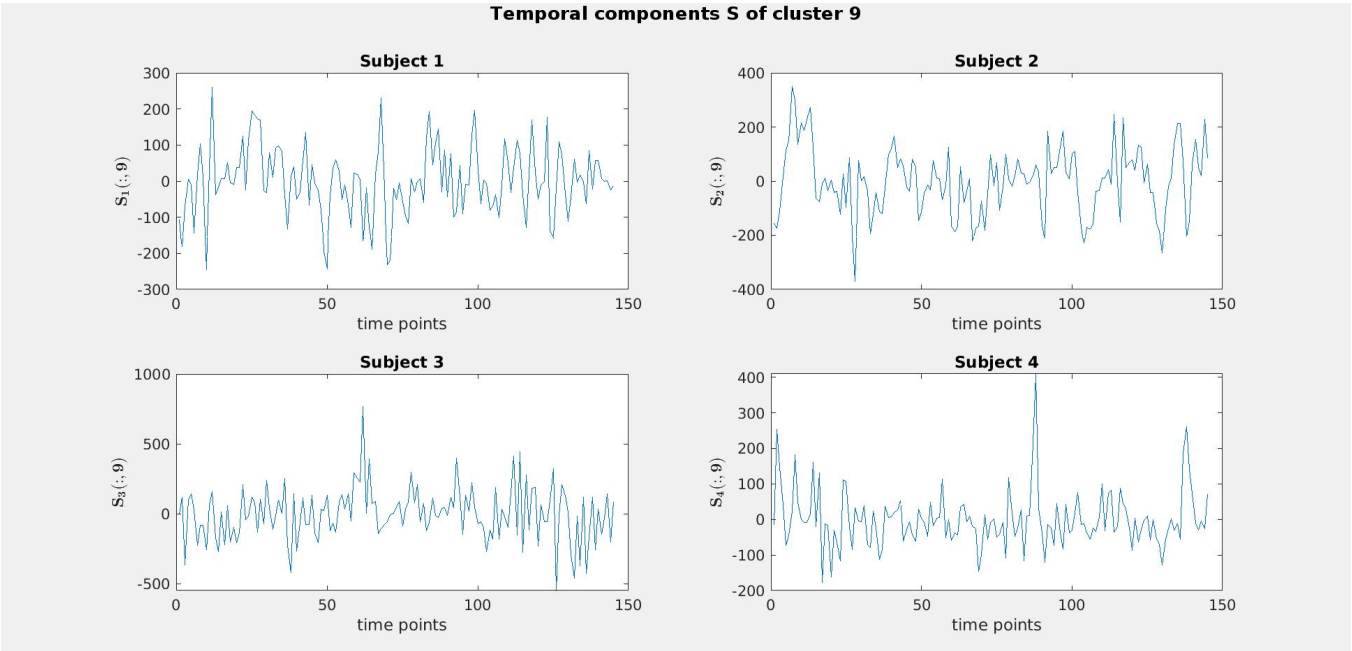


Figure B.556: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 1, \ldots, 4$ and R=42.

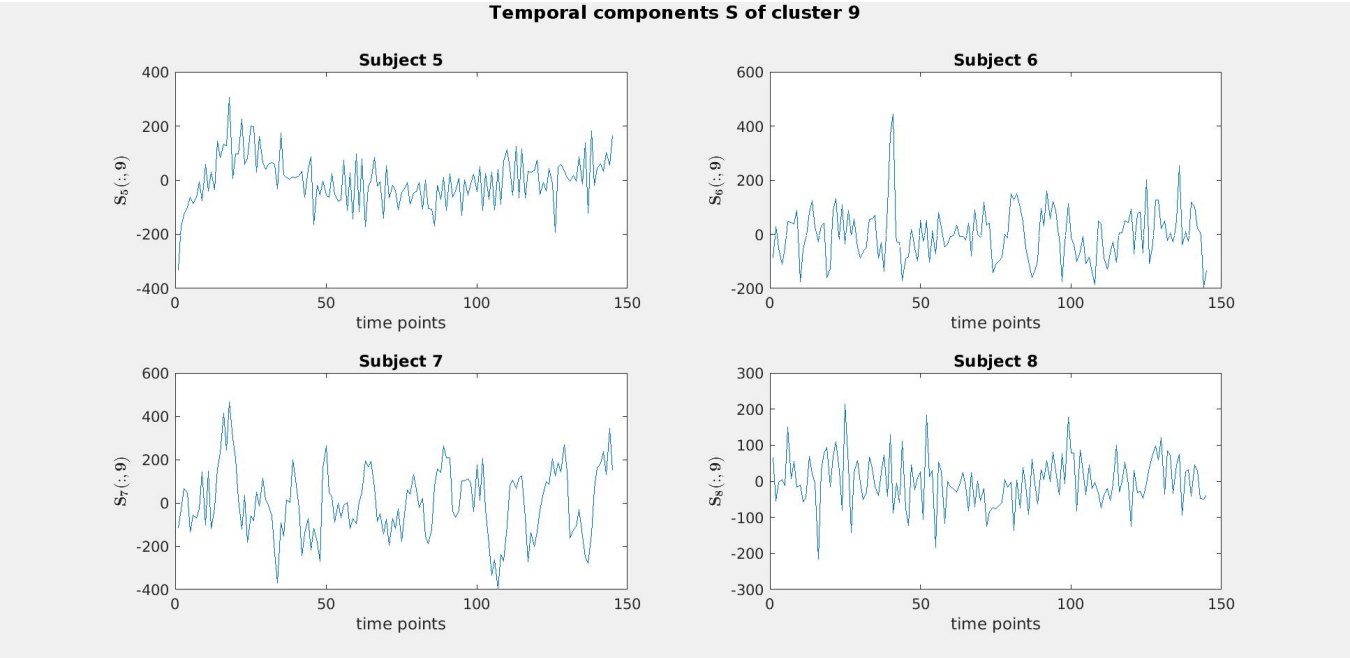


Figure B.557: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 5, \ldots, 8$ and R=42.

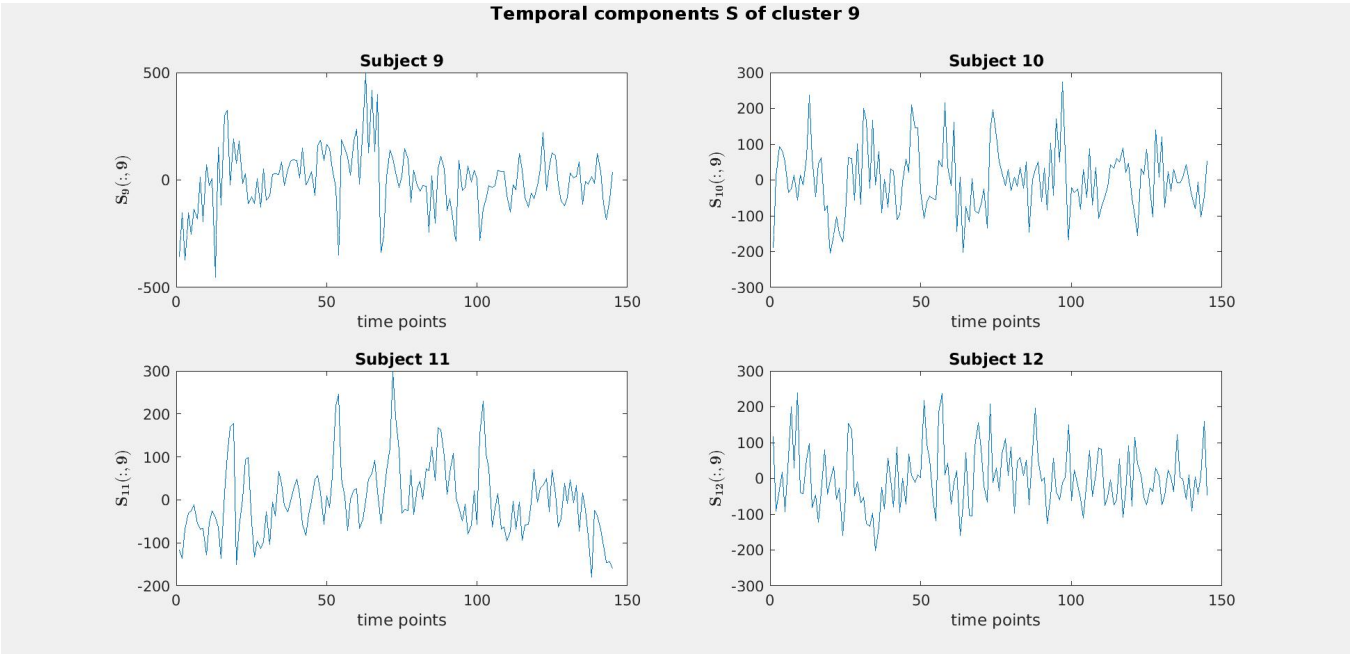


Figure B.558: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 9, \ldots, 12$ and R=42.

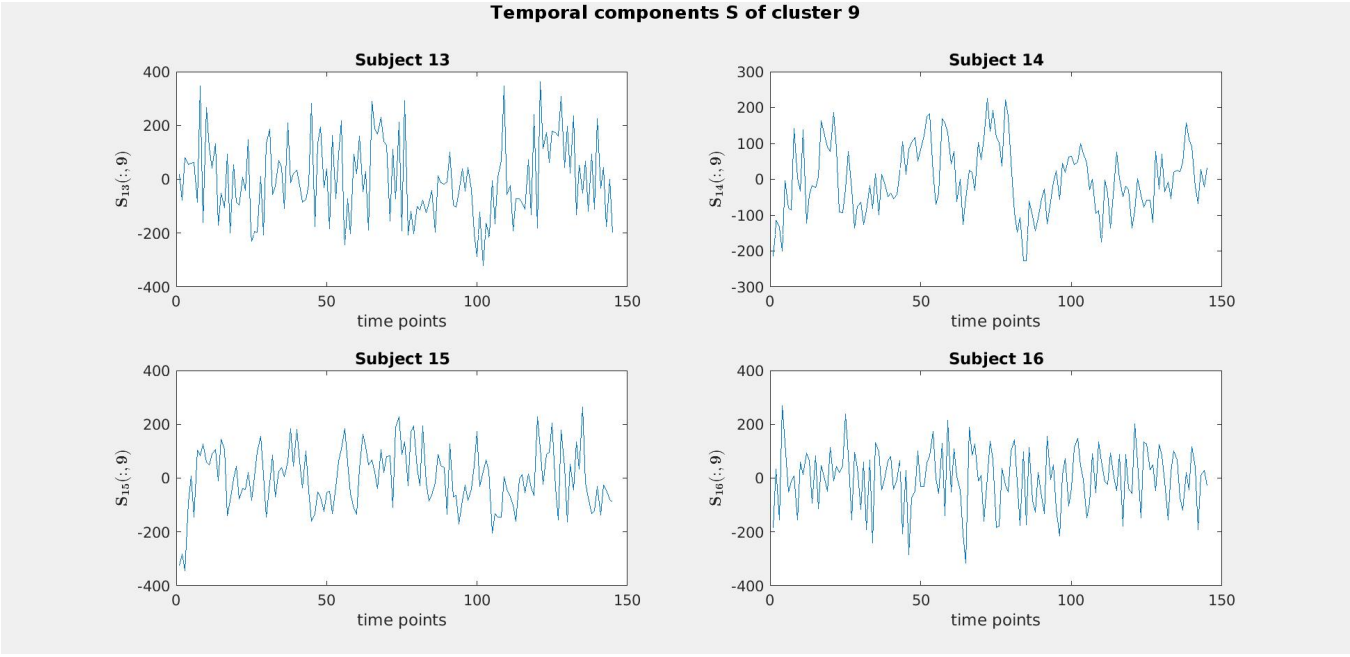


Figure B.559: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 13, \ldots, 16$ and R=42.

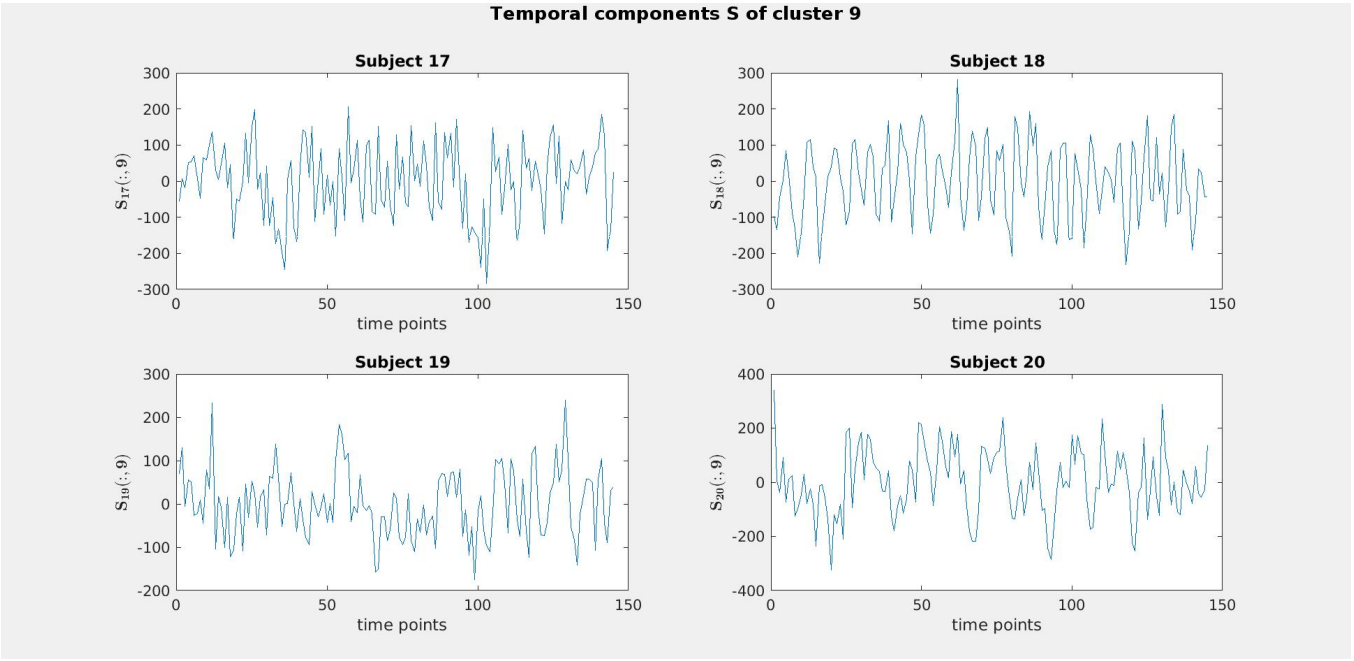


Figure B.560: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 17, \ldots, 20$ and R=42.

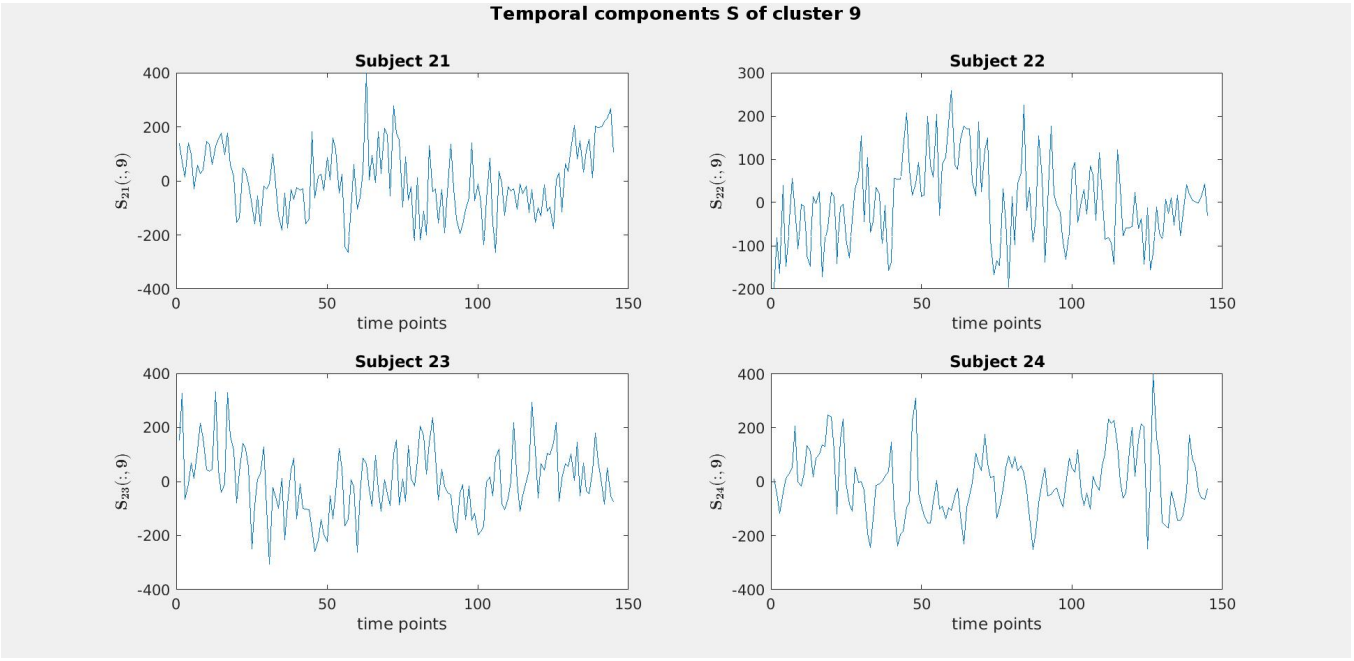


Figure B.561: Waveform of temporal components $\mathbf{S}_k(:, 19)$ for $k = 21, \ldots, 24$ and R=42.

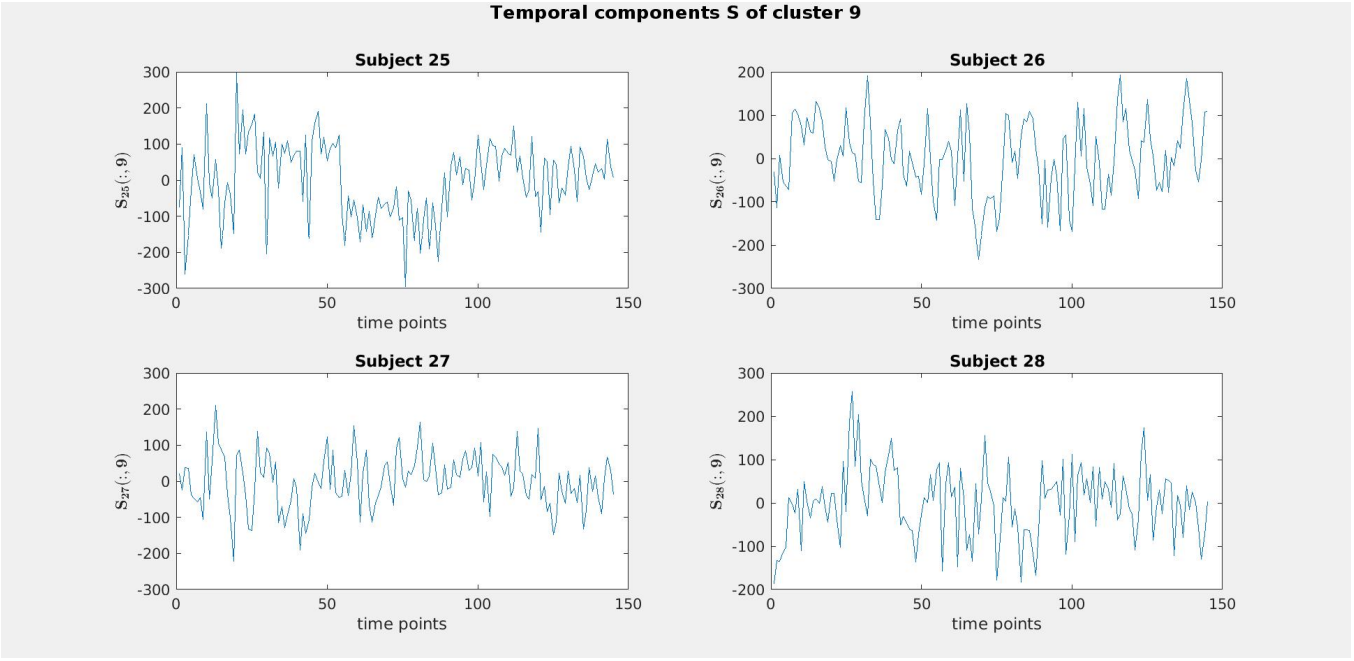


Figure B.562: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 25, \ldots, 28$ and R=42.

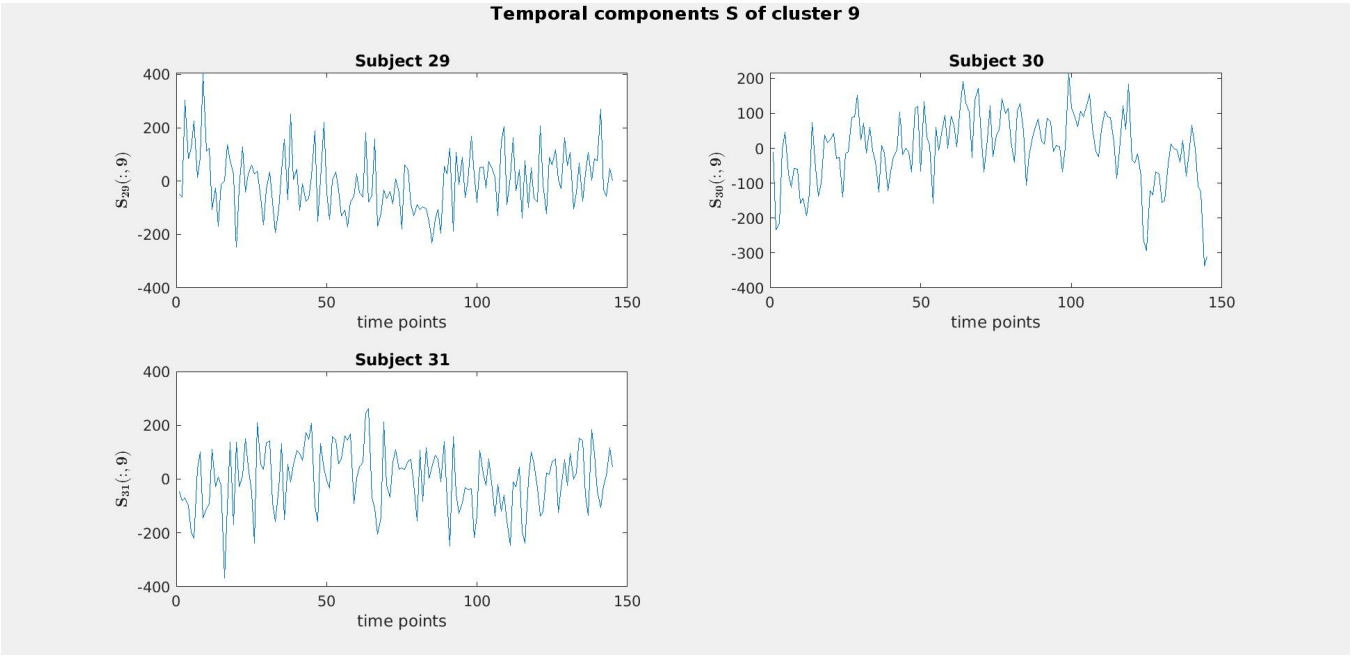


Figure B.563: Waveform of temporal components $\mathbf{S}_k(:, 9)$ for $k = 29, \ldots, 31$ and R=42.

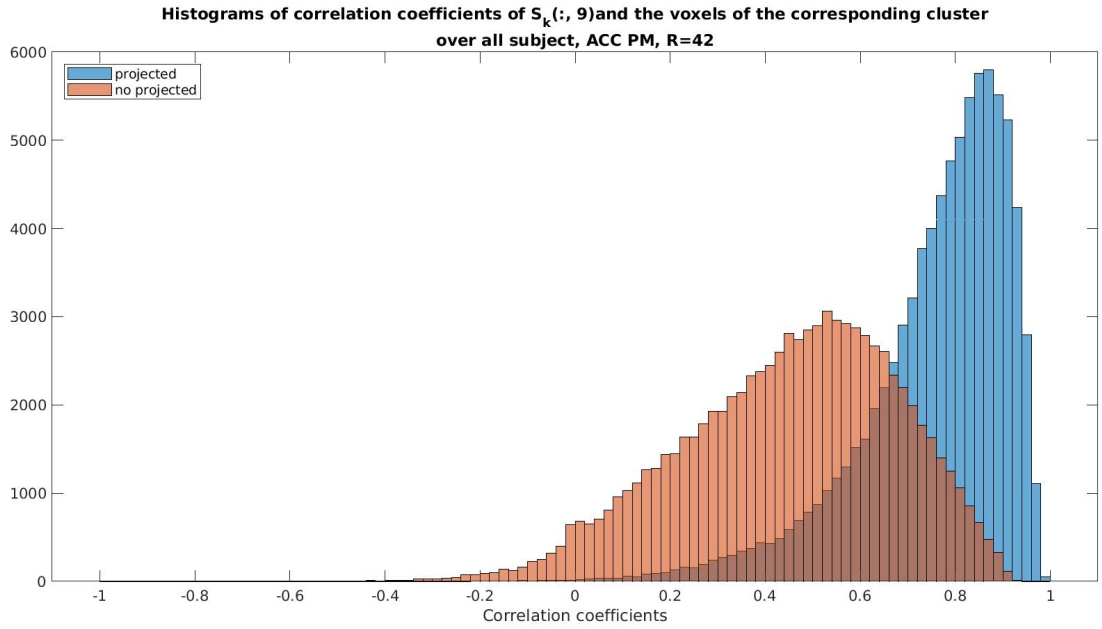


Figure B.564: For the column 9 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 9)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

28. **Cluster 15**

Cluster 15 contains 3335 voxels, which correspond to 1.476003% of the whole brain (background, soft tissue, bone are excluded).

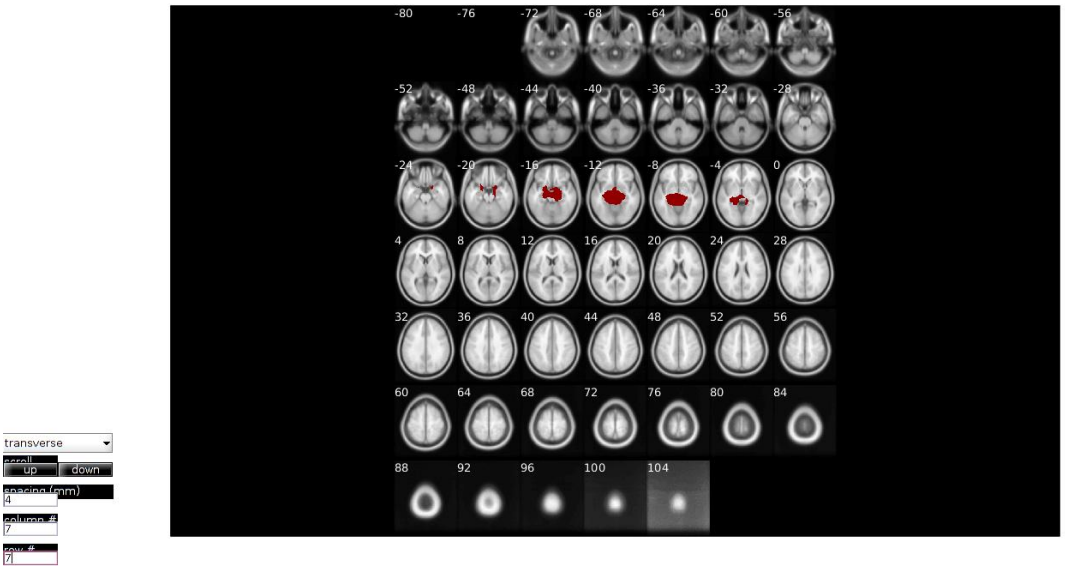


Figure B.565: Cluster 15. Red areas indicate the voxels which belong to cluster 15.

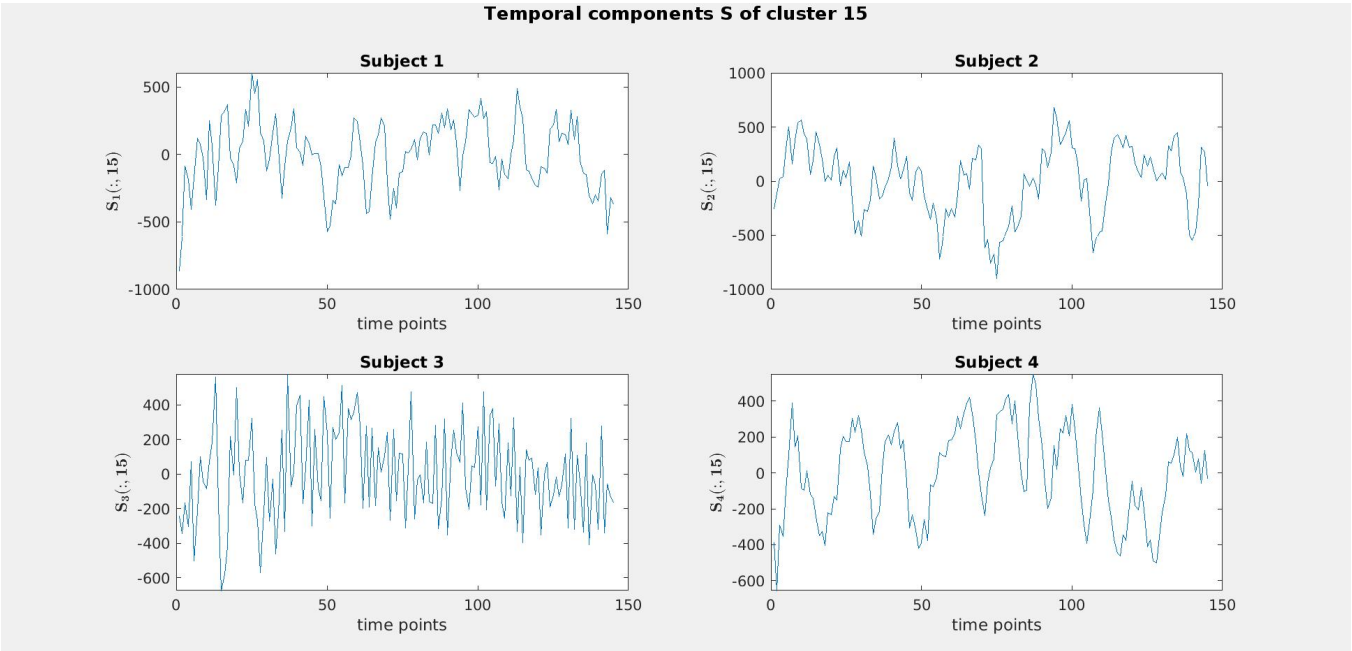


Figure B.566: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 1, \ldots, 4$ and R=42.

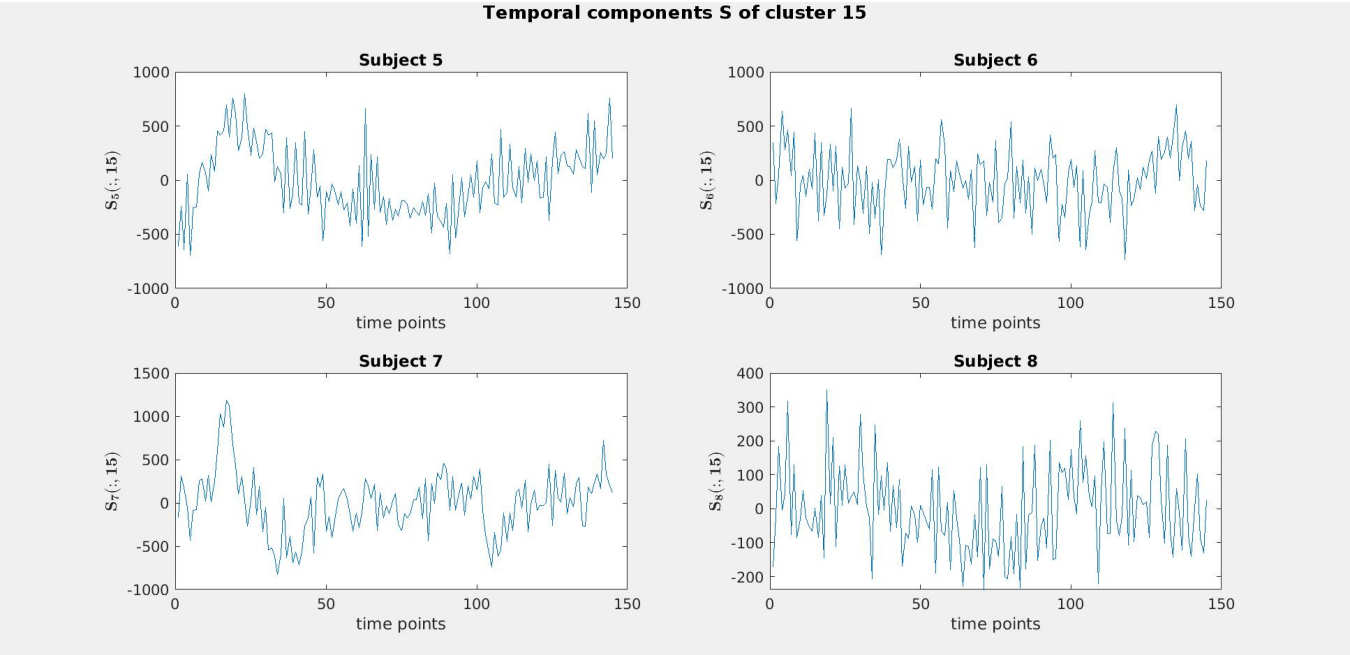


Figure B.567: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 5, \ldots, 8$ and R=42.

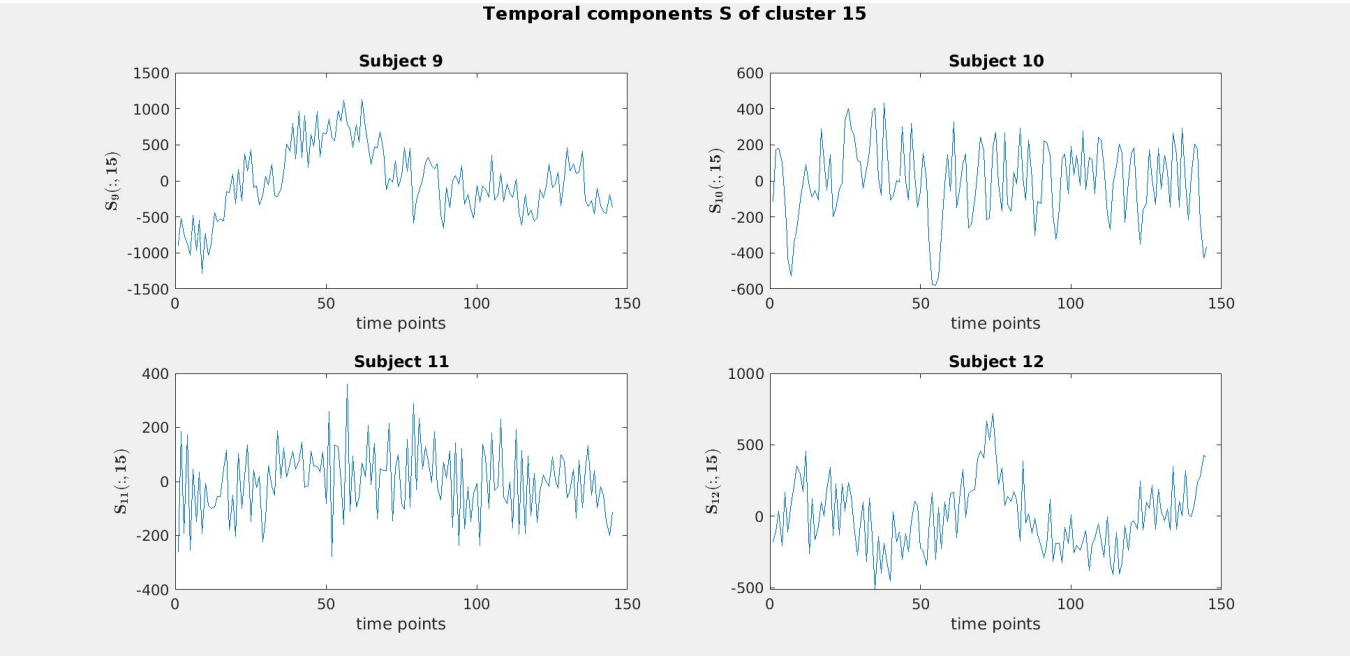


Figure B.568: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 9, \ldots, 12$ and R=42.

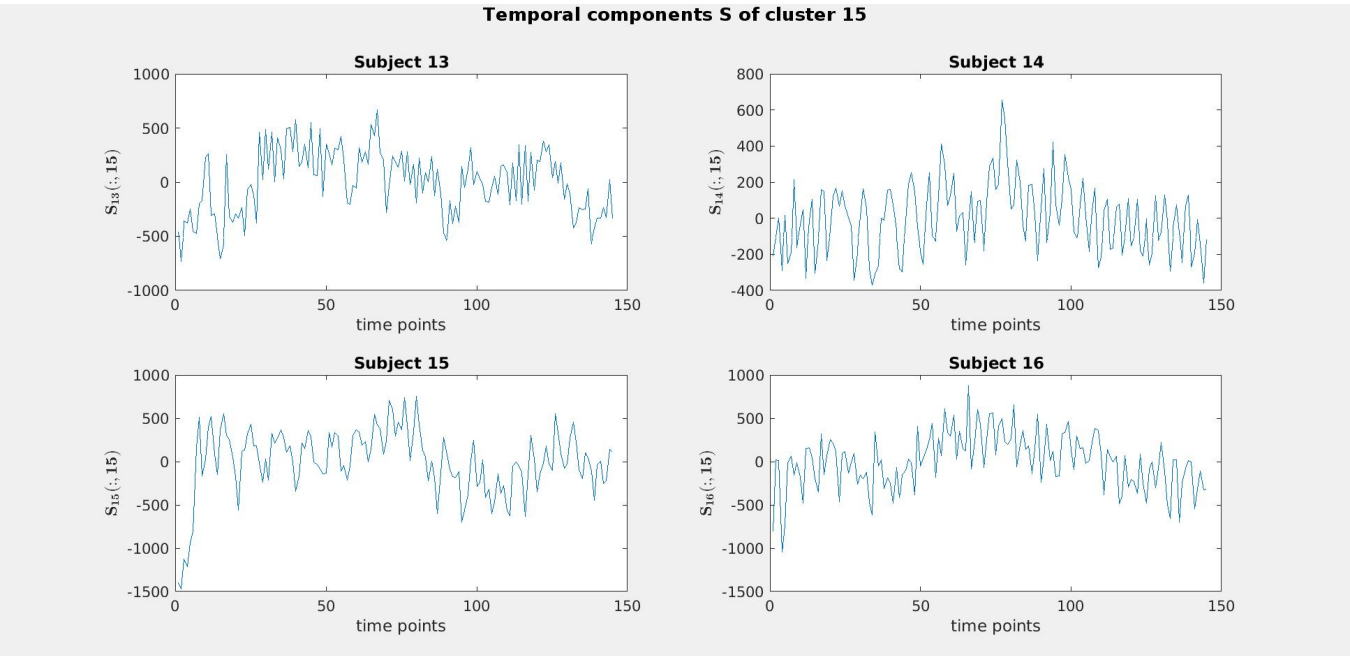


Figure B.569: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 13, \ldots, 16$ and R=42.

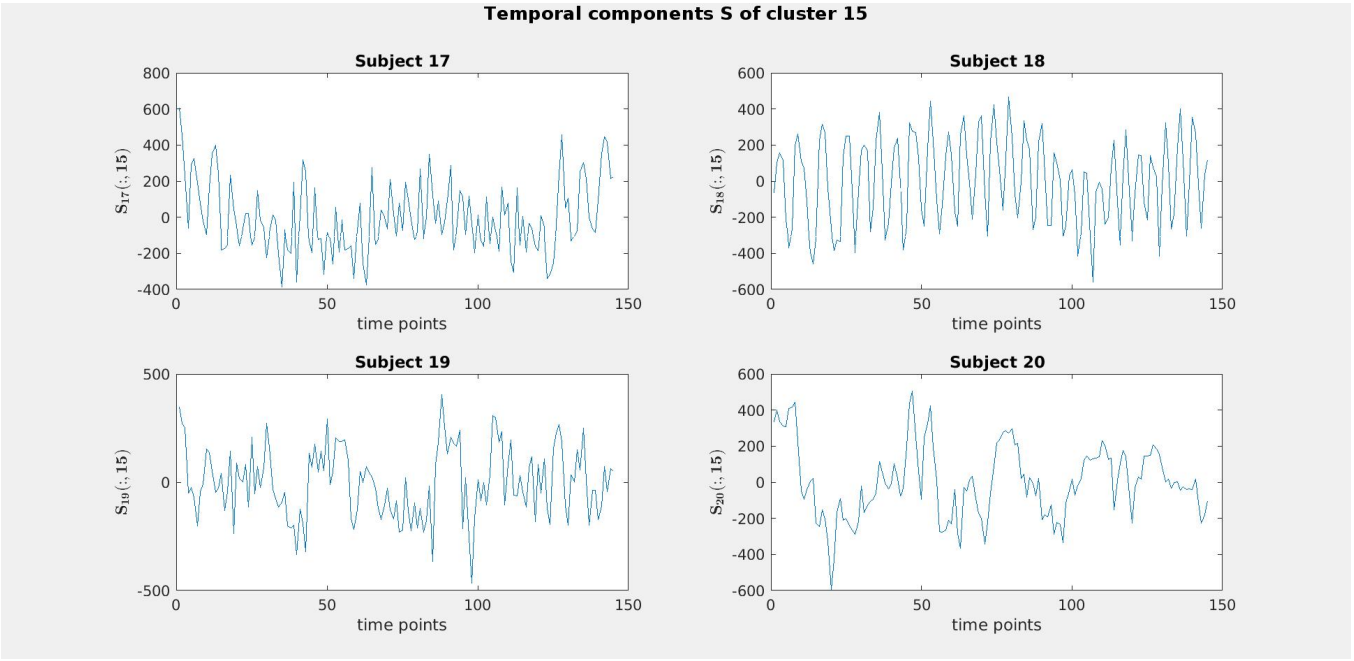


Figure B.570: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 17, \ldots, 20$ and R=42.

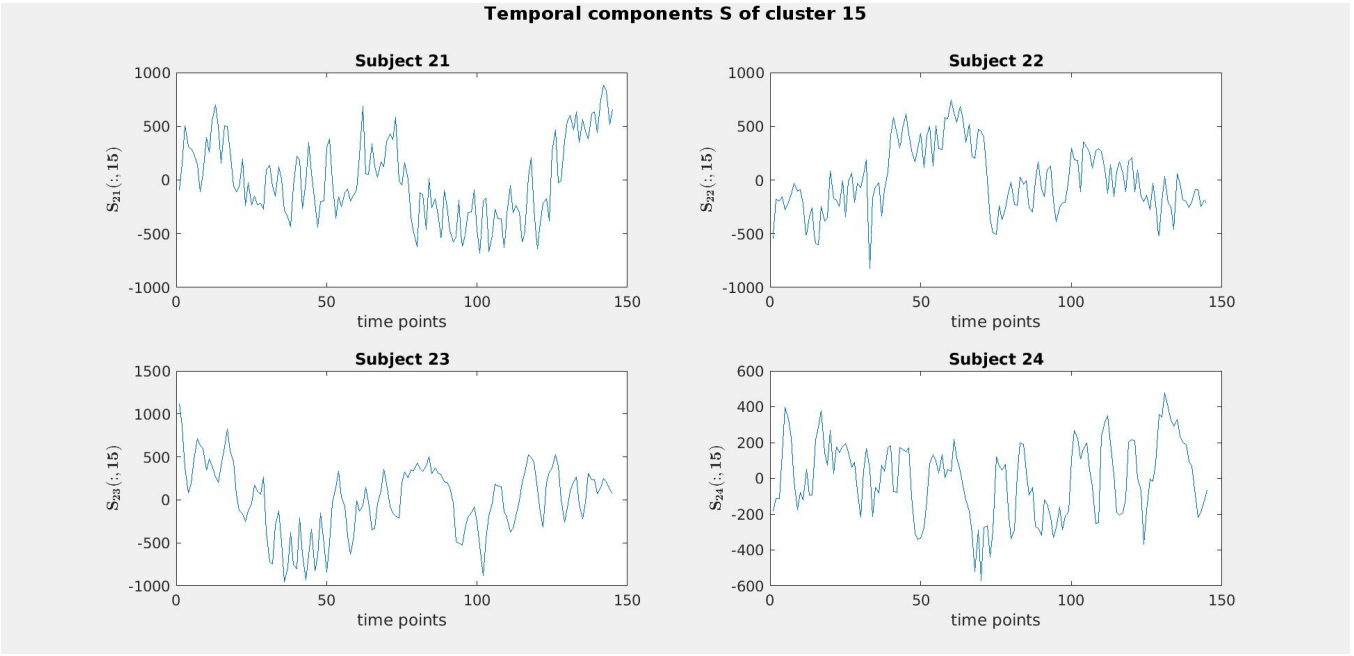


Figure B.571: Waveform of temporal components $\mathbf{S}_k(:, 115)$ for $k = 21, \ldots, 24$ and R=42.

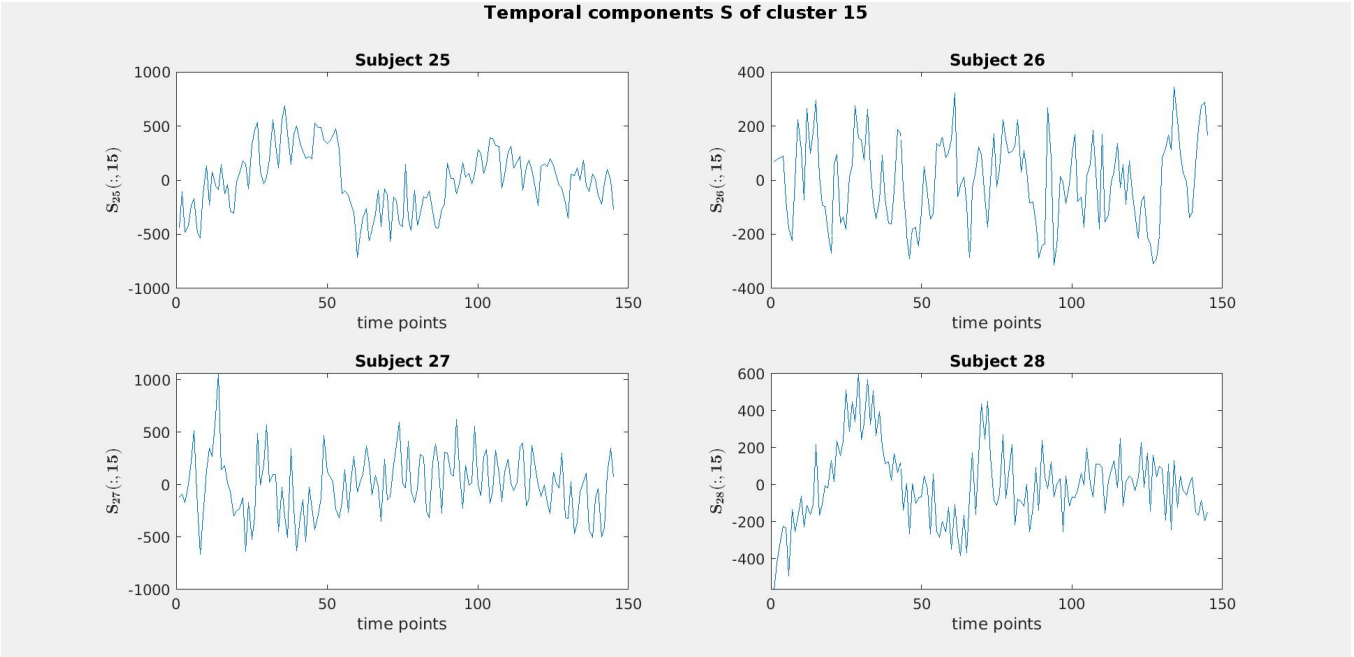


Figure B.572: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 25, \ldots, 28$ and R=42.

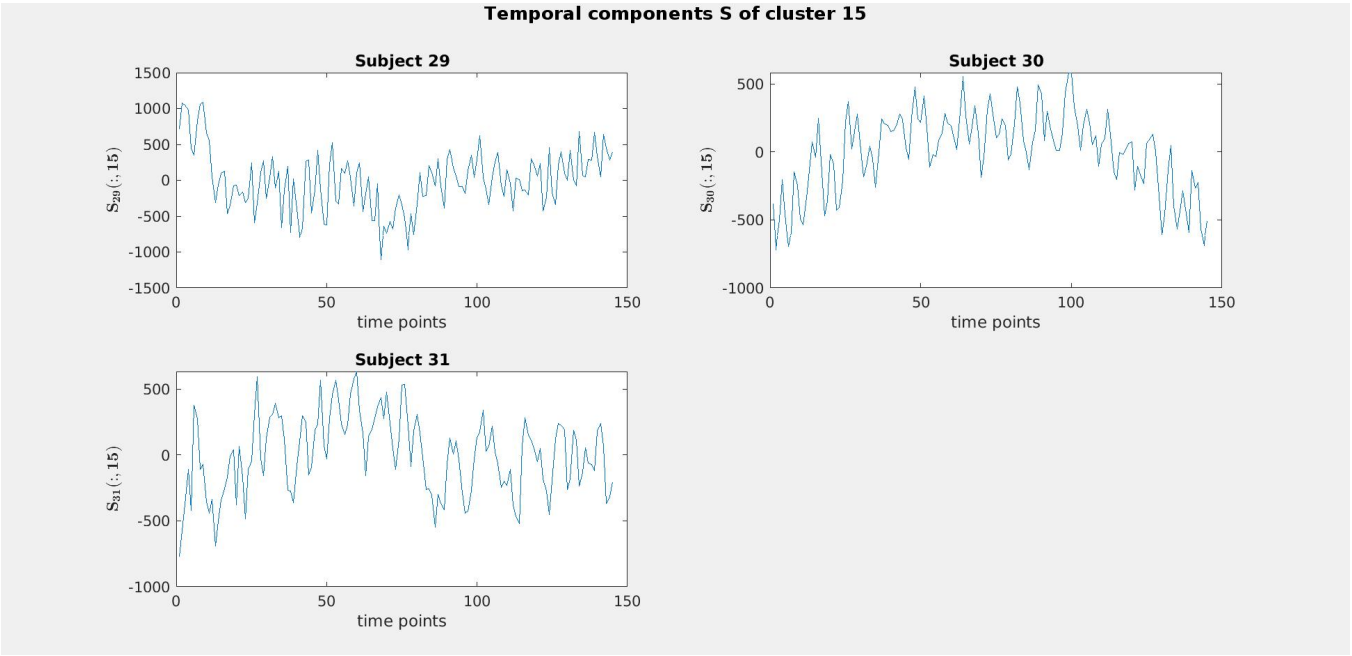


Figure B.573: Waveform of temporal components $\mathbf{S}_k(:, 15)$ for $k = 29, \ldots, 31$ and R=42.

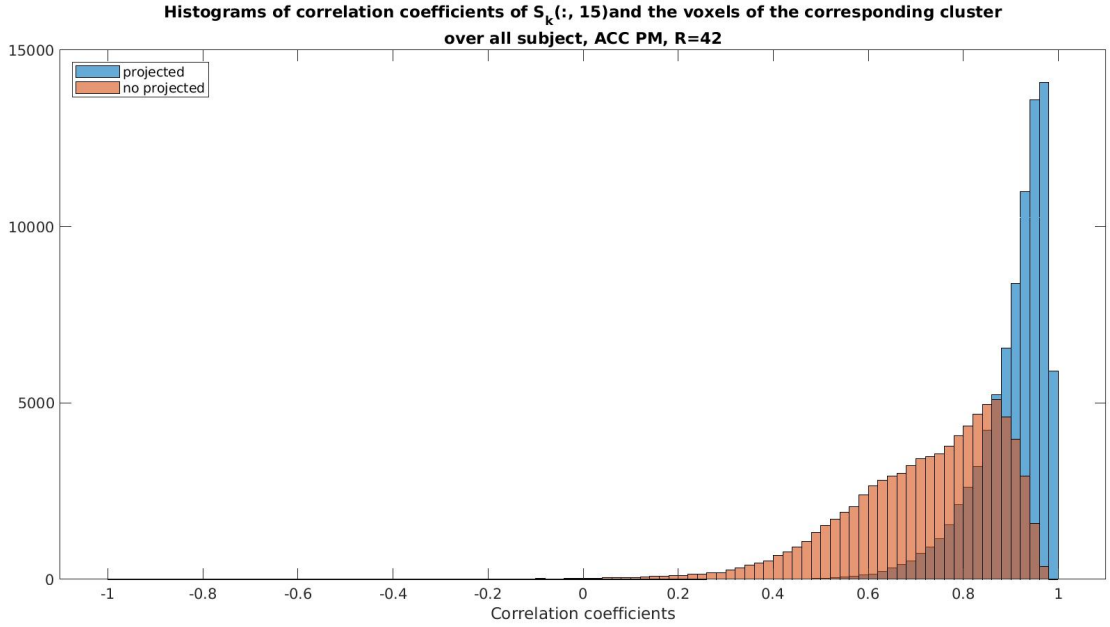


Figure B.574: For the column 15 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 15)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

29. **Cluster 6**

Cluster 6 contains 3312 voxels, which correspond to 1.465824% of the whole brain (background, soft tissue, bone are excluded).

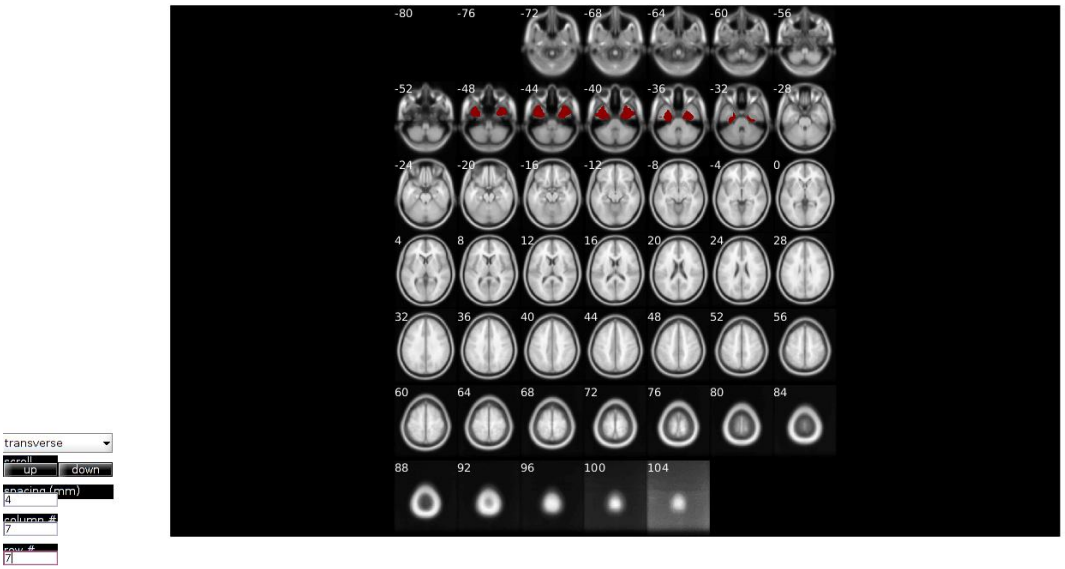


Figure B.575: Cluster 6. Red areas indicate the voxels which belong to cluster 6.

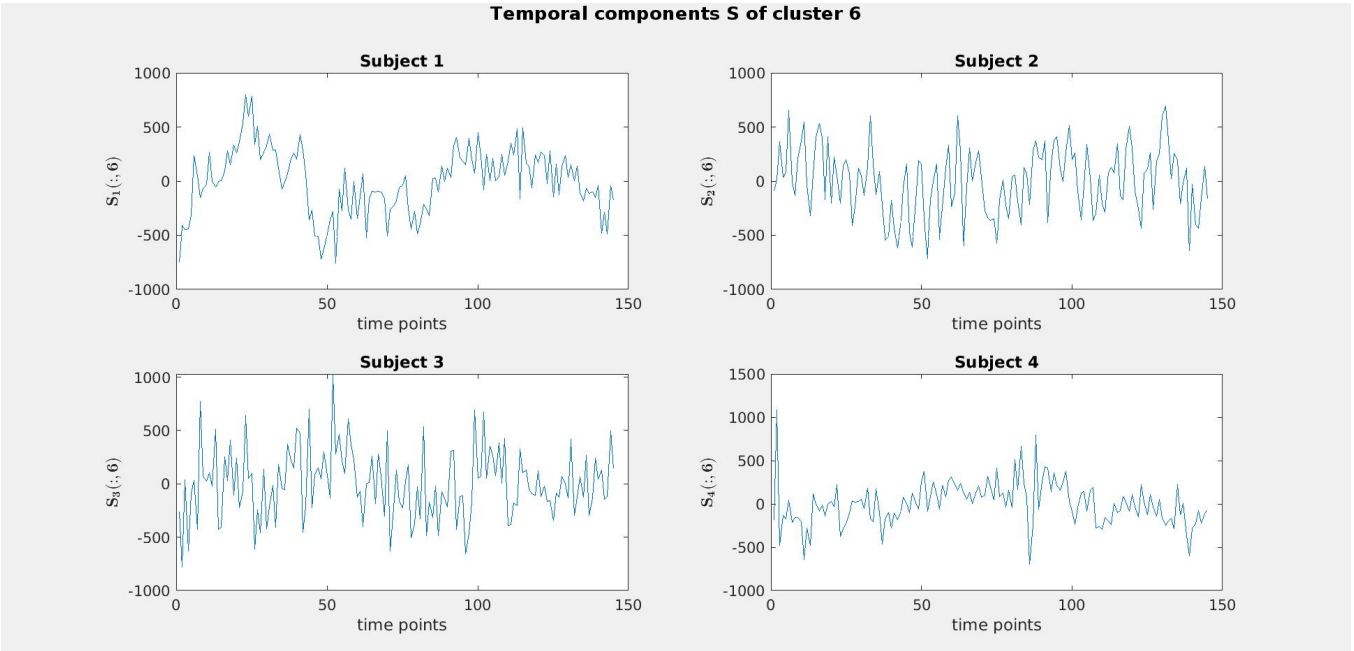


Figure B.576: Waveform of temporal components $\mathbf{S}_k(:,6)$ for $k = 1,\ldots,4$ and R=42.

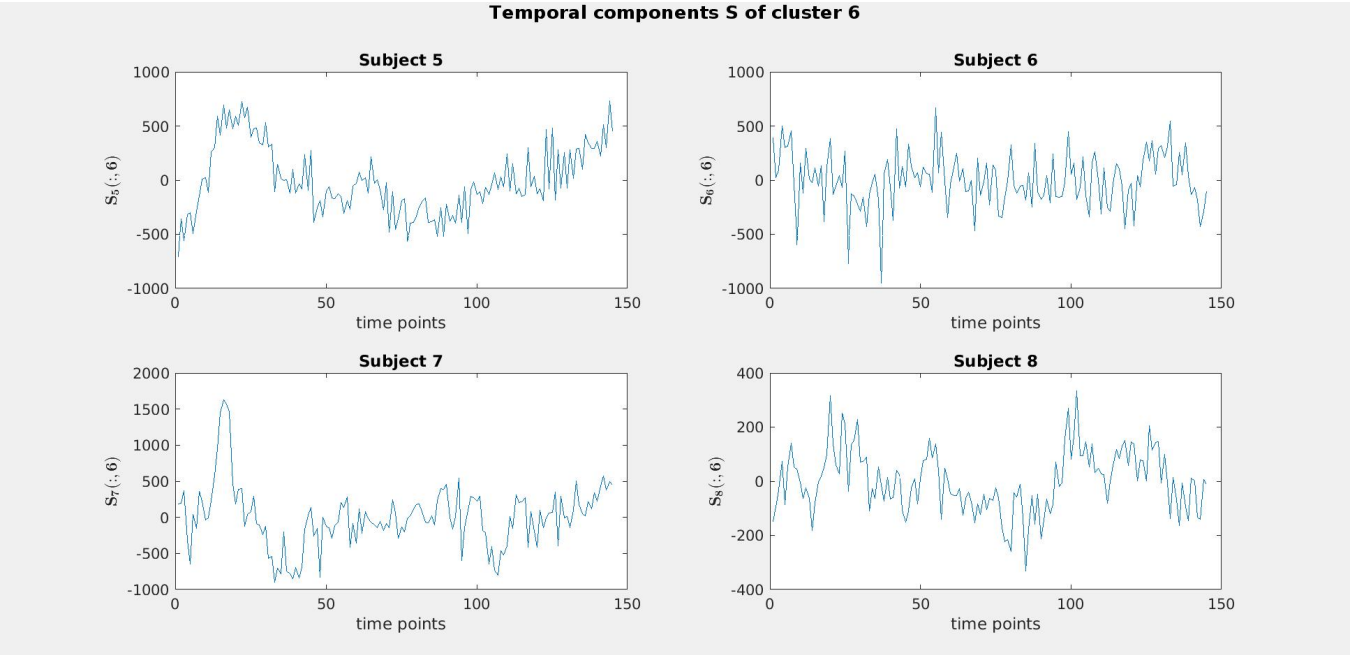


Figure B.577: Waveform of temporal components $\mathbf{S}_k(:,6)$ for $k = 5,\ldots,8$ and R=42.

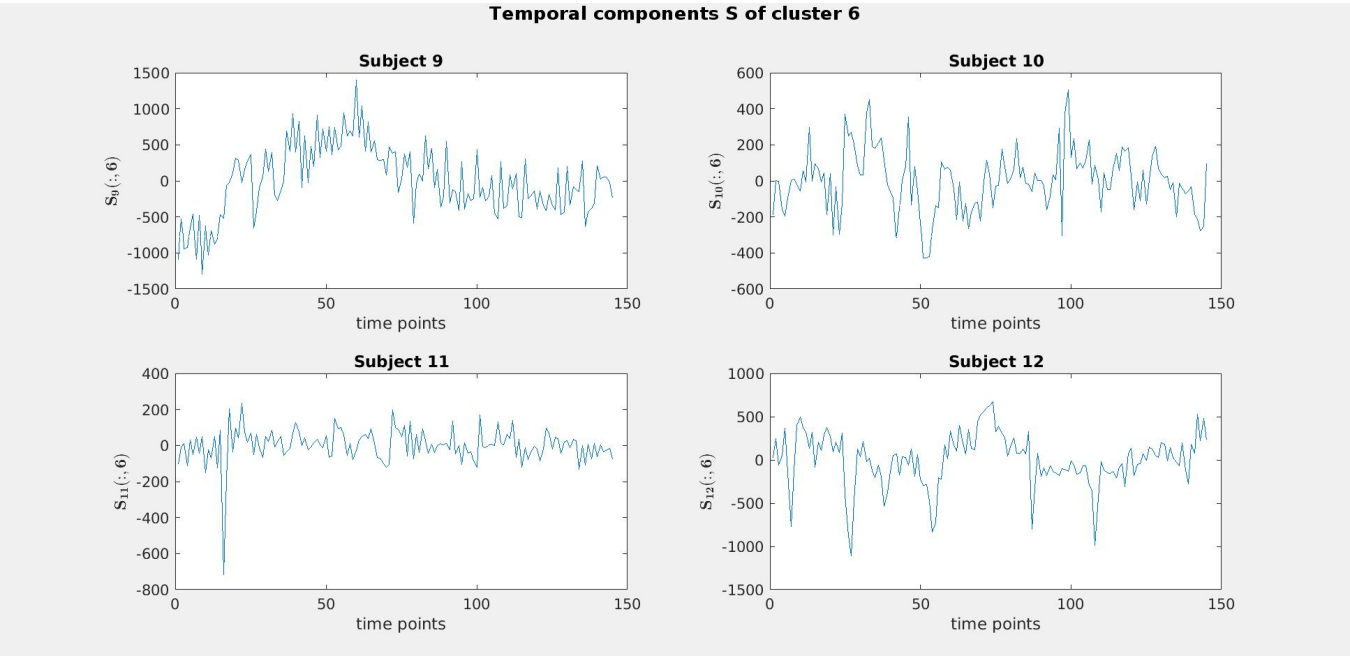


Figure B.578: Waveform of temporal components $\mathbf{S}_k(:,6)$ for $k = 9,\ldots,12$ and R=42.

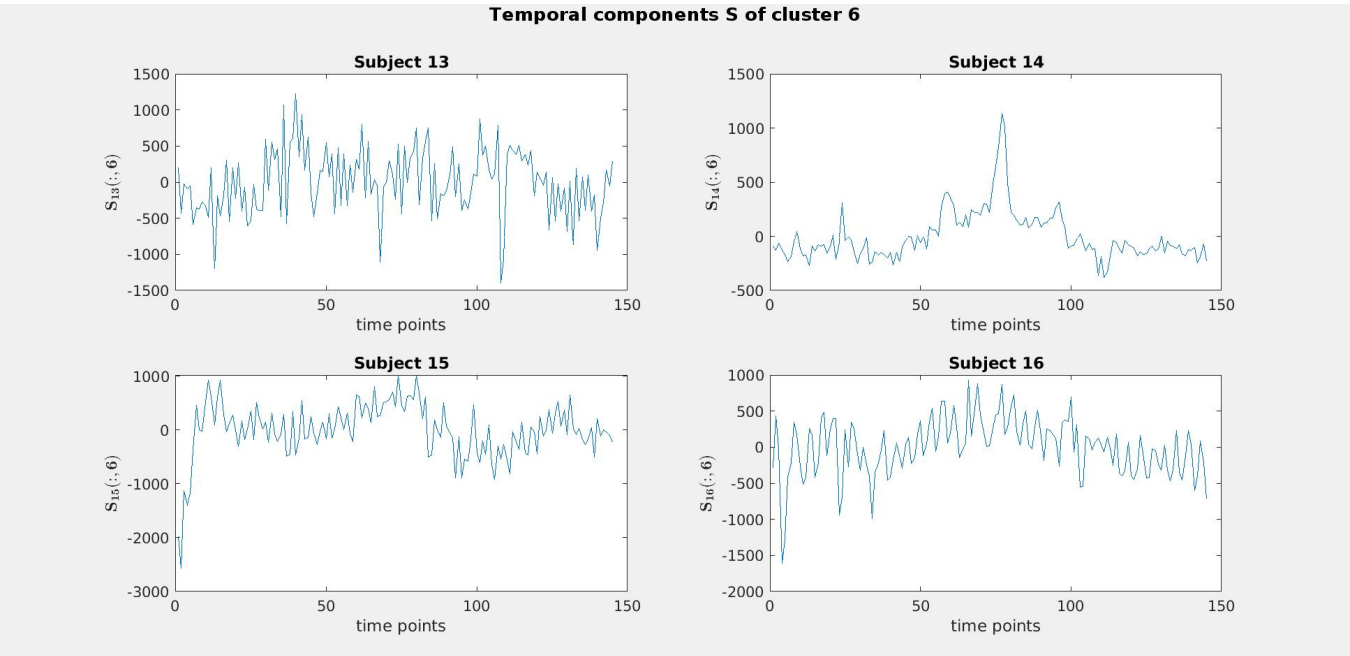


Figure B.579: Waveform of temporal components $\mathbf{S}_k(:,6)$ for $k = 13,\ldots,16$ and R=42.

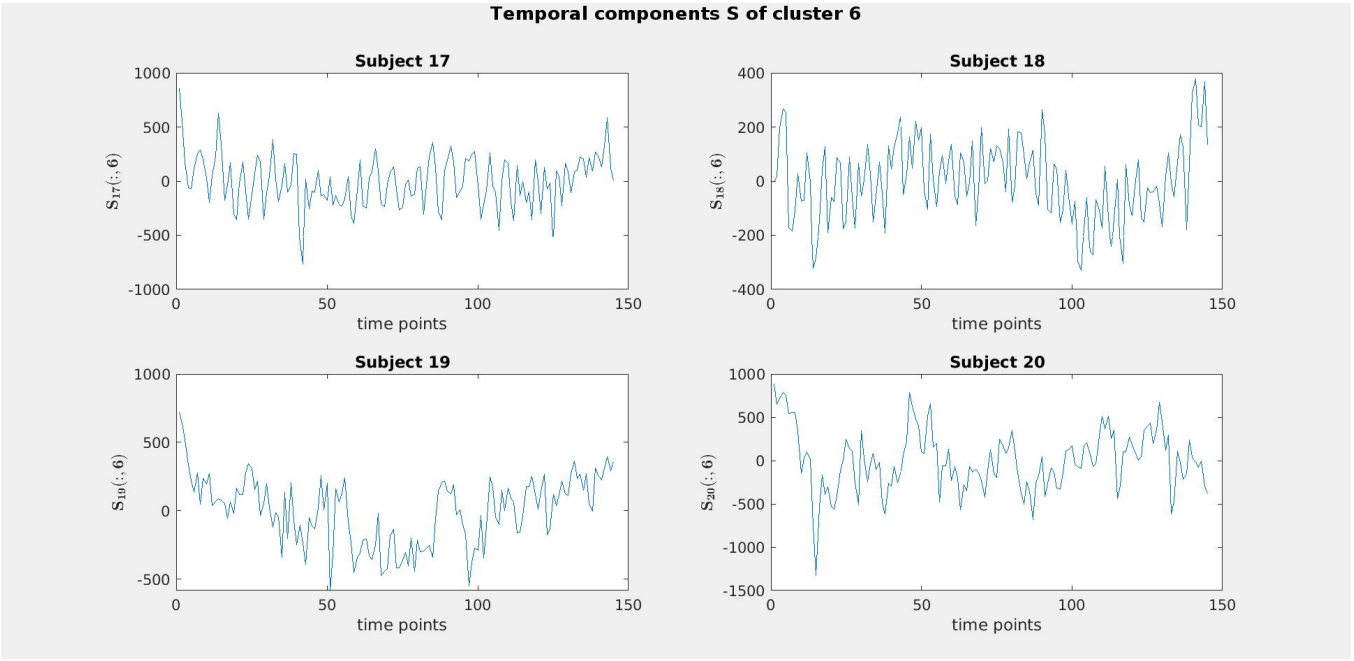


Figure B.580: Waveform of temporal components $\mathbf{S}_k(:, 6)$ for $k = 17, \ldots, 20$ and R=42.

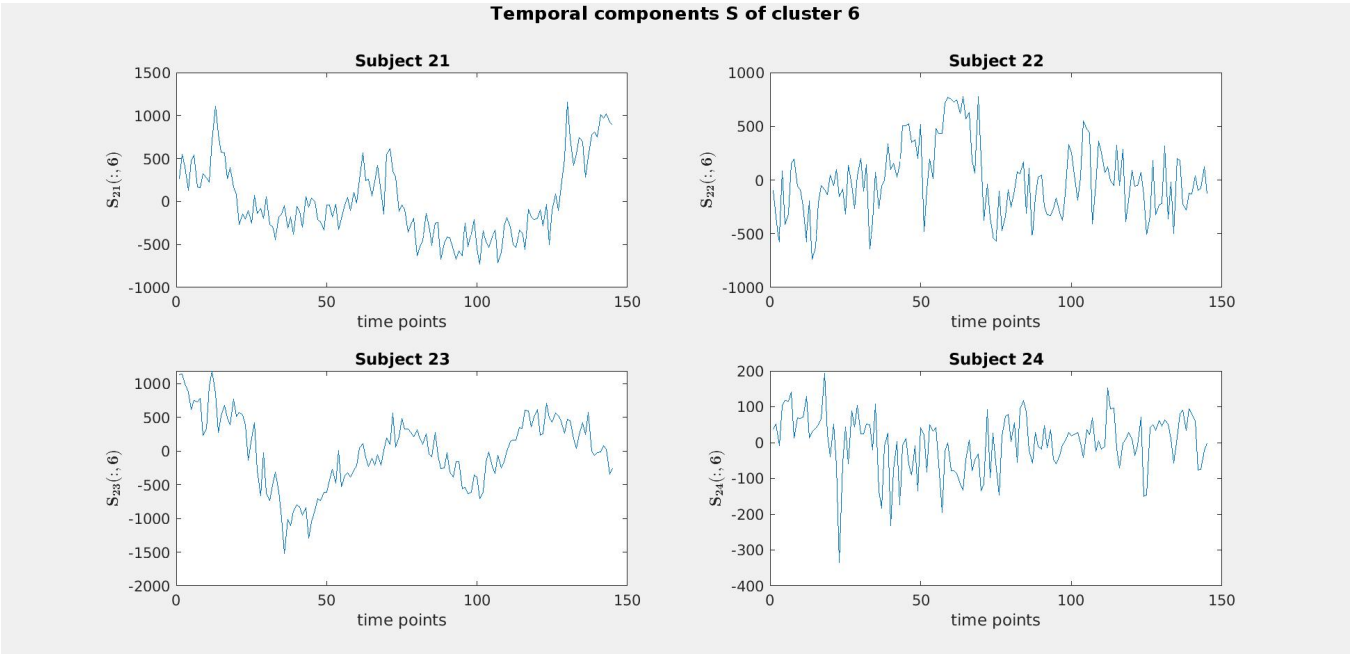


Figure B.581: Waveform of temporal components $\mathbf{S}_k(:, 16)$ for $k = 21, \ldots, 24$ and R=42.

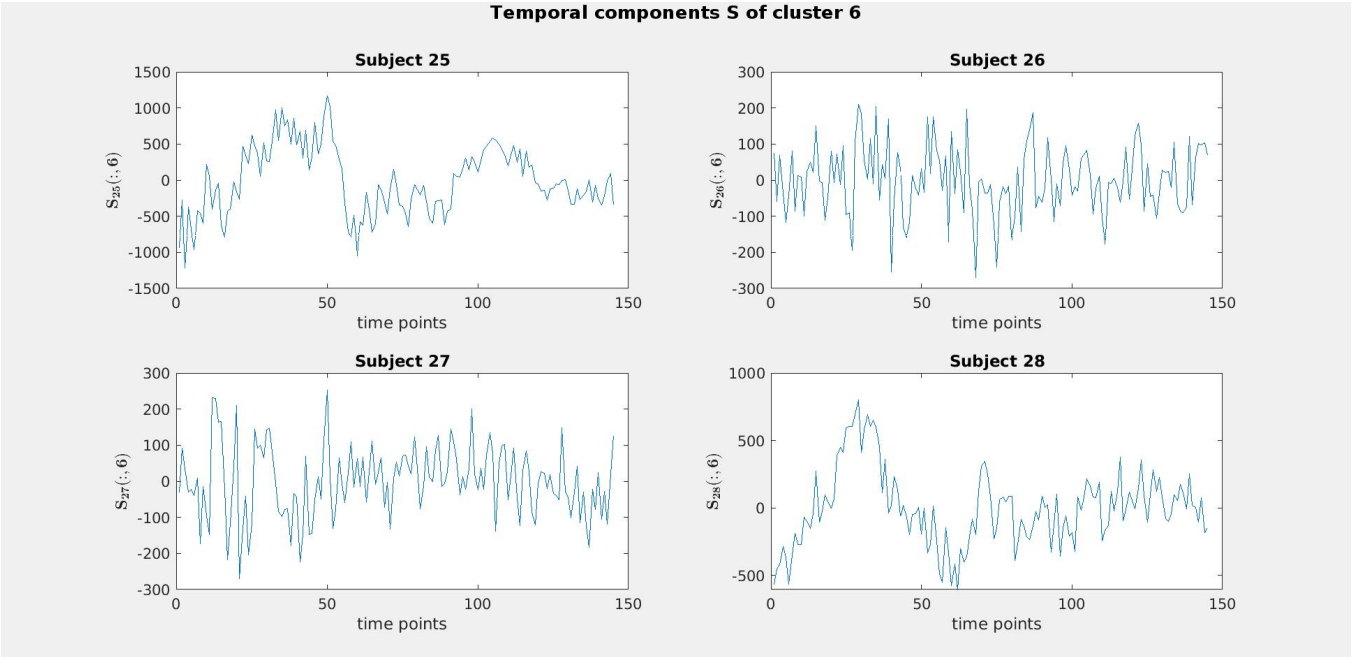


Figure B.582: Waveform of temporal components $\mathbf{S}_k(:, 6)$ for $k = 25, \ldots, 28$ and R=42.

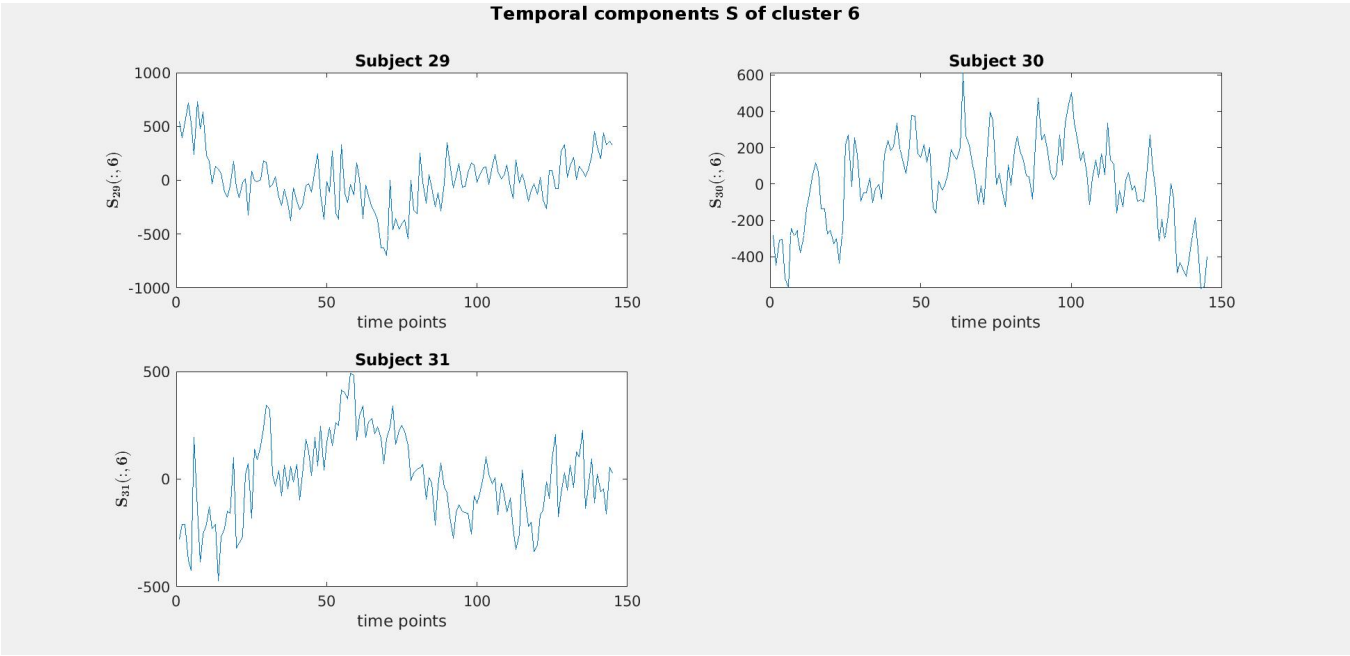


Figure B.583: Waveform of temporal components $\mathbf{S}_k(:, 6)$ for $k = 29, \ldots, 31$ and R=42.

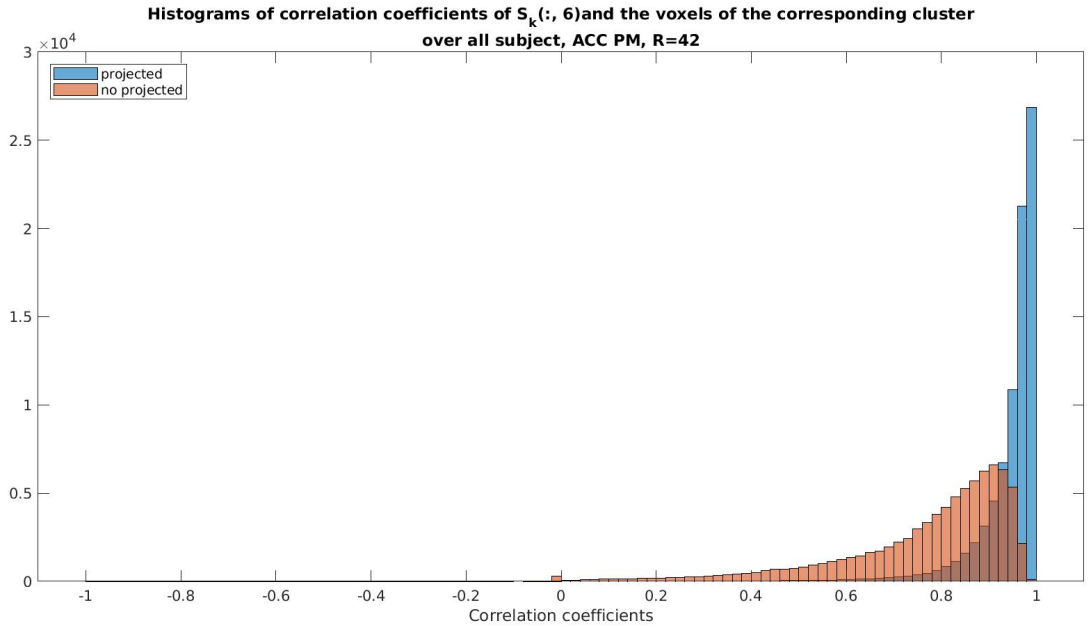


Figure B.584: For the column 6 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 6)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

30. **Cluster 39**

Cluster 39 contains 3179 voxels, which correspond to 1.406961% of the whole brain (background, soft tissue, bone are excluded).

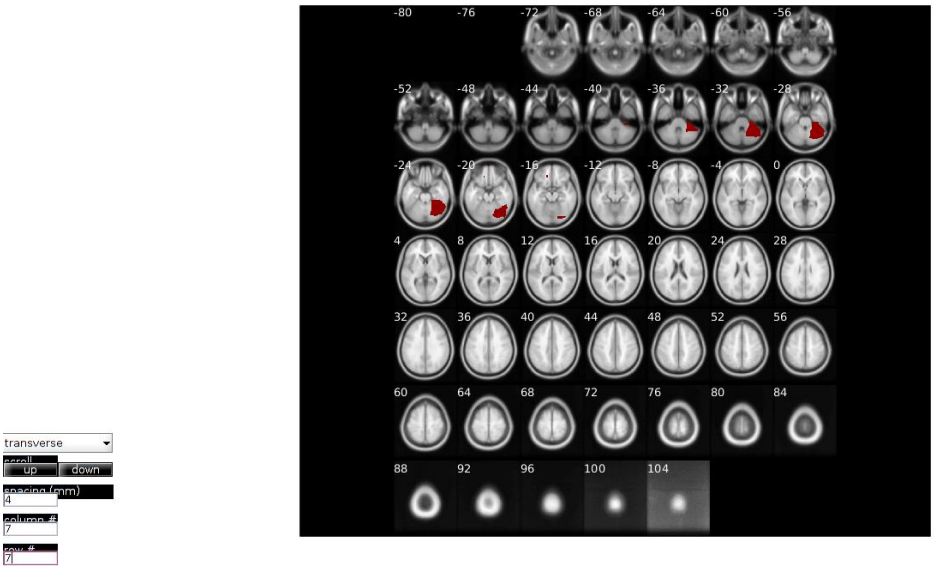


Figure B.585: Cluster 39. Red areas indicate the voxels which belong to cluster 39.

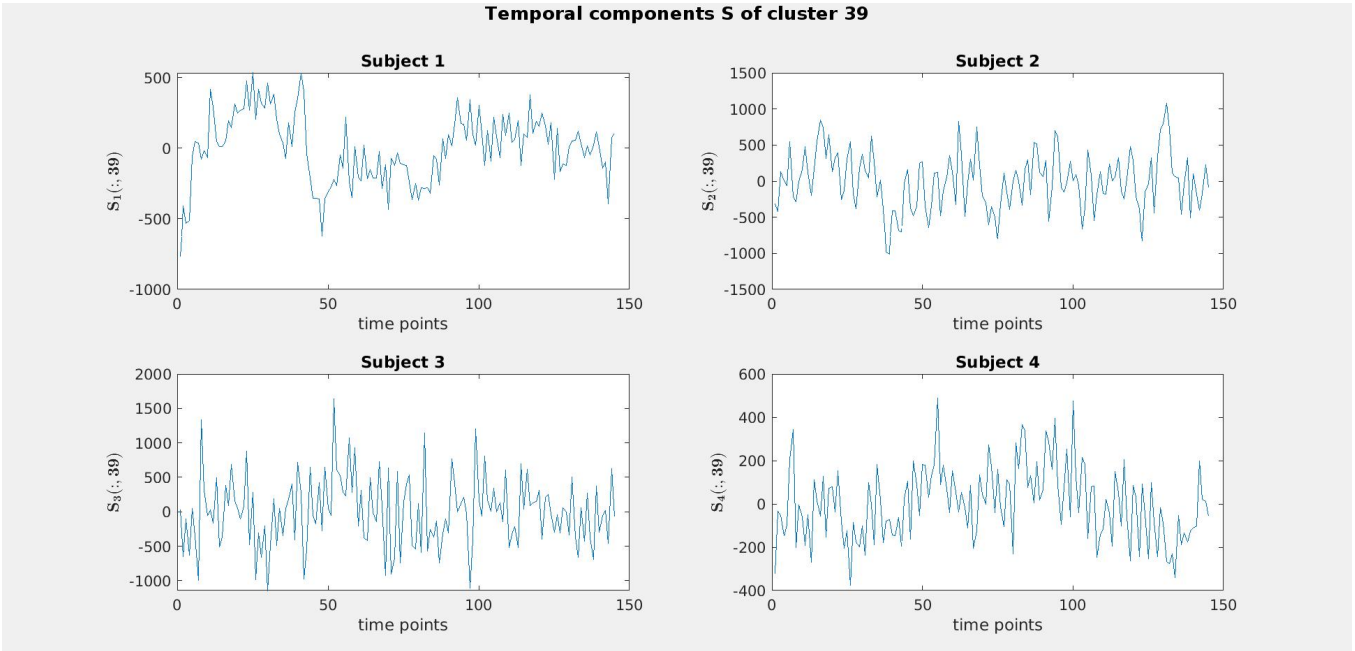


Figure B.586: Waveform of temporal components $\mathbf{S}_k(:, 39)$ for $k = 1, \ldots, 4$ and R=42.

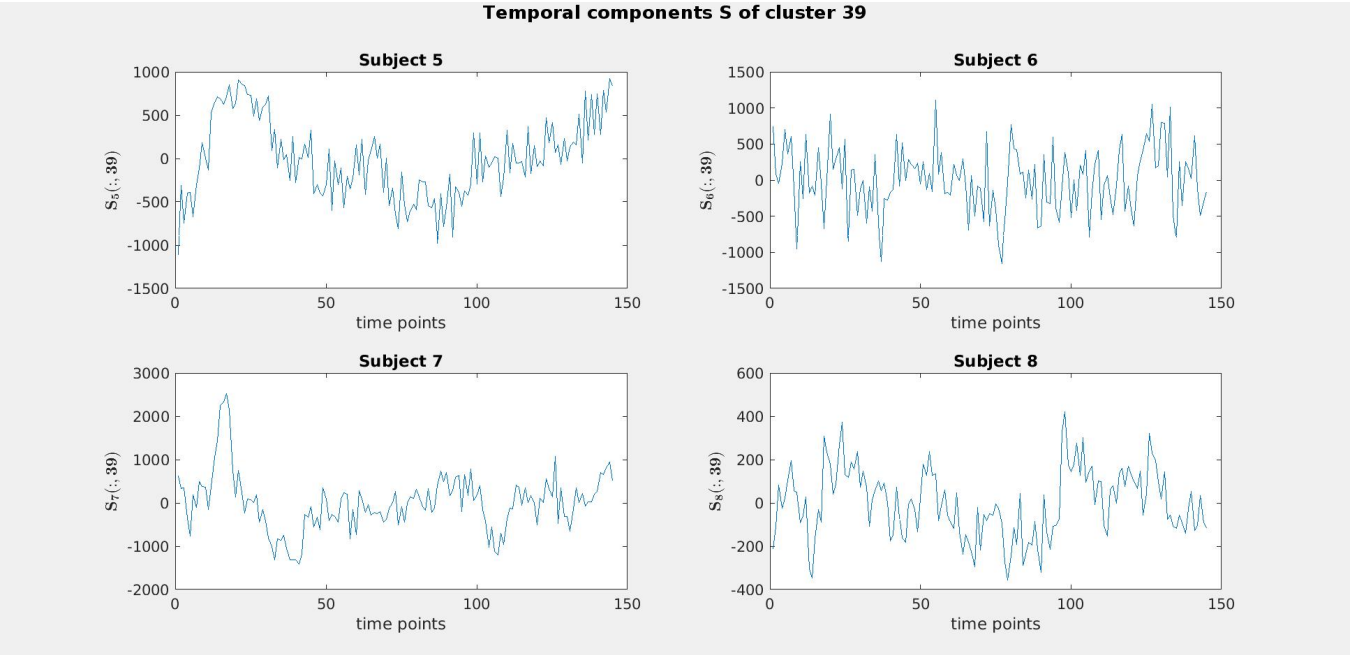


Figure B.587: Waveform of temporal components $\mathbf{S}_k(:, 39)$ for $k = 5, \ldots, 8$ and R=42.

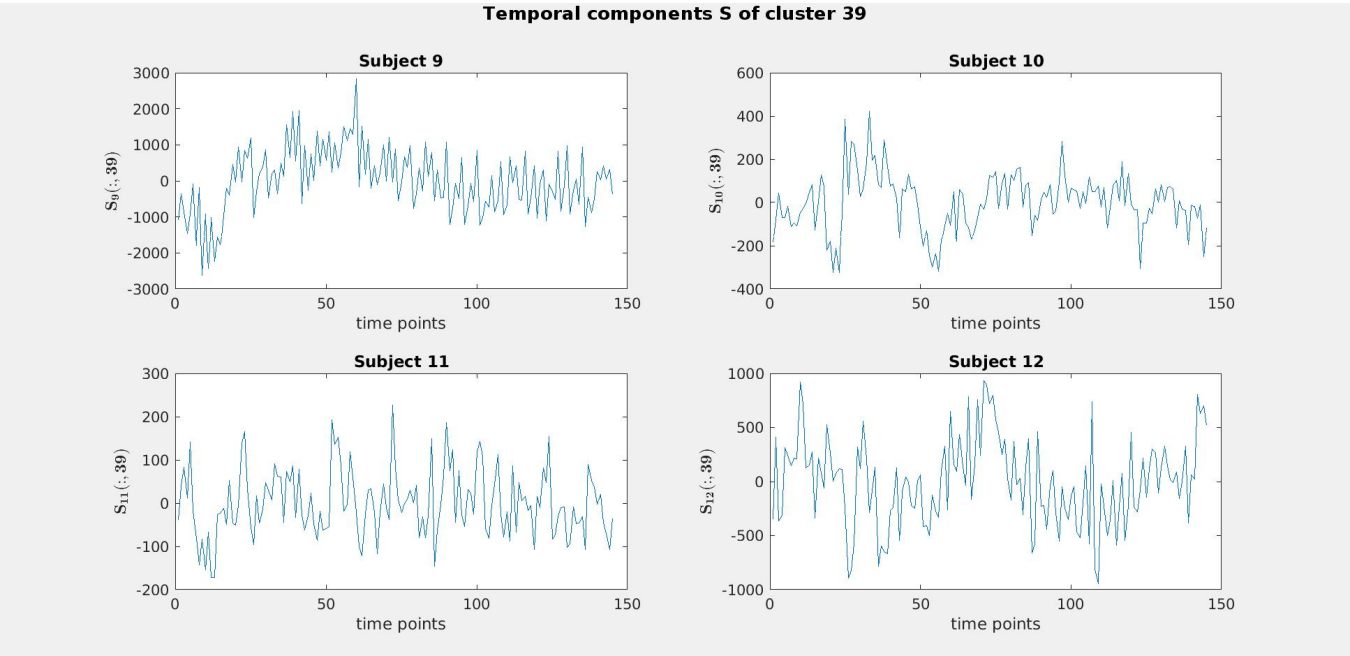


Figure B.588: Waveform of temporal components $\mathbf{S}_k(:, 39)$ for $k = 9, \ldots, 12$ and R=42.

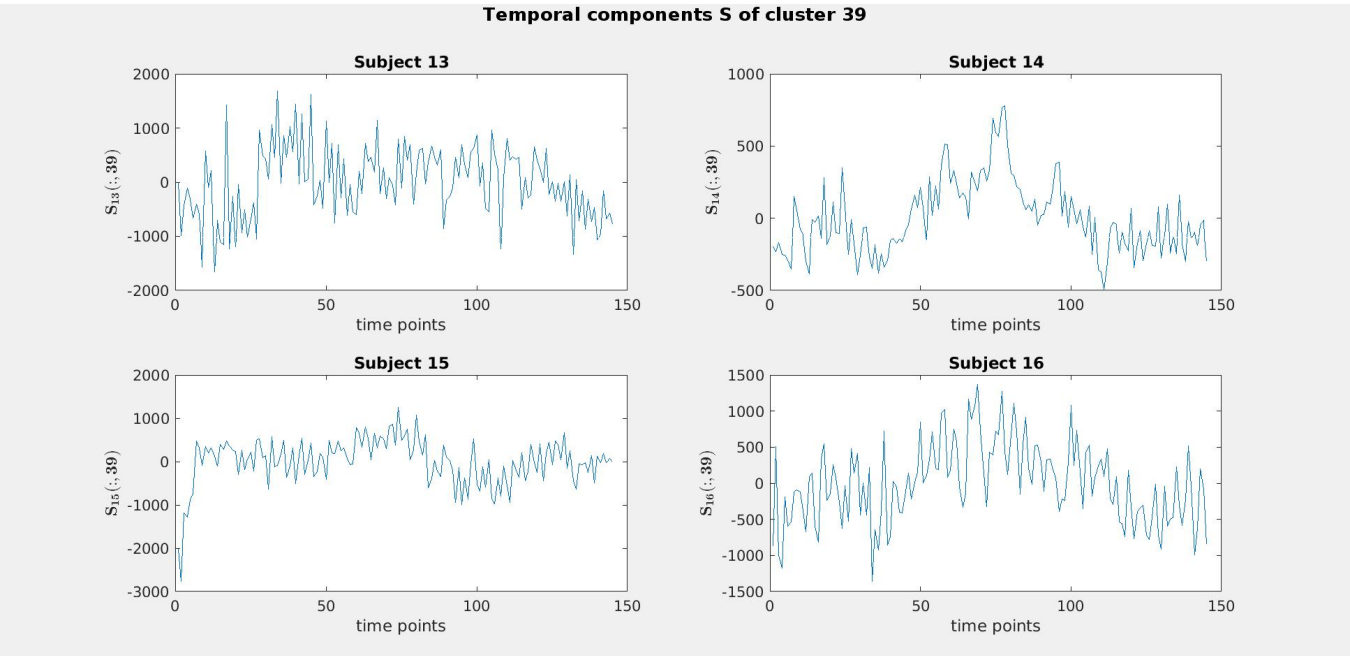


Figure B.589: Waveform of temporal components $\mathbf{S}_k(:, 39)$ for $k = 13, \ldots, 16$ and R=42.

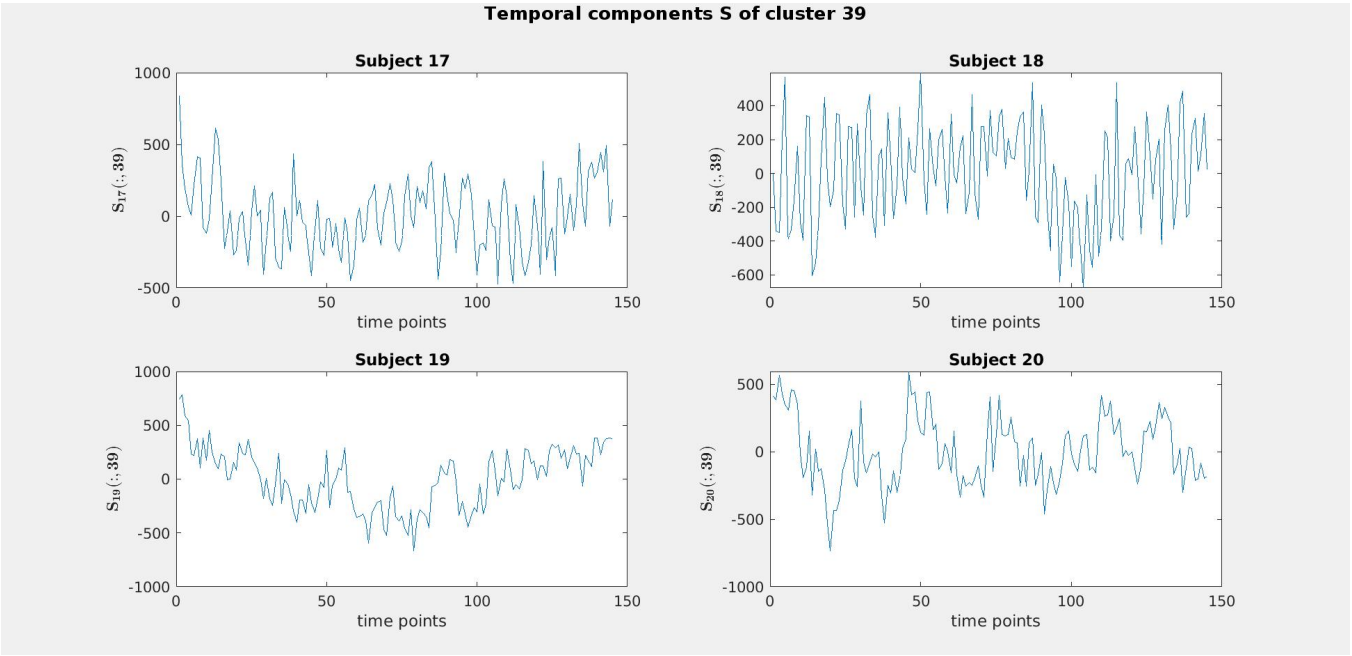


Figure B.590: Waveform of temporal components $\mathbf{S}_k(:, 39)$ for $k = 17, \ldots, 20$ and R=42.

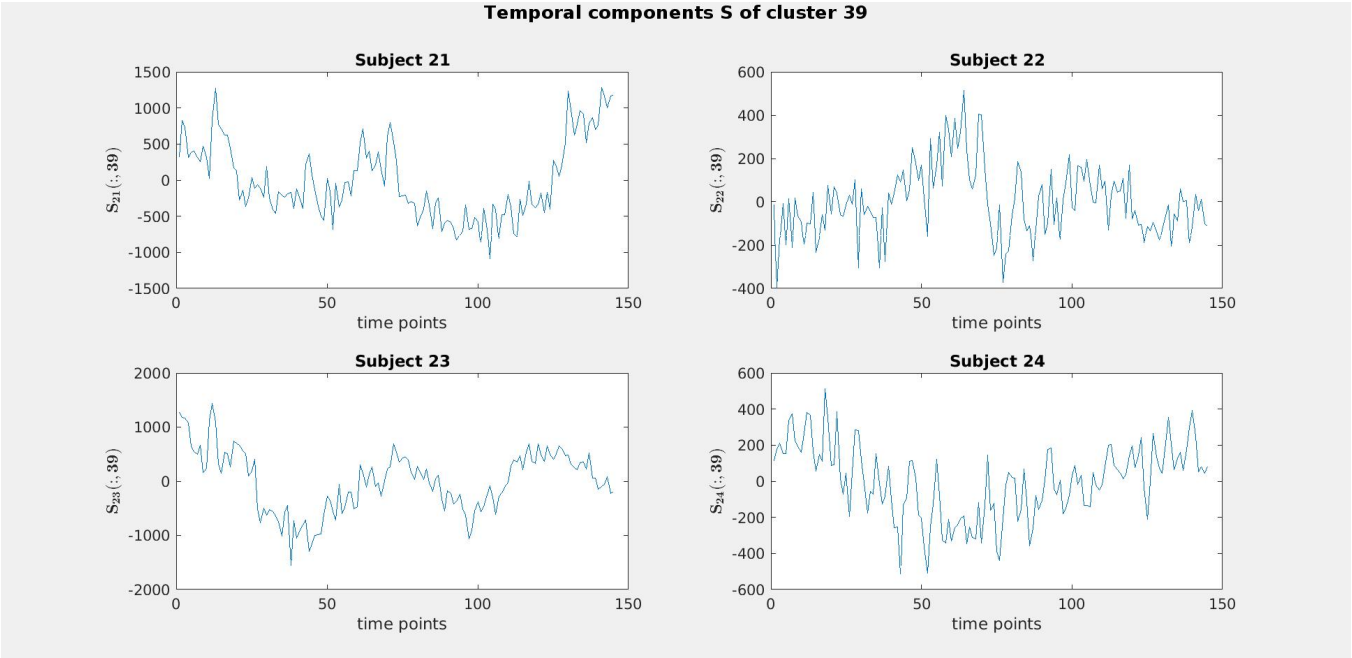


Figure B.591: Waveform of temporal components $\mathbf{S}_k(:, 139)$ for $k = 21, \ldots, 24$ and R=42.

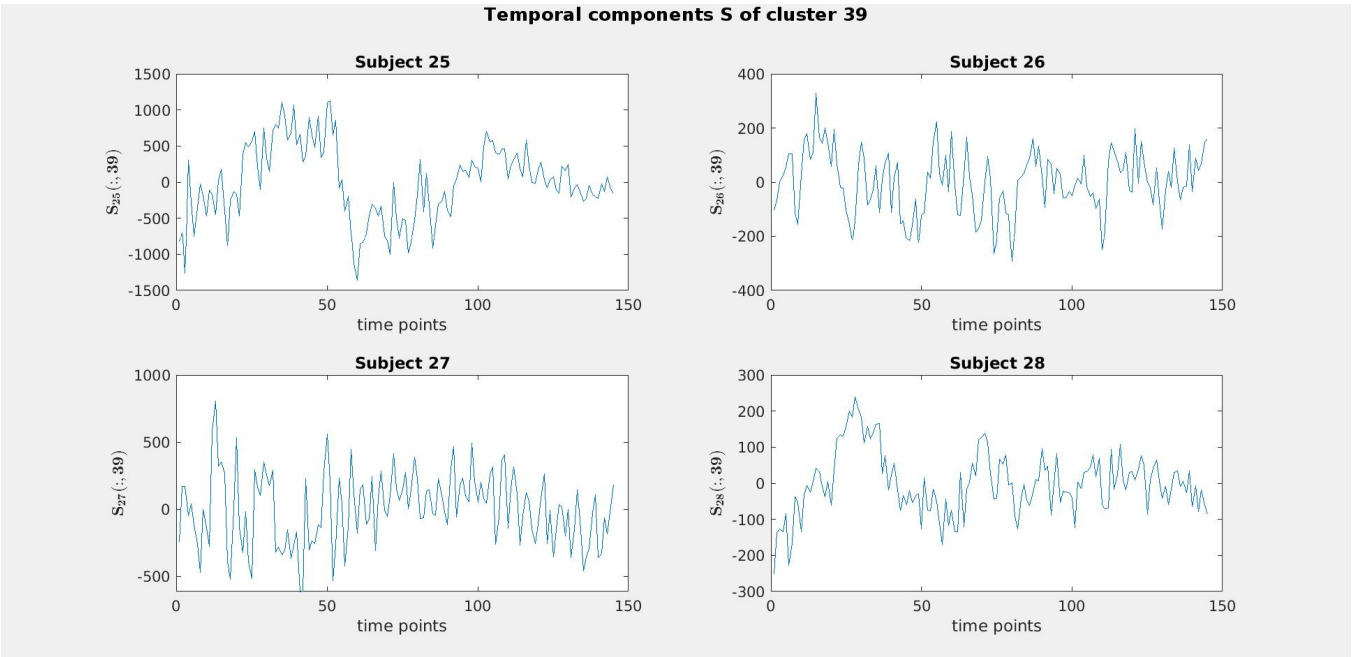


Figure B.592: Waveform of temporal components $\mathbf{S}_k(:, 39)$ for $k = 25, \ldots, 28$ and R=42.

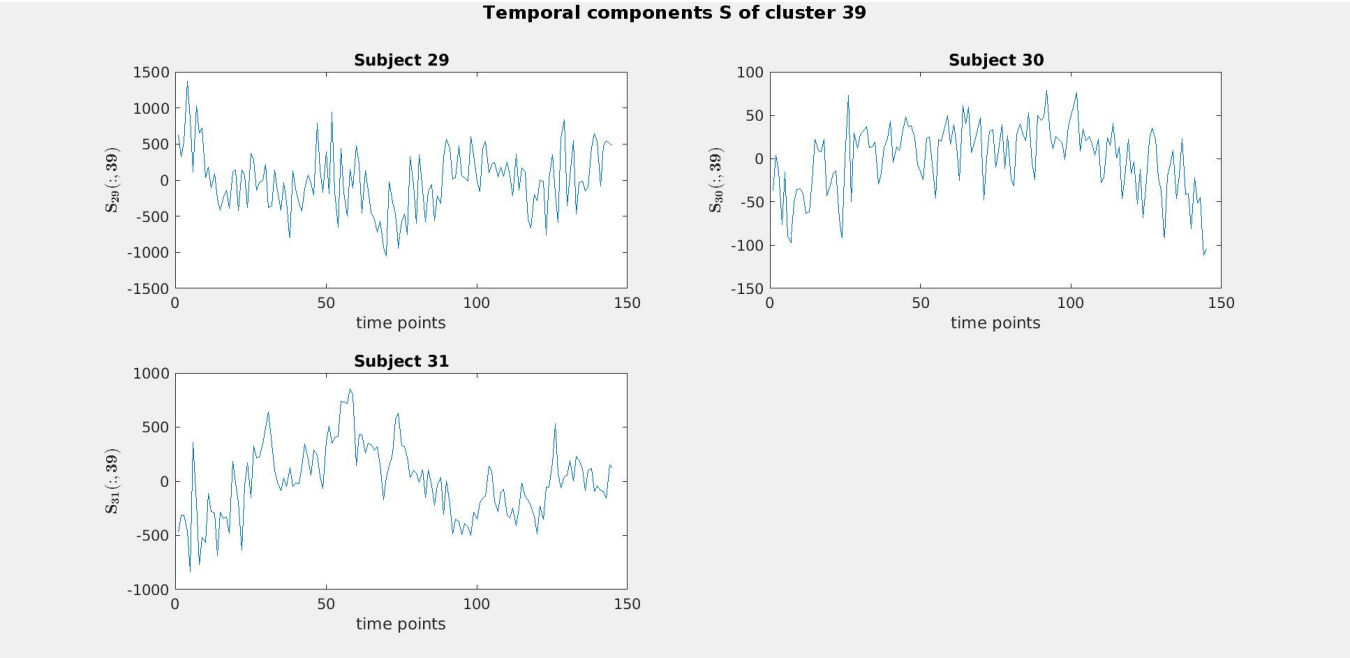


Figure B.593: Waveform of temporal components $\mathbf{S}_k(:, 39)$ for $k = 29, \ldots, 31$ and R=42.

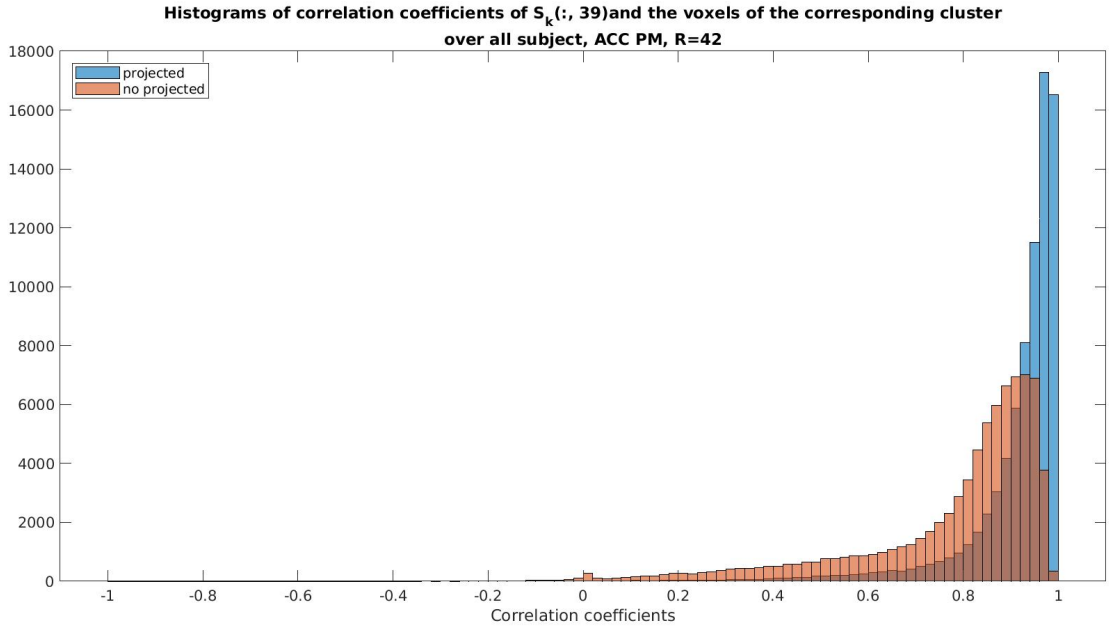


Figure B.594: For the column 39 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 39)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

31. **Cluster 31**

Cluster 31 contains 3032 voxels, which correspond to 1.341902% of the whole brain (background, soft tissue, bone are excluded).

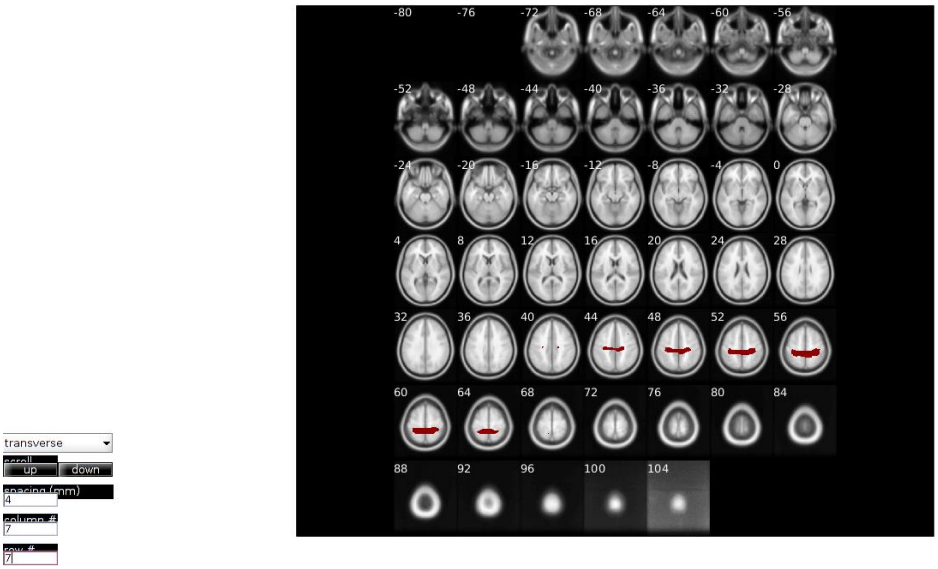


Figure B.595: Cluster 31. Red areas indicate the voxels which belong to cluster 31.

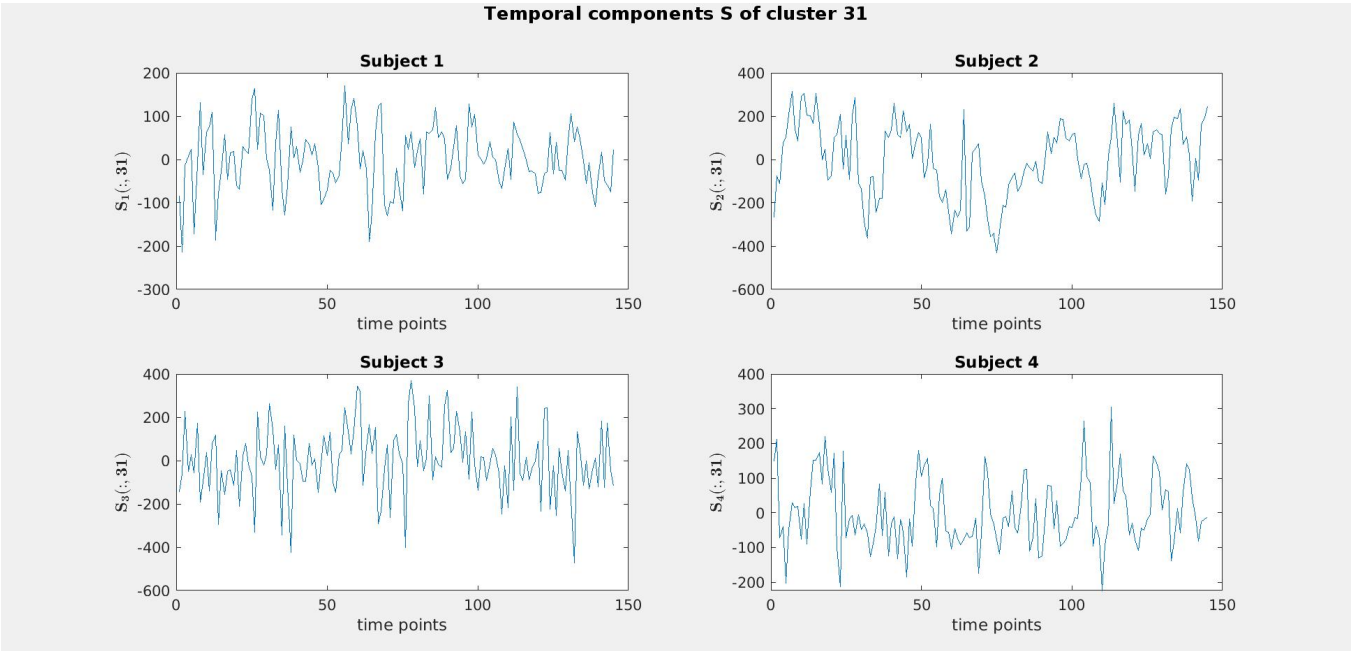


Figure B.596: Waveform of temporal components $\mathbf{S}_k(:, 31)$ for $k = 1, \ldots, 4$ and R=42.

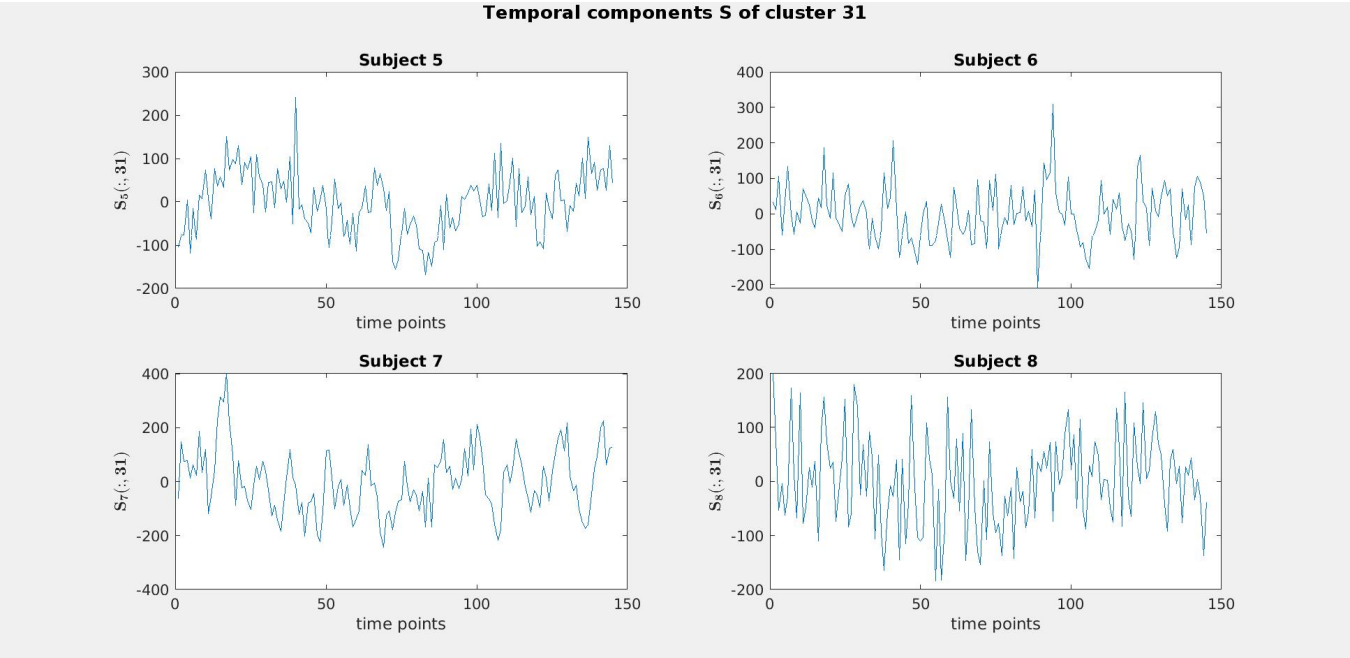


Figure B.597: Waveform of temporal components $\mathbf{S}_k(:, 31)$ for $k = 5, \ldots, 8$ and R=42.

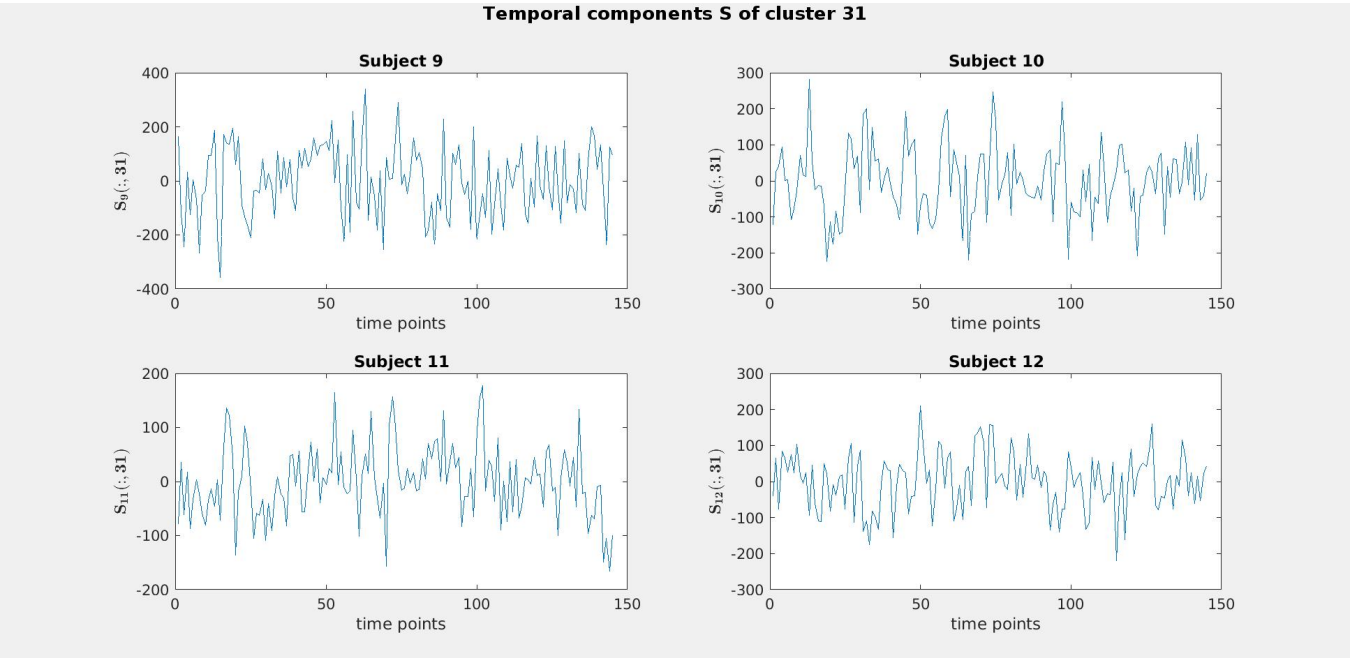


Figure B.598: Waveform of temporal components $\mathbf{S}_k(:, 31)$ for $k = 9, \ldots, 12$ and R=42.

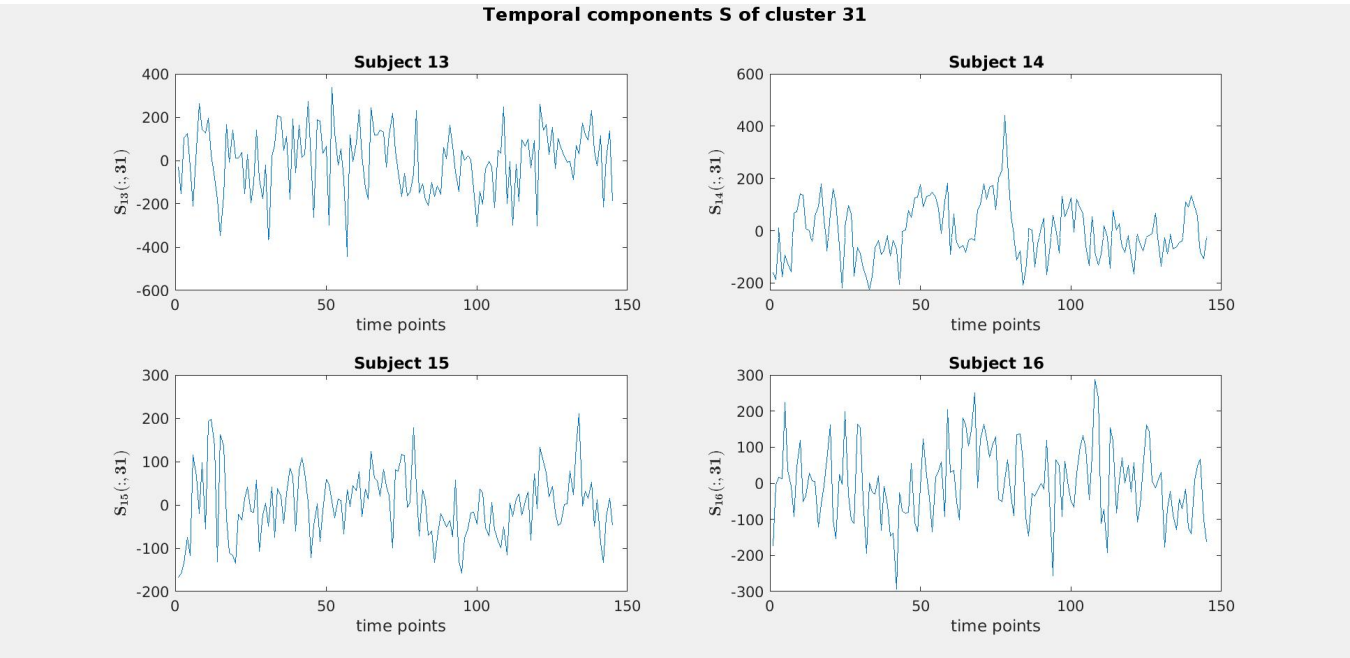


Figure B.599: Waveform of temporal components $\mathbf{S}_k(:, 31)$ for $k = 13, \ldots, 16$ and R=42.

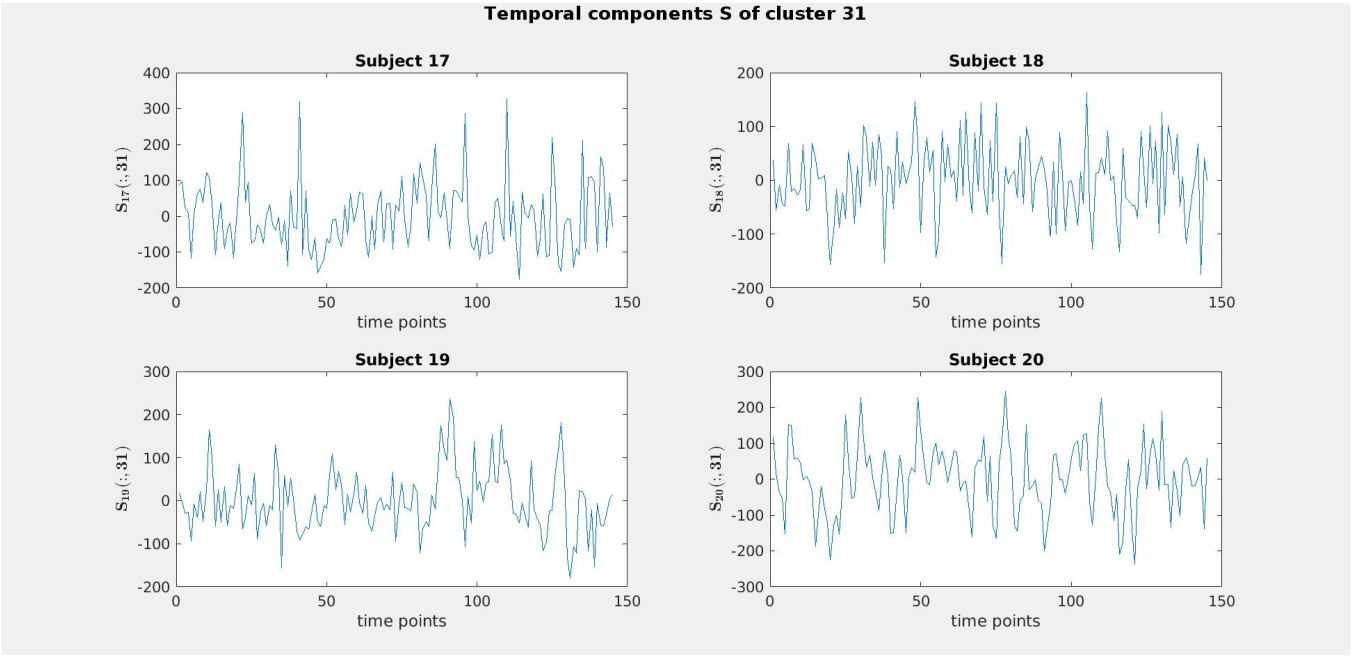


Figure B.600: Waveform of temporal components $\mathbf{S}_k(:, 31)$ for $k = 17, \ldots, 20$ and R=42.

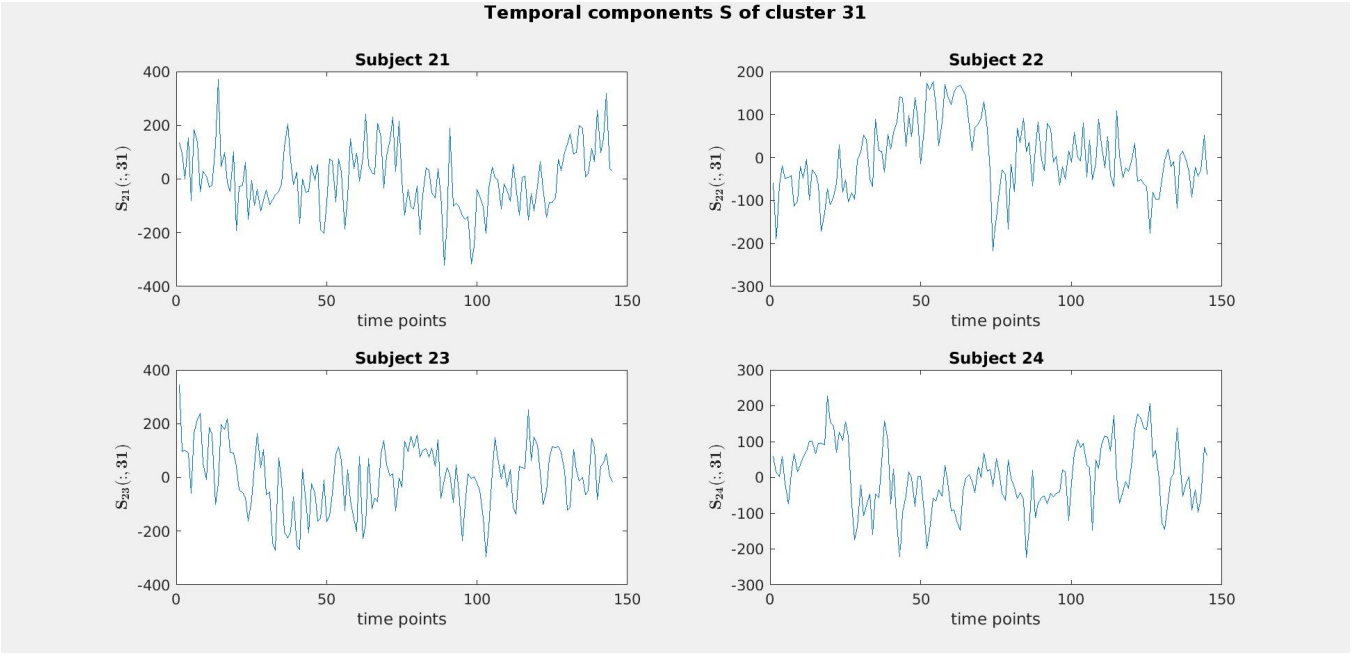


Figure B.601: Waveform of temporal components $\mathbf{S}_k(:, 131)$ for $k = 21, \ldots, 24$ and R=42.

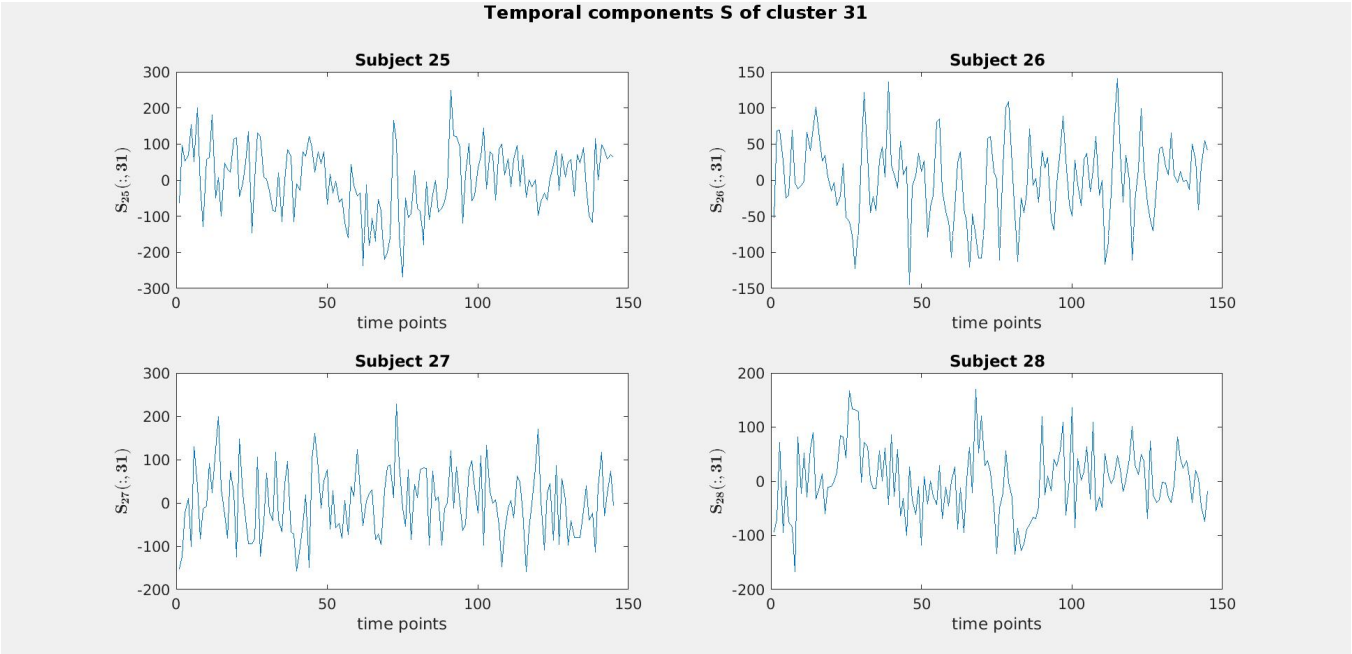


Figure B.602: Waveform of temporal components $\mathbf{S}_k(:, 31)$ for $k = 25, \ldots, 28$ and R=42.

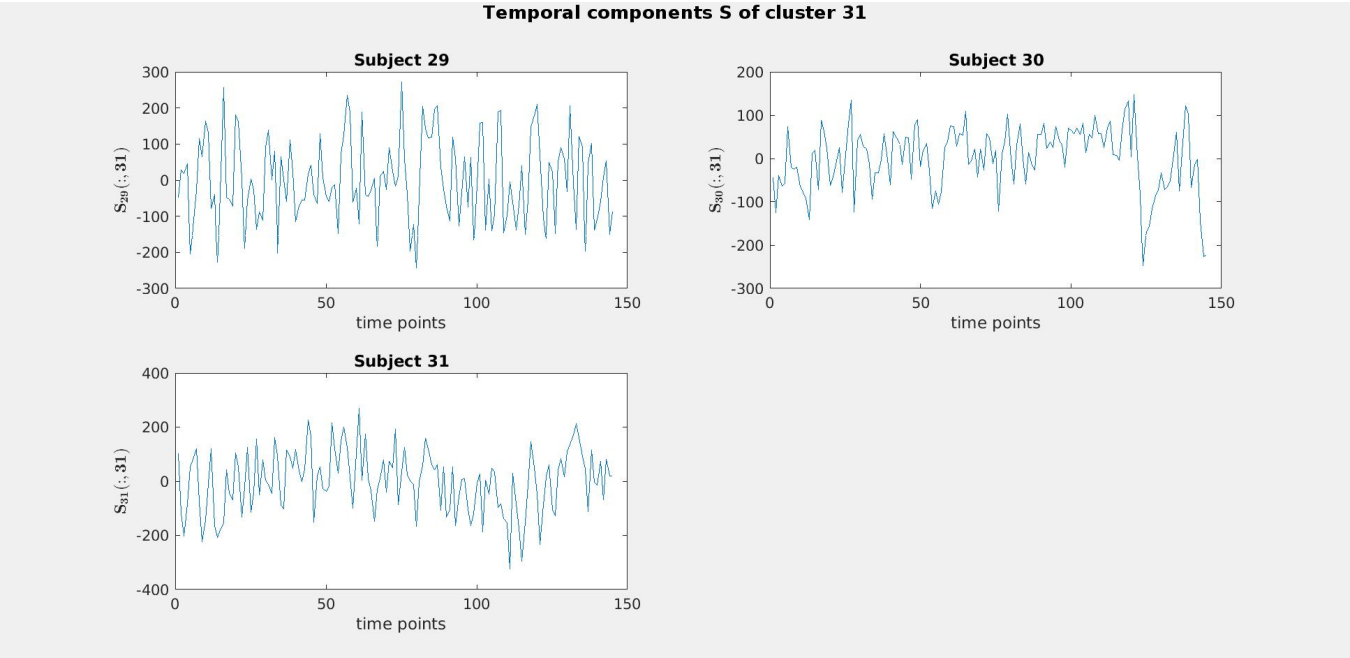


Figure B.603: Waveform of temporal components $\mathbf{S}_k(:, 31)$ for $k = 29, \ldots, 31$ and R=42.

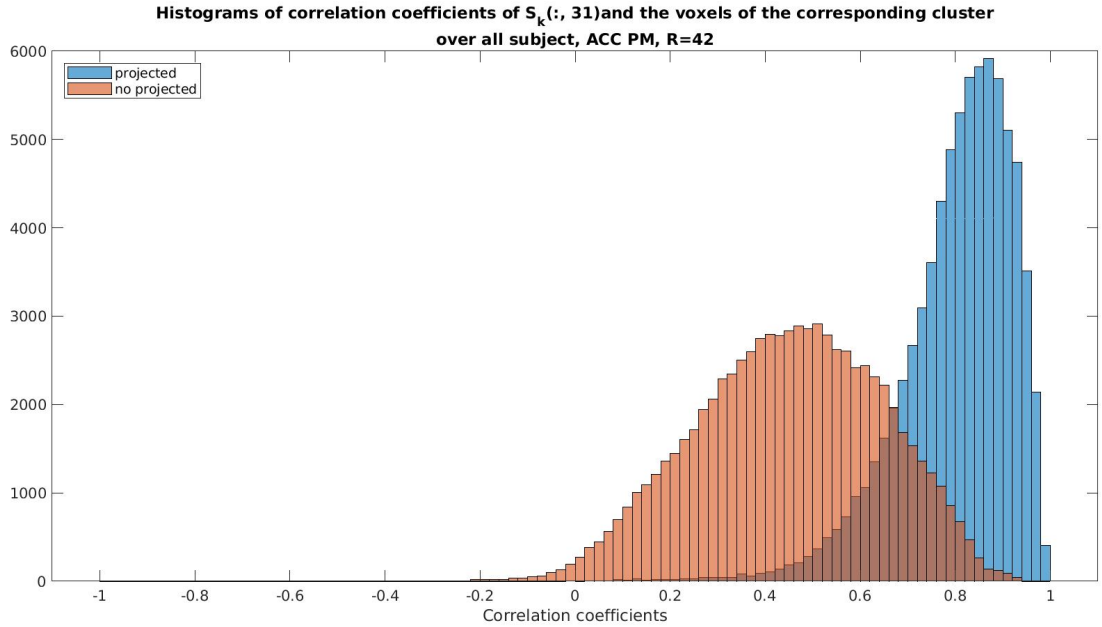


Figure B.604: For the column 31 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 31)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

32. **Cluster 38**

Cluster 38 contains 2988 voxels, which correspond to 1.322428% of the whole brain (background, soft tissue, bone are excluded).

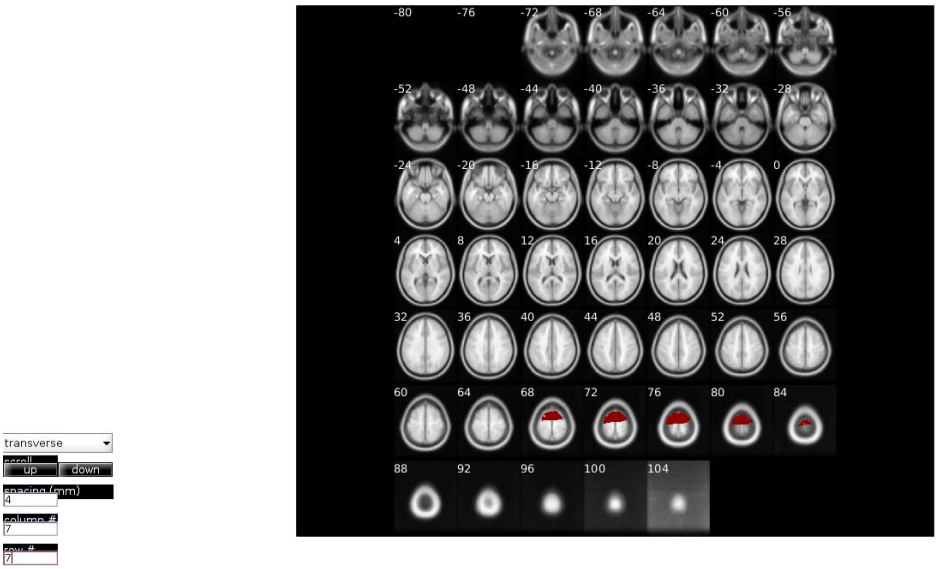


Figure B.605: Cluster 38. Red areas indicate the voxels which belong to cluster 38.

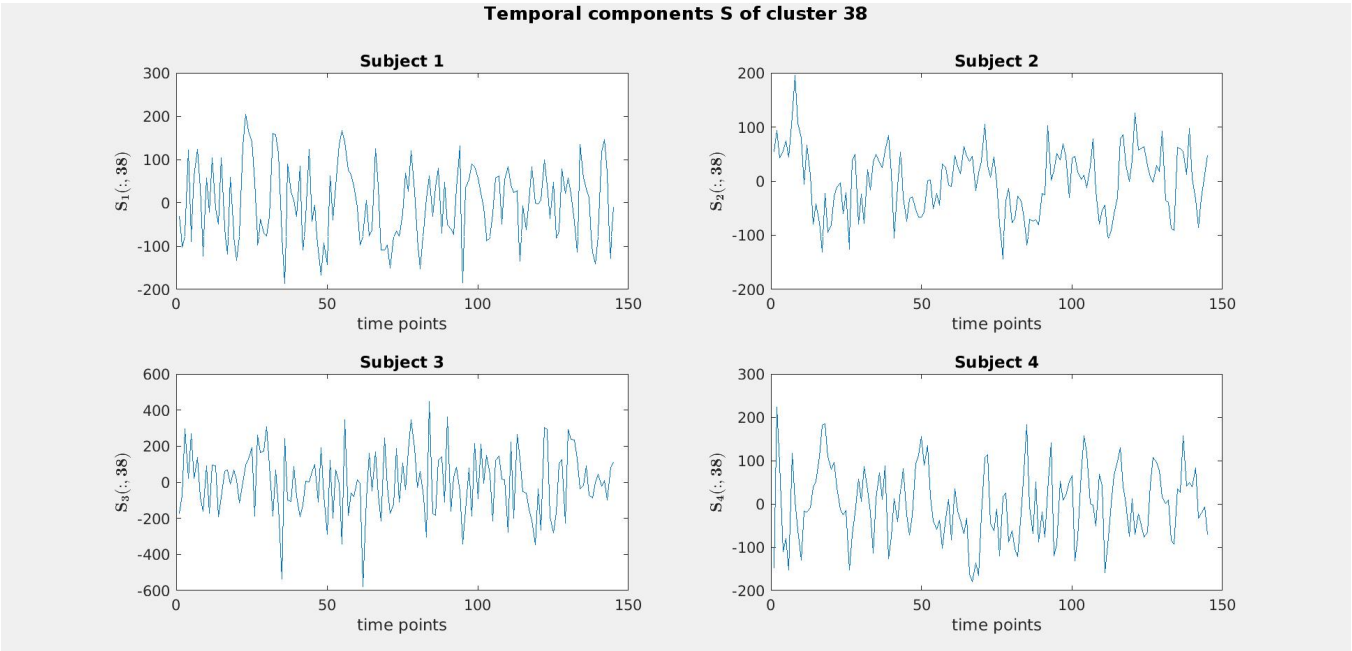


Figure B.606: Waveform of temporal components $\mathbf{S}_k(:, 38)$ for $k = 1, \ldots, 4$ and R=42.

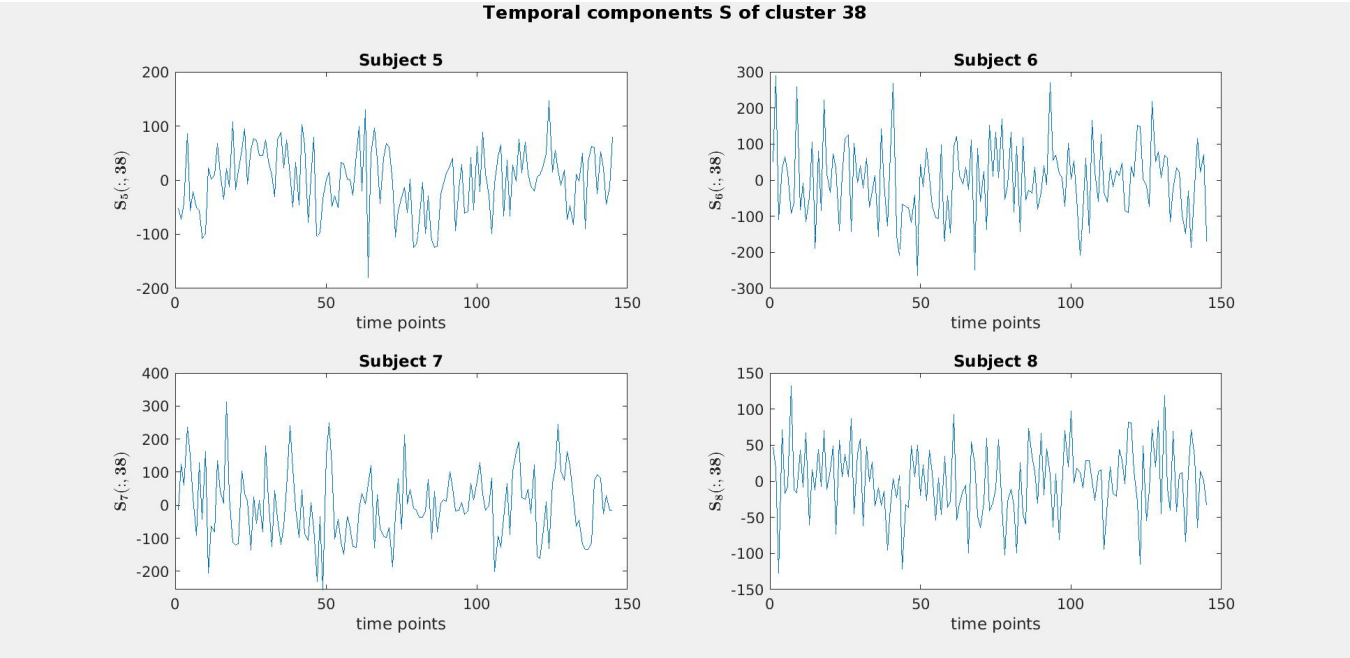


Figure B.607: Waveform of temporal components $\mathbf{S}_k(:, 38)$ for $k = 5, \ldots, 8$ and R=42.

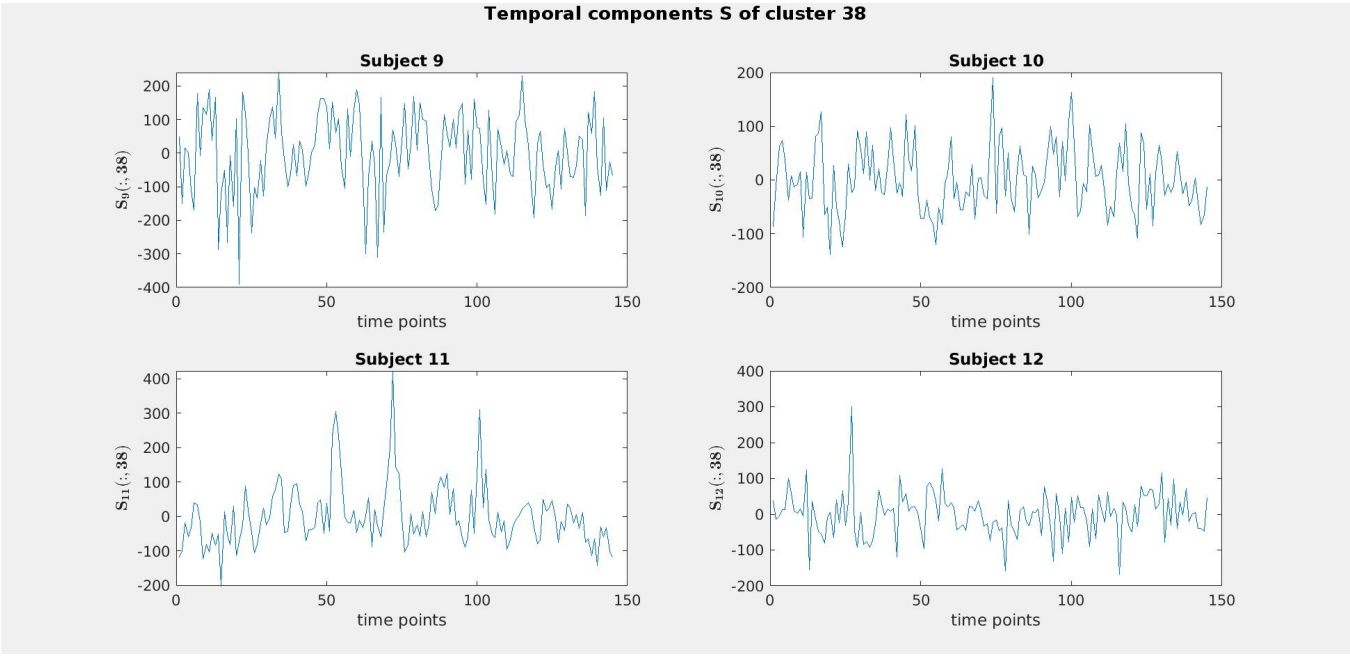


Figure B.608: Waveform of temporal components $\mathbf{S}_k(:, 38)$ for $k = 9, \ldots, 12$ and R=42.

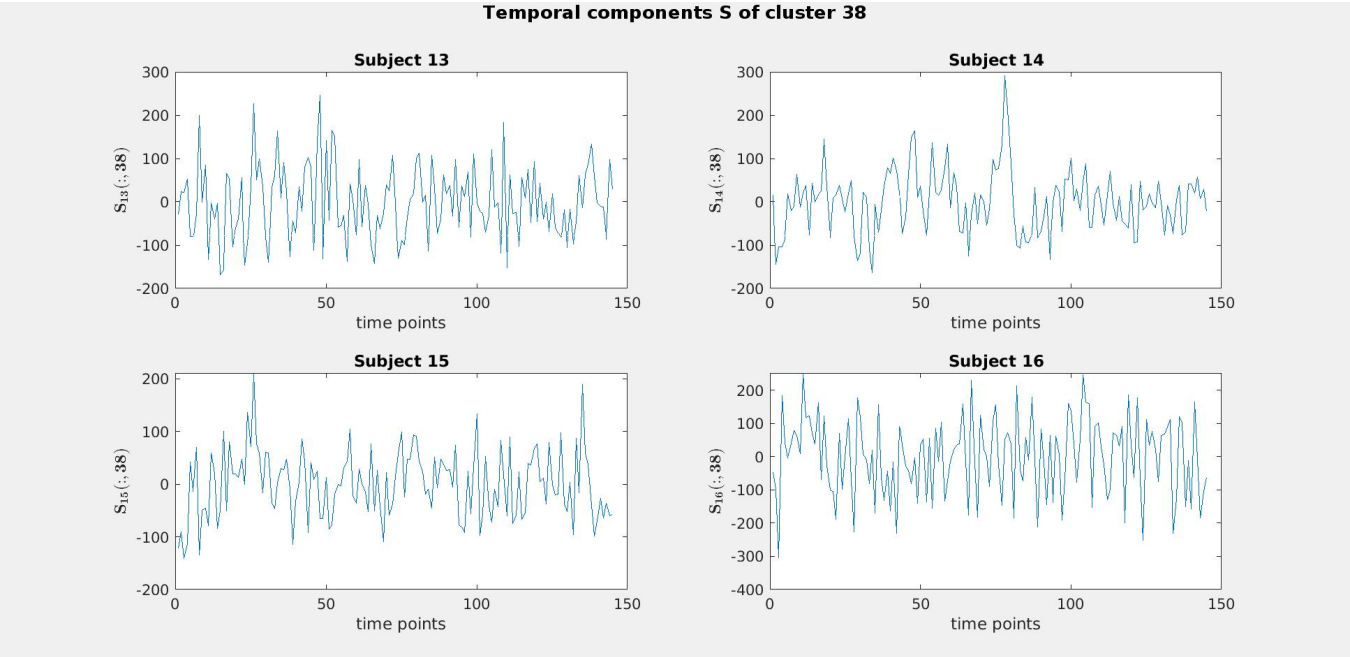


Figure B.609: Waveform of temporal components $\mathbf{S}_k(:, 38)$ for $k = 13, \ldots, 16$ and R=42.

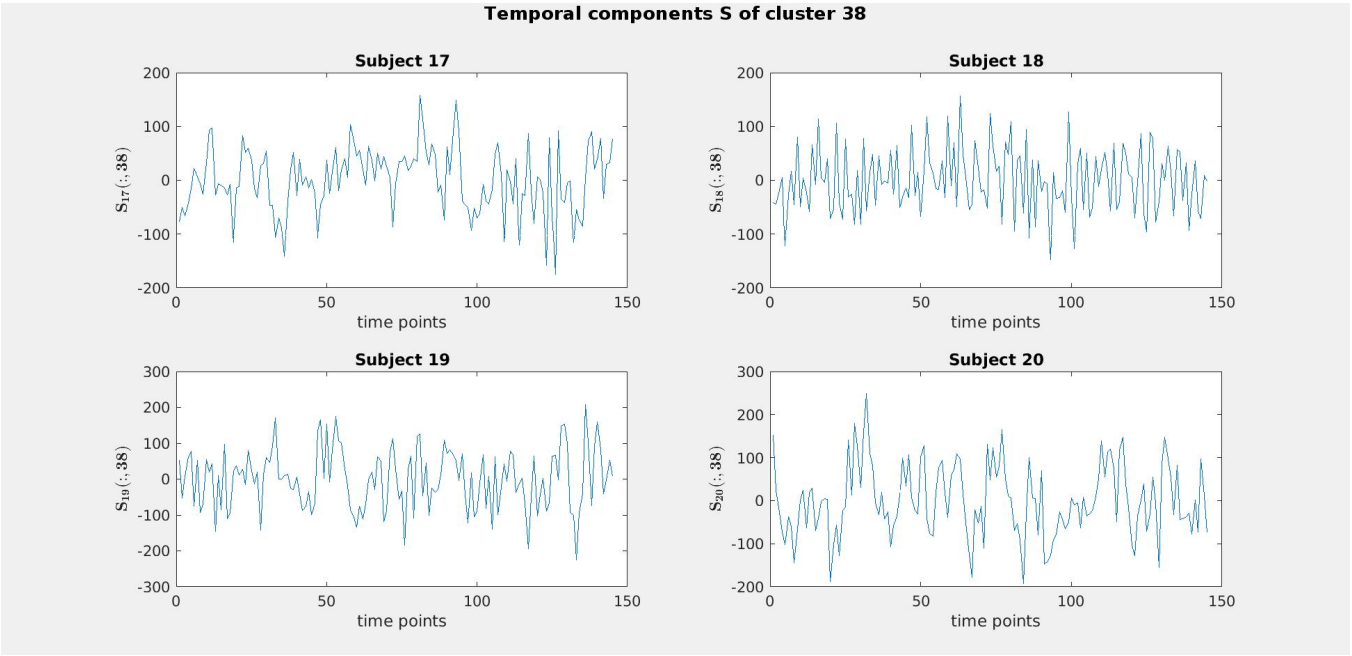


Figure B.610: Waveform of temporal components $\mathbf{S}_k(:, 38)$ for $k = 17, \ldots, 20$ and R=42.

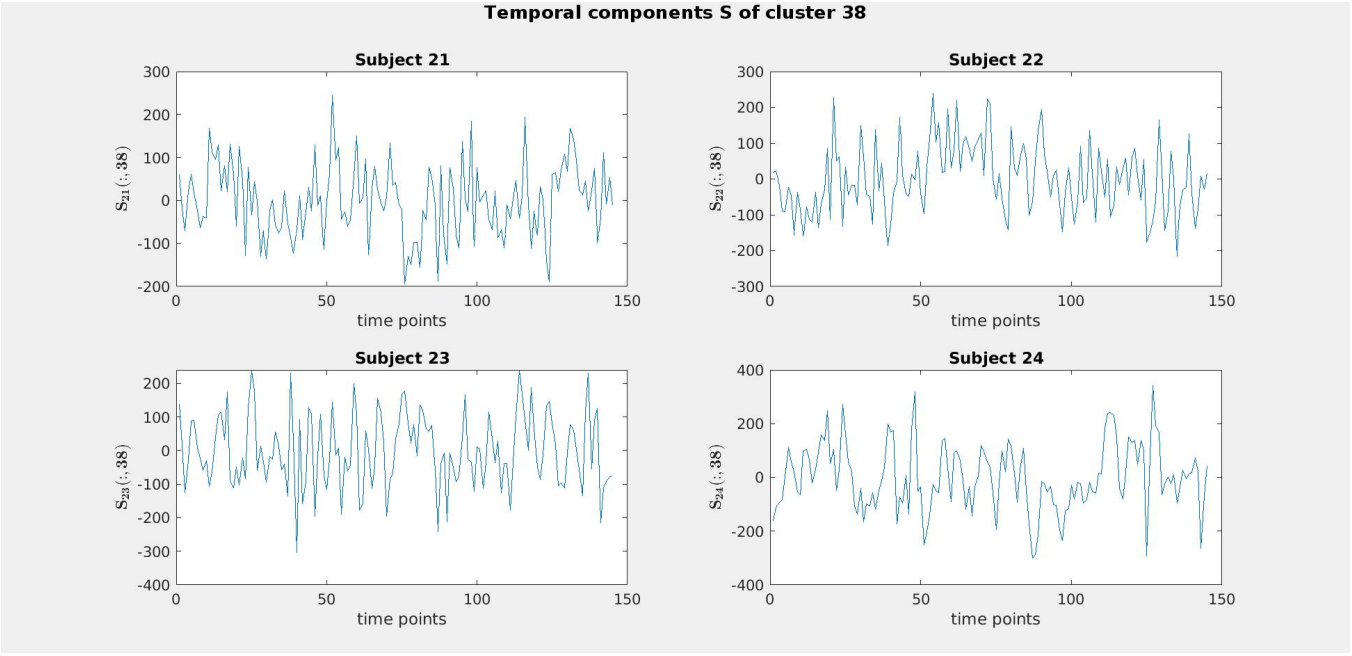


Figure B.611: Waveform of temporal components $\mathbf{S}_k(:, 138)$ for $k = 21, \ldots, 24$ and R=42.

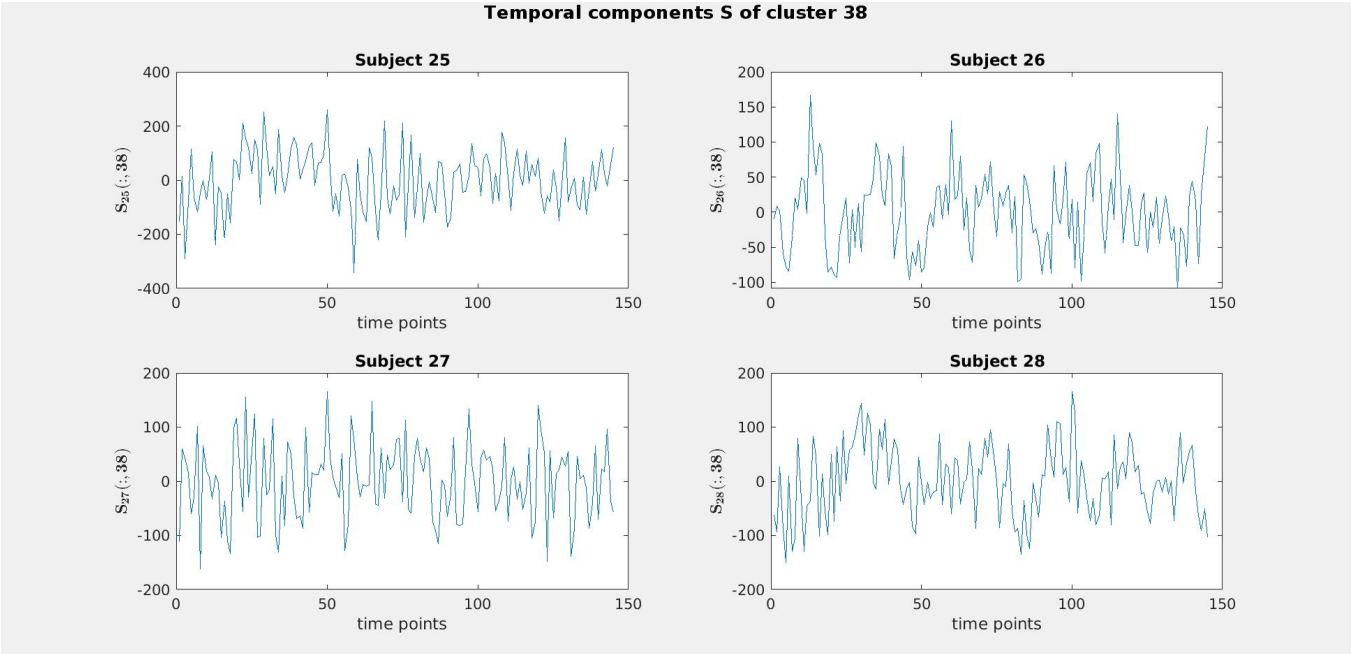


Figure B.612: Waveform of temporal components $\mathbf{S}_k(:, 38)$ for $k = 25, \ldots, 28$ and R=42.

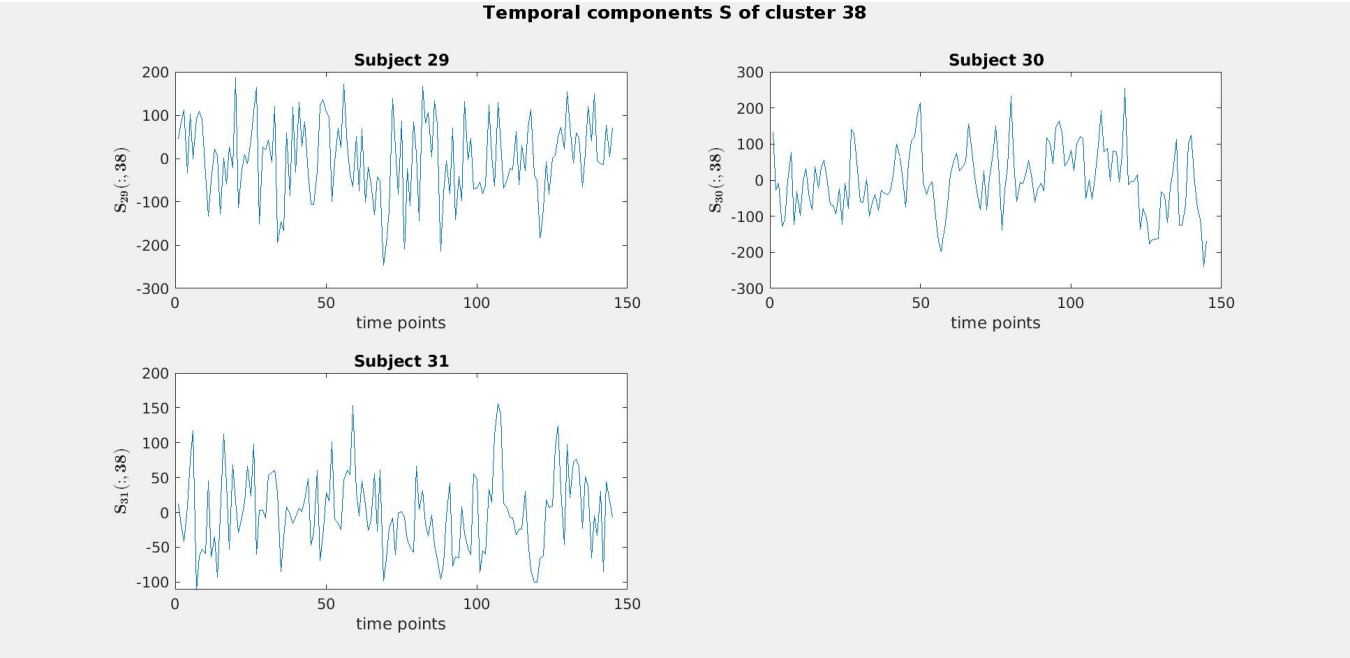


Figure B.613: Waveform of temporal components $\mathbf{S}_k(:, 38)$ for $k = 29, \ldots, 31$ and R=42.

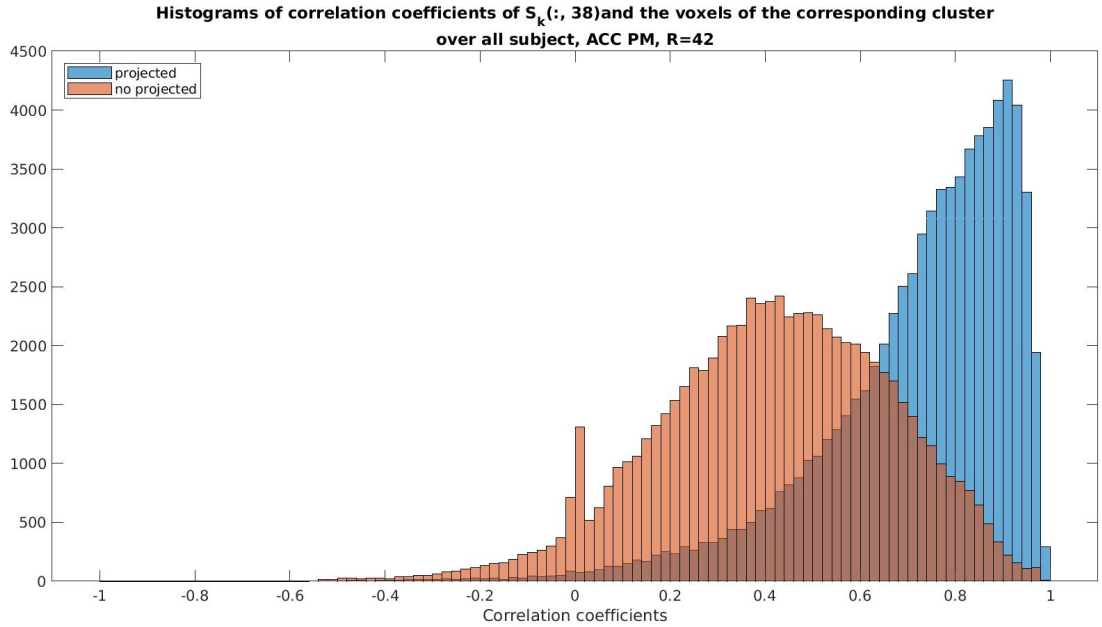


Figure B.614: For the column 38 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 38)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

33. **Cluster 11**

Cluster 11 contains 2665 voxels, which correspond to 1.179475% of the whole brain (background, soft tissue, bone are excluded).

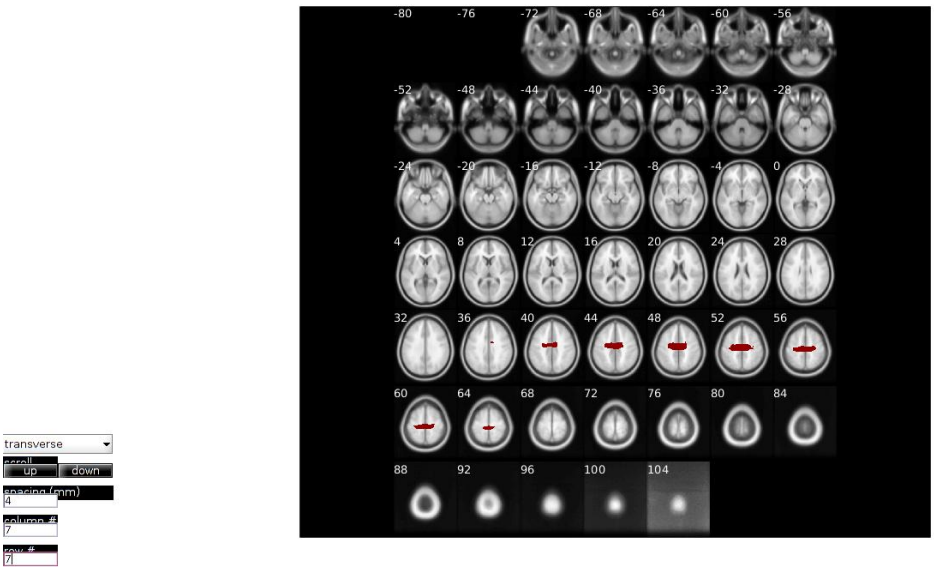


Figure B.615: Cluster 11. Red areas indicate the voxels which belong to cluster 11.

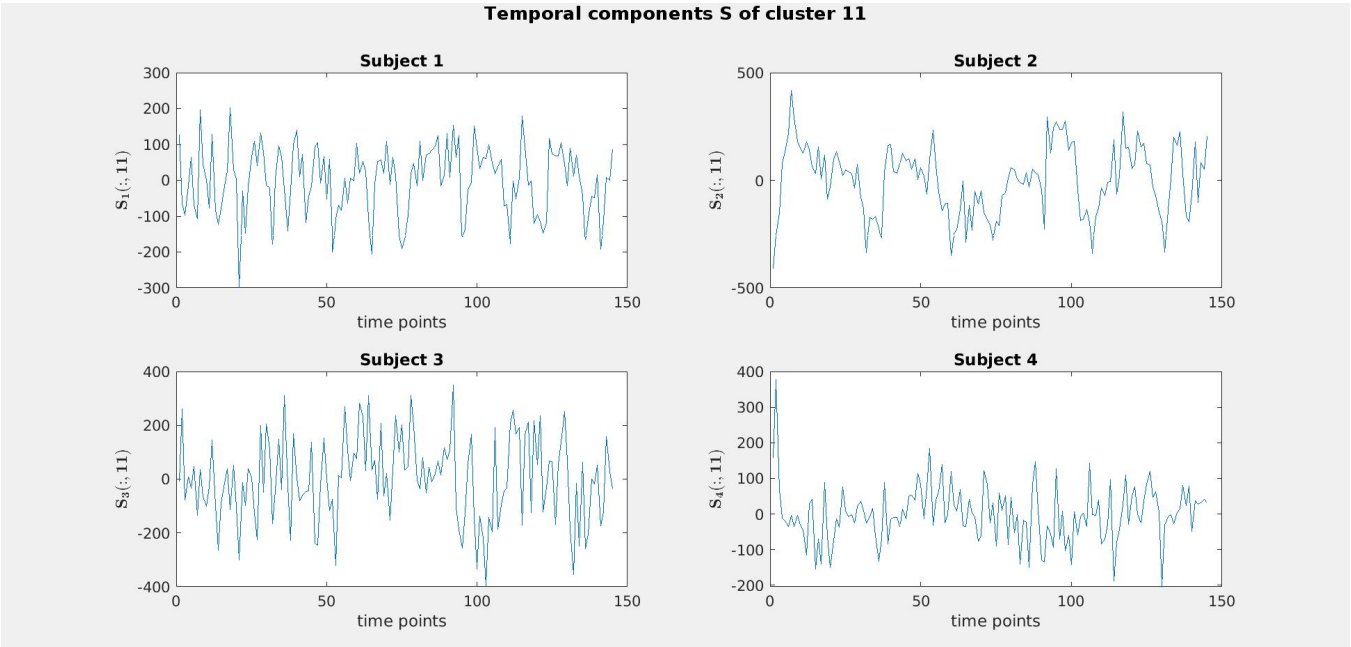


Figure B.616: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 1, \ldots, 4$ and R=42.

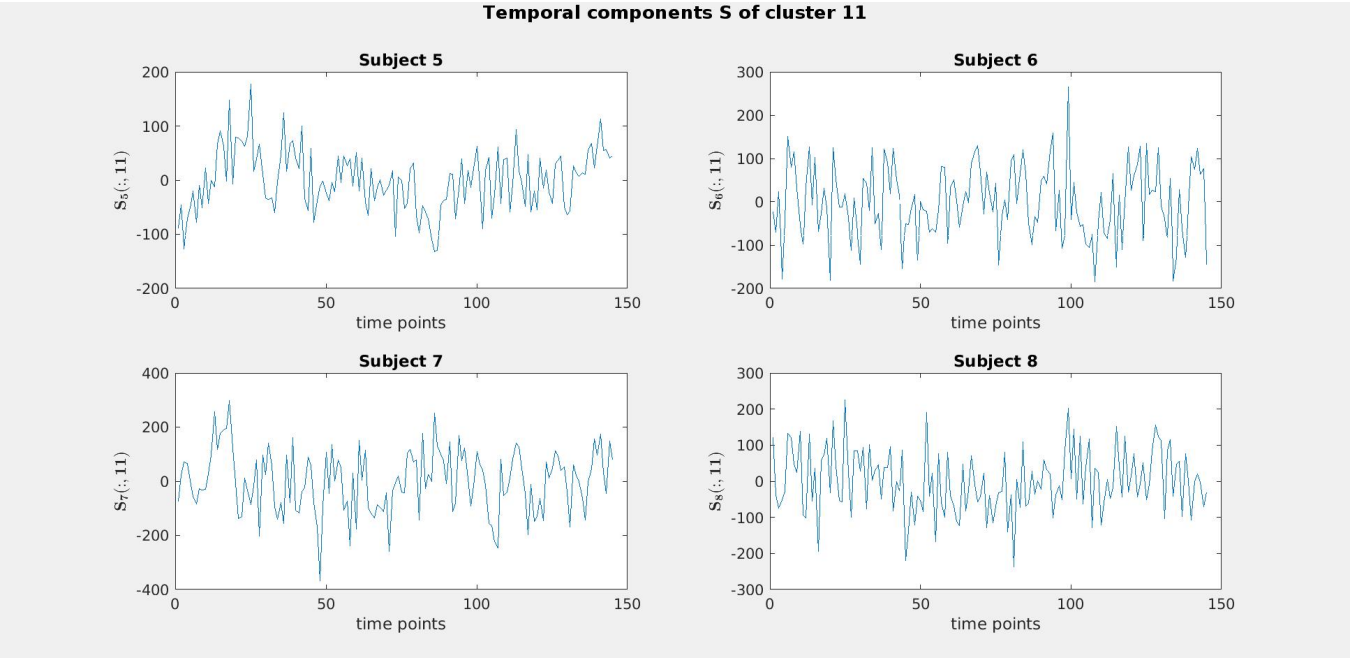


Figure B.617: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 5, \ldots, 8$ and R=42.

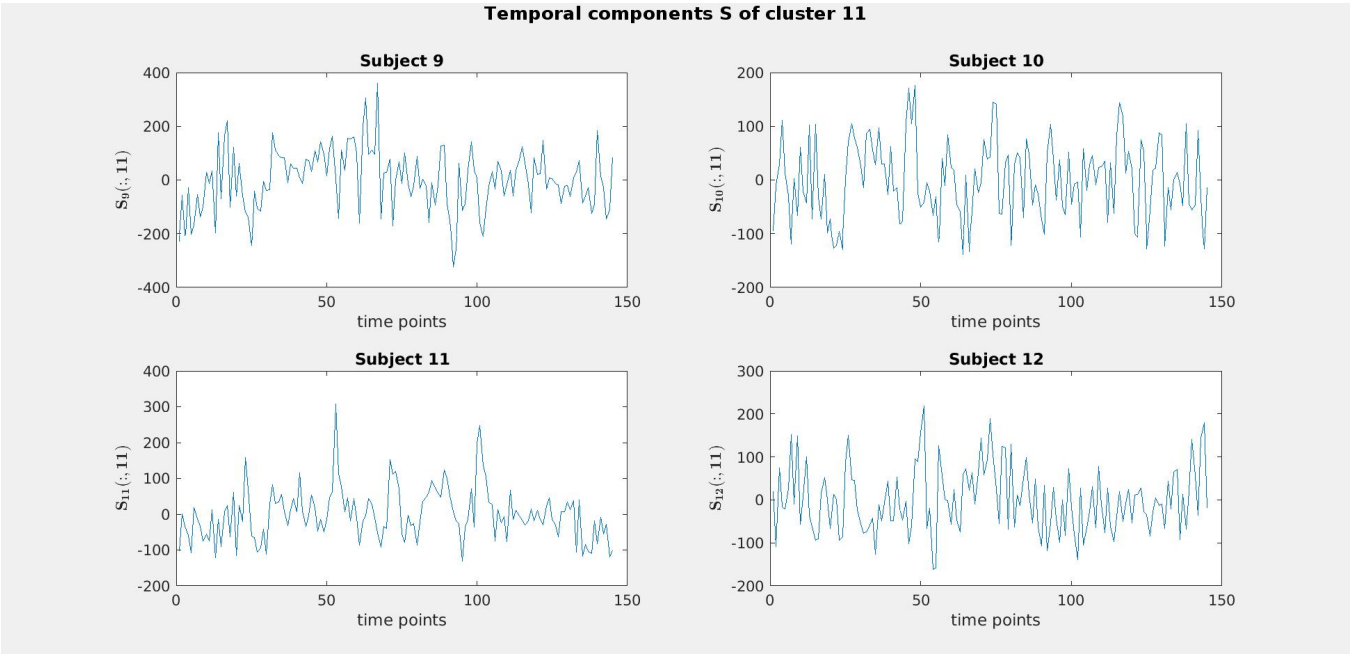


Figure B.618: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 9, \ldots, 12$ and R=42.

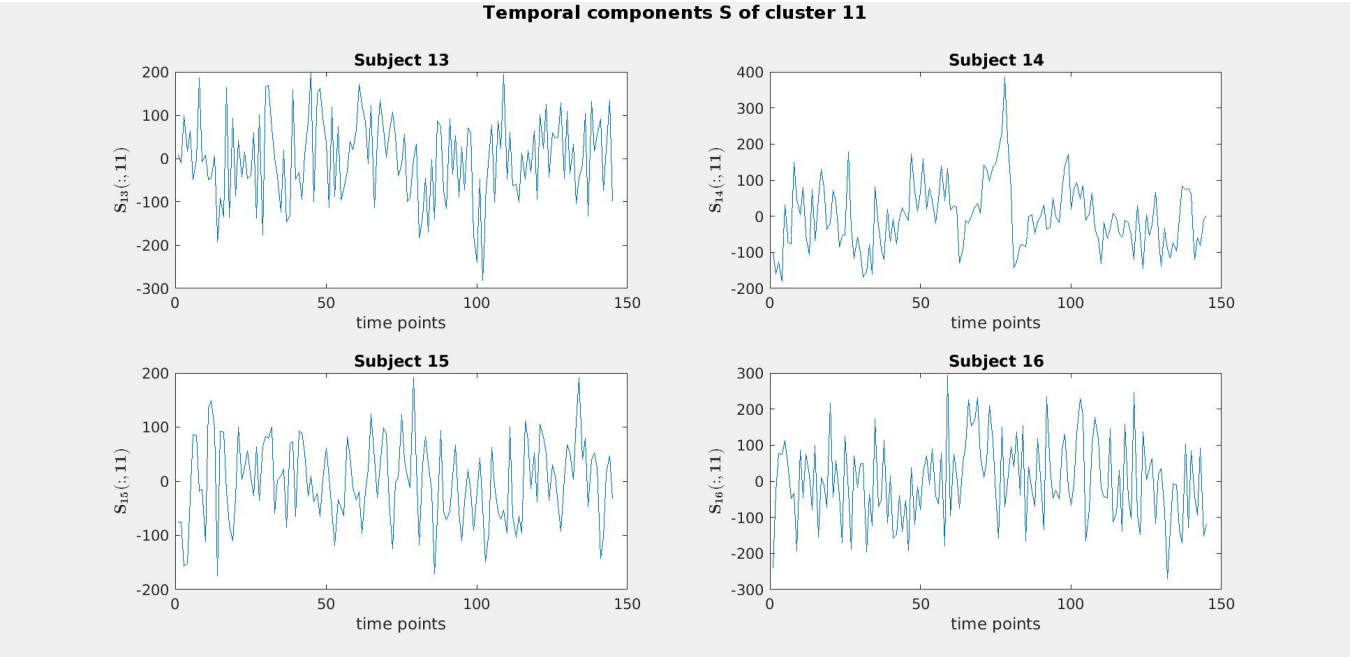


Figure B.619: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 13, \ldots, 16$ and R=42.

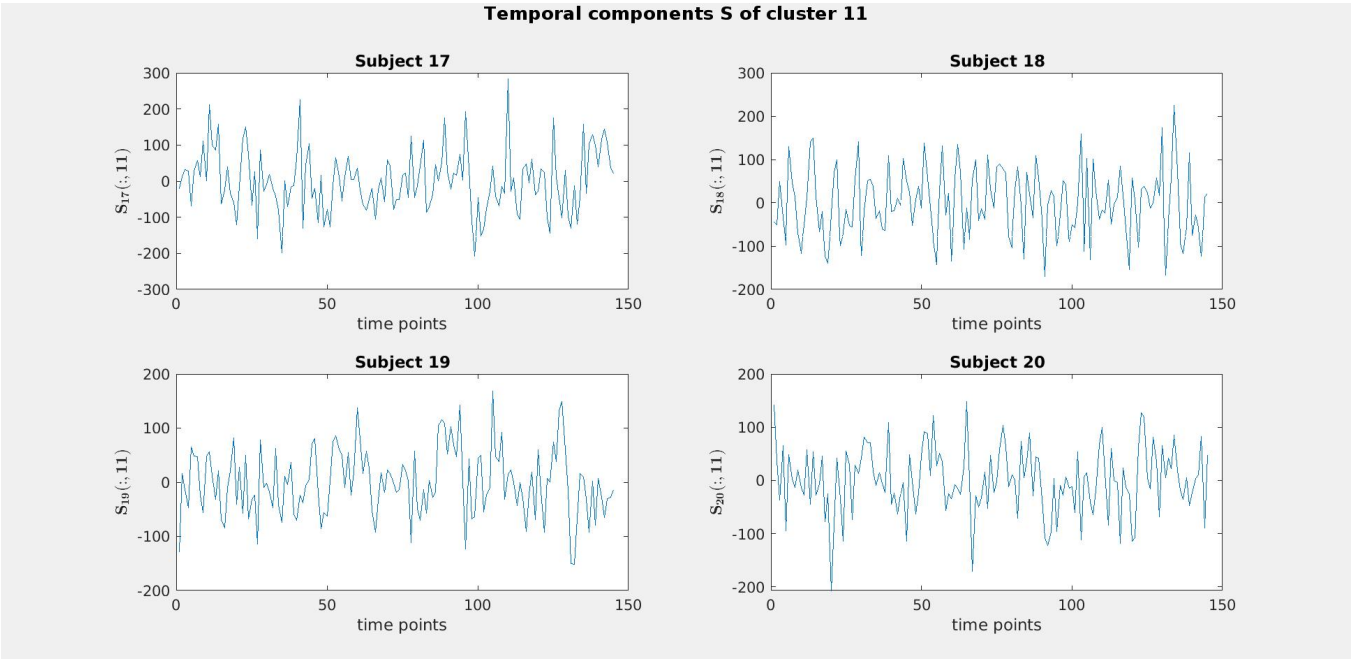


Figure B.620: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 17, \ldots, 20$ and R=42.

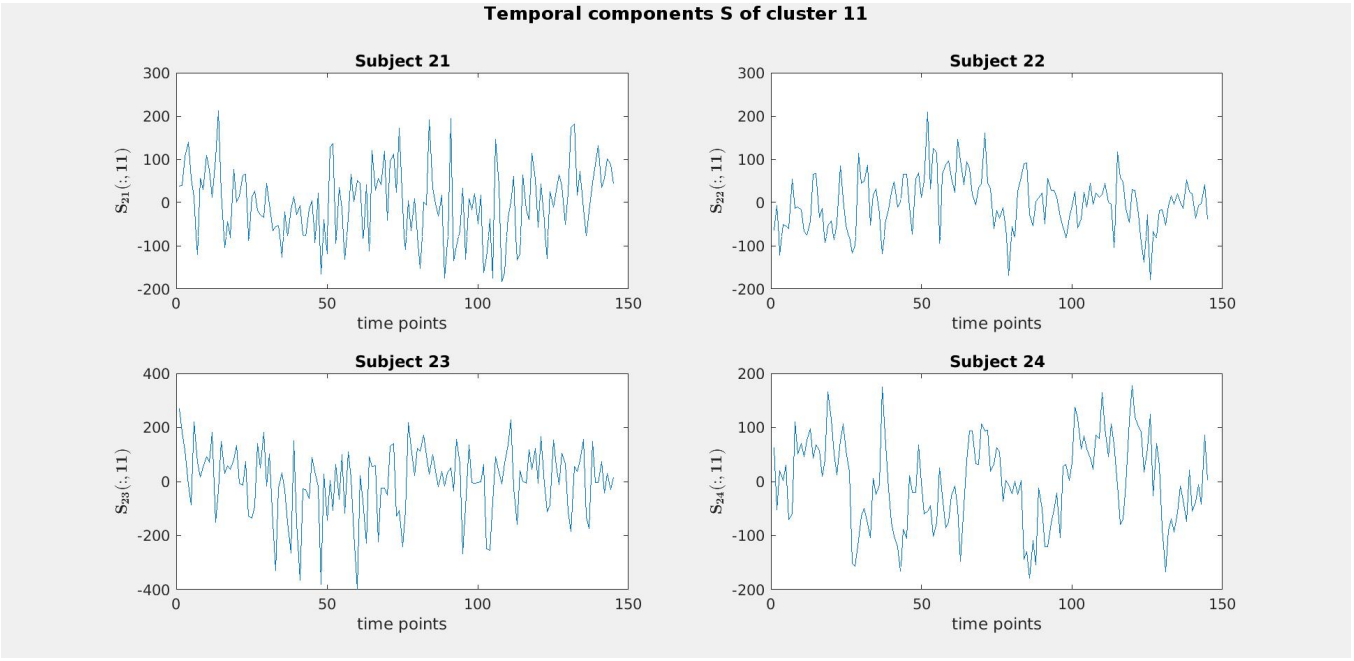


Figure B.621: Waveform of temporal components $\mathbf{S}_k(:, 111)$ for $k = 21, \ldots, 24$ and R=42.

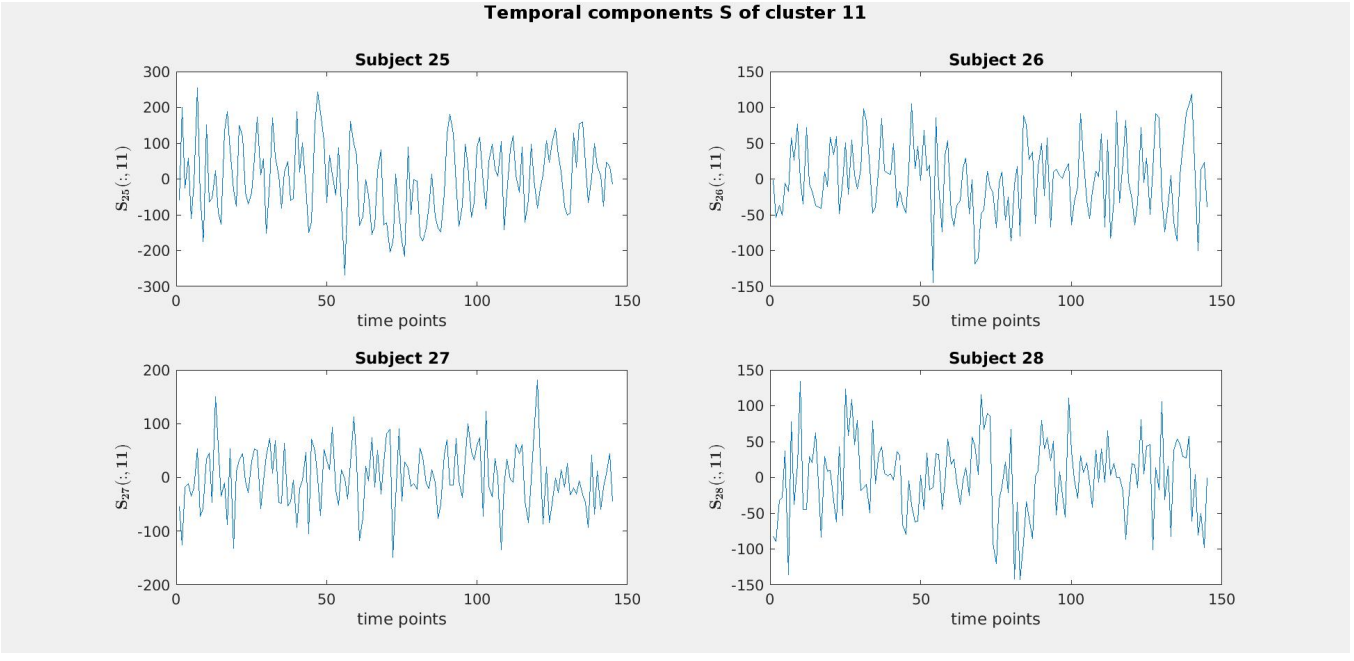


Figure B.622: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 25, \ldots, 28$ and R=42.

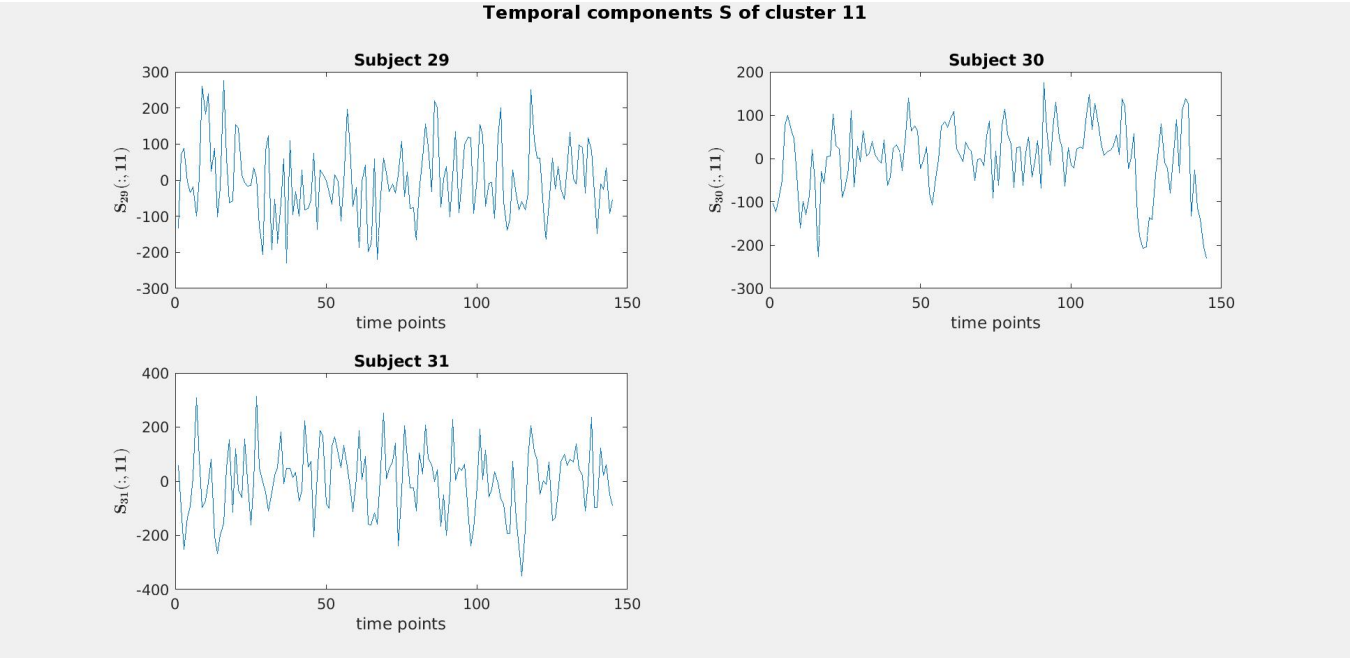


Figure B.623: Waveform of temporal components $\mathbf{S}_k(:, 11)$ for $k = 29, \ldots, 31$ and R=42.

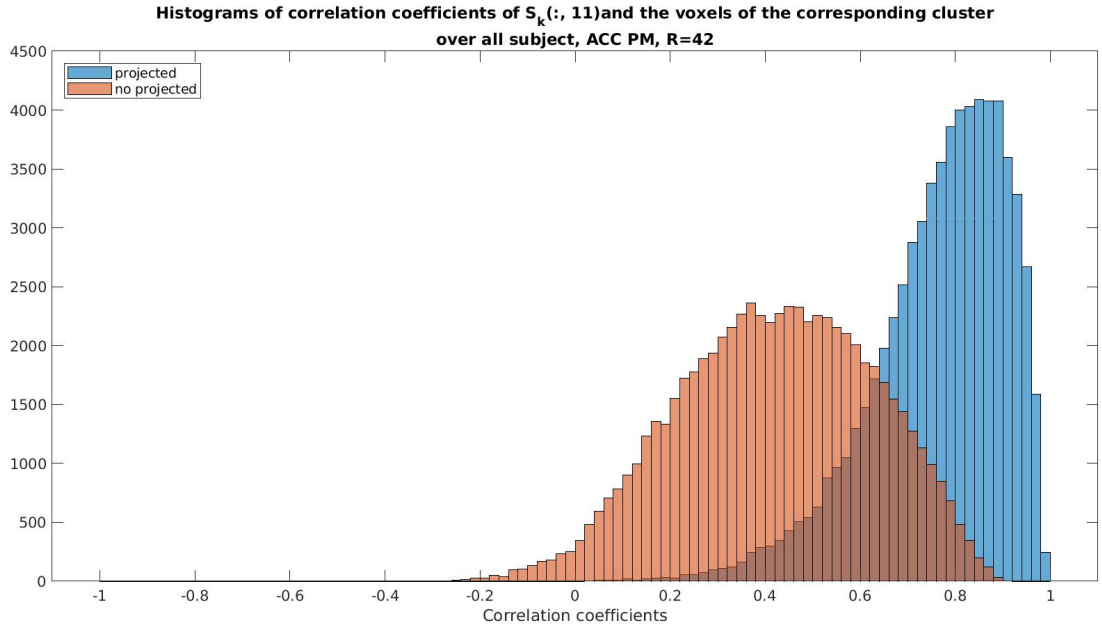


Figure B.624: For the column 11 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 11)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

34. **Cluster 36**

Cluster 36 contains 2209 voxels, which correspond to 0.977659% of the whole brain (background, soft tissue, bone are excluded).

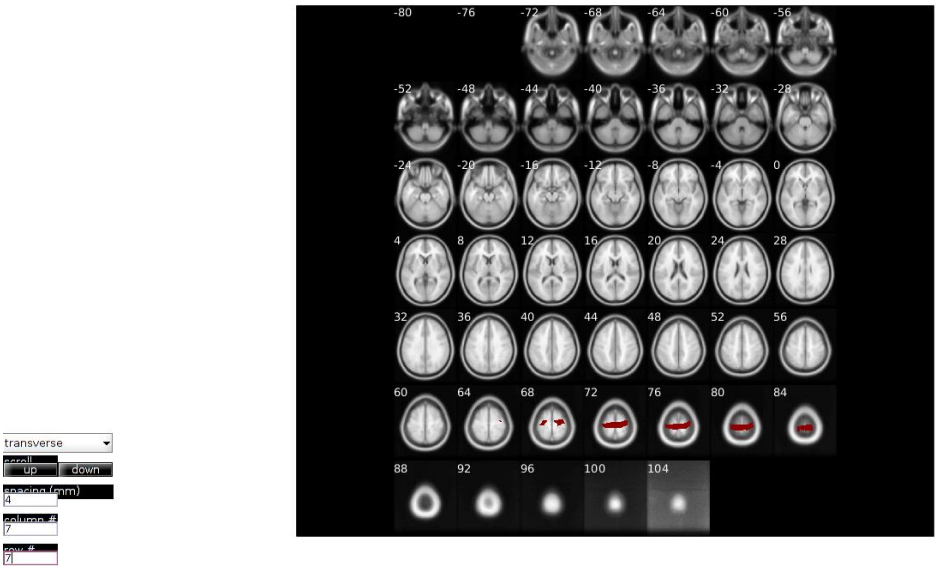


Figure B.625: Cluster 36. Red areas indicate the voxels which belong to cluster 36.

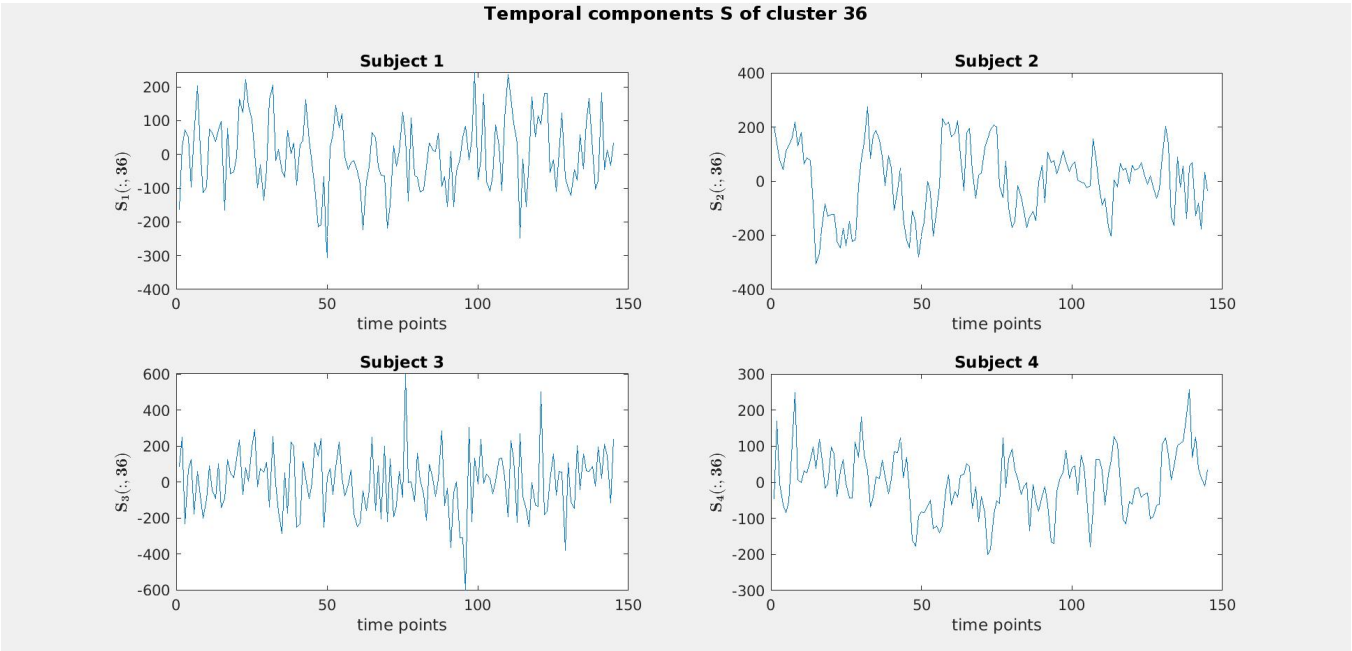


Figure B.626: Waveform of temporal components $\mathbf{S}_k(:, 36)$ for $k = 1, \ldots, 4$ and R=42.

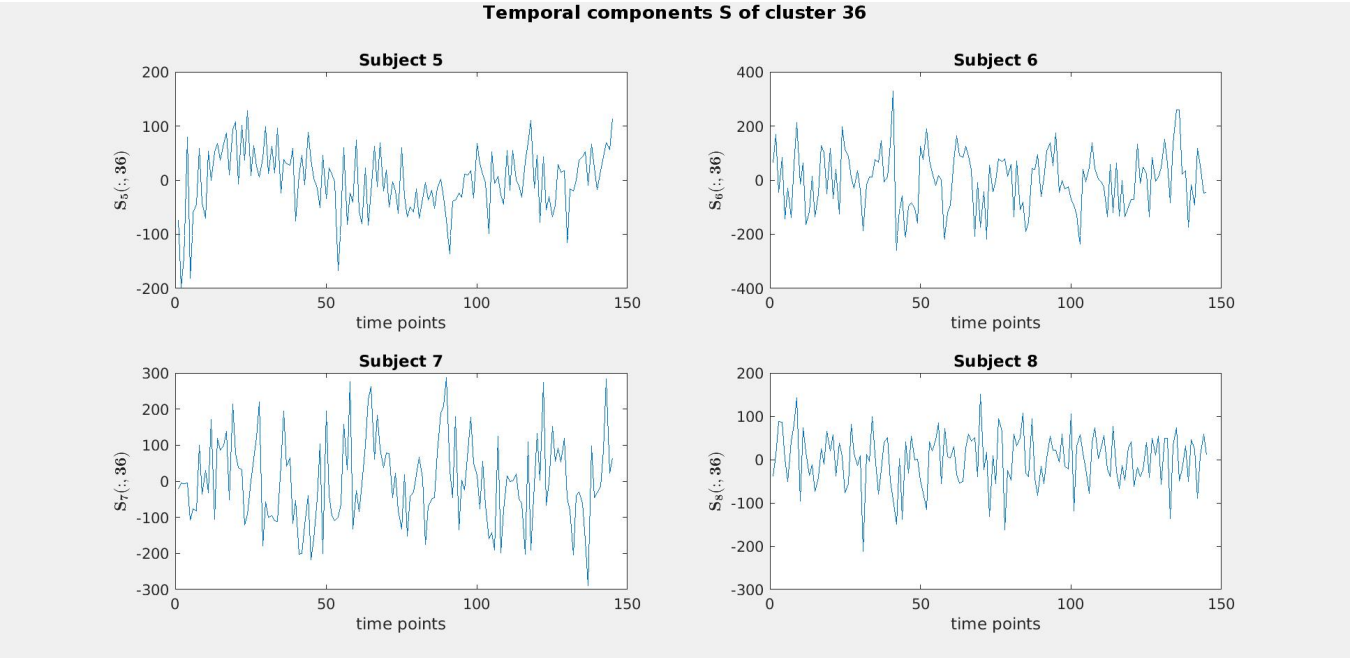


Figure B.627: Waveform of temporal components $\mathbf{S}_k(:, 36)$ for $k = 5, \ldots, 8$ and R=42.

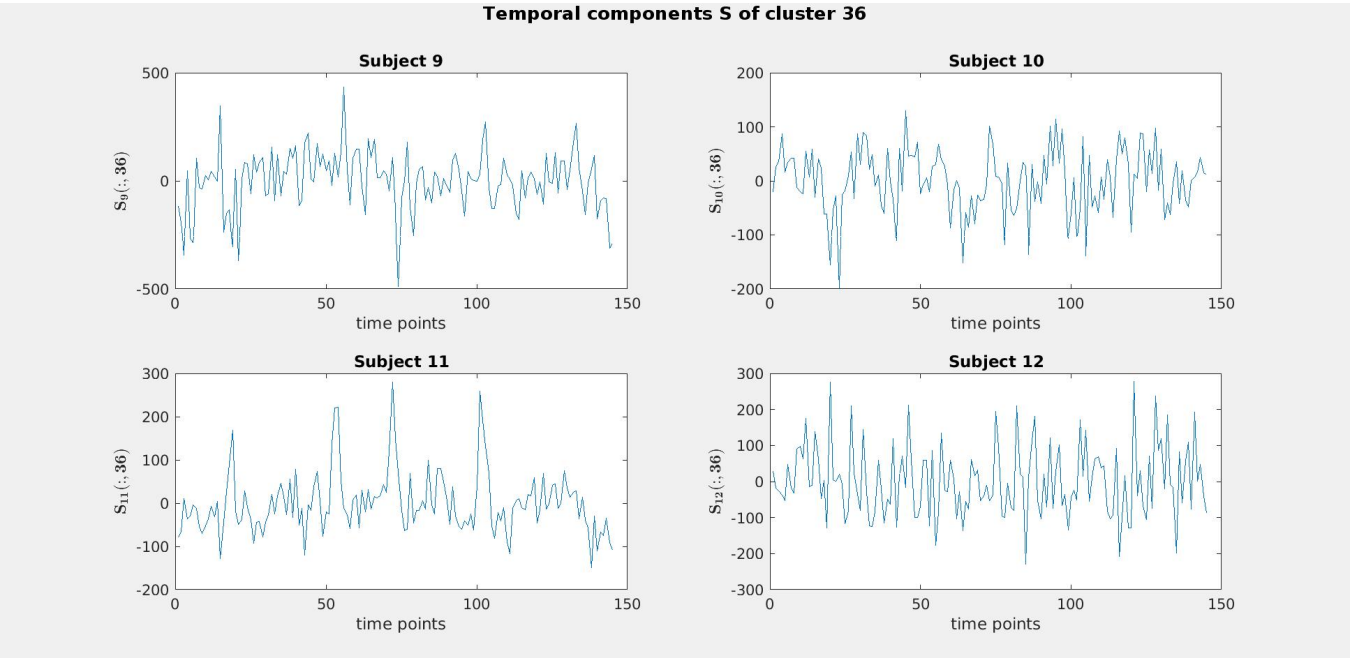


Figure B.628: Waveform of temporal components $\mathbf{S}_k(:, 36)$ for $k = 9, \ldots, 12$ and R=42.

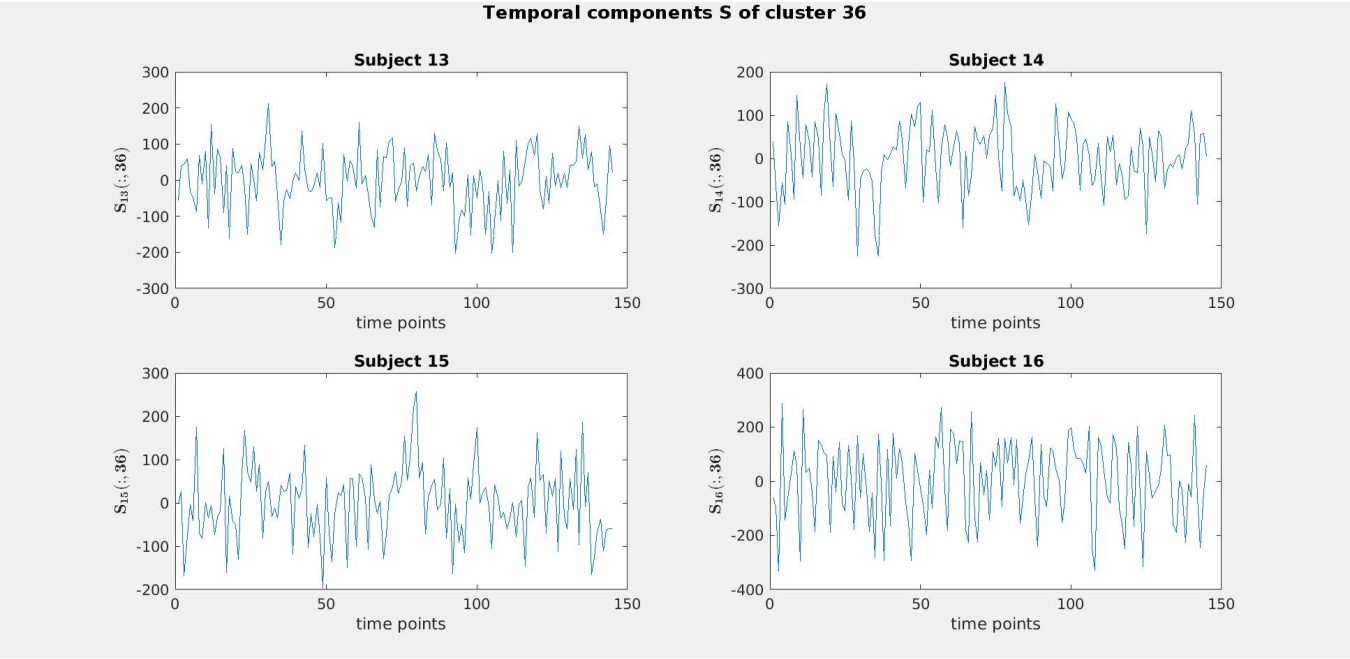


Figure B.629: Waveform of temporal components $\mathbf{S}_k(:, 36)$ for $k = 13, \ldots, 16$ and R=42.

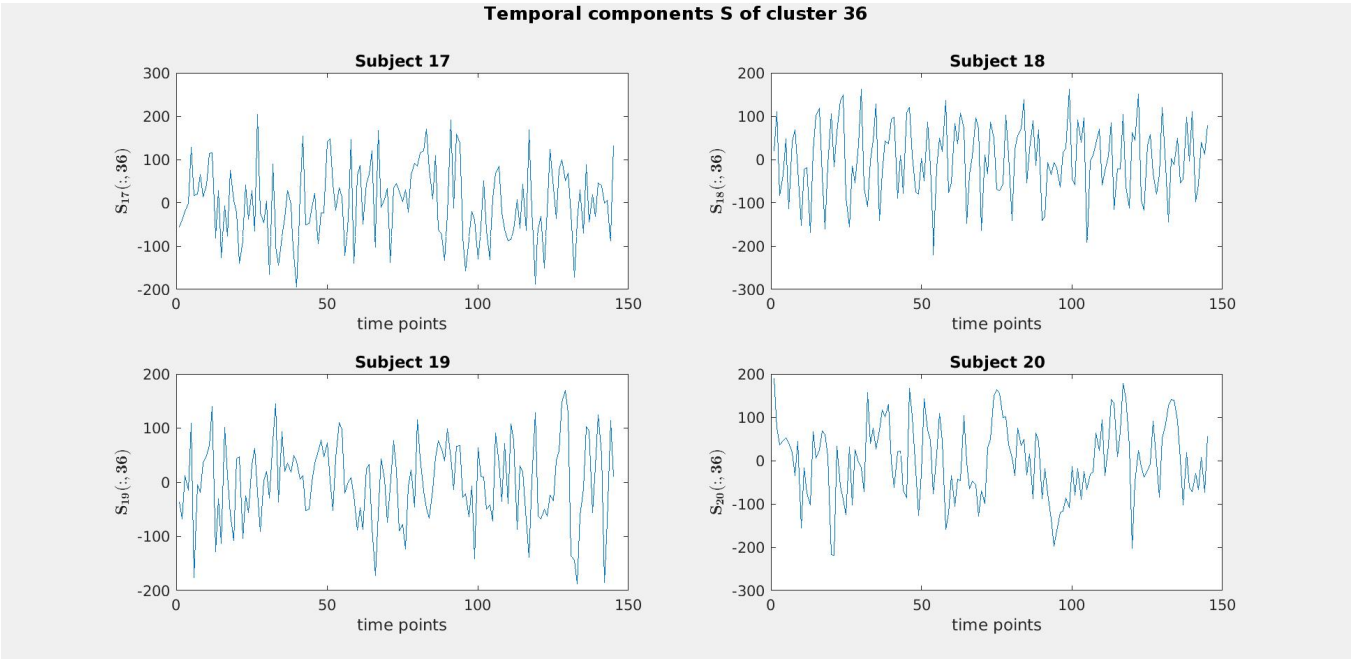


Figure B.630: Waveform of temporal components $\mathbf{S}_k(:, 36)$ for $k = 17, \ldots, 20$ and R=42.

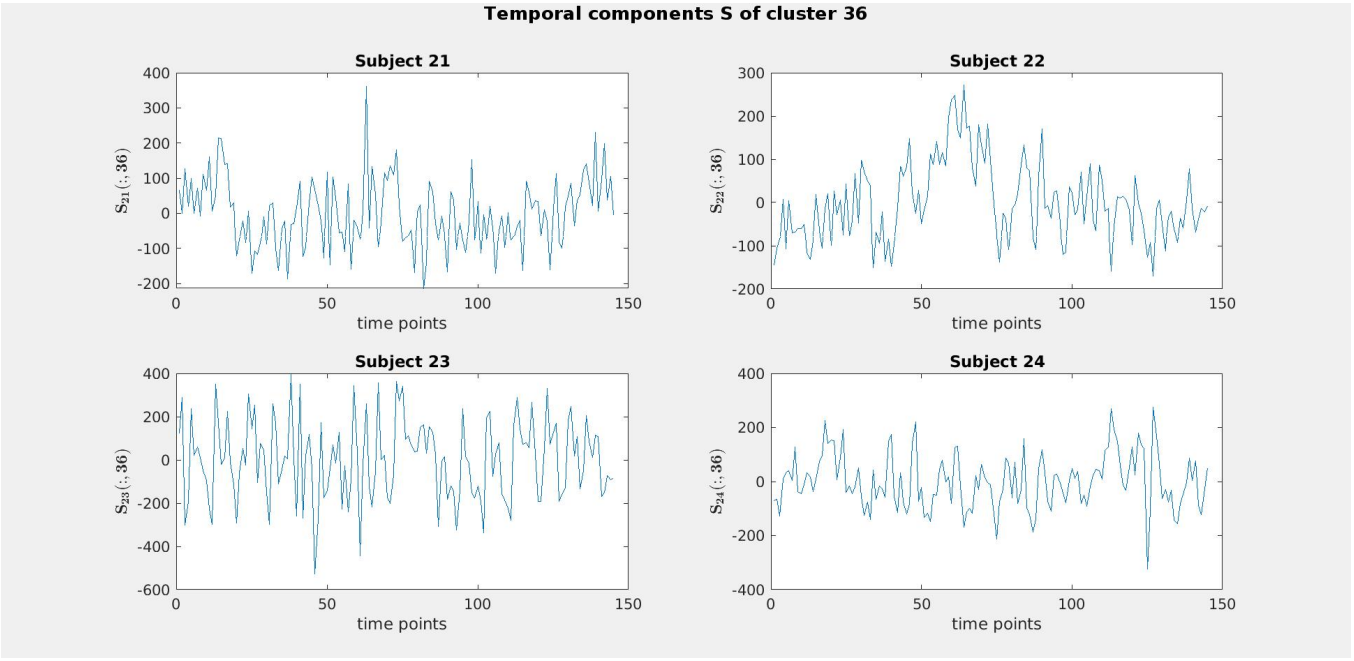


Figure B.631: Waveform of temporal components $\mathbf{S}_k(:, 136)$ for $k = 21, \ldots, 24$ and R=42.

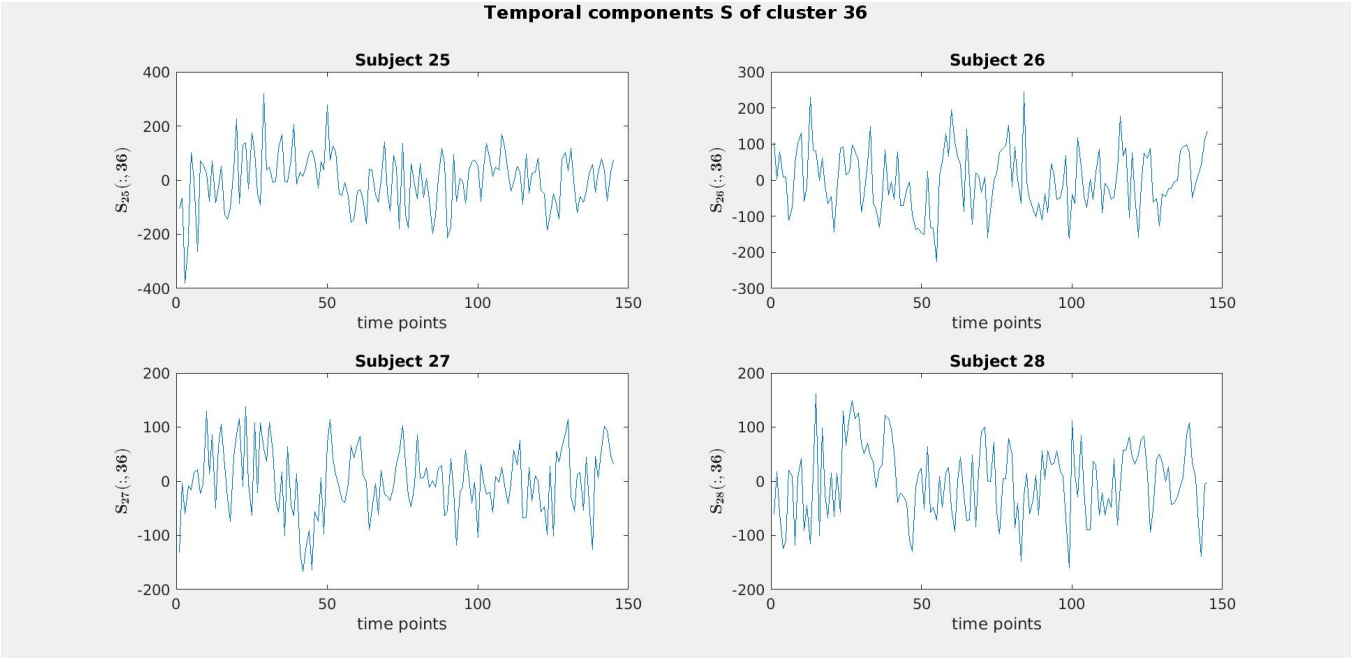


Figure B.632: Waveform of temporal components $\mathbf{S}_k(:, 36)$ for $k = 25, \ldots, 28$ and R=42.

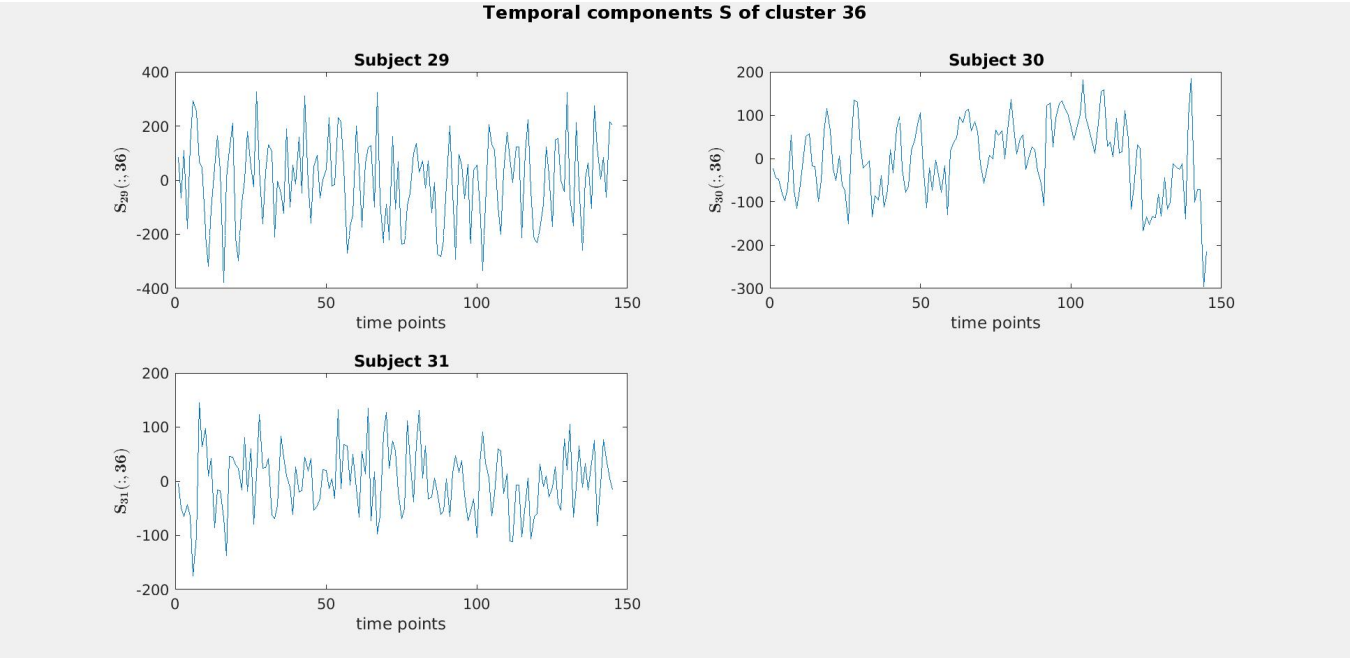


Figure B.633: Waveform of temporal components $\mathbf{S}_k(:, 36)$ for $k = 29, \ldots, 31$ and R=42.

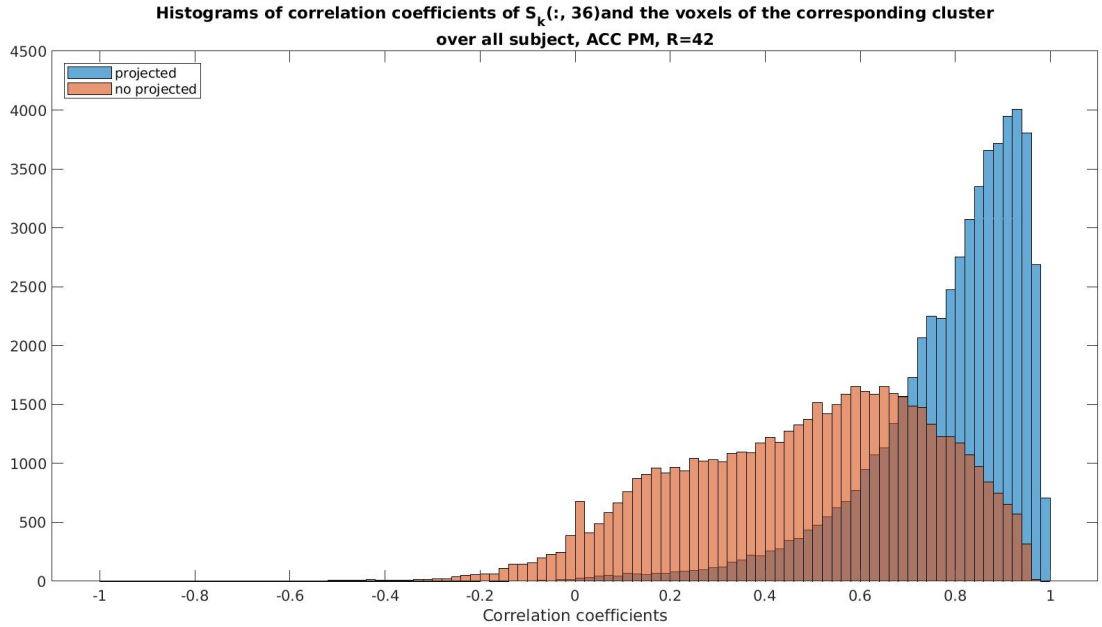


Figure B.634: For the column 36 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 36)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

35. **Cluster 12**

Cluster 12 contains 2206 voxels, which correspond to 0.976331% of the whole brain (background, soft tissue, bone are excluded).

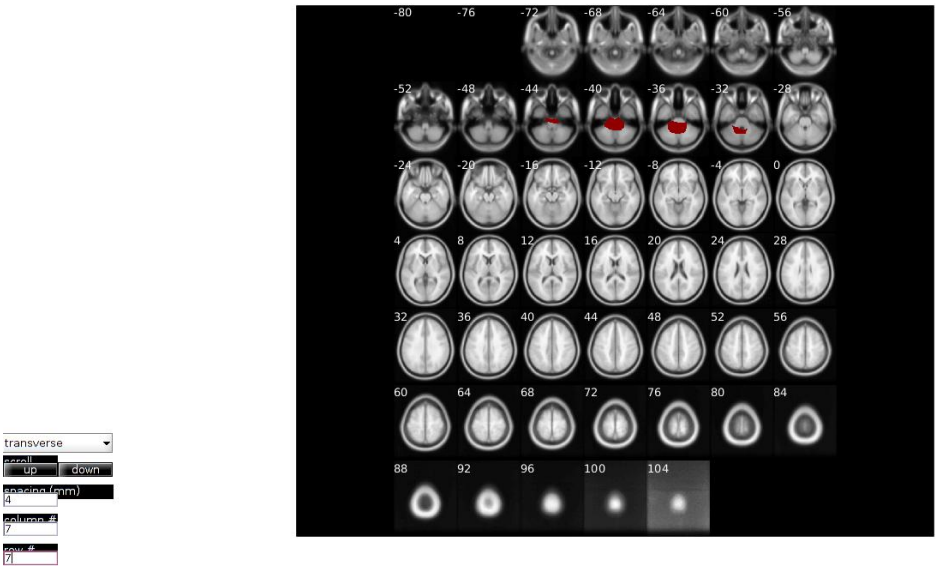


Figure B.635: Cluster 12. Red areas indicate the voxels which belong to cluster 12.

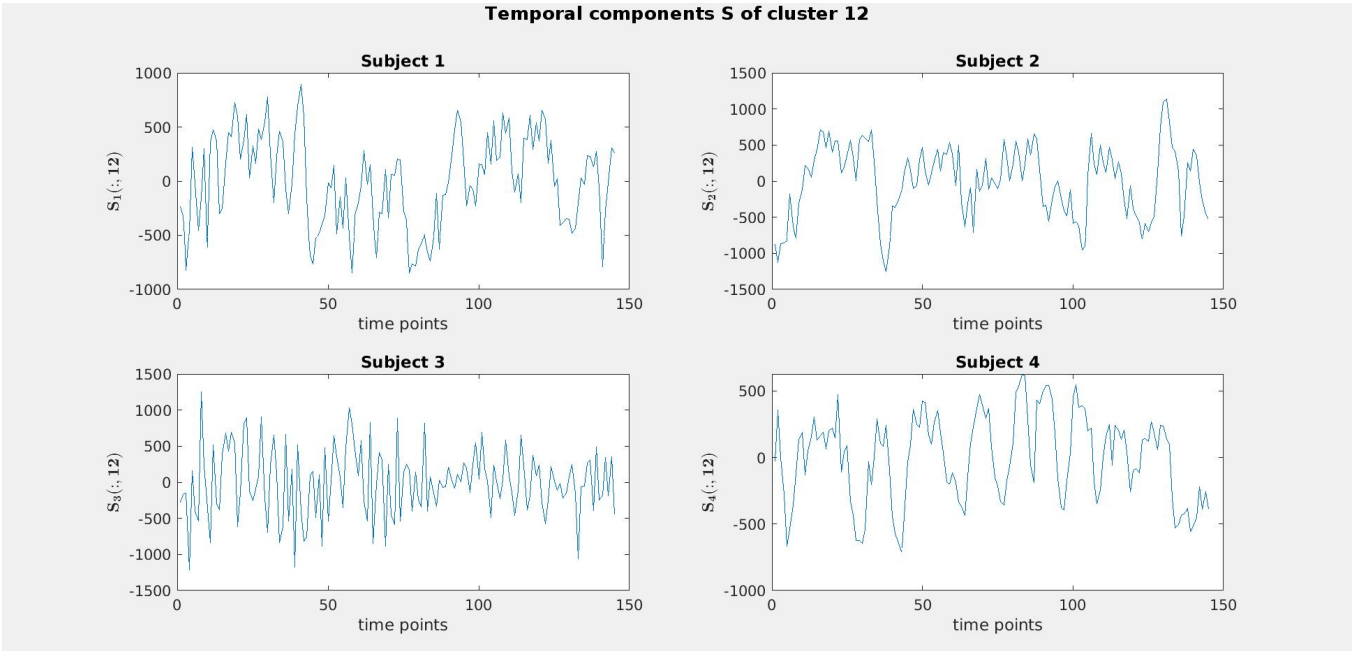


Figure B.636: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 1, \ldots, 4$ and R=42.

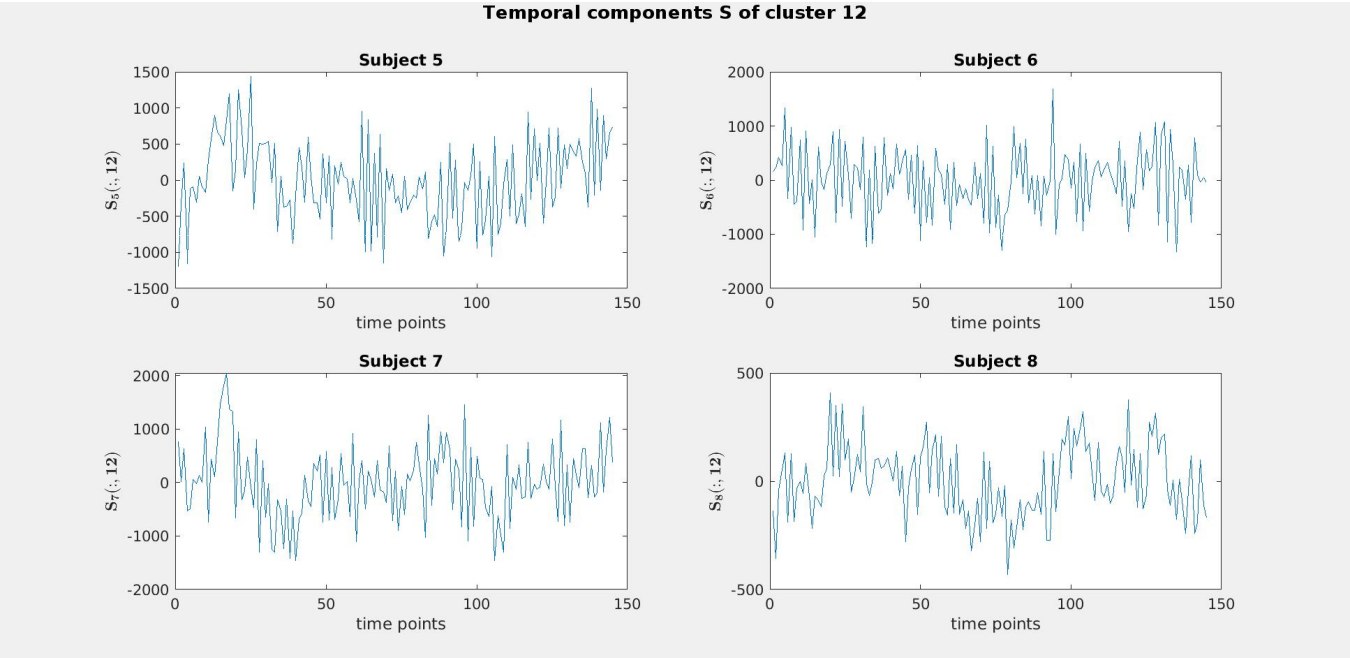


Figure B.637: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 5, \ldots, 8$ and R=42.

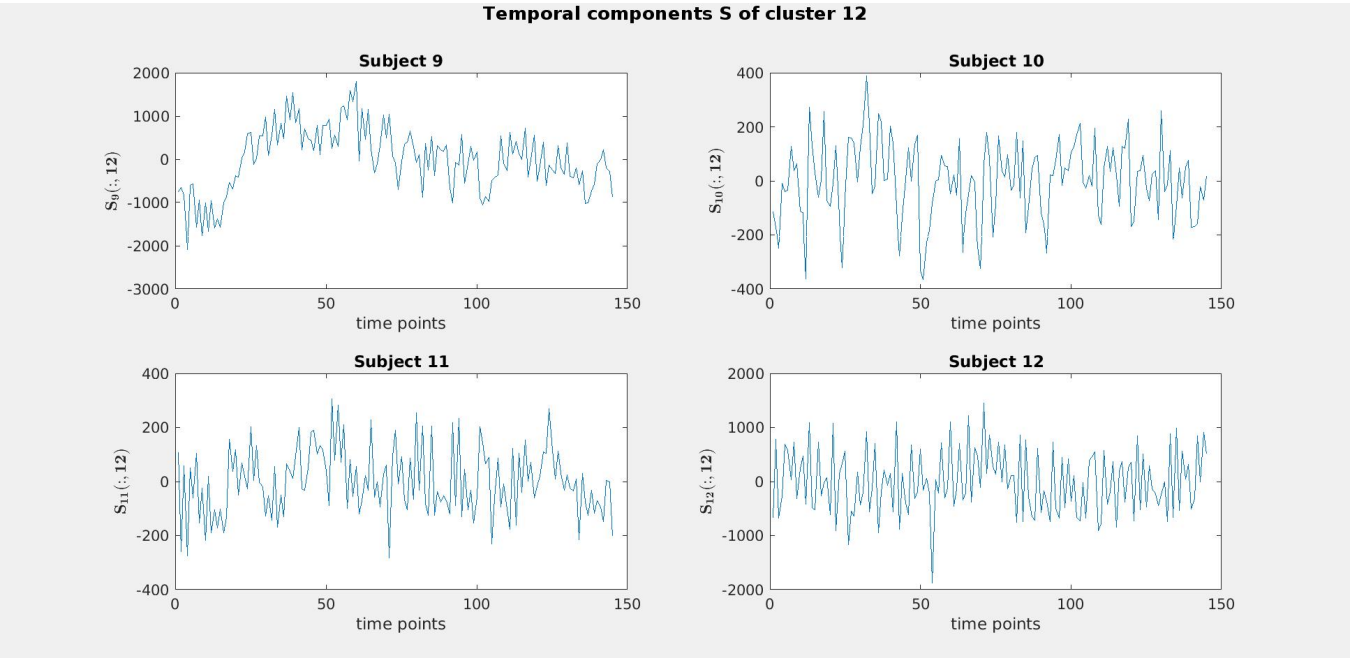


Figure B.638: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 9, \ldots, 12$ and R=42.

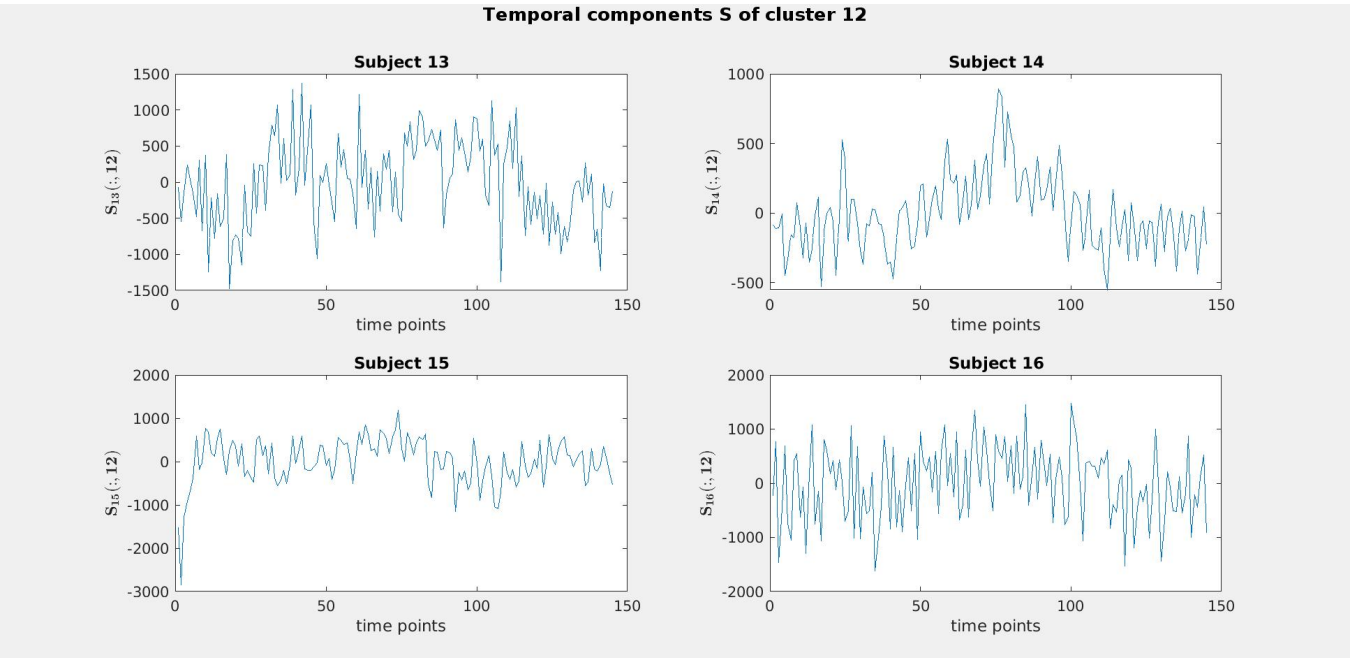


Figure B.639: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 13, \ldots, 16$ and R=42.

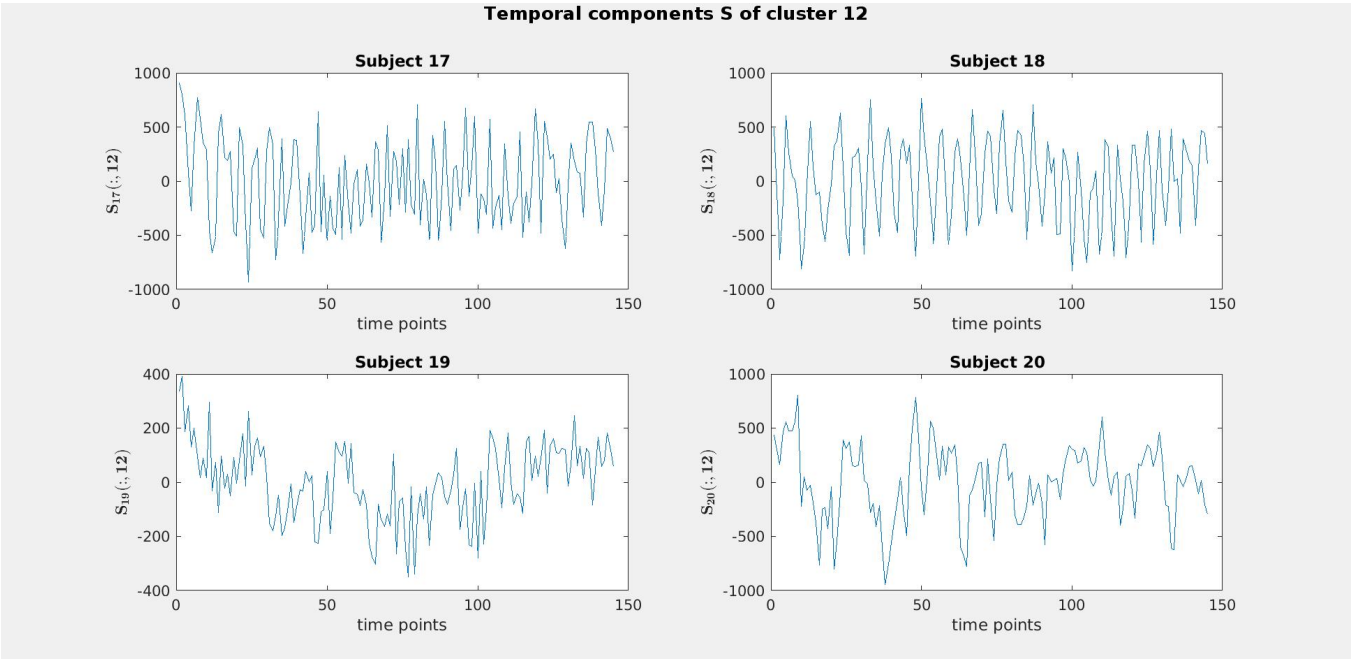


Figure B.640: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 17, \ldots, 20$ and R=42.

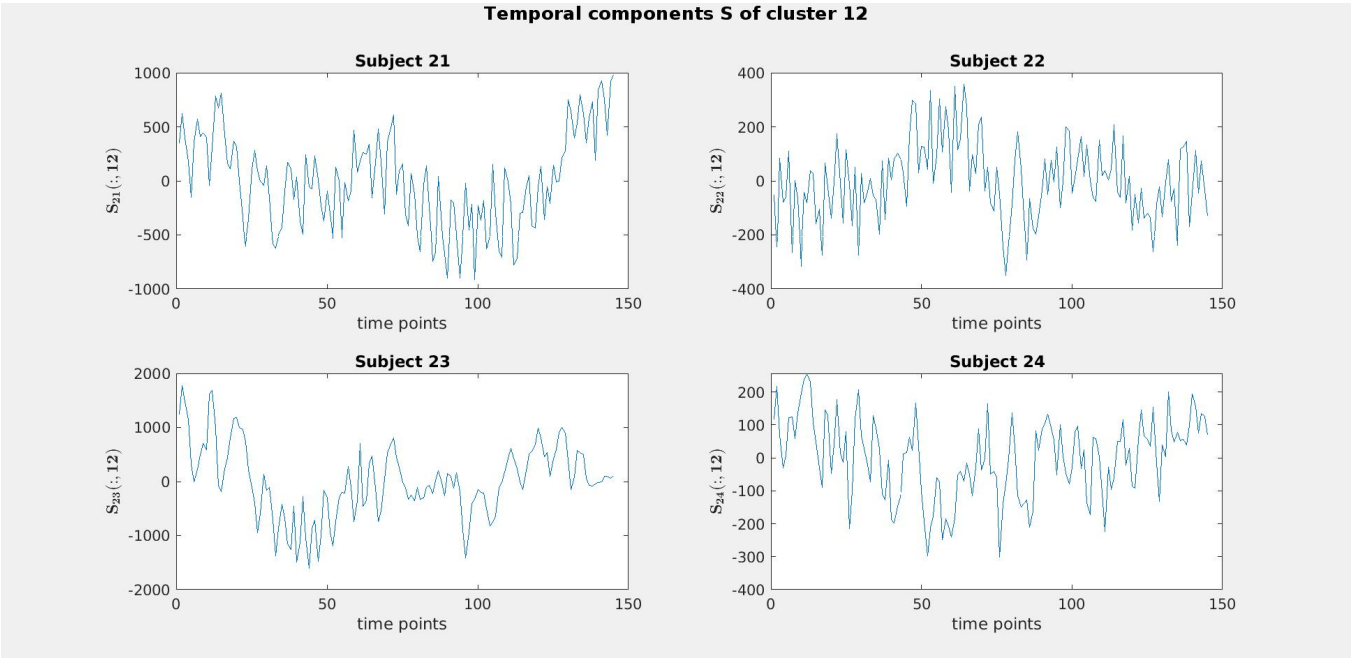


Figure B.641: Waveform of temporal components $\mathbf{S}_k(:, 112)$ for $k = 21, \ldots, 24$ and R=42.

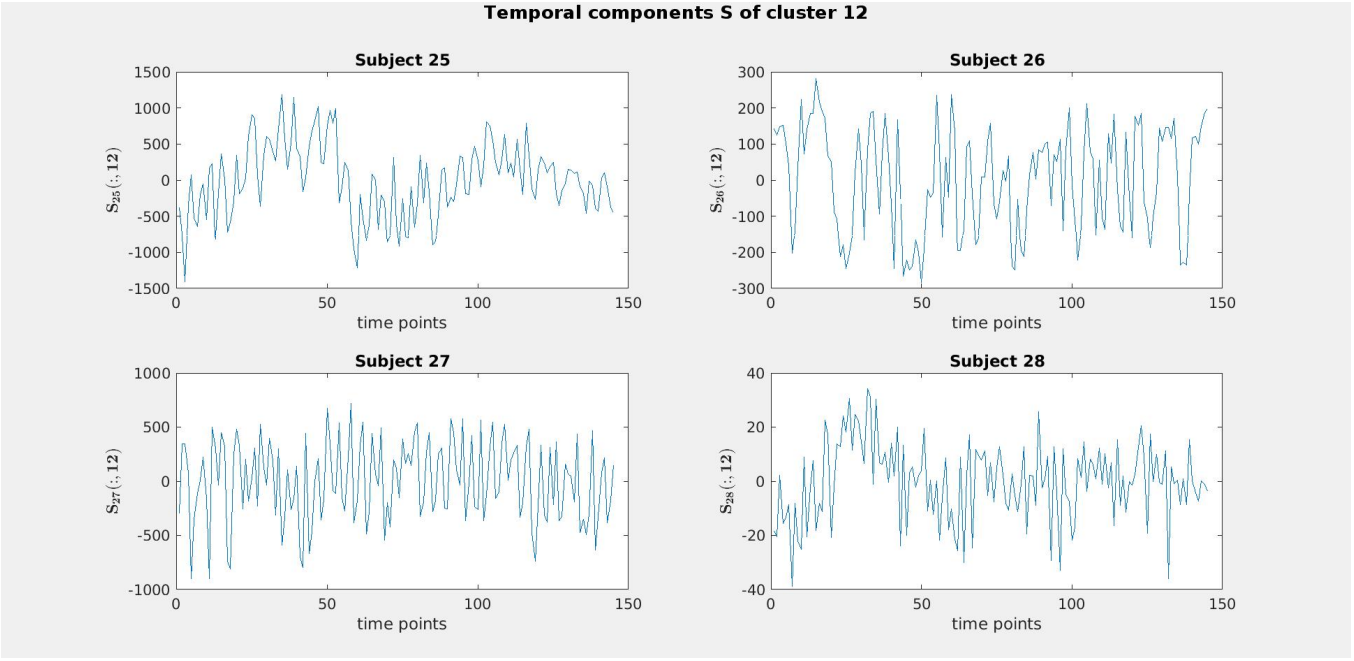


Figure B.642: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 25, \ldots, 28$ and R=42.

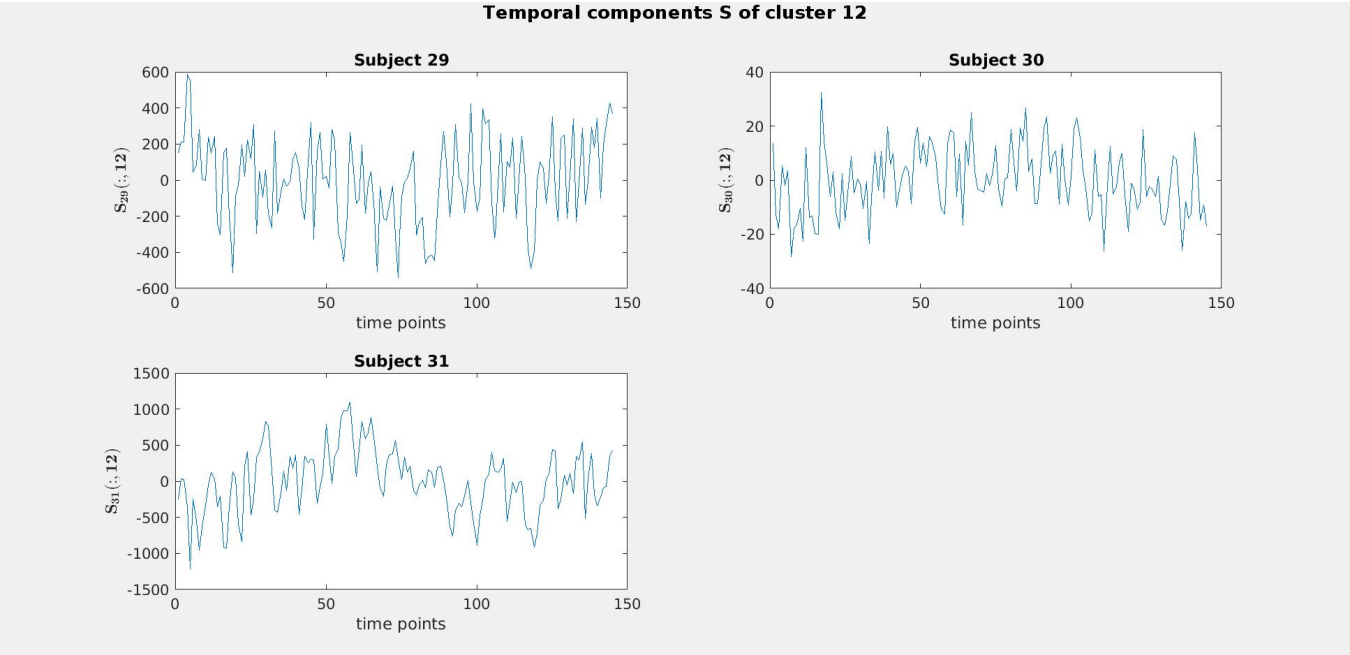


Figure B.643: Waveform of temporal components $\mathbf{S}_k(:, 12)$ for $k = 29, \ldots, 31$ and R=42.

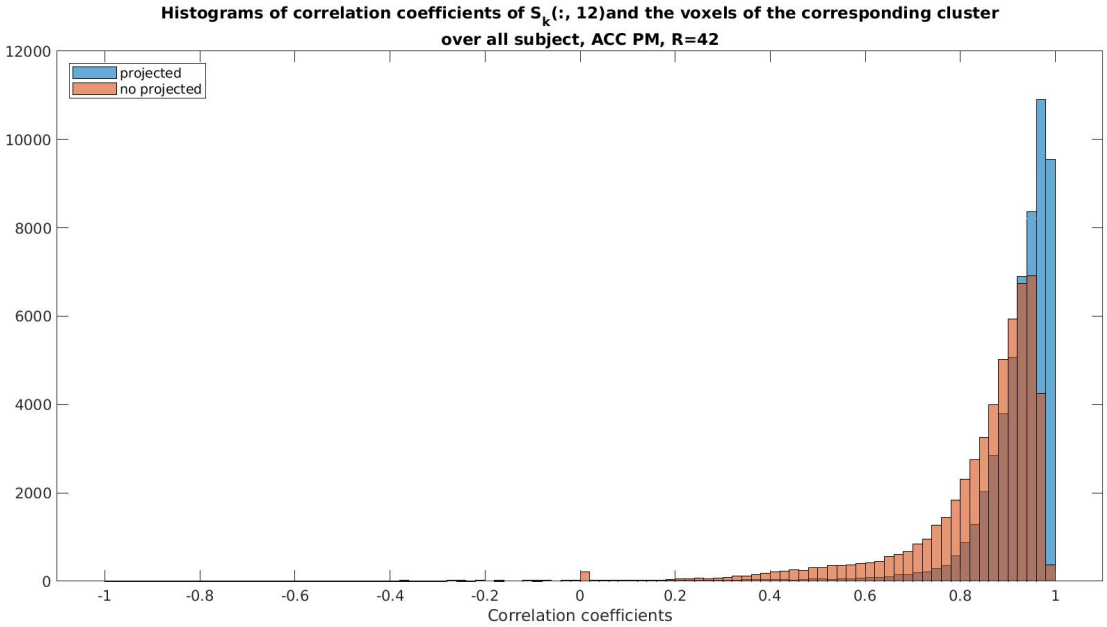


Figure B.644: For the column 12 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 12)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

36. **Cluster 26**

Cluster 26 contains 2135 voxels, which correspond to 0.944908% of the whole brain (background, soft tissue, bone are excluded).

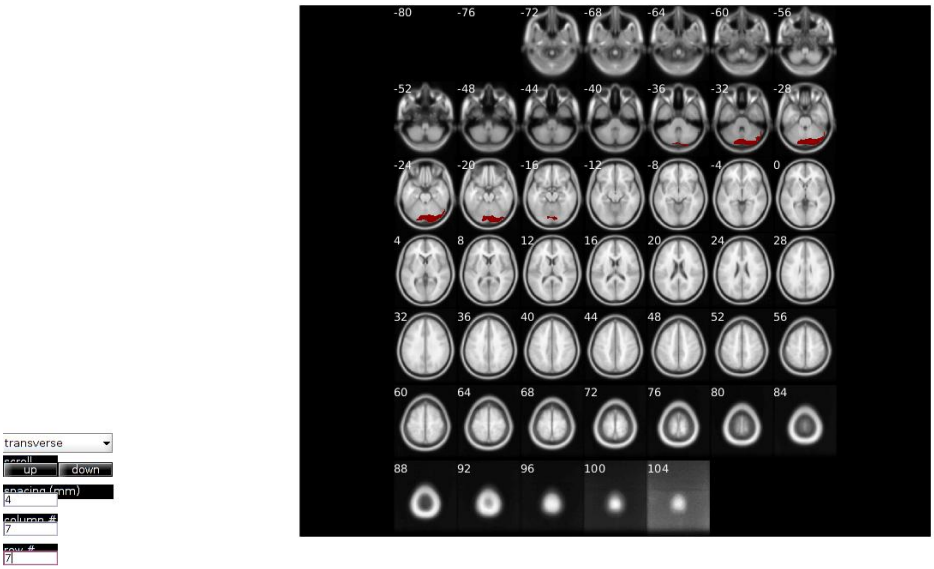


Figure B.645: Cluster 26. Red areas indicate the voxels which belong to cluster 26.

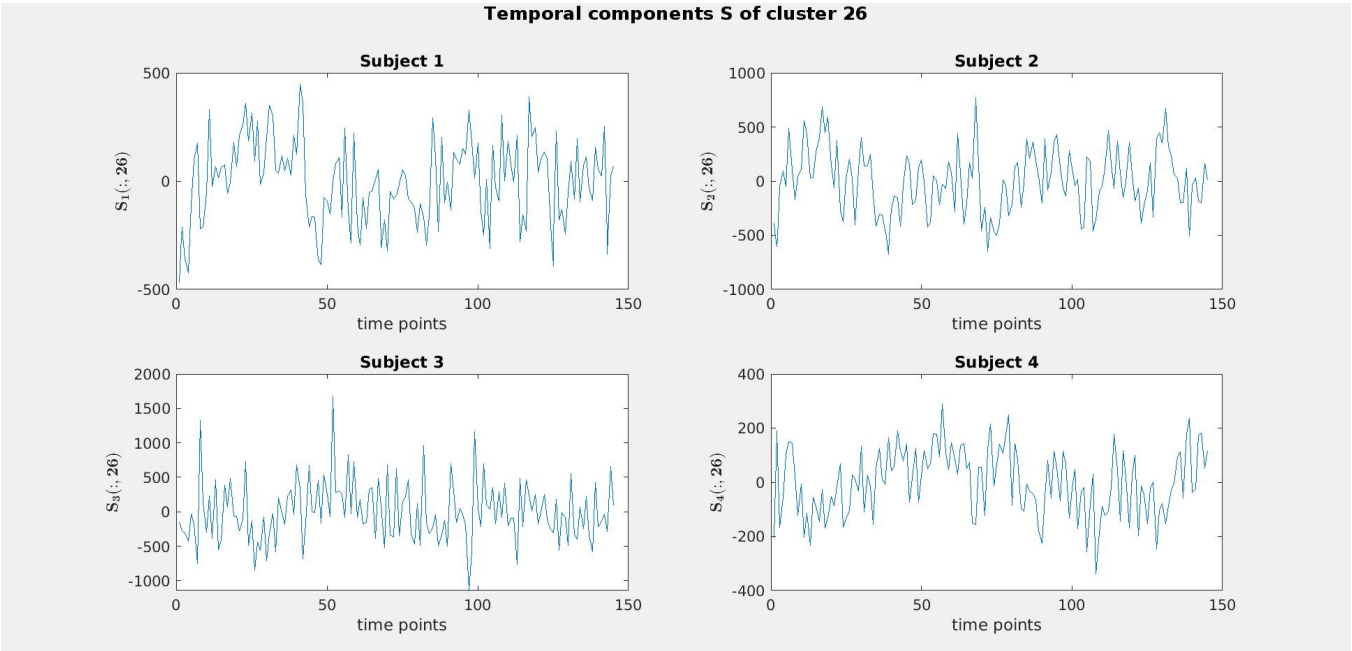


Figure B.646: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 1, \ldots, 4$ and R=42.

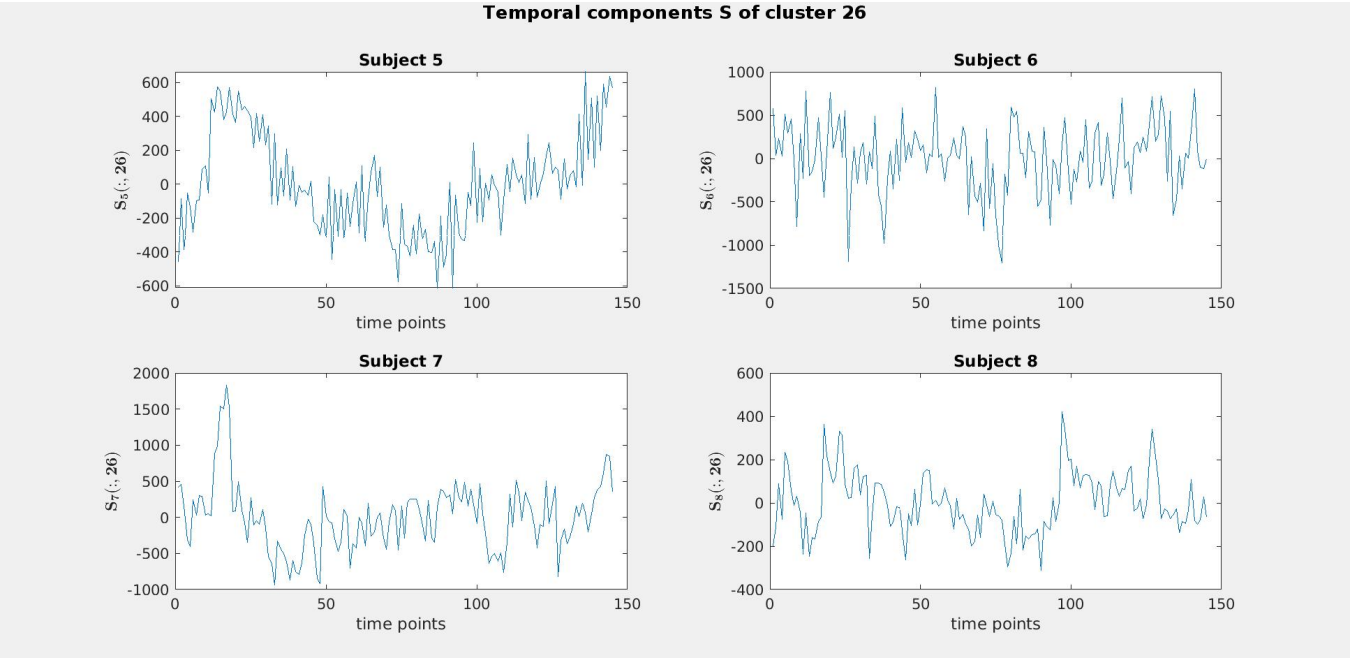


Figure B.647: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 5, \ldots, 8$ and R=42.

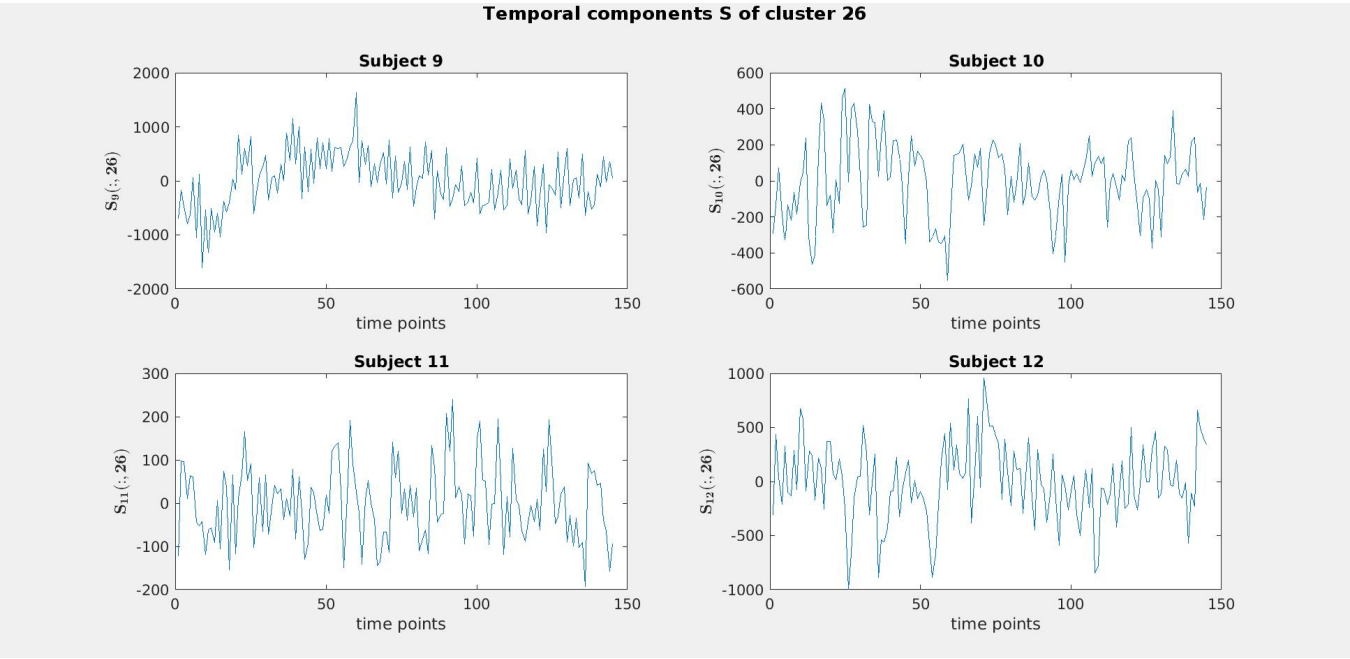


Figure B.648: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 9, \ldots, 12$ and R=42.

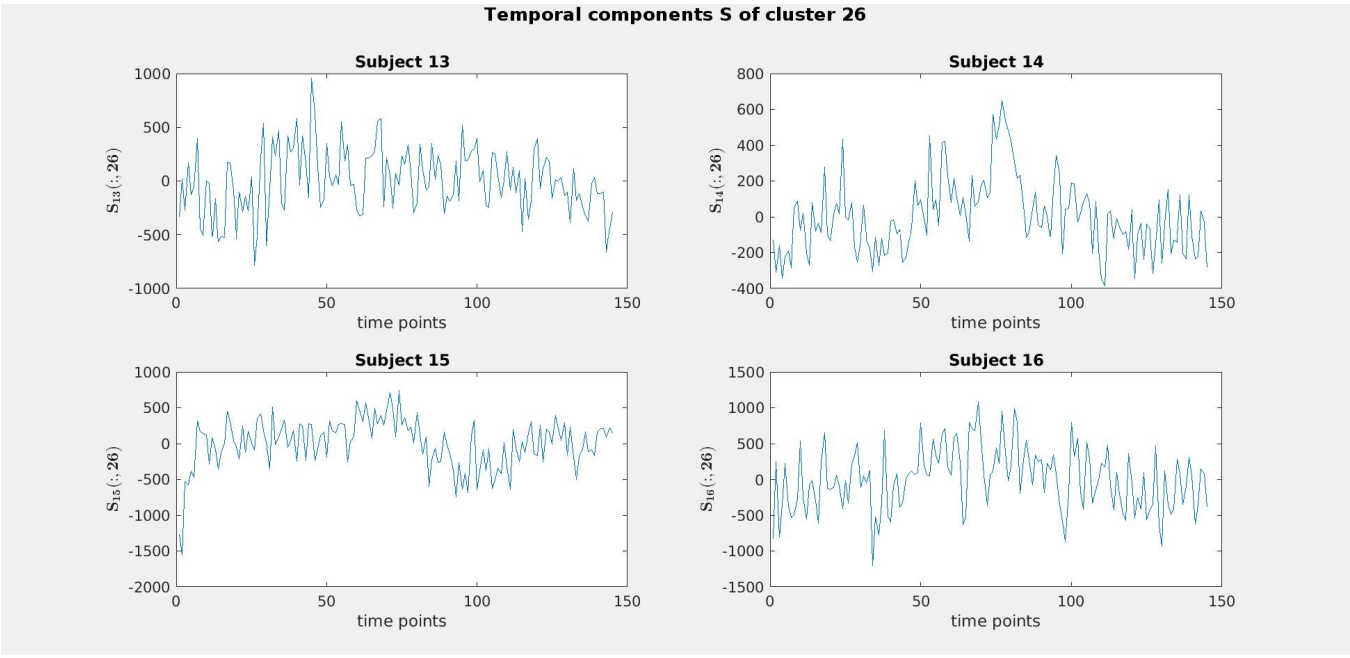


Figure B.649: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 13, \ldots, 16$ and R=42.

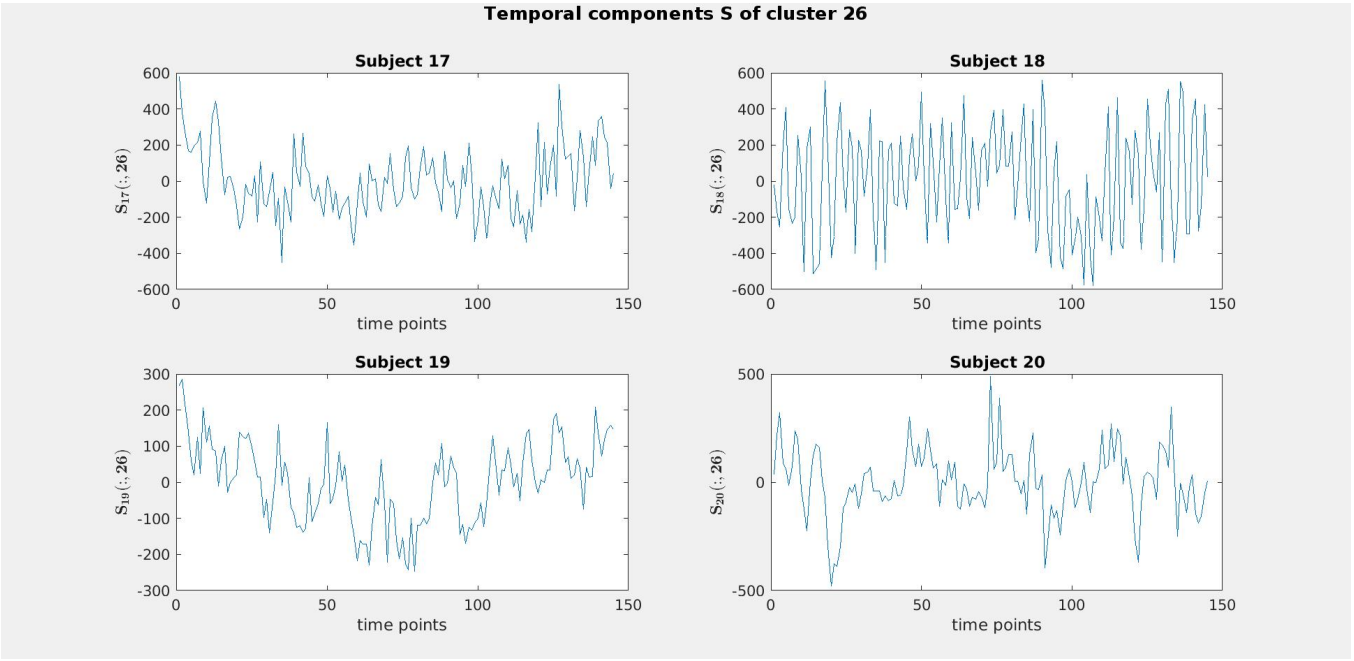


Figure B.650: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 17, \ldots, 20$ and R=42.

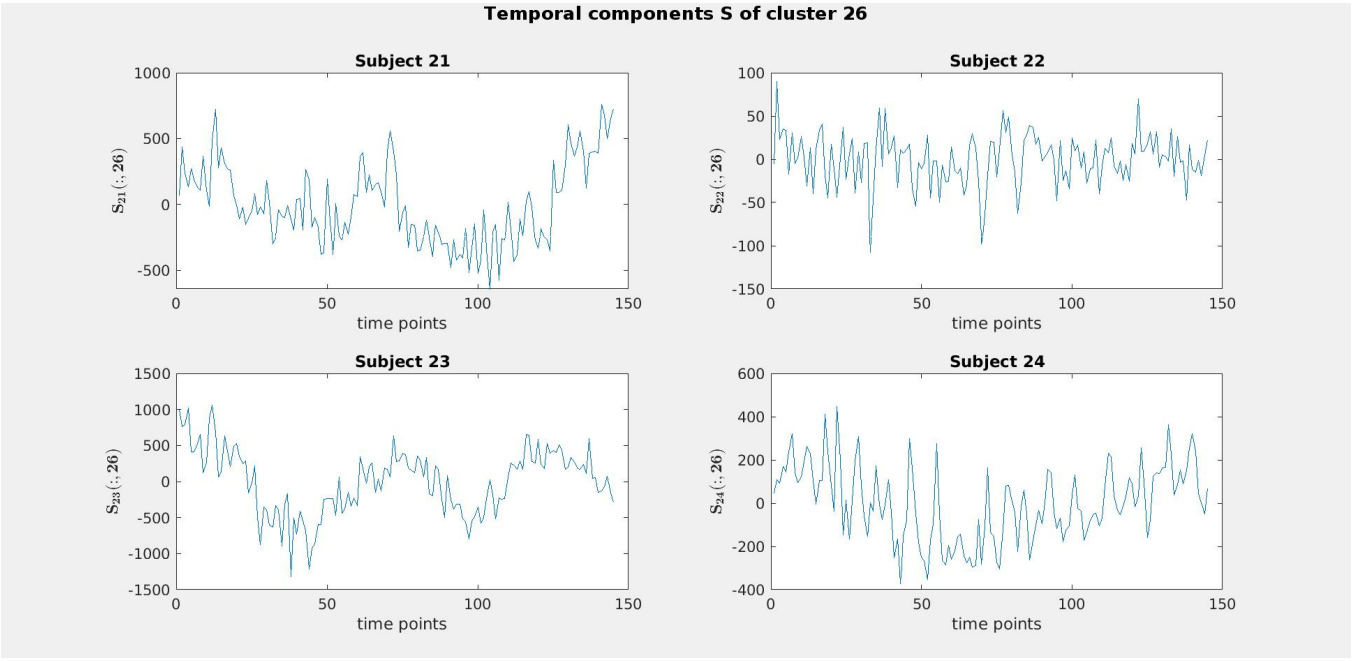


Figure B.651: Waveform of temporal components $\mathbf{S}_k(:, 126)$ for $k = 21, \ldots, 24$ and R=42.

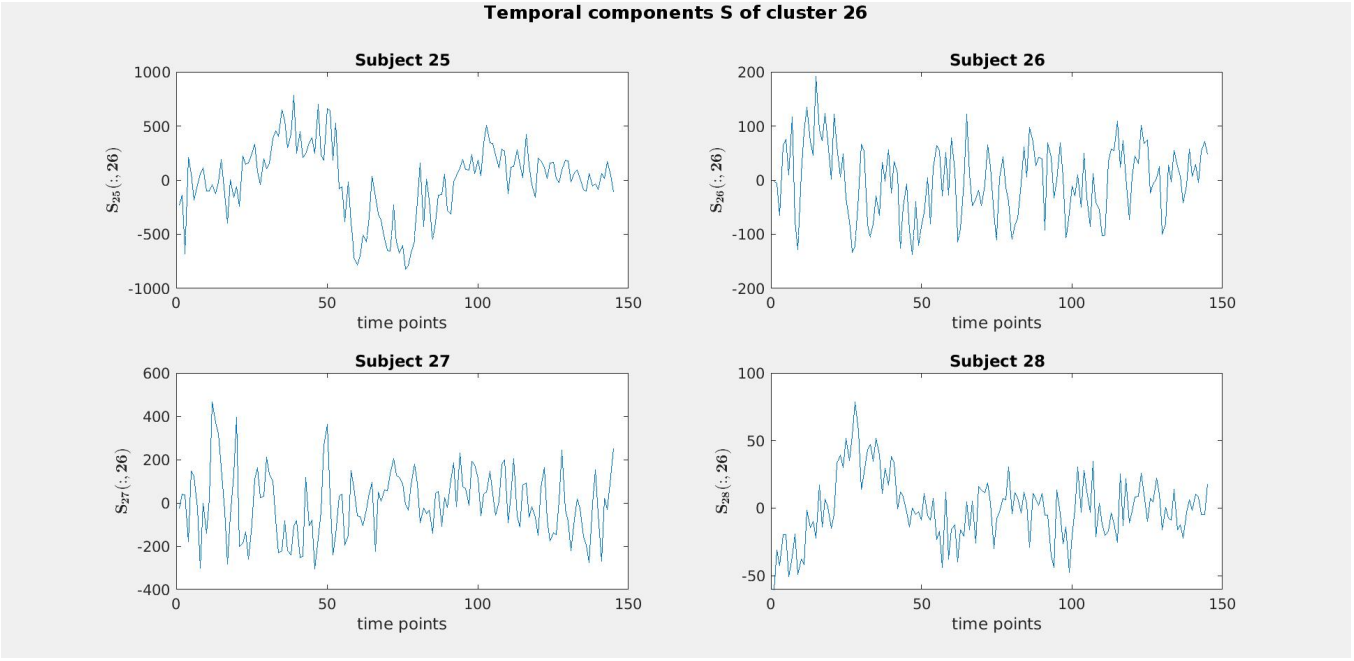


Figure B.652: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 25, \ldots, 28$ and R=42.

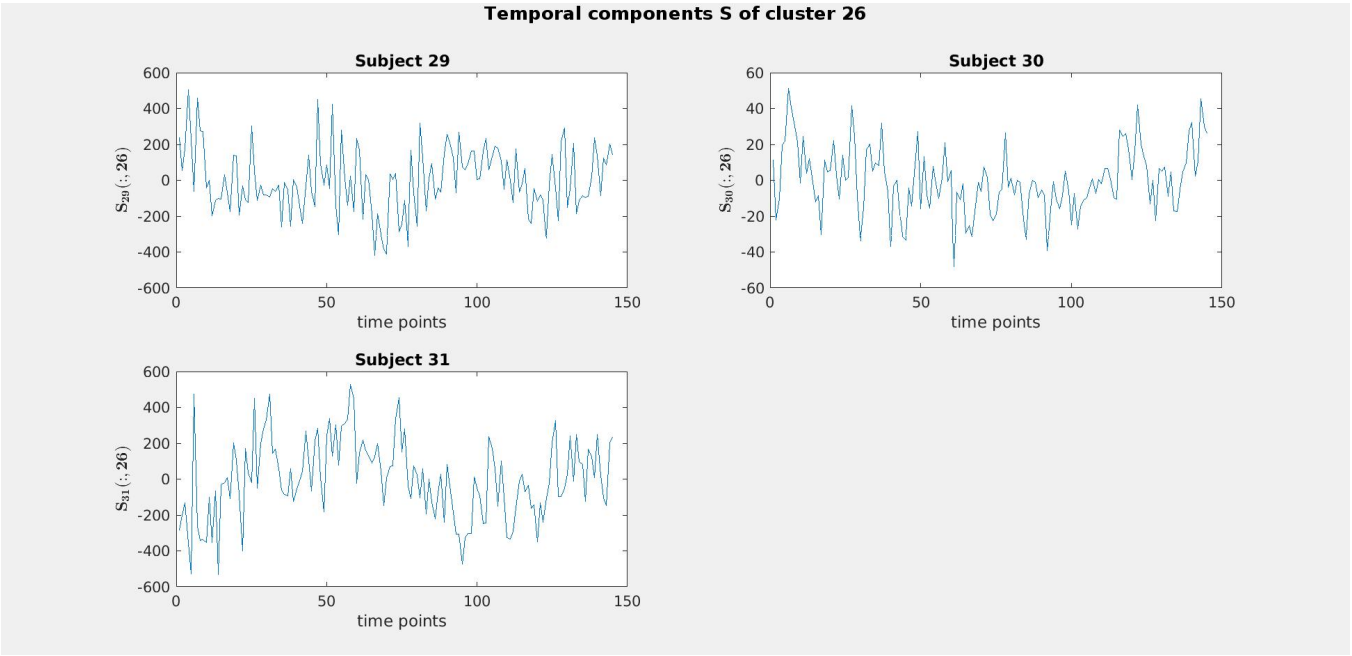


Figure B.653: Waveform of temporal components $\mathbf{S}_k(:, 26)$ for $k = 29, \ldots, 31$ and R=42.

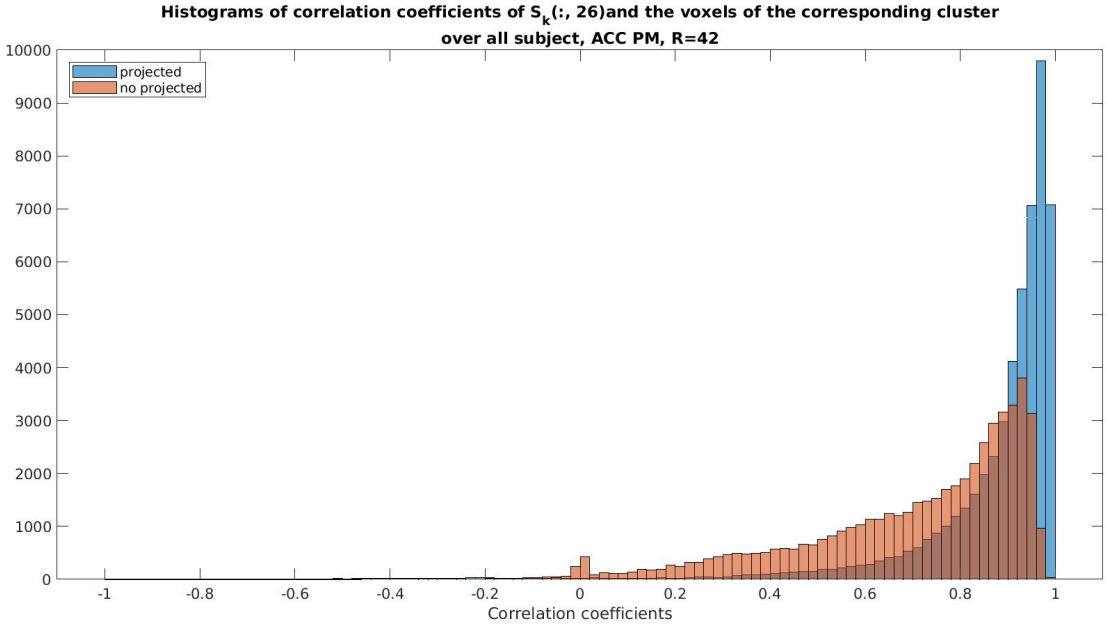


Figure B.654: For the column 26 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 26)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

37. **Cluster 4**

Cluster 4 contains 2043 voxels, which correspond to 0.904190% of the whole brain (background, soft tissue, bone are excluded).

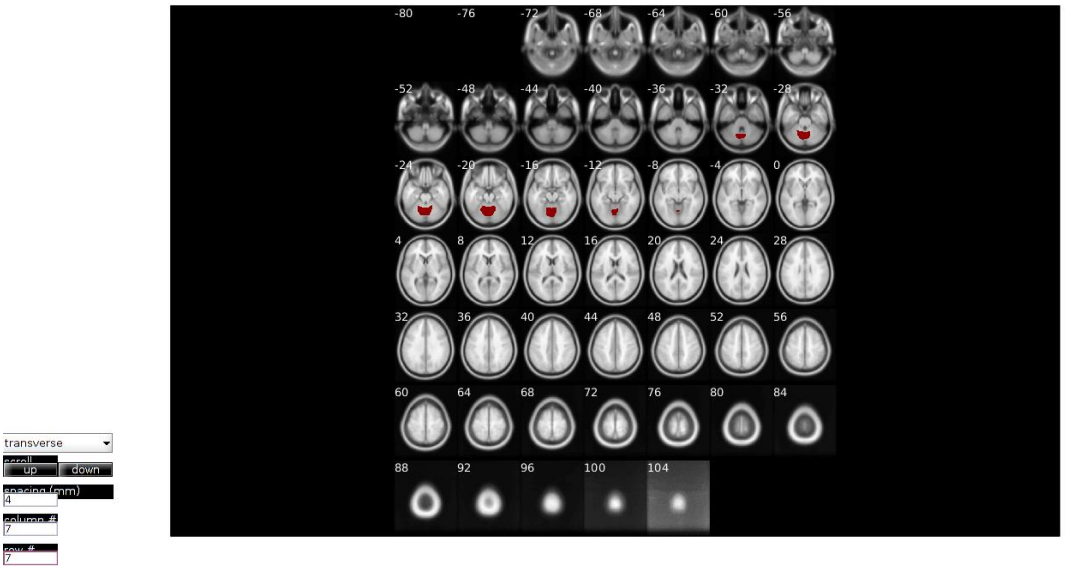


Figure B.655: Cluster 4. Red areas indicate the voxels which belong to cluster 4.

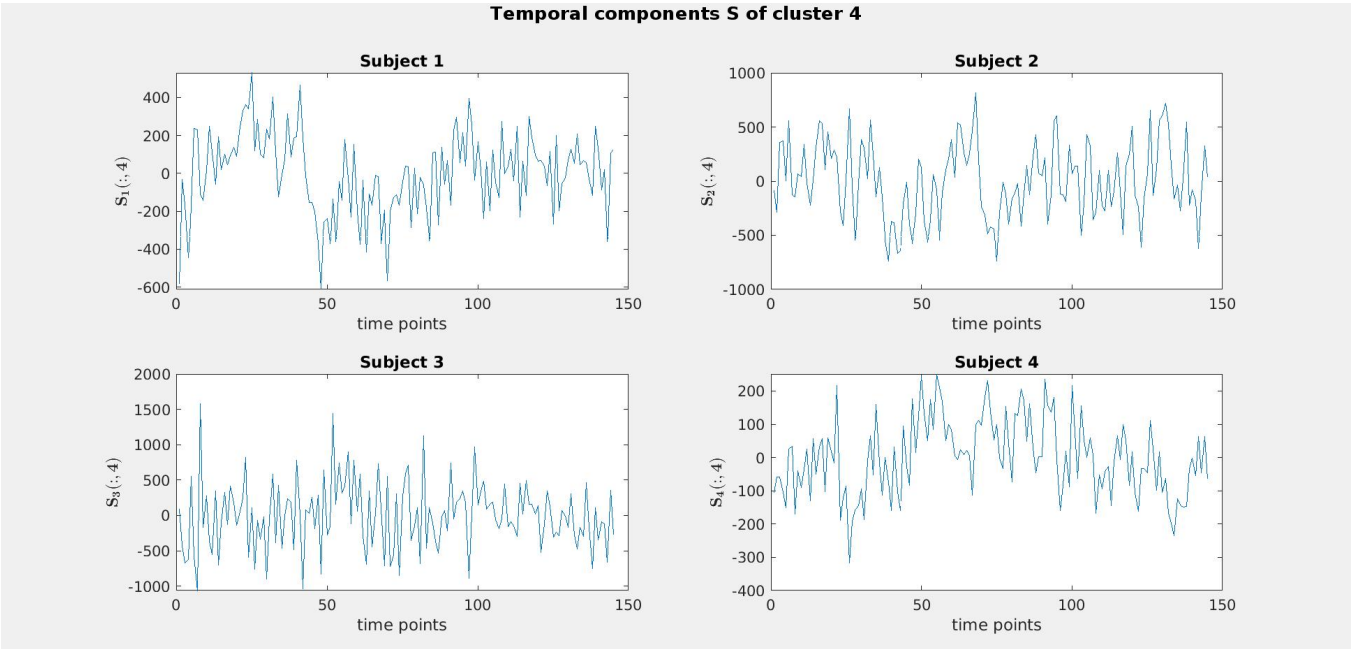


Figure B.656: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 1,\ldots,4$ and R=42.

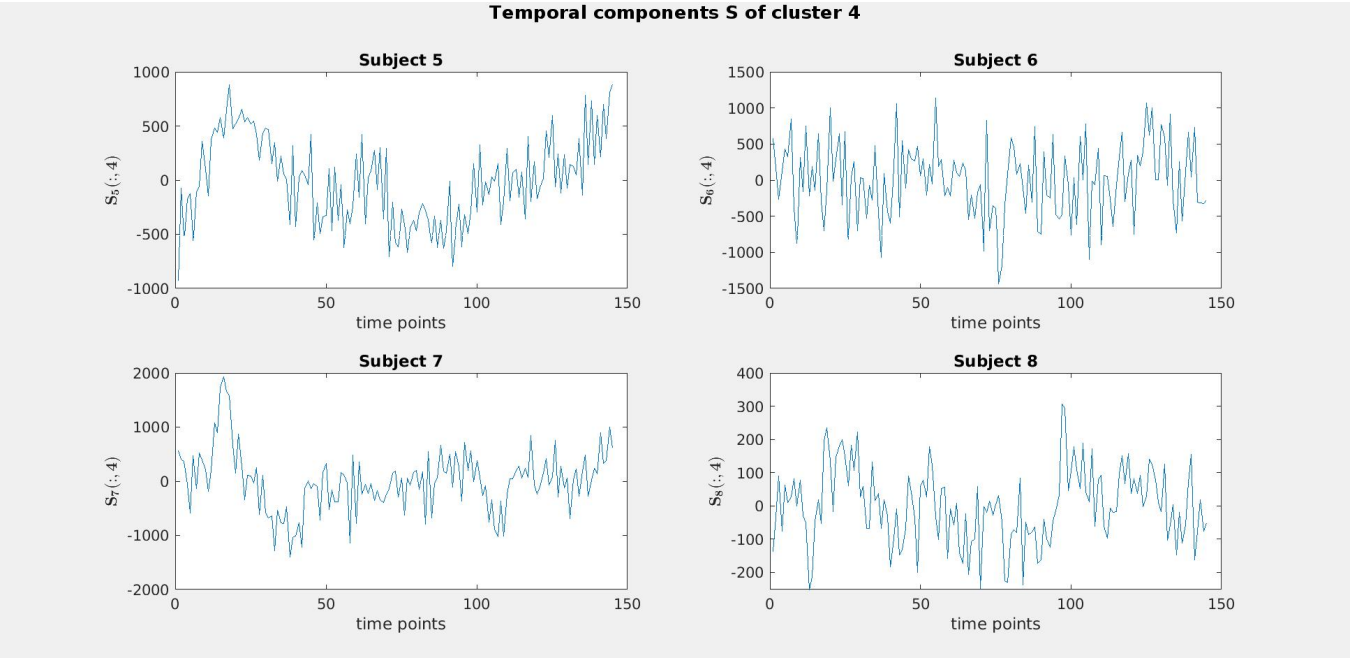


Figure B.657: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 5,\ldots,8$ and R=42.

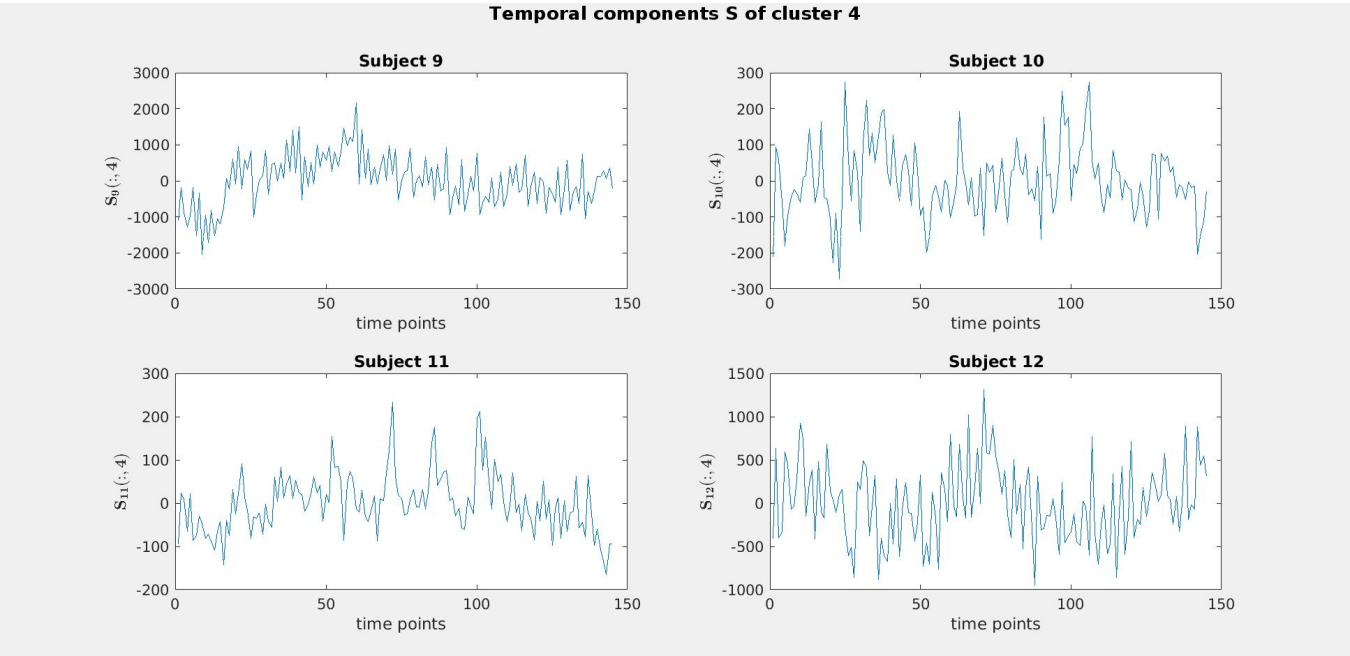


Figure B.658: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 9,\ldots,12$ and R=42.

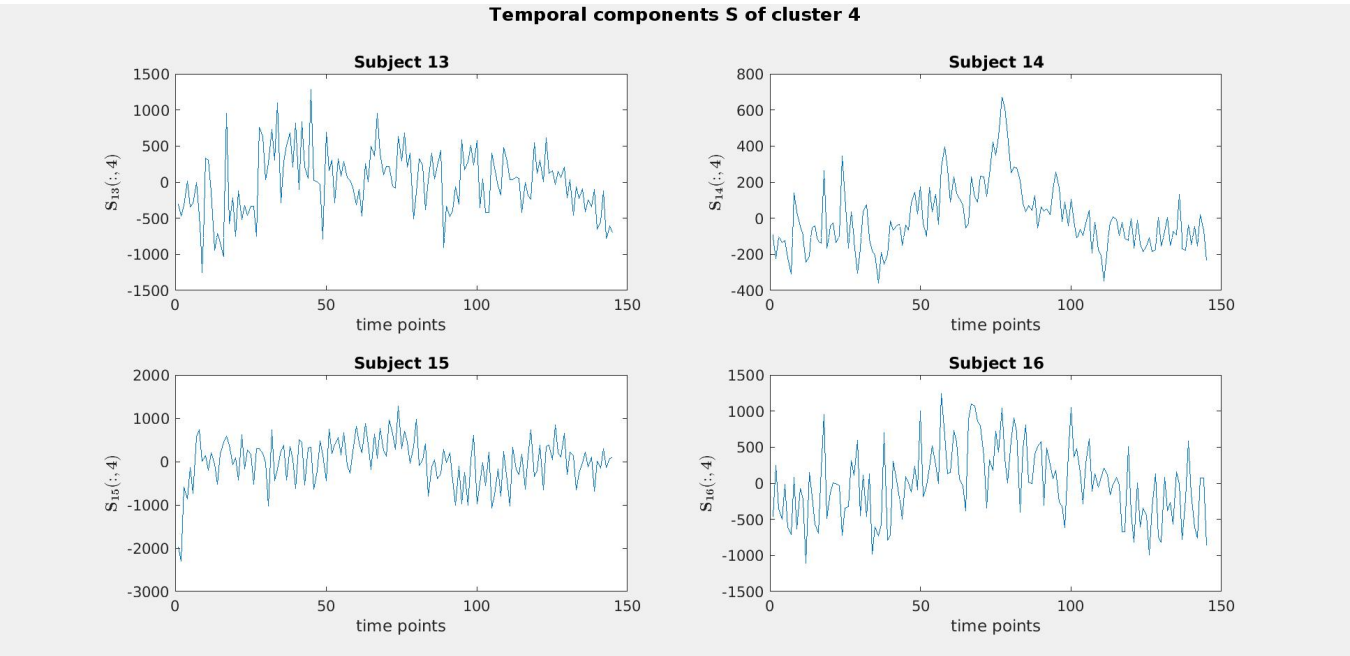


Figure B.659: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 13,\ldots,16$ and R=42.

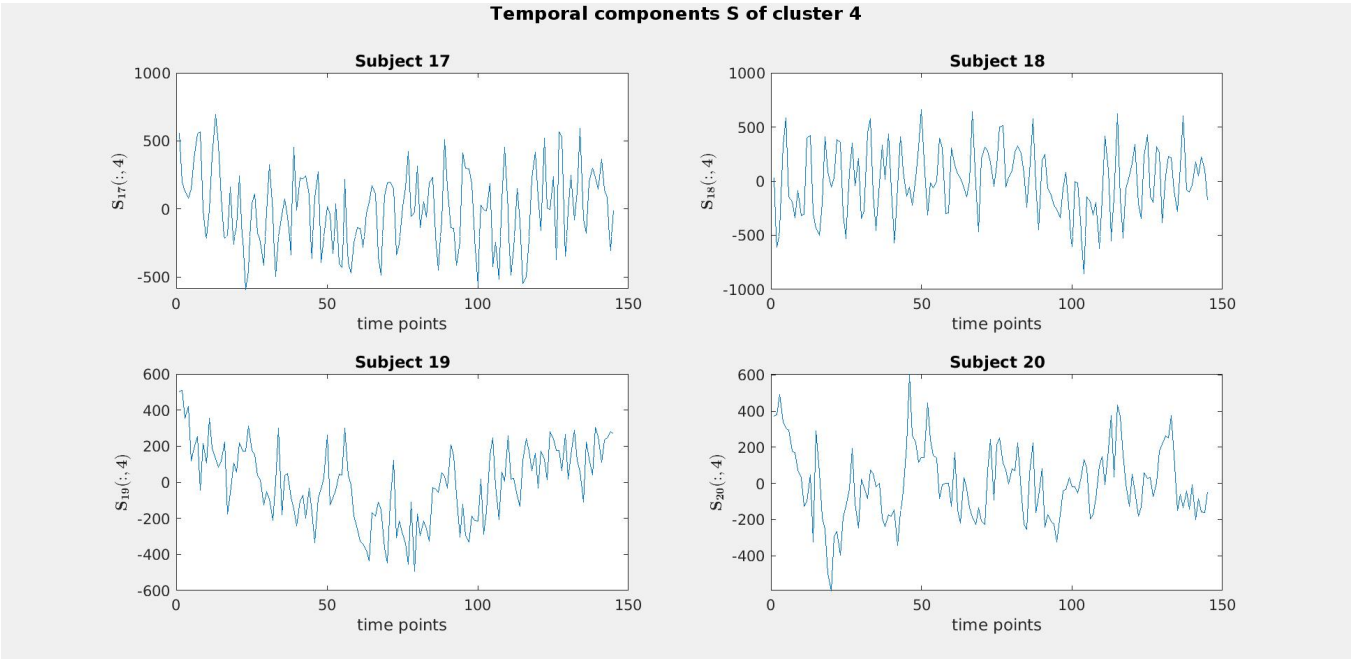


Figure B.660: Waveform of temporal components $\mathbf{S}_k(:, 4)$ for $k = 17, \ldots, 20$ and R=42.

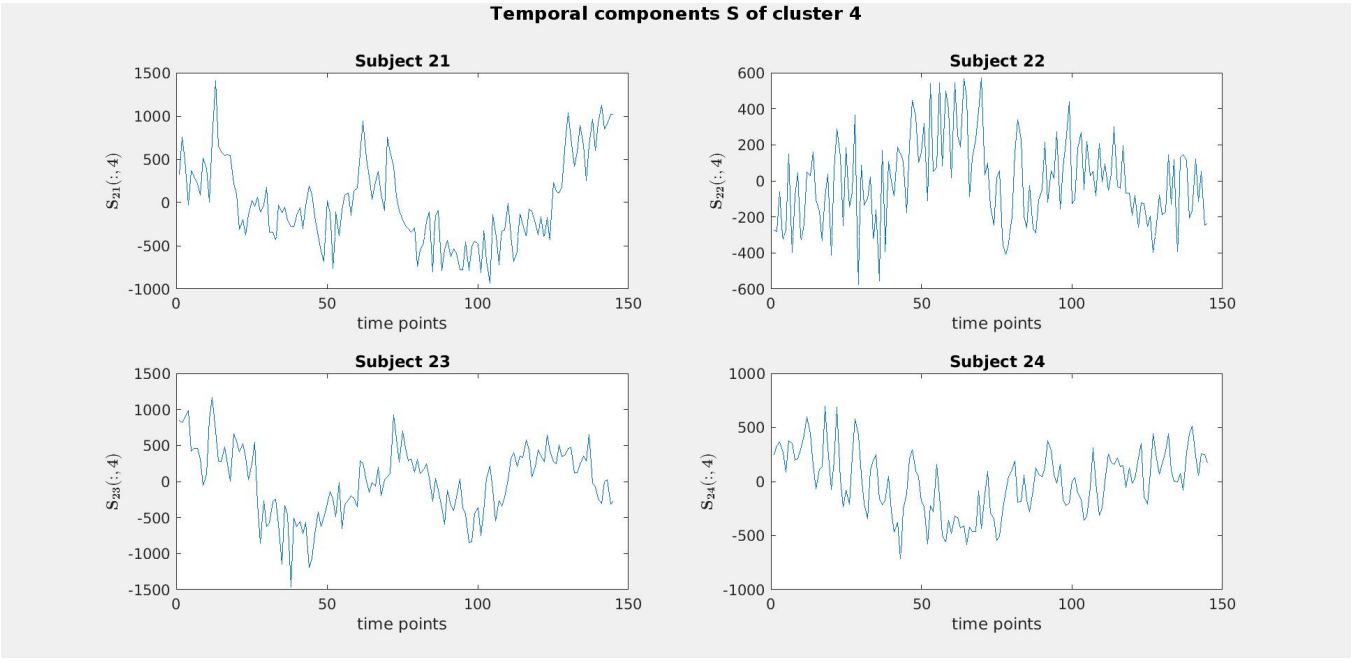


Figure B.661: Waveform of temporal components $\mathbf{S}_k(:, 14)$ for $k = 21, \ldots, 24$ and R=42.

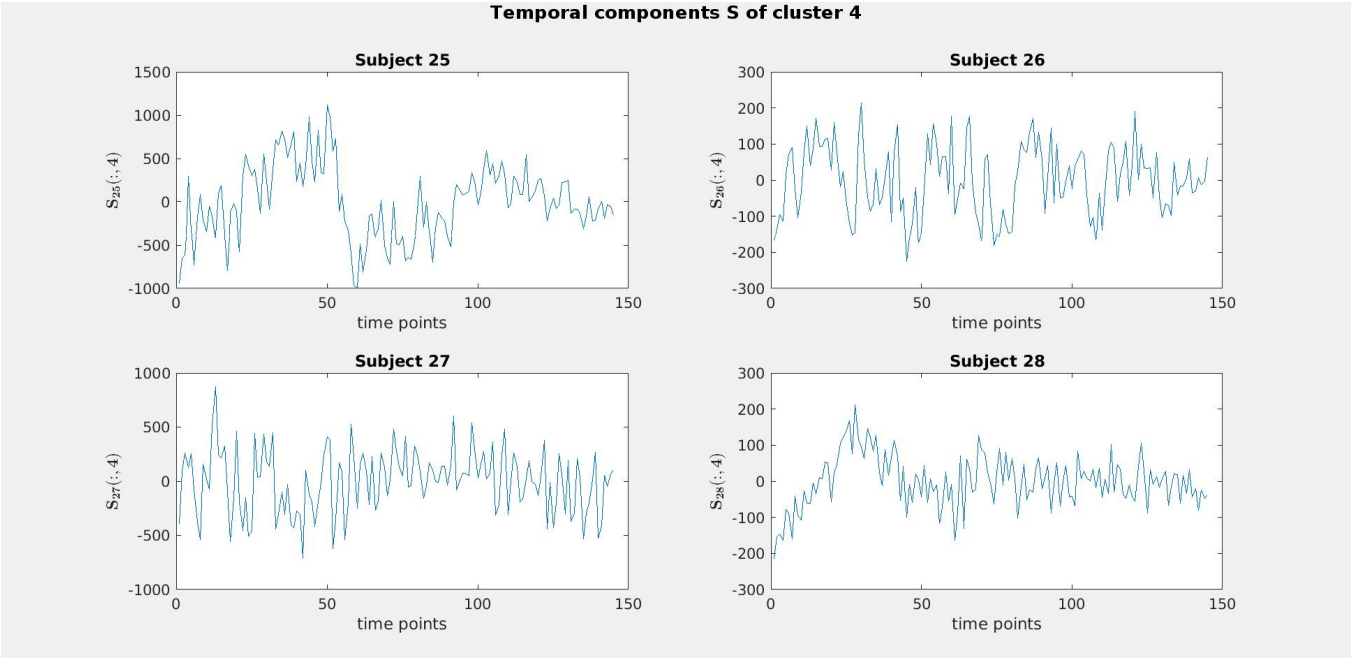


Figure B.662: Waveform of temporal components $\mathbf{S}_k(:, 4)$ for $k = 25, \ldots, 28$ and R=42.

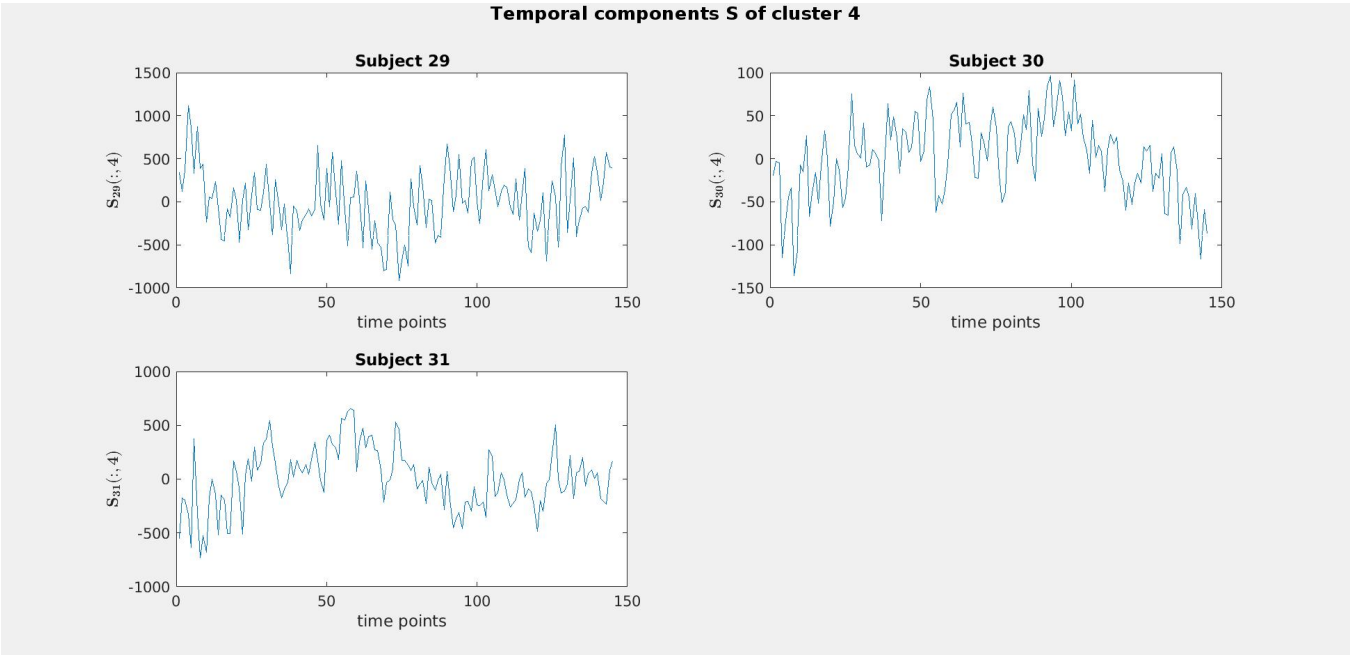


Figure B.663: Waveform of temporal components $\mathbf{S}_k(:,4)$ for $k = 29, \ldots, 31$ and R=42.

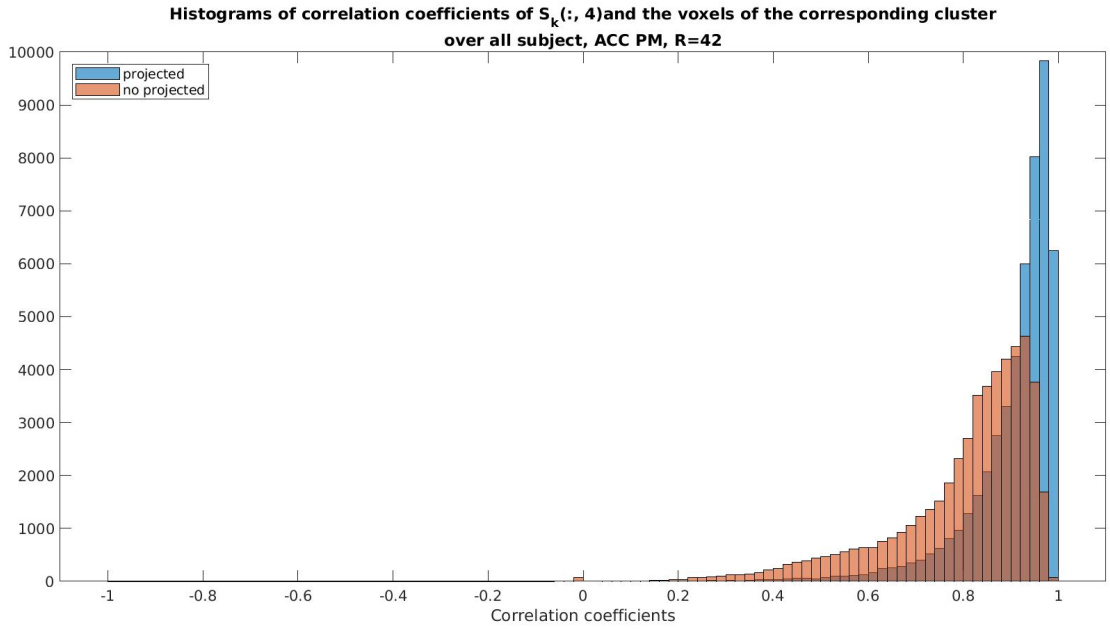


Figure B.664: For the column 4 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:,4)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

38. **Cluster 27**

Cluster 27 contains 2019 voxels, which correspond to 0.893568% of the whole brain (background, soft tissue, bone are excluded).

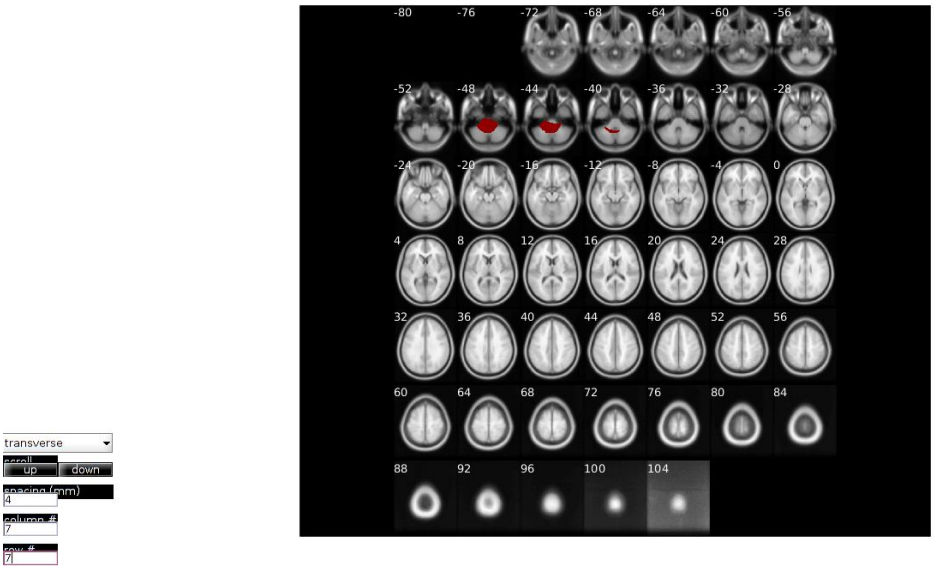


Figure B.665: Cluster 27. Red areas indicate the voxels which belong to cluster 27.

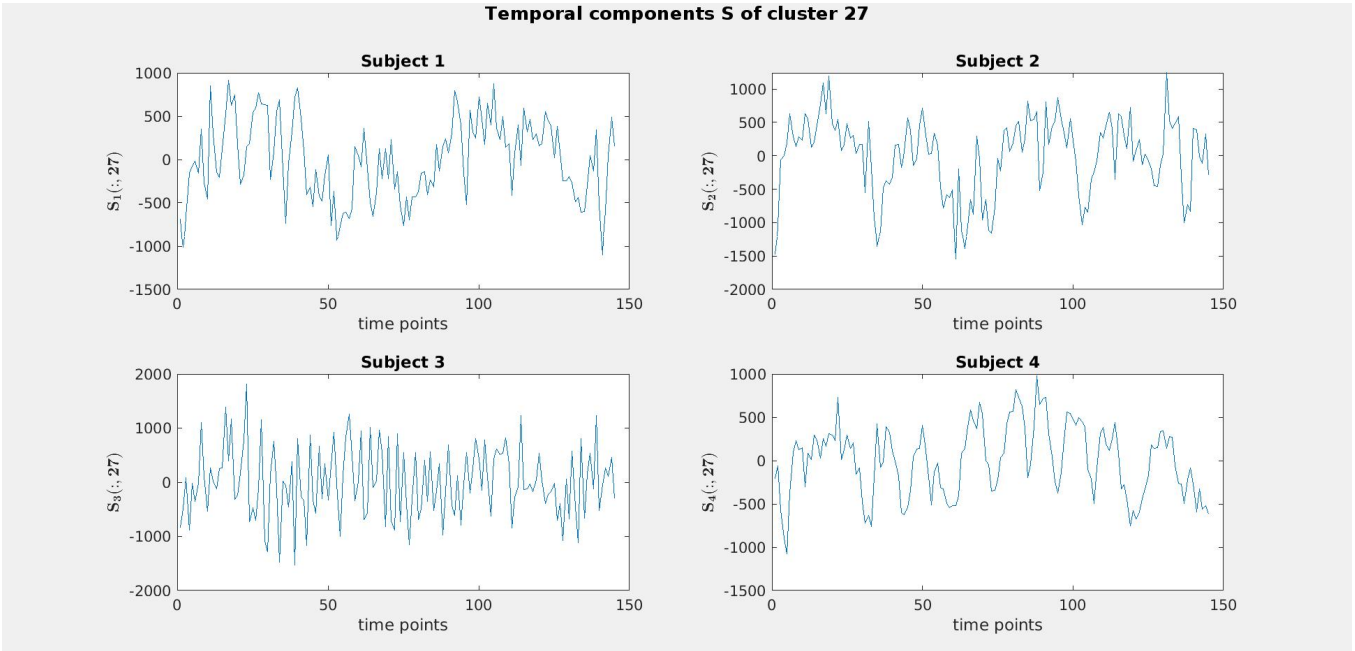


Figure B.666: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 1, \ldots, 4$ and R=42.

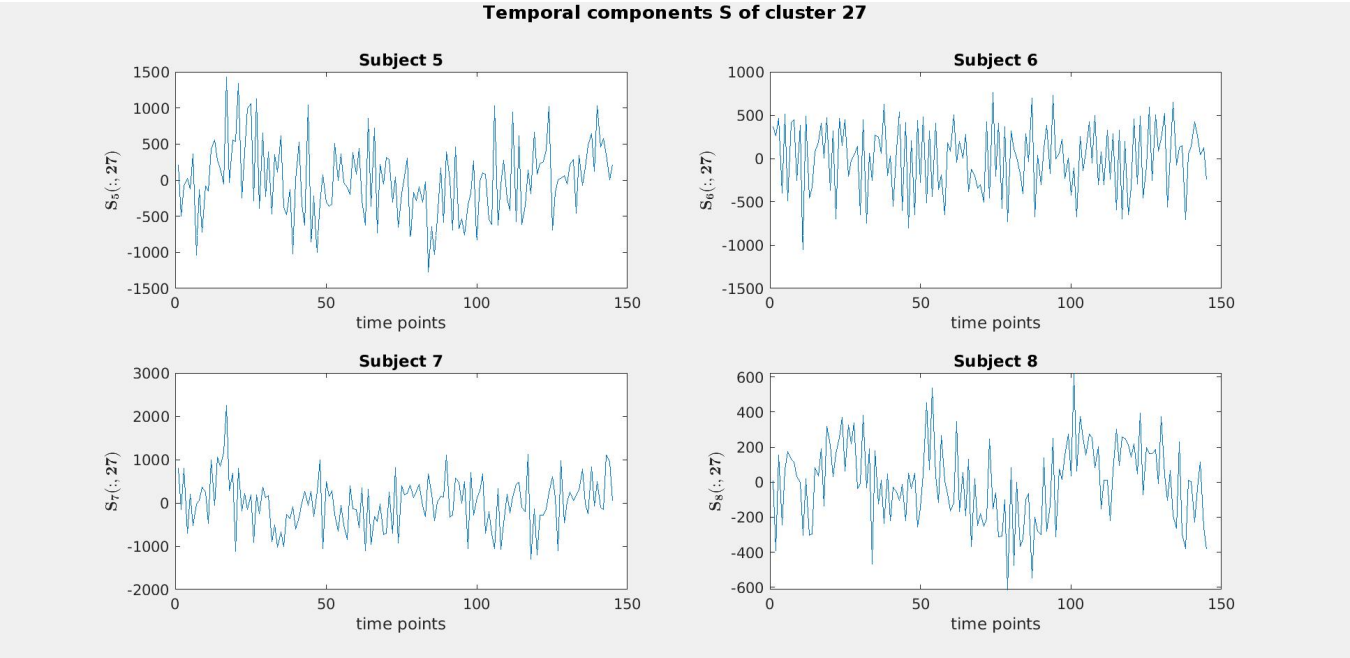


Figure B.667: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 5, \ldots, 8$ and R=42.

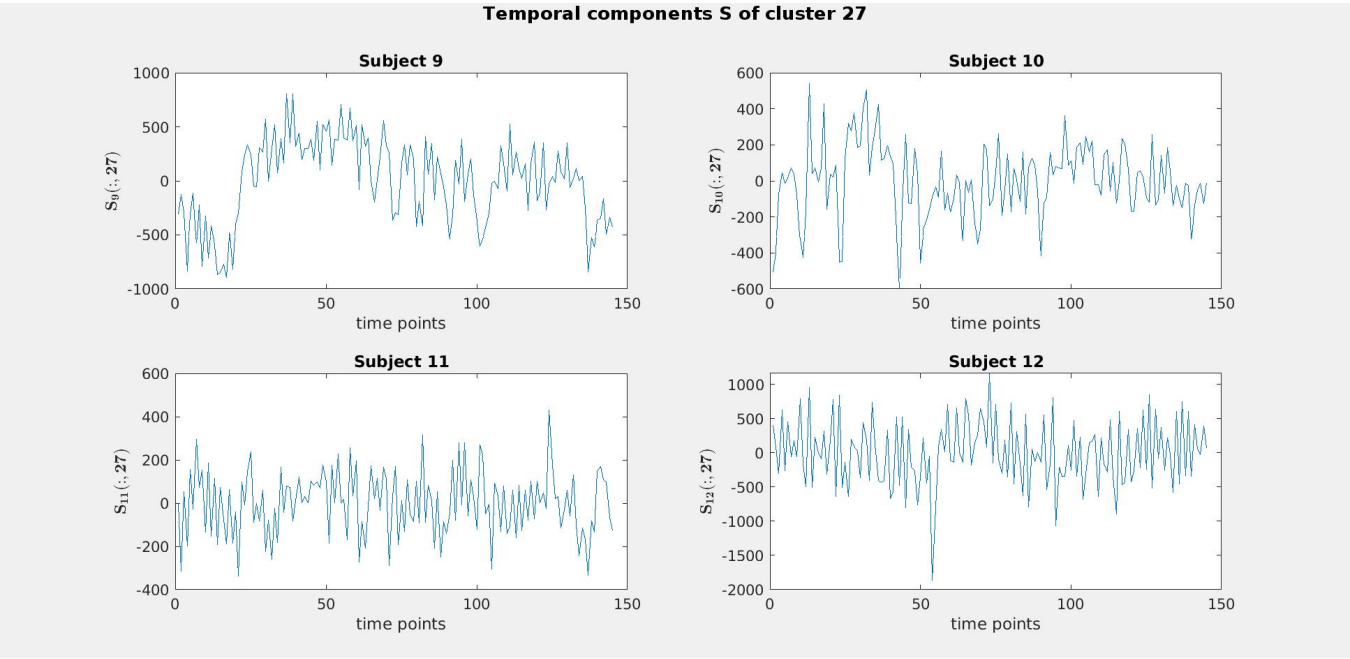


Figure B.668: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 9, \ldots, 12$ and R=42.

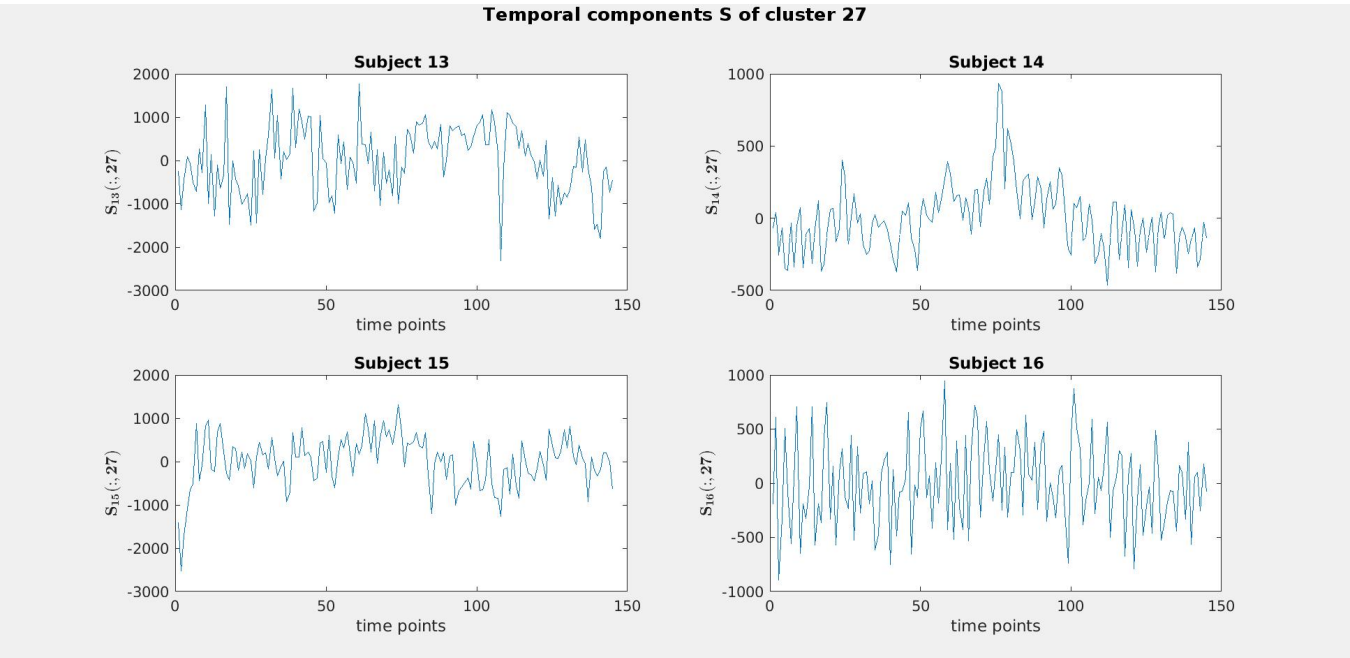


Figure B.669: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 13, \ldots, 16$ and R=42.

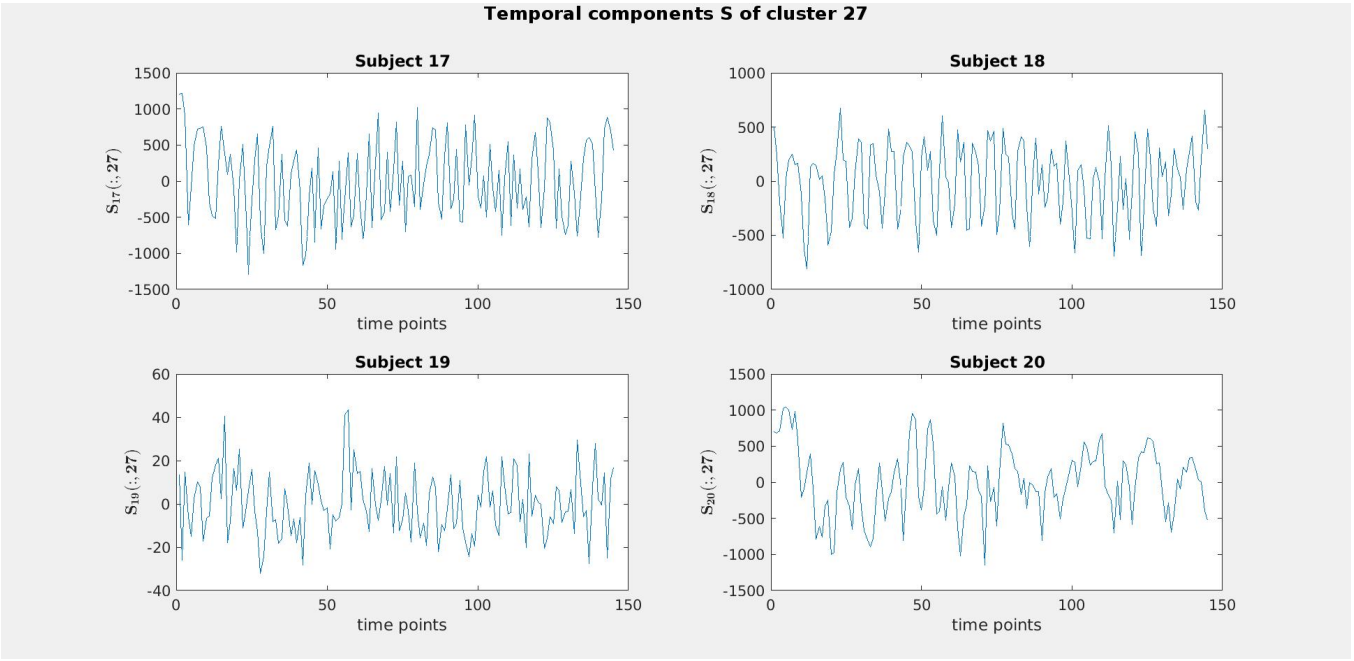


Figure B.670: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 17, \ldots, 20$ and R=42.

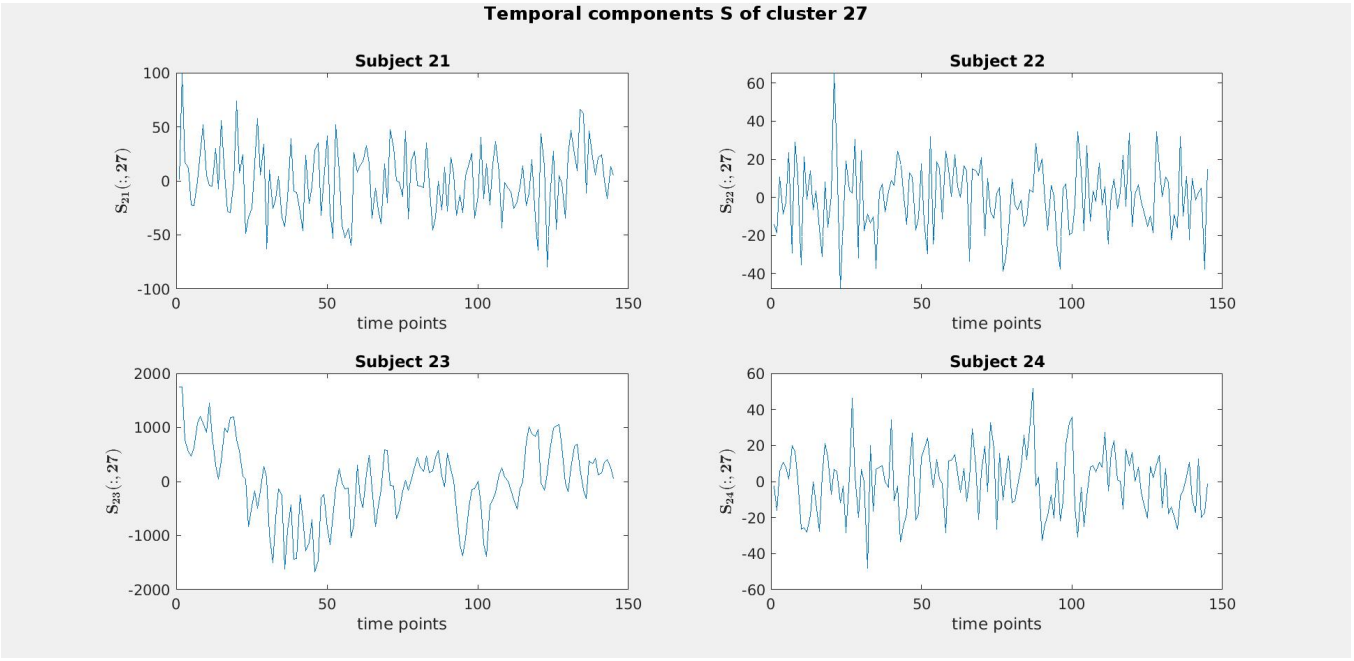


Figure B.671: Waveform of temporal components $\mathbf{S}_k(:, 127)$ for $k = 21, \ldots, 24$ and R=42.

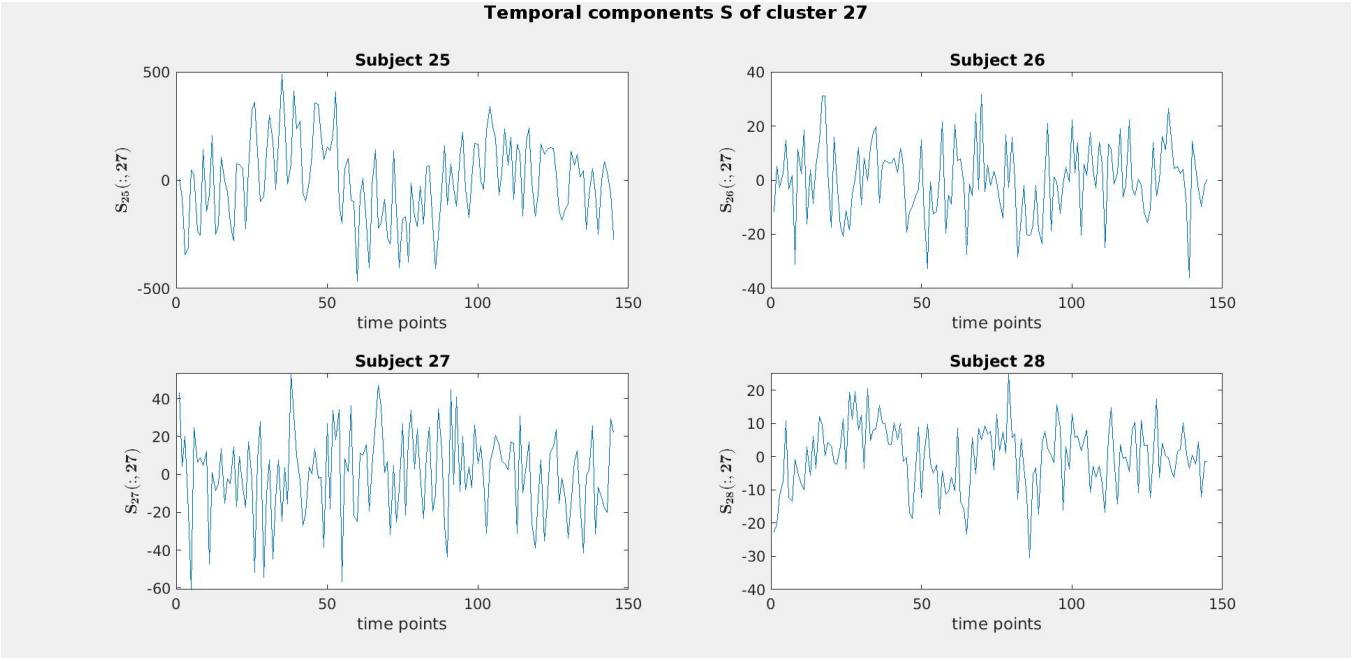


Figure B.672: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 25, \ldots, 28$ and R=42.

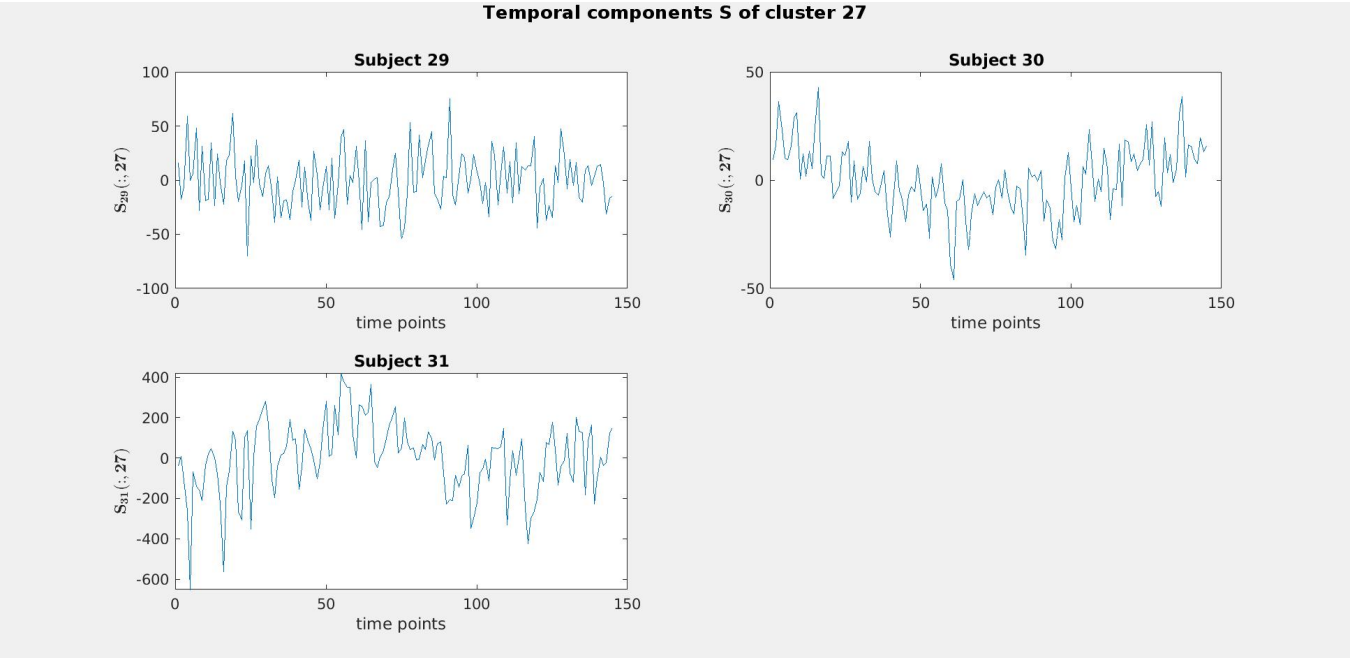


Figure B.673: Waveform of temporal components $\mathbf{S}_k(:, 27)$ for $k = 29, \ldots, 31$ and R=42.

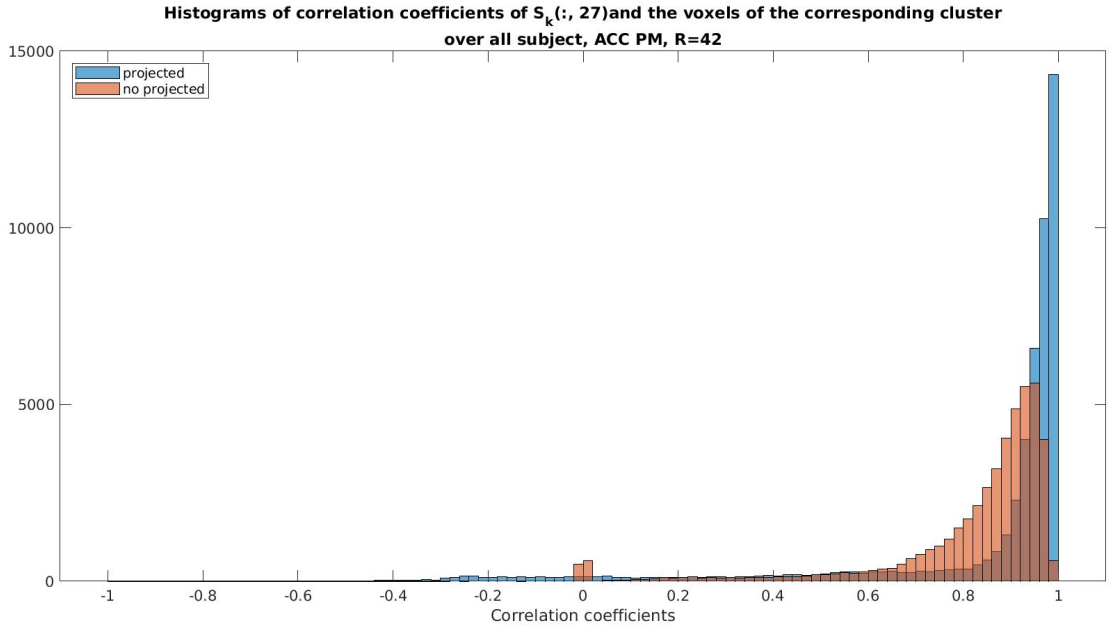


Figure B.674: For the column 27 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 27)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

39. **Cluster 24**

Cluster 24 contains 1948 voxels, which correspond to 0.862145% of the whole brain (background, soft tissue, bone are excluded).

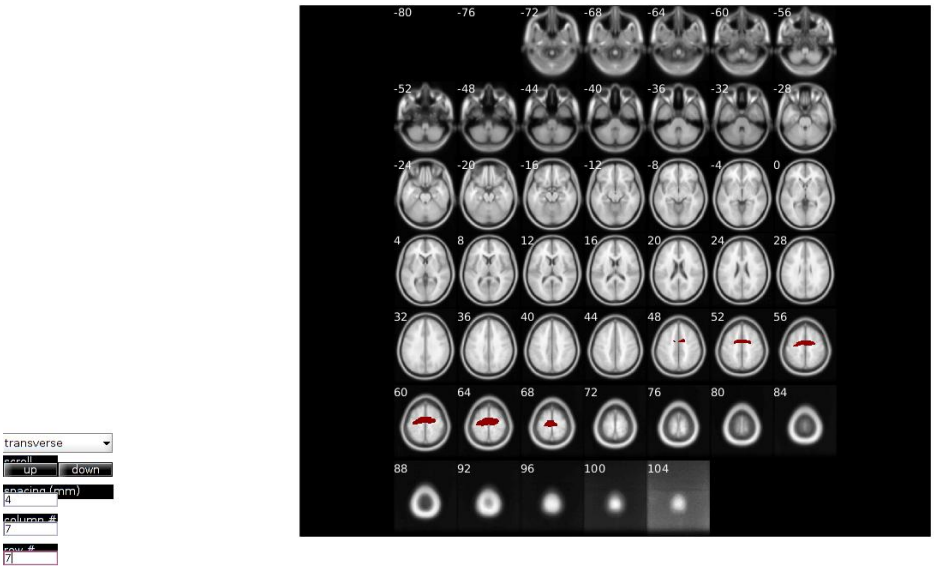


Figure B.675: Cluster 24. Red areas indicate the voxels which belong to cluster 24.

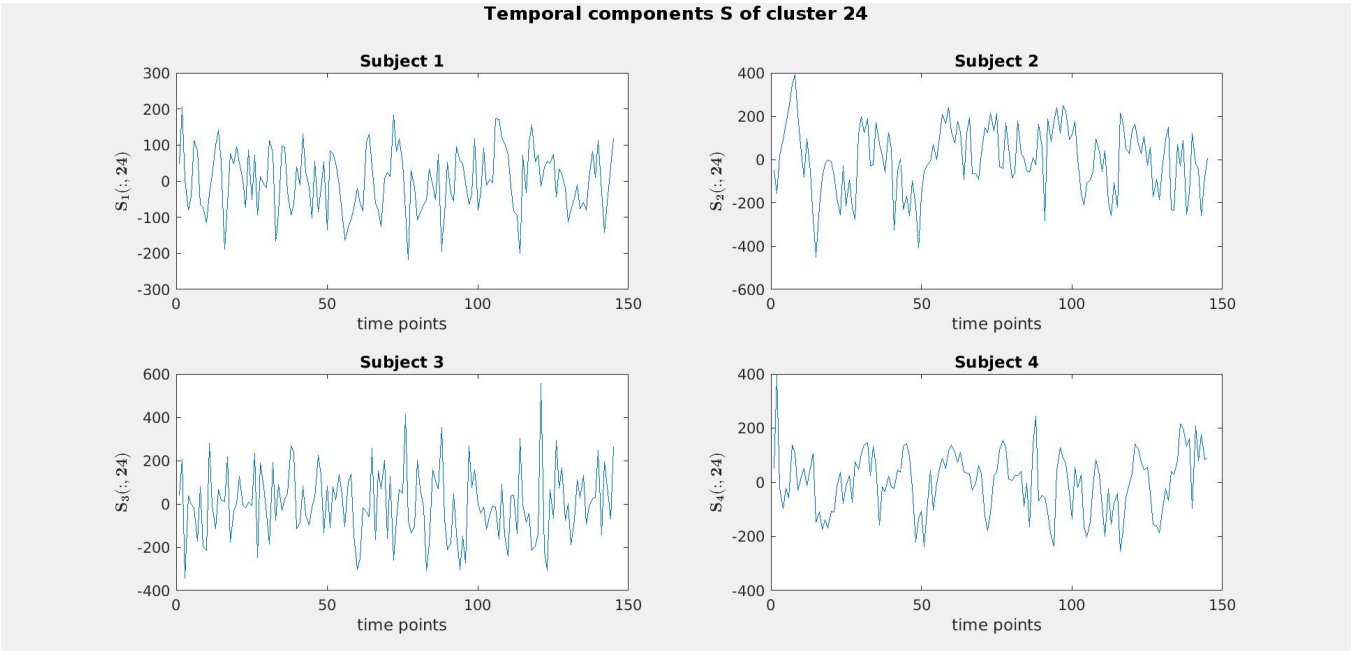


Figure B.676: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 1, \ldots, 4$ and R=42.

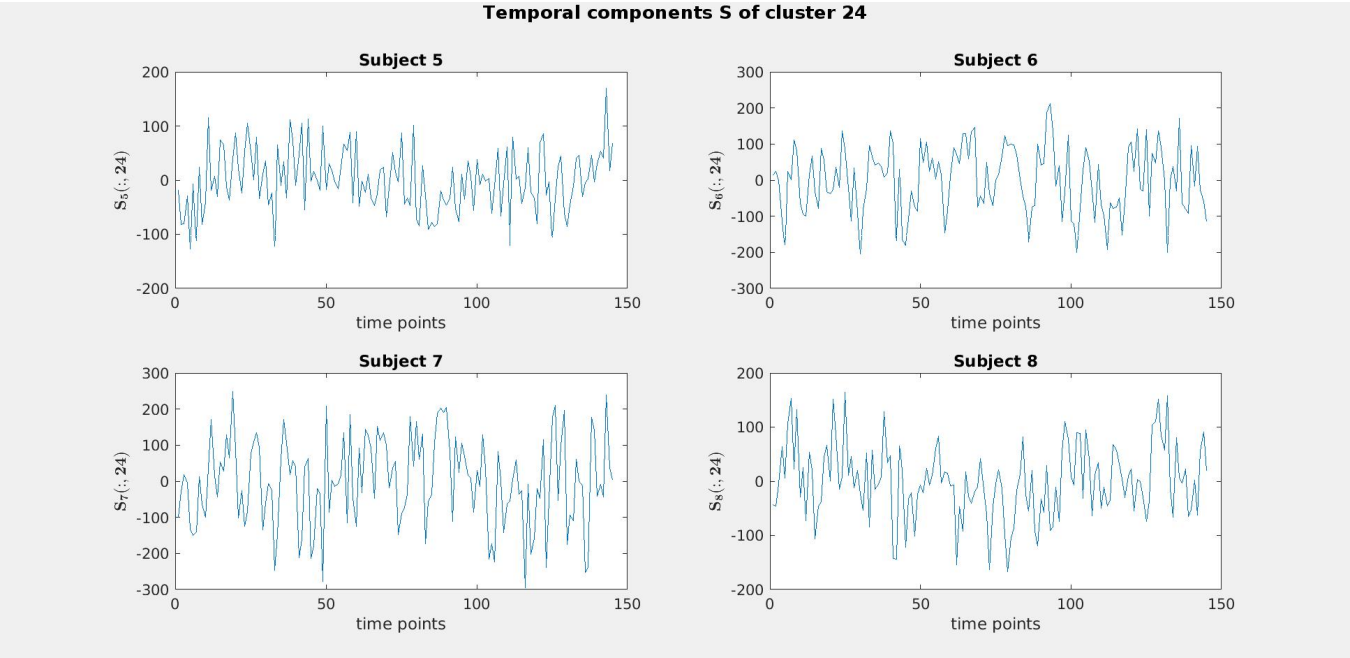


Figure B.677: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 5, \ldots, 8$ and R=42.

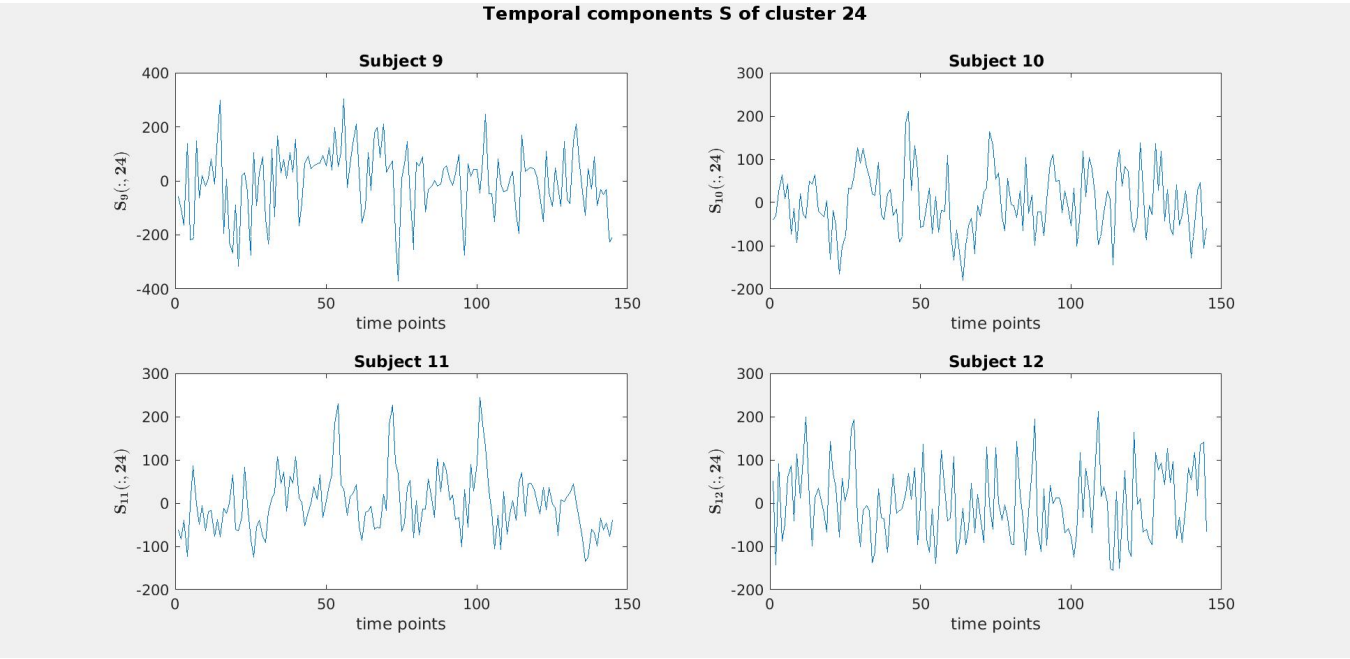


Figure B.678: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 9, \ldots, 12$ and R=42.

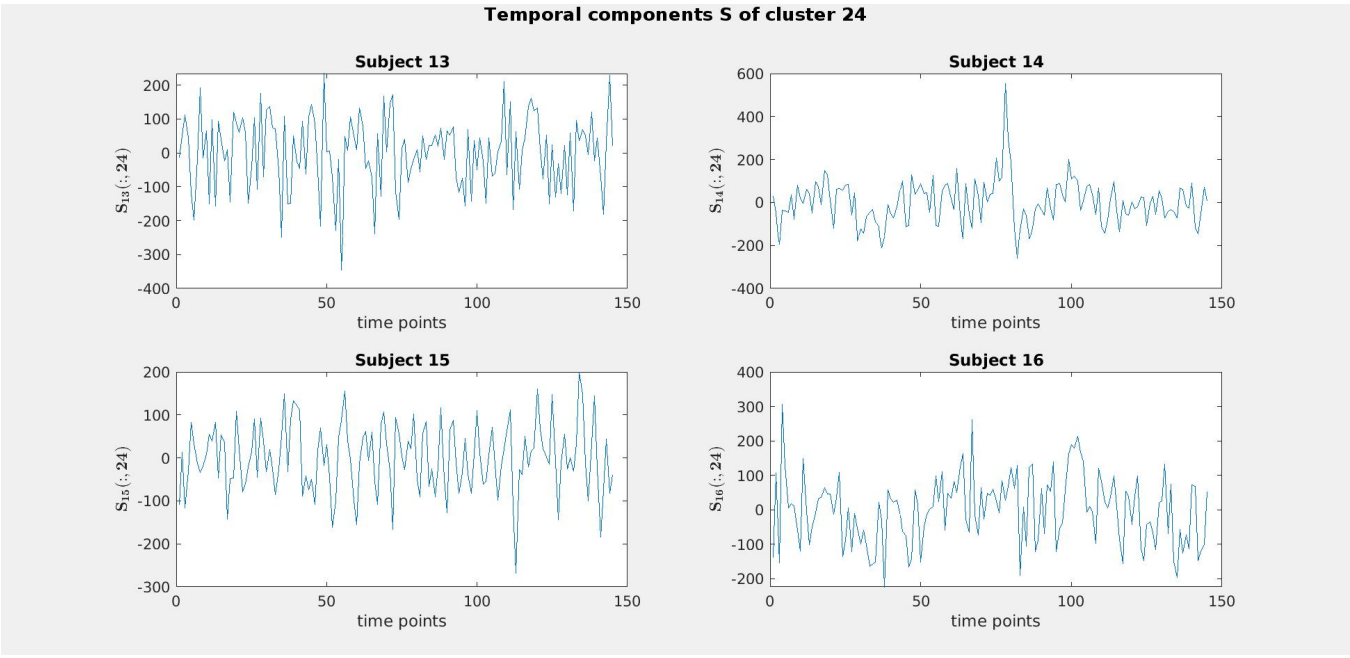


Figure B.679: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 13, \ldots, 16$ and R=42.

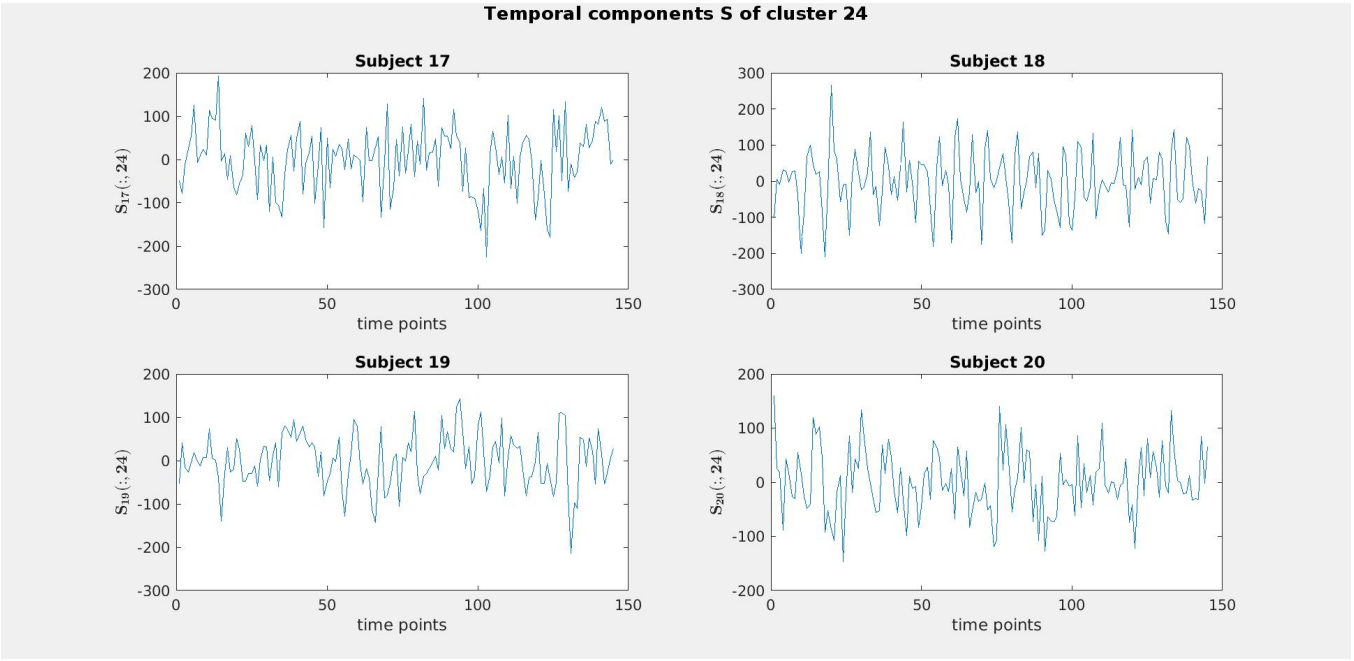


Figure B.680: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 17, \ldots, 20$ and R=42.

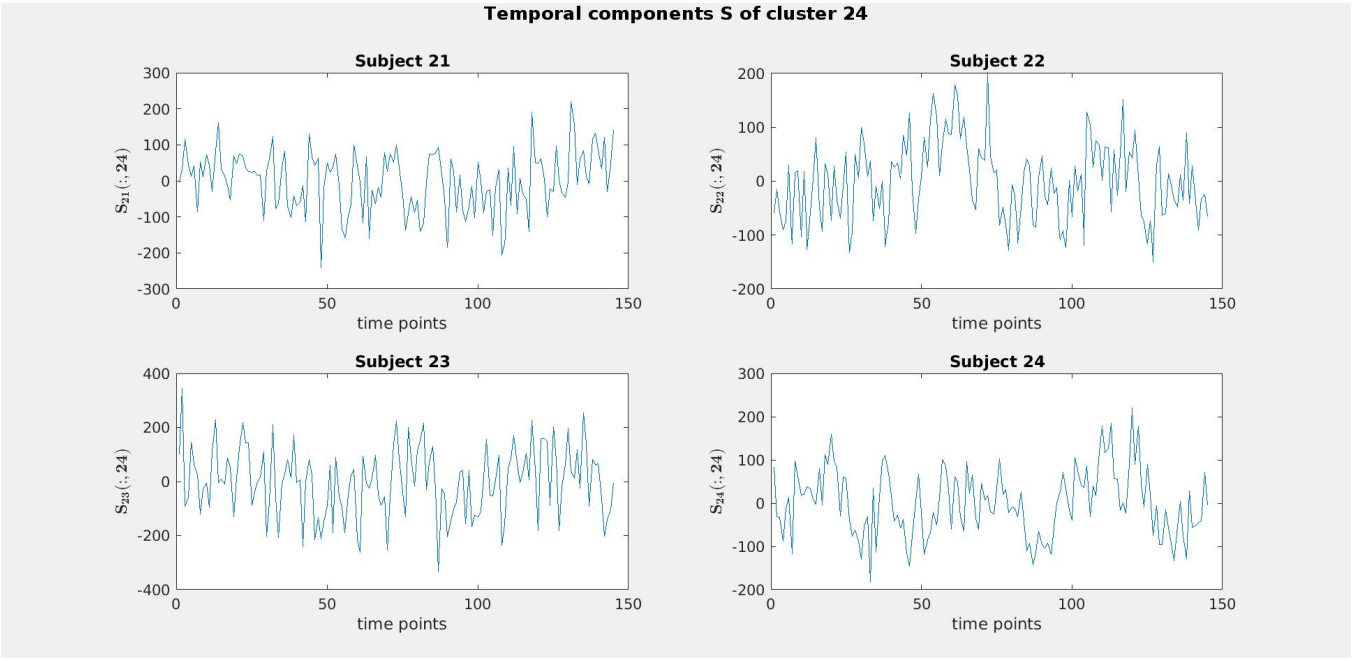


Figure B.681: Waveform of temporal components $\mathbf{S}_k(:, 124)$ for $k = 21, \ldots, 24$ and R=42.

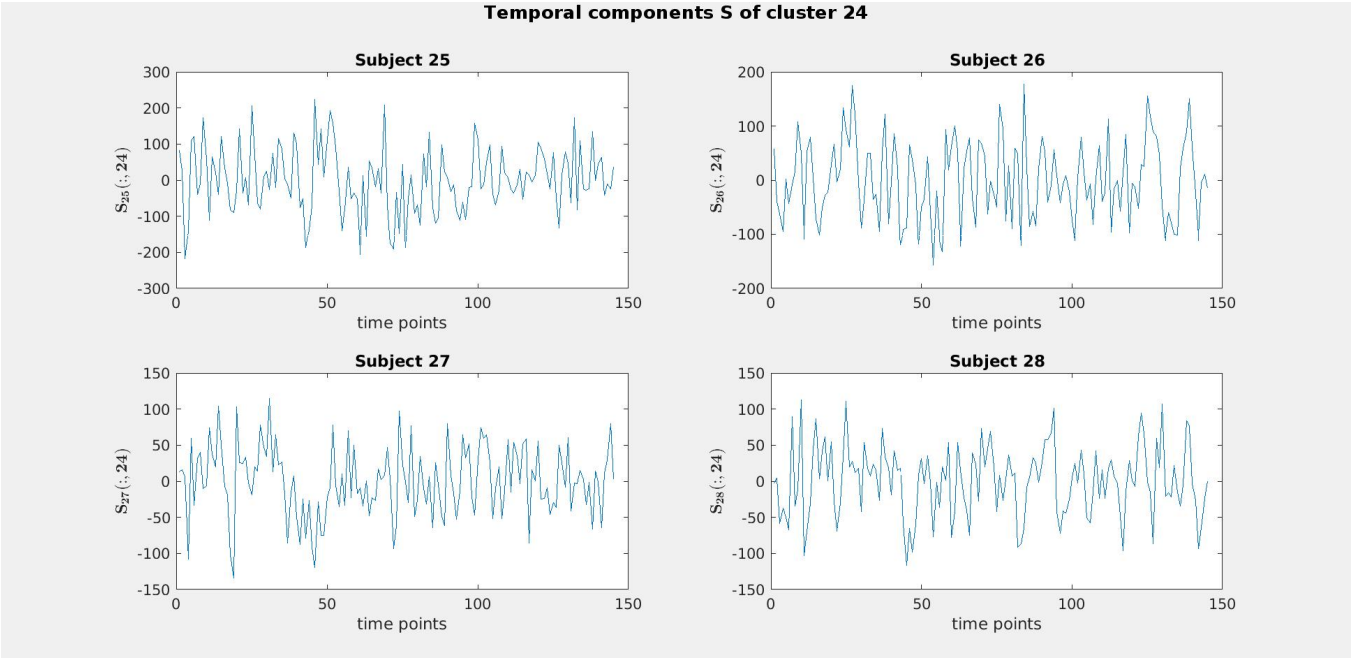


Figure B.682: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 25, \ldots, 28$ and R=42.

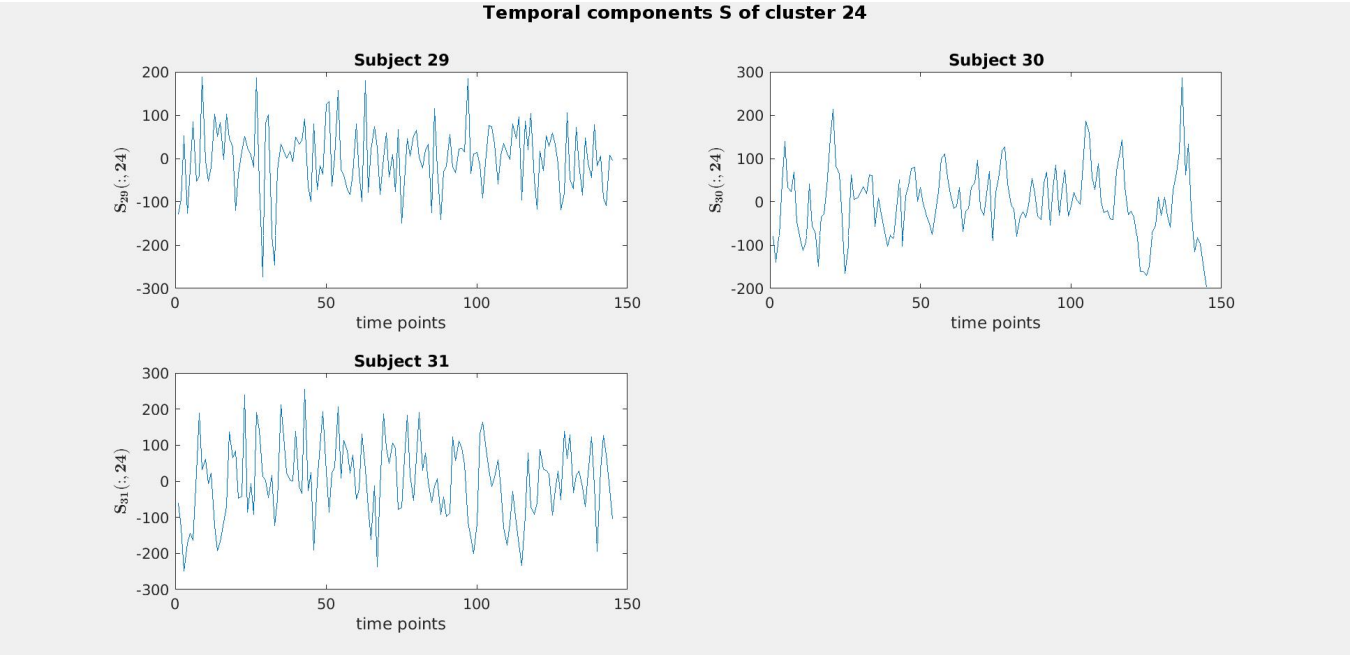


Figure B.683: Waveform of temporal components $\mathbf{S}_k(:, 24)$ for $k = 29, \ldots, 31$ and R=42.

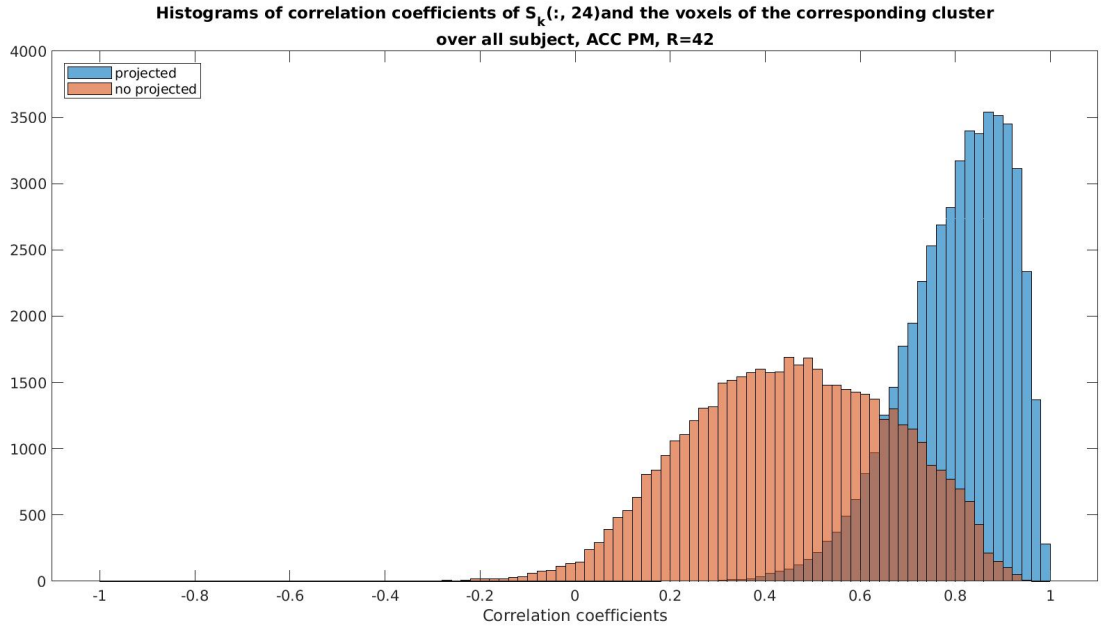


Figure B.684: For the column 24 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 24)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

40. **Cluster 8**

Cluster 8 contains 1897 voxels, which correspond to 0.839574% of the whole brain (background, soft tissue, bone are excluded).

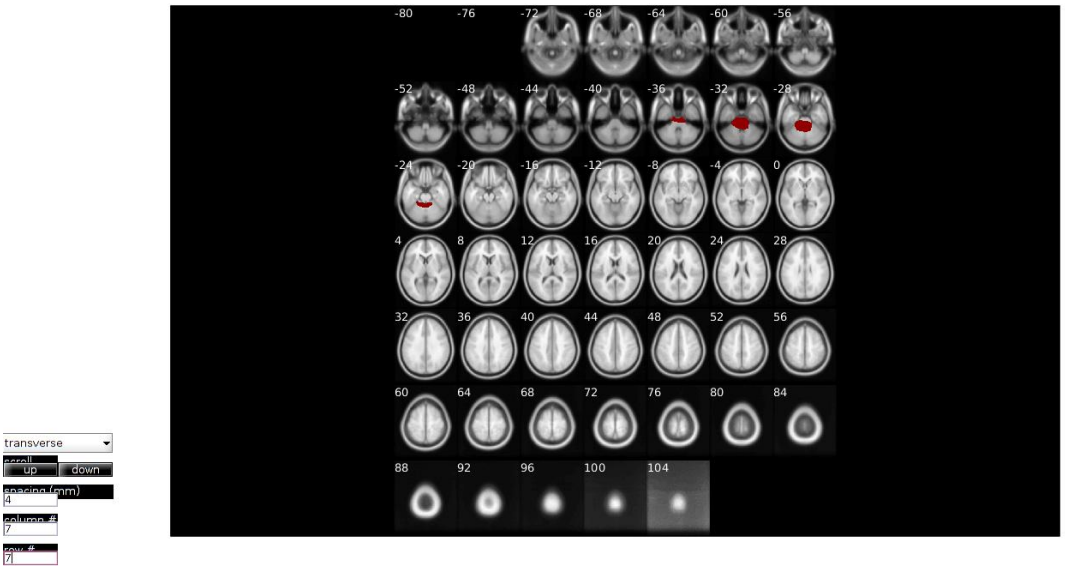


Figure B.685: Cluster 8. Red areas indicate the voxels which belong to cluster 8.

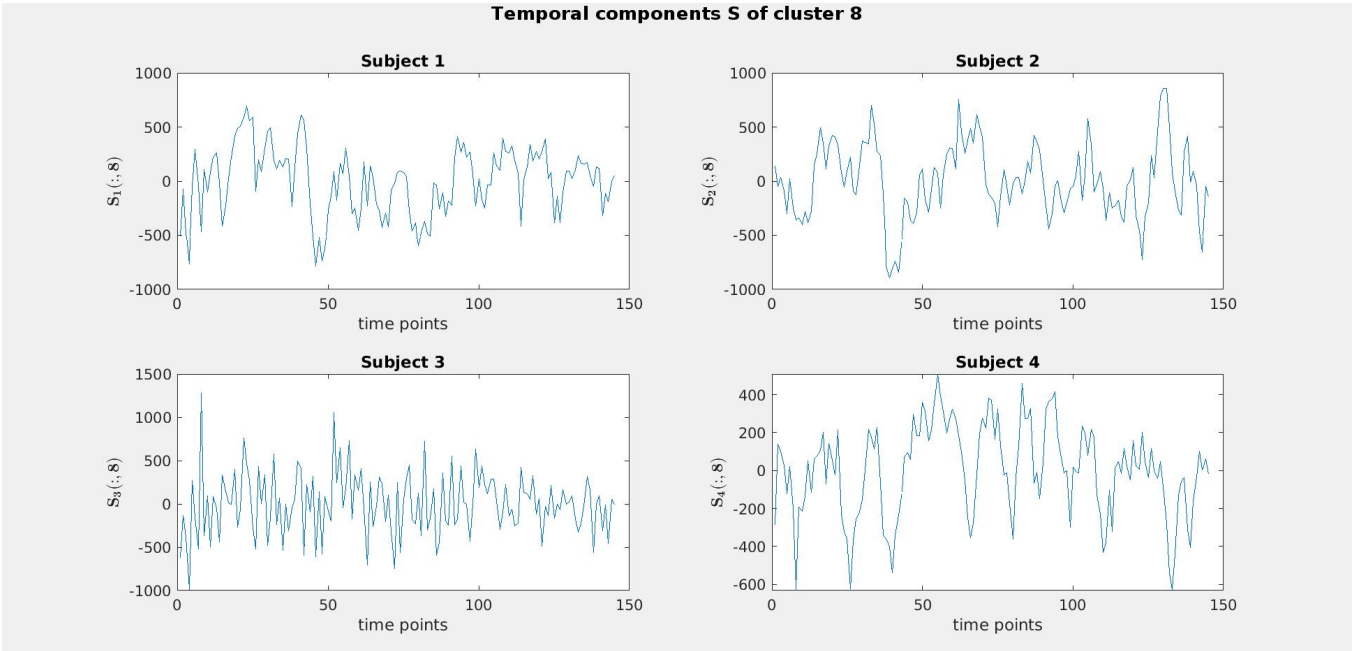


Figure B.686: Waveform of temporal components $\mathbf{S}_k(:,8)$ for $k = 1,\ldots,4$ and R=42.

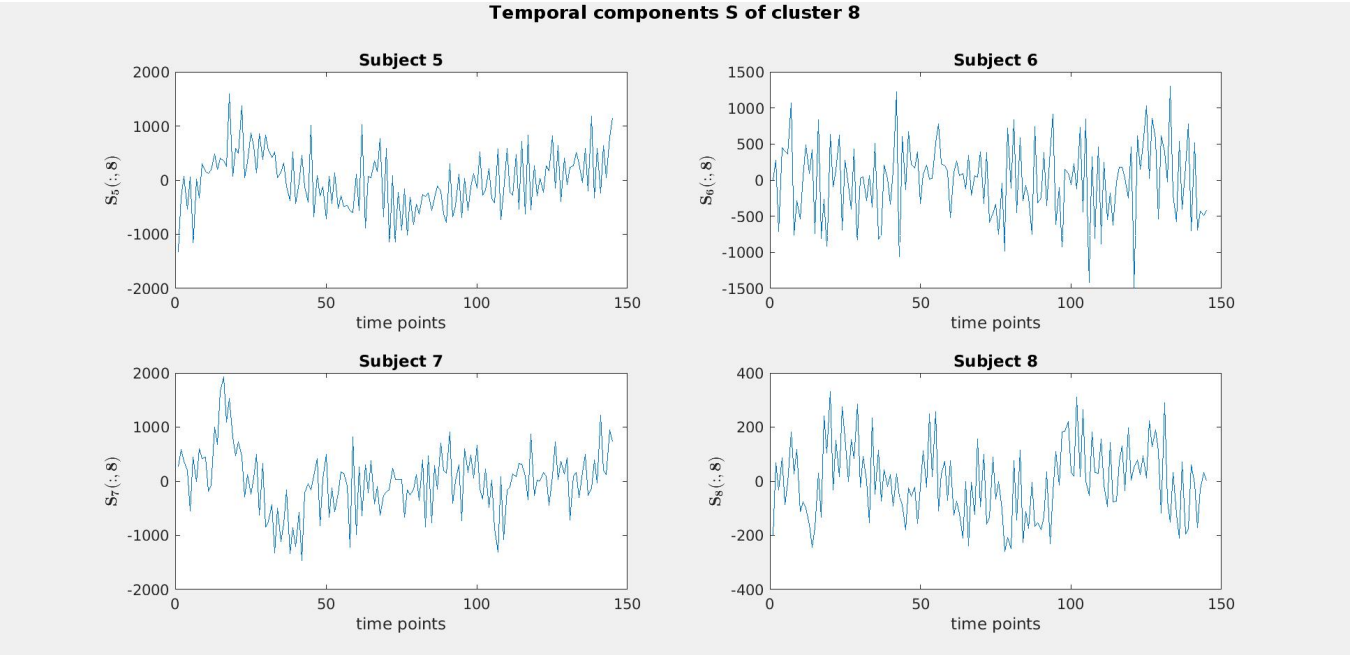


Figure B.687: Waveform of temporal components $\mathbf{S}_k(:,8)$ for $k = 5,\ldots,8$ and R=42.

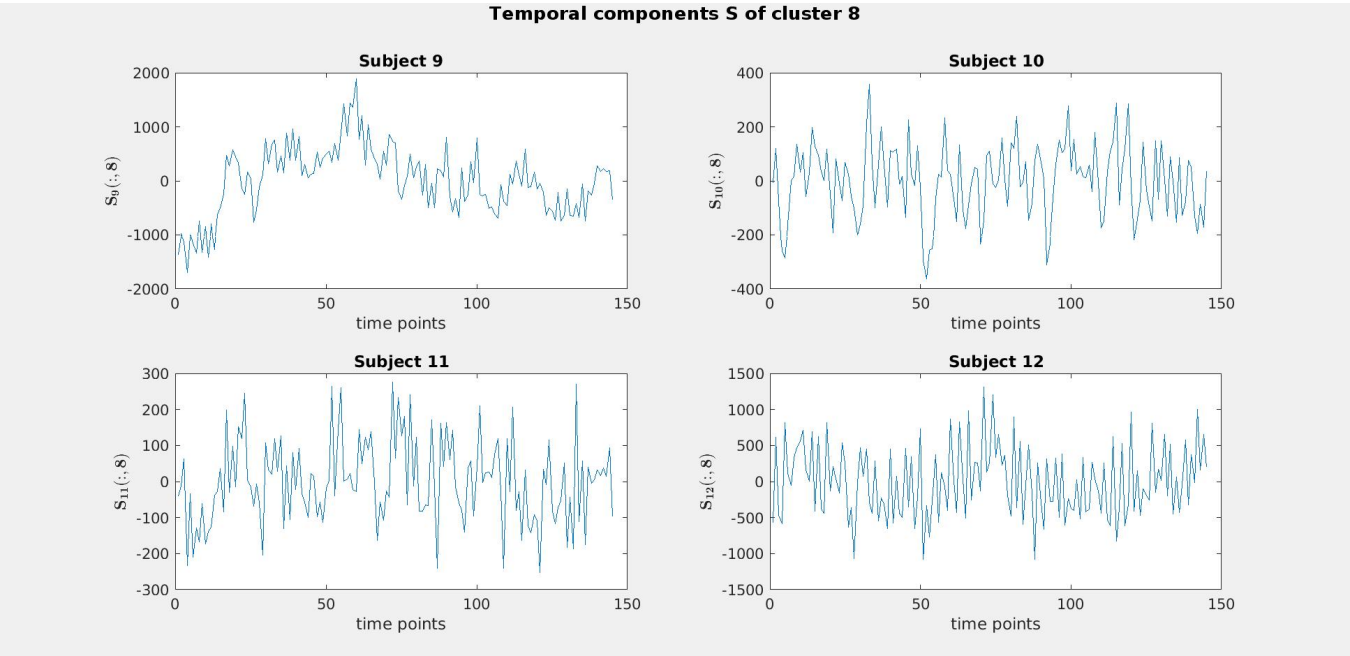


Figure B.688: Waveform of temporal components $\mathbf{S}_k(:,8)$ for $k = 9,\ldots,12$ and R=42.

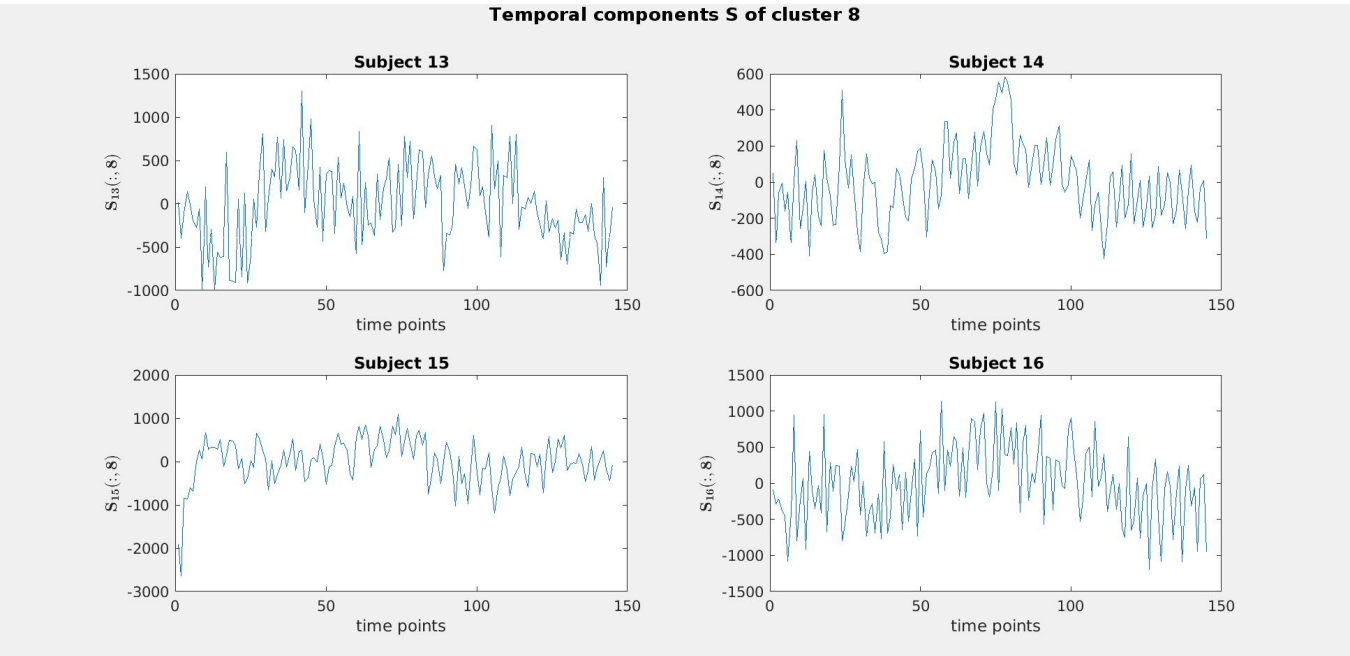


Figure B.689: Waveform of temporal components $\mathbf{S}_k(:,8)$ for $k = 13,\ldots,16$ and R=42.

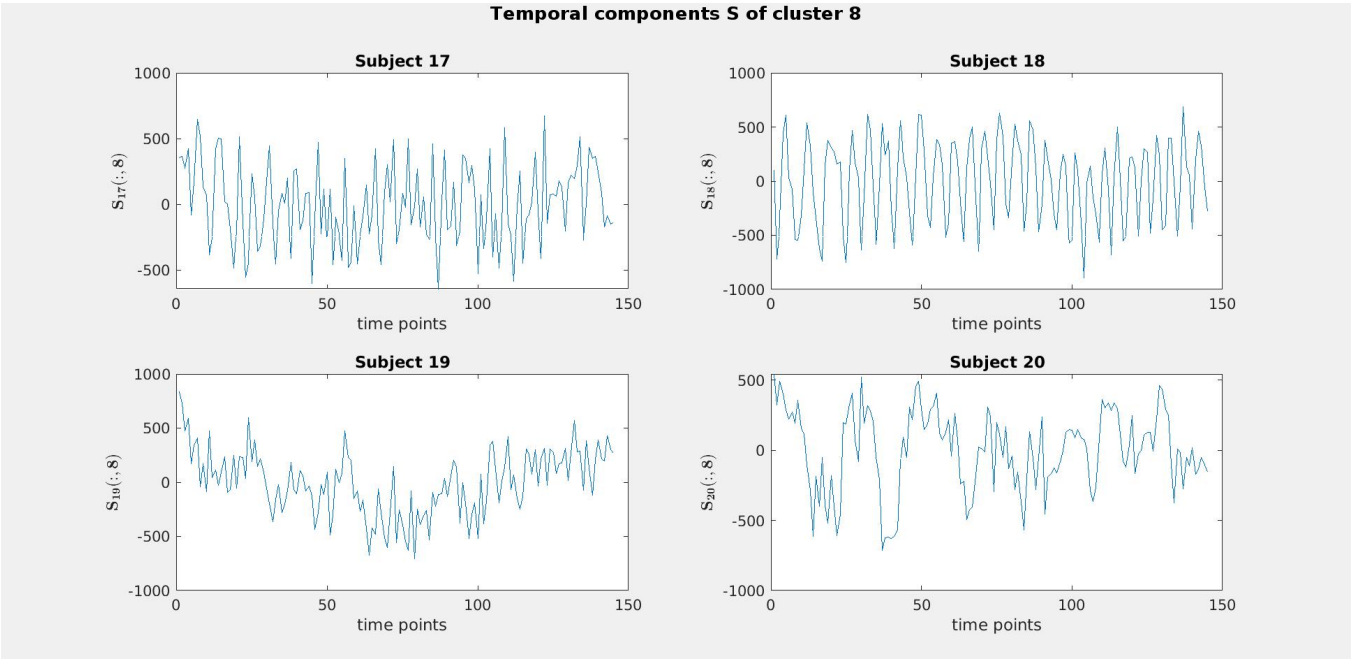


Figure B.690: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 17, \ldots, 20$ and R=42.

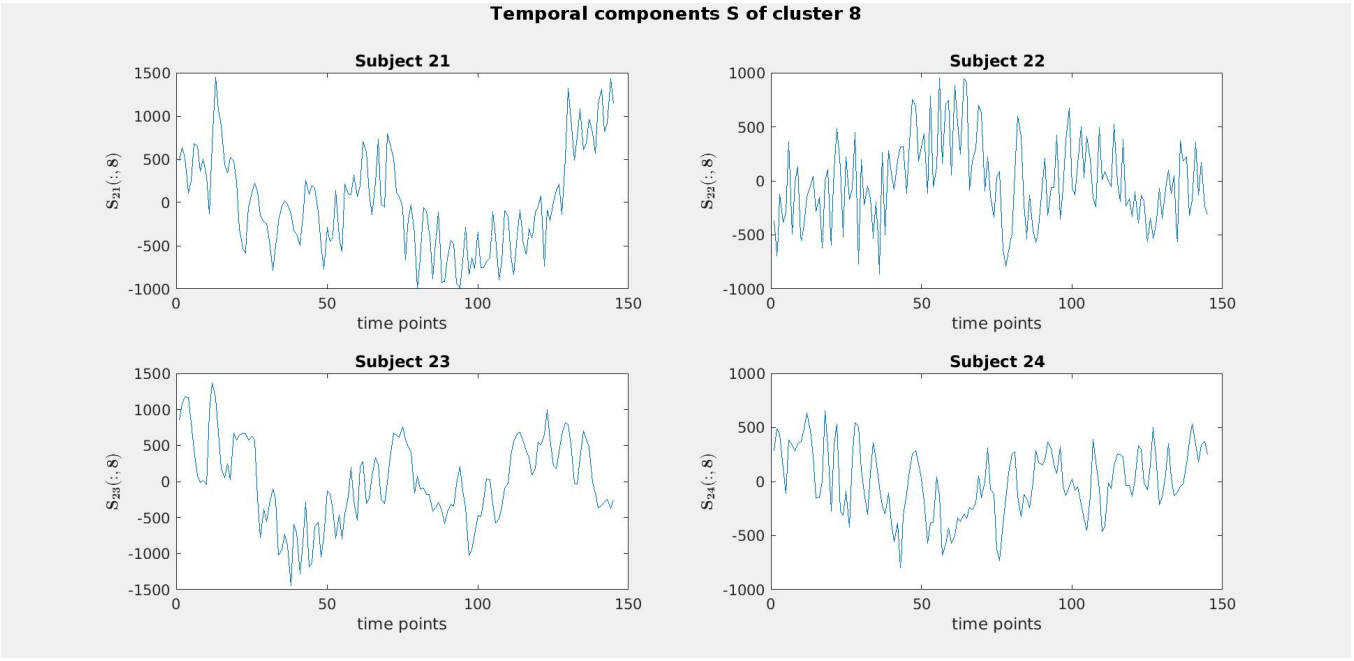


Figure B.691: Waveform of temporal components $\mathbf{S}_k(:, 18)$ for $k = 21, \ldots, 24$ and R=42.

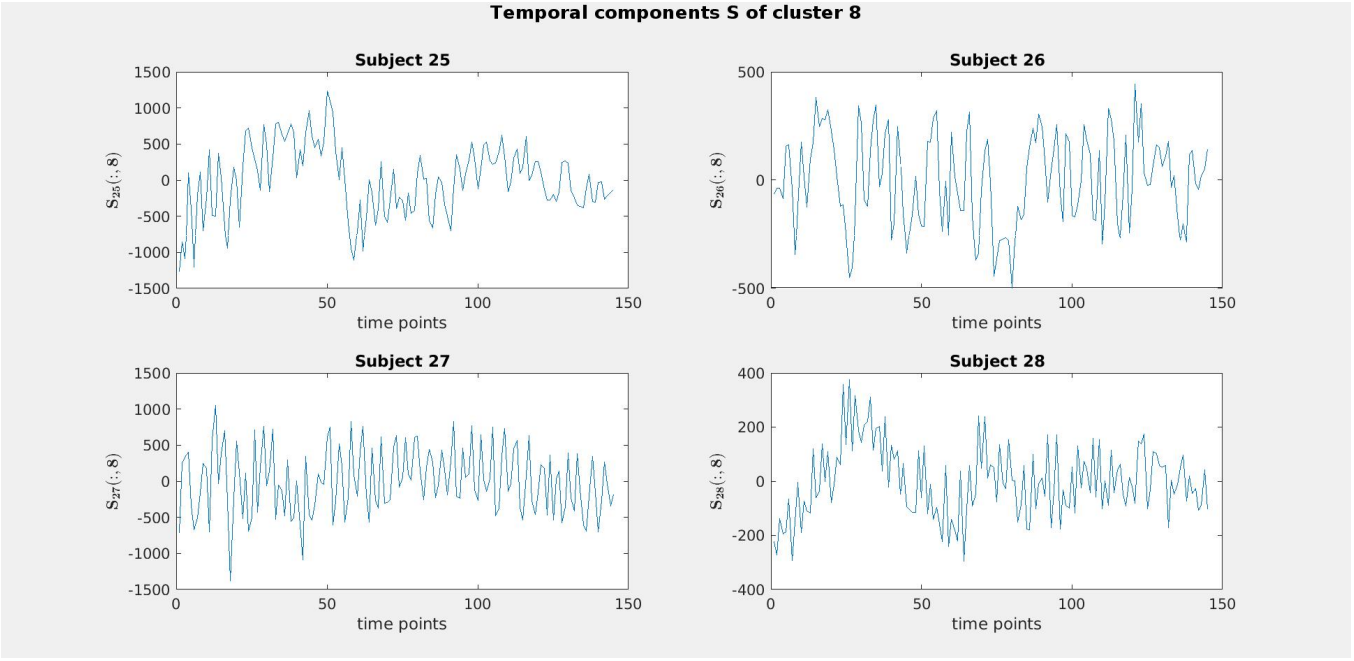


Figure B.692: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 25, \ldots, 28$ and R=42.

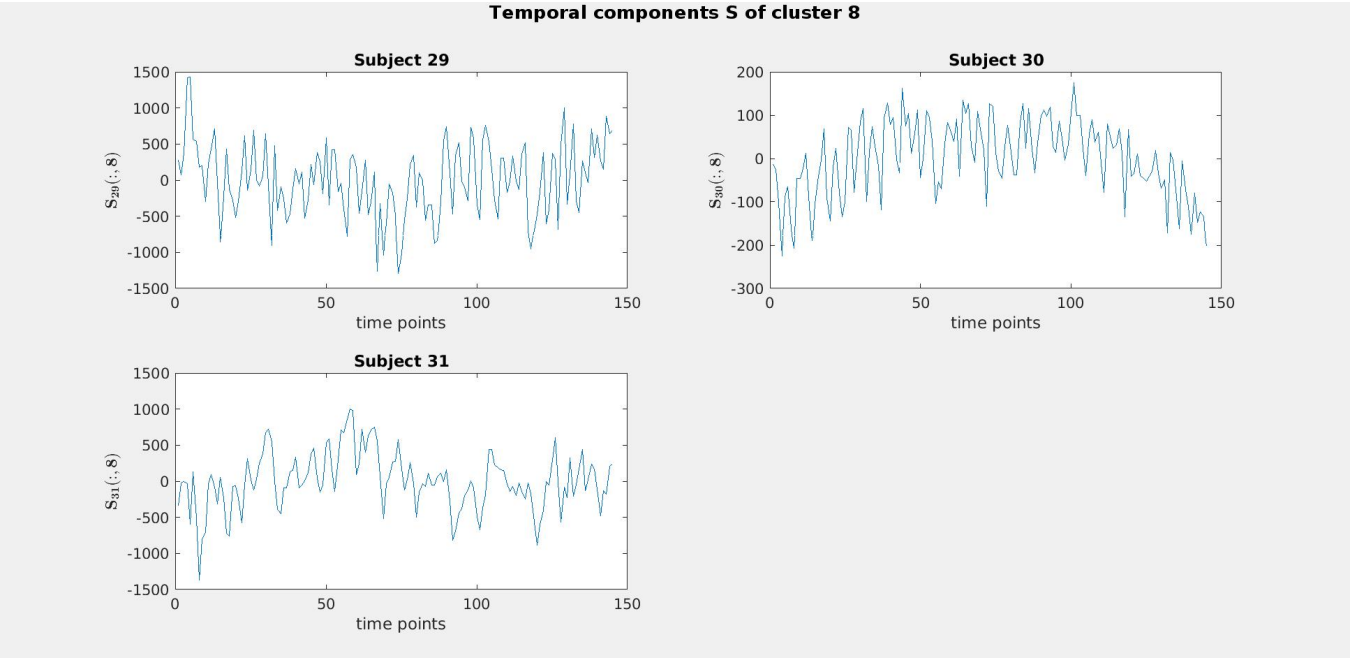


Figure B.693: Waveform of temporal components $\mathbf{S}_k(:, 8)$ for $k = 29, \ldots, 31$ and R=42.

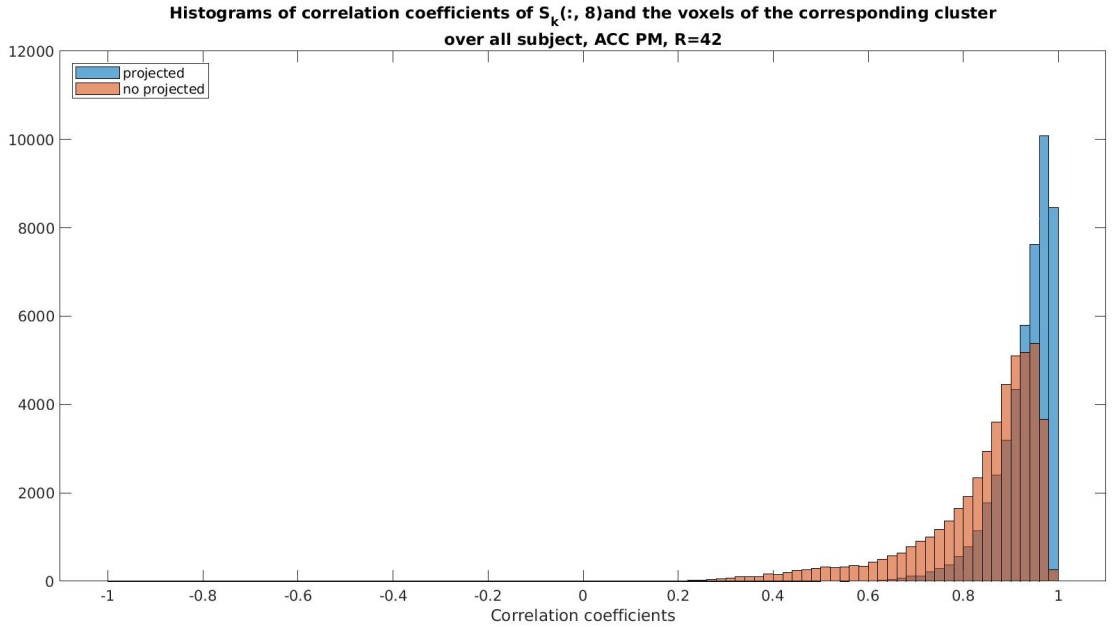


Figure B.694: For the column 8 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 8)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

41. **Cluster 32**

Cluster 32 contains 1823 voxels, which correspond to 0.806823% of the whole brain (background, soft tissue, bone are excluded).

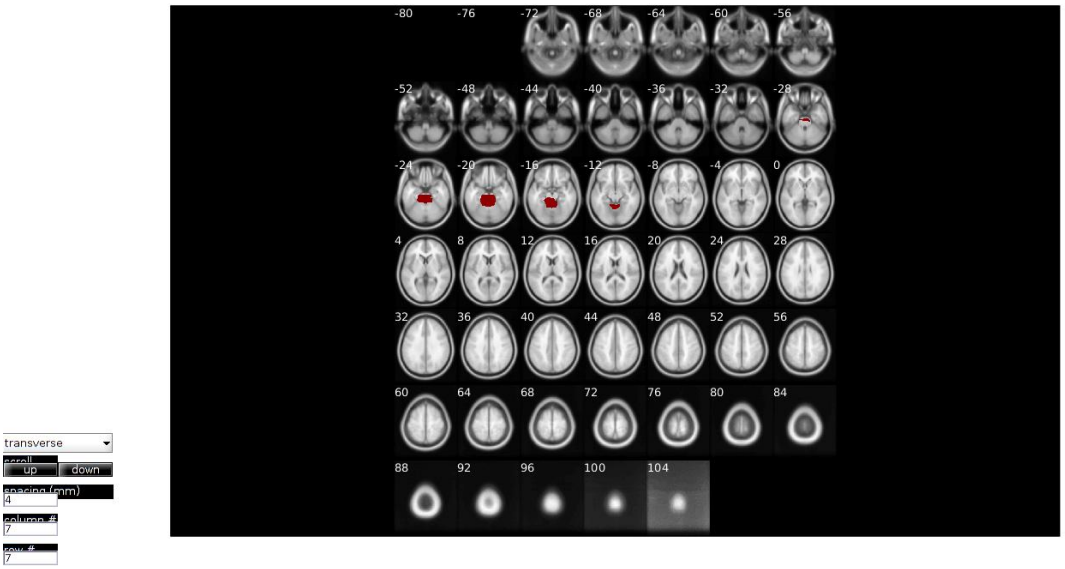


Figure B.695: Cluster 32. Red areas indicate the voxels which belong to cluster 32.

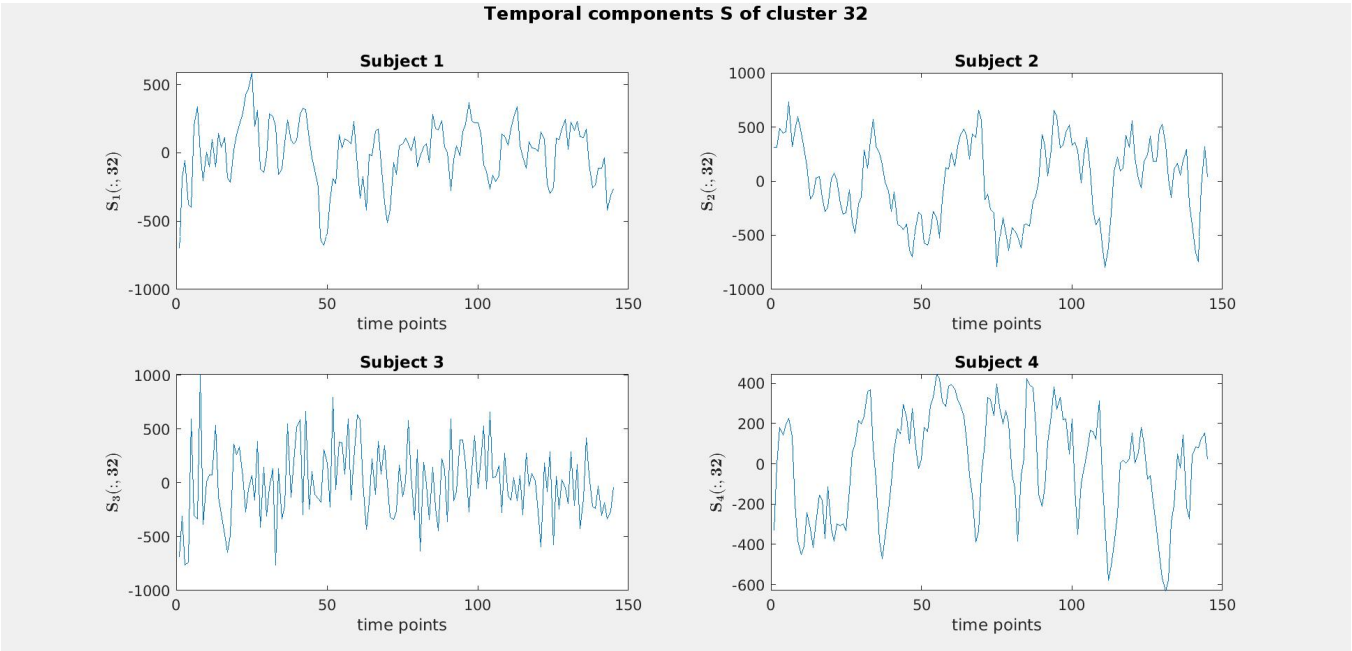


Figure B.696: Waveform of temporal components $\mathbf{S}_k(:, 32)$ for $k = 1, \ldots, 4$ and R=42.

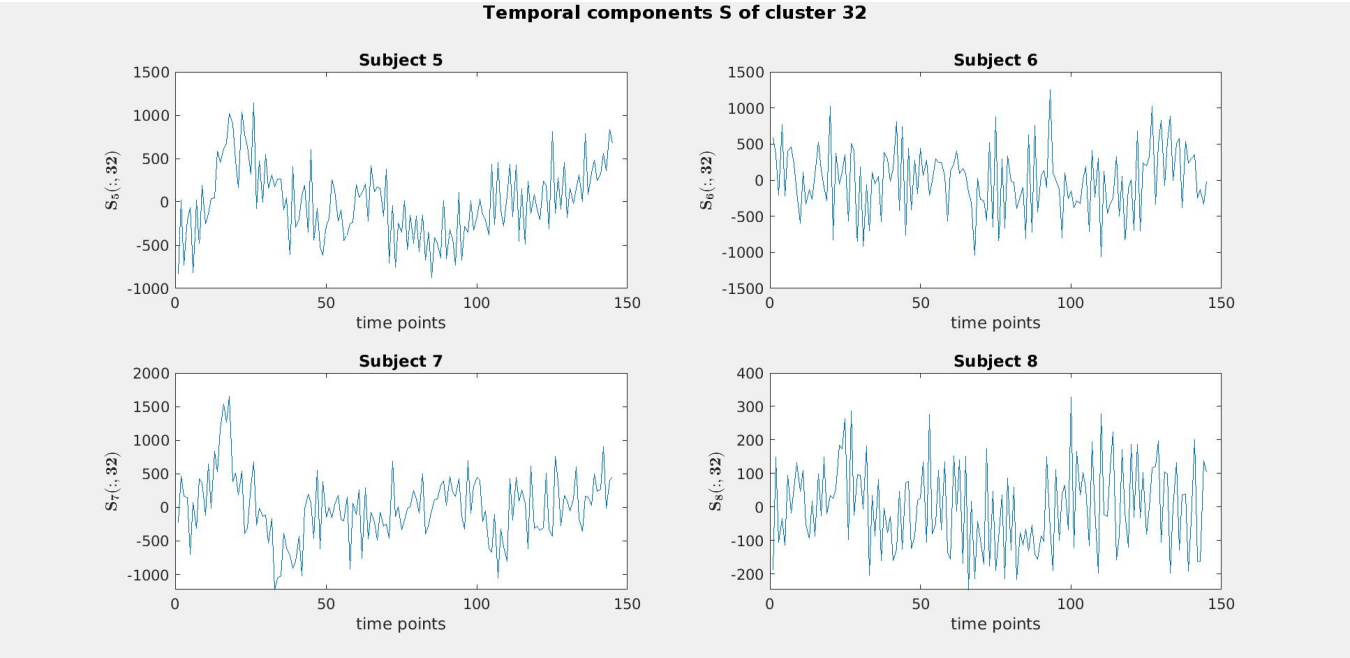


Figure B.697: Waveform of temporal components $\mathbf{S}_k(:, 32)$ for $k = 5, \ldots, 8$ and R=42.

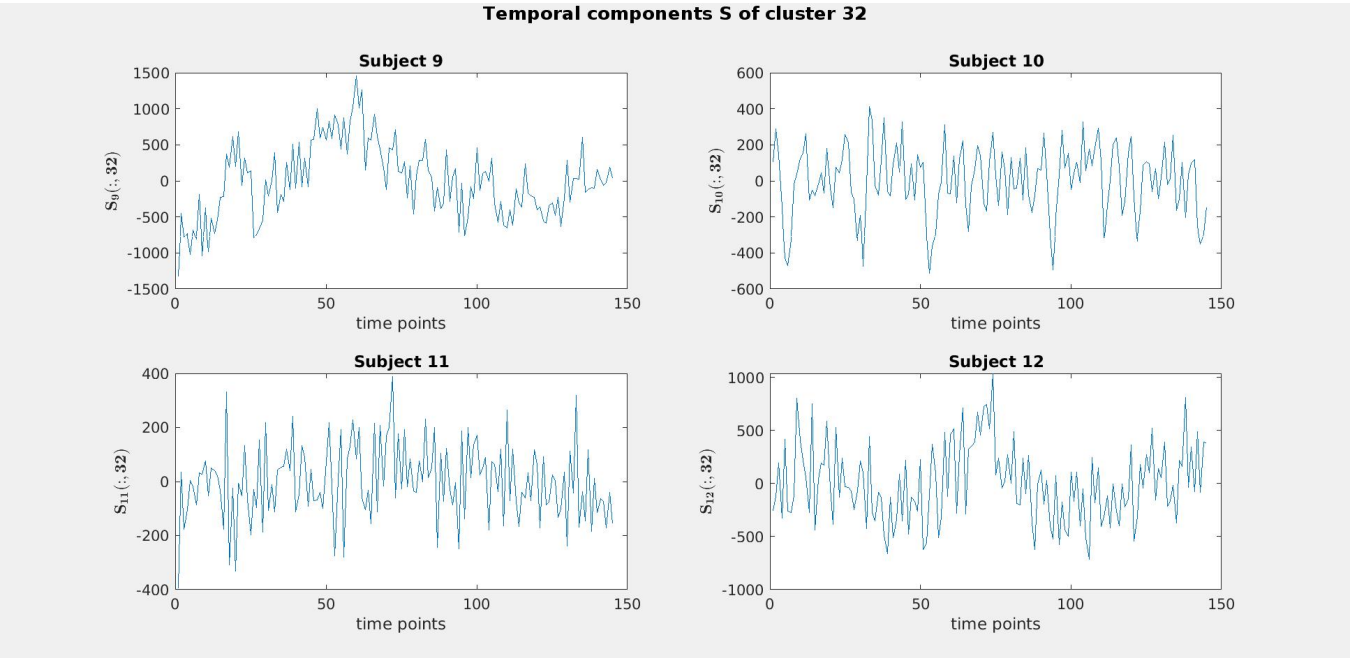


Figure B.698: Waveform of temporal components $\mathbf{S}_k(:, 32)$ for $k = 9, \ldots, 12$ and R=42.

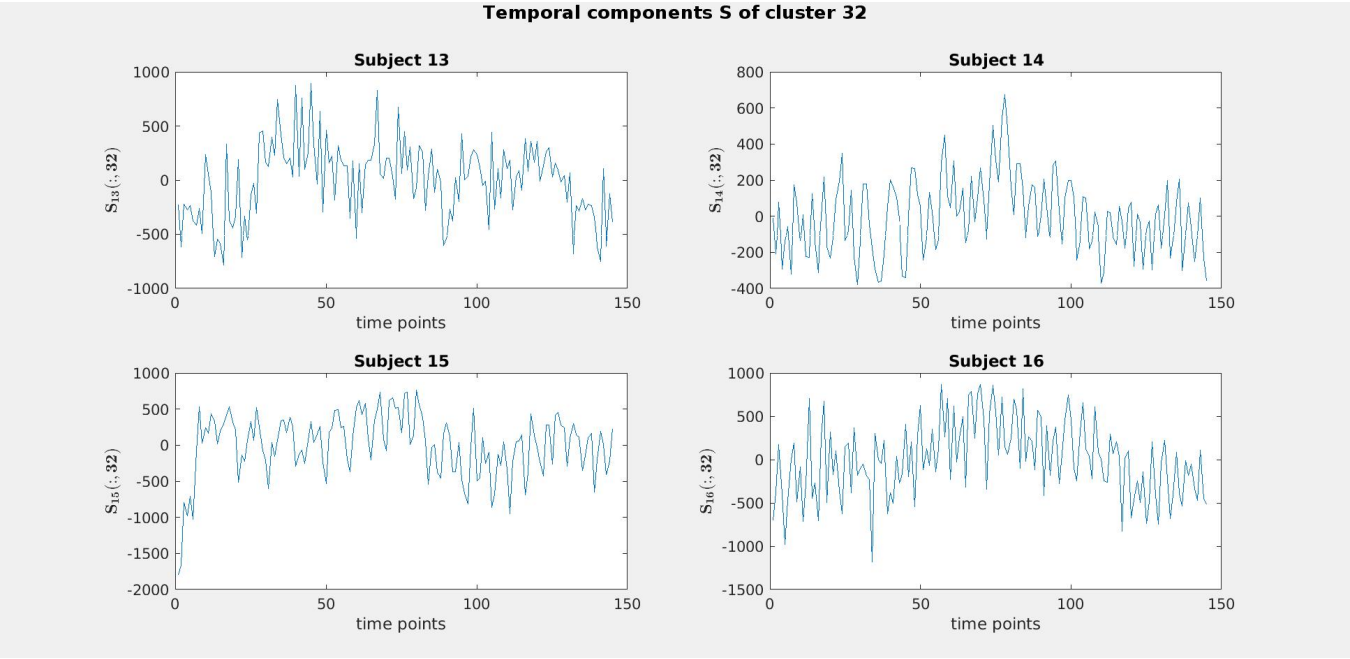


Figure B.699: Waveform of temporal components $\mathbf{S}_k(:, 32)$ for $k = 13, \ldots, 16$ and R=42.

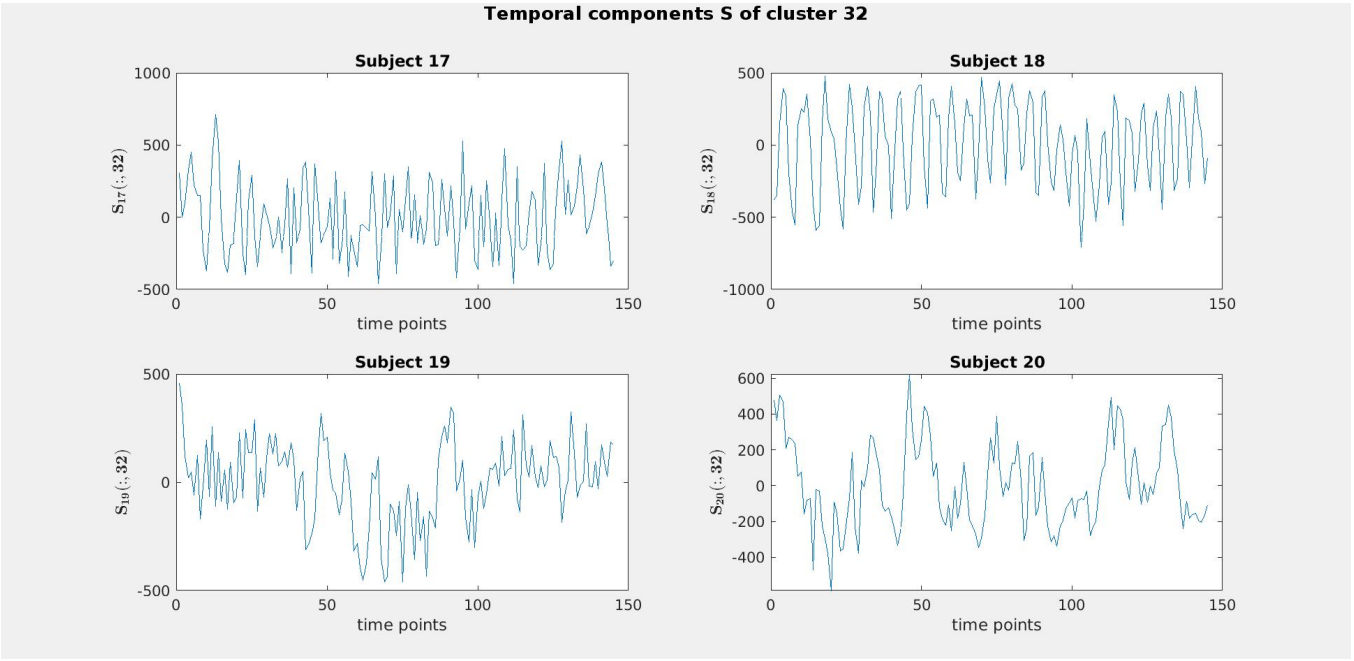


Figure B.700: Waveform of temporal components $\mathbf{S}_k(:, 32)$ for $k = 17, \ldots, 20$ and R=42.

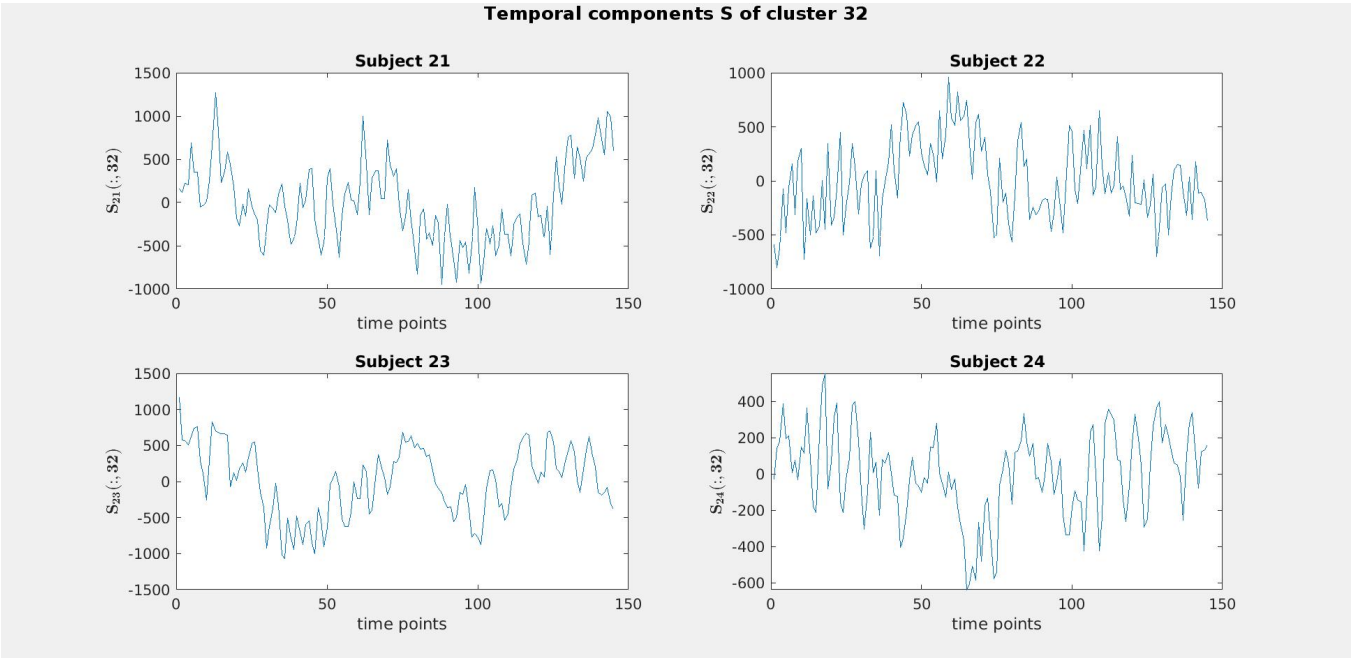


Figure B.701: Waveform of temporal components $\mathbf{S}_k(:, 132)$ for $k = 21, \ldots, 24$ and R=42.

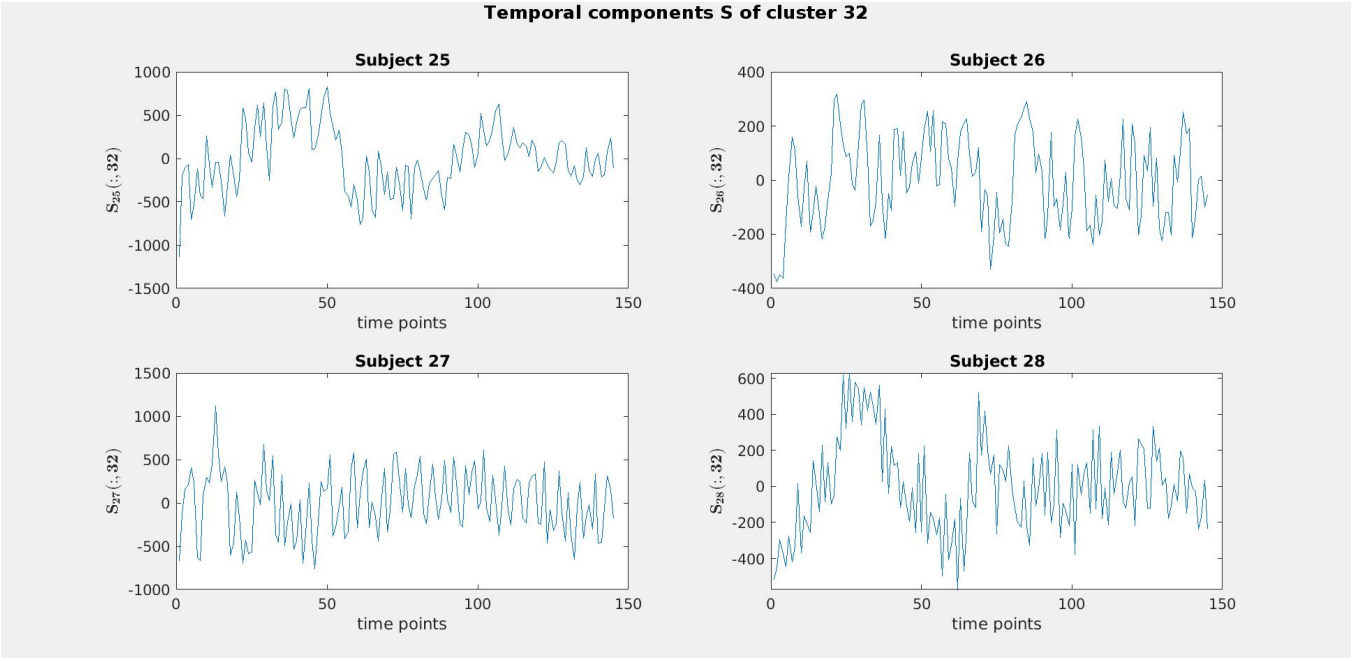


Figure B.702: Waveform of temporal components $\mathbf{S}_k(:, 32)$ for $k = 25, \ldots, 28$ and R=42.

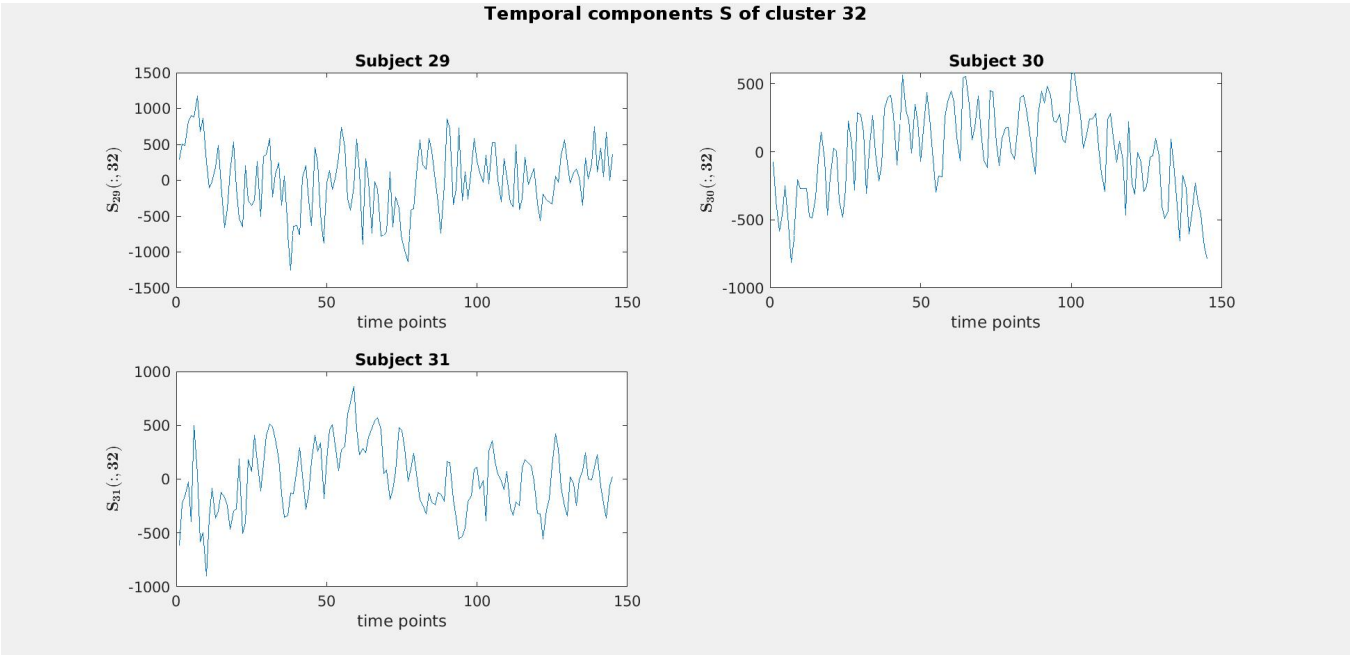


Figure B.703: Waveform of temporal components $\mathbf{S}_k(:, 32)$ for $k = 29, \ldots, 31$ and R=42.

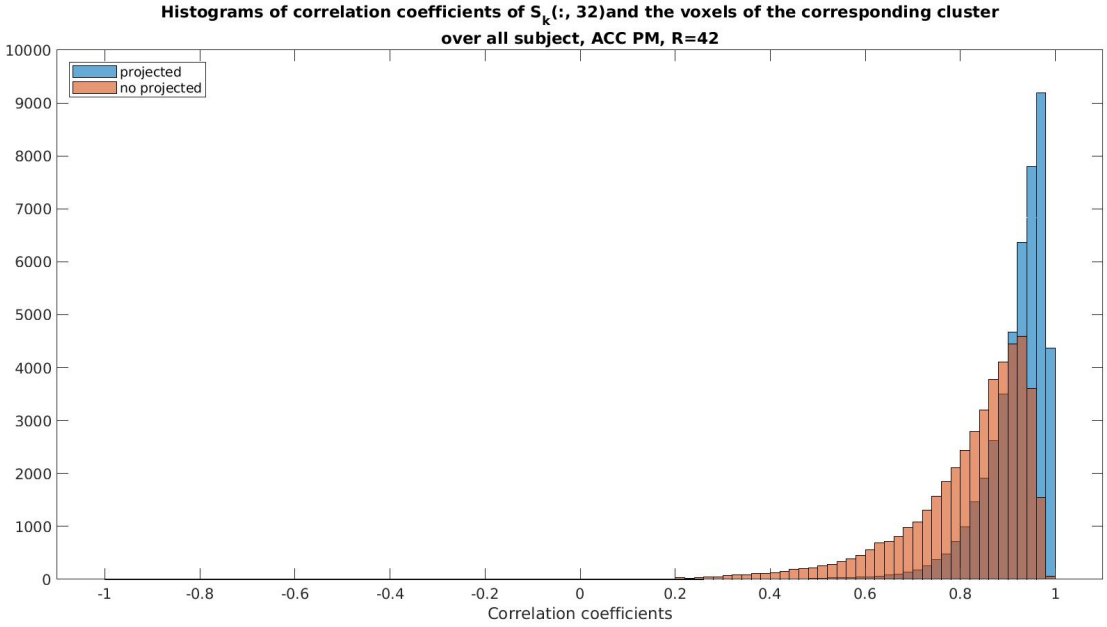


Figure B.704: For the column 32 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 32)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

42. **Cluster 34**

Cluster 34 contains 1170 voxels, which correspond to 0.517818% of the whole brain (background, soft tissue, bone are excluded).

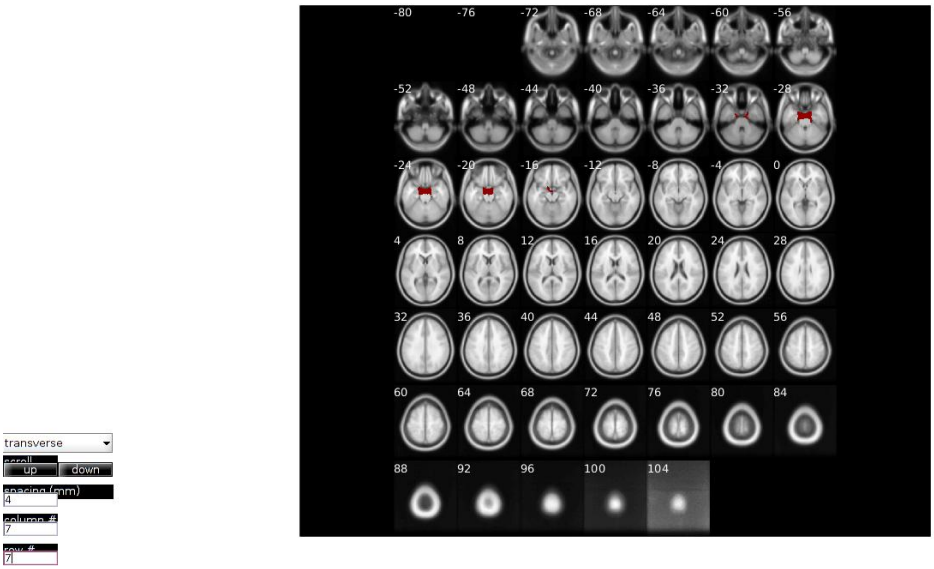


Figure B.705: Cluster 34. Red areas indicate the voxels which belong to cluster 34.

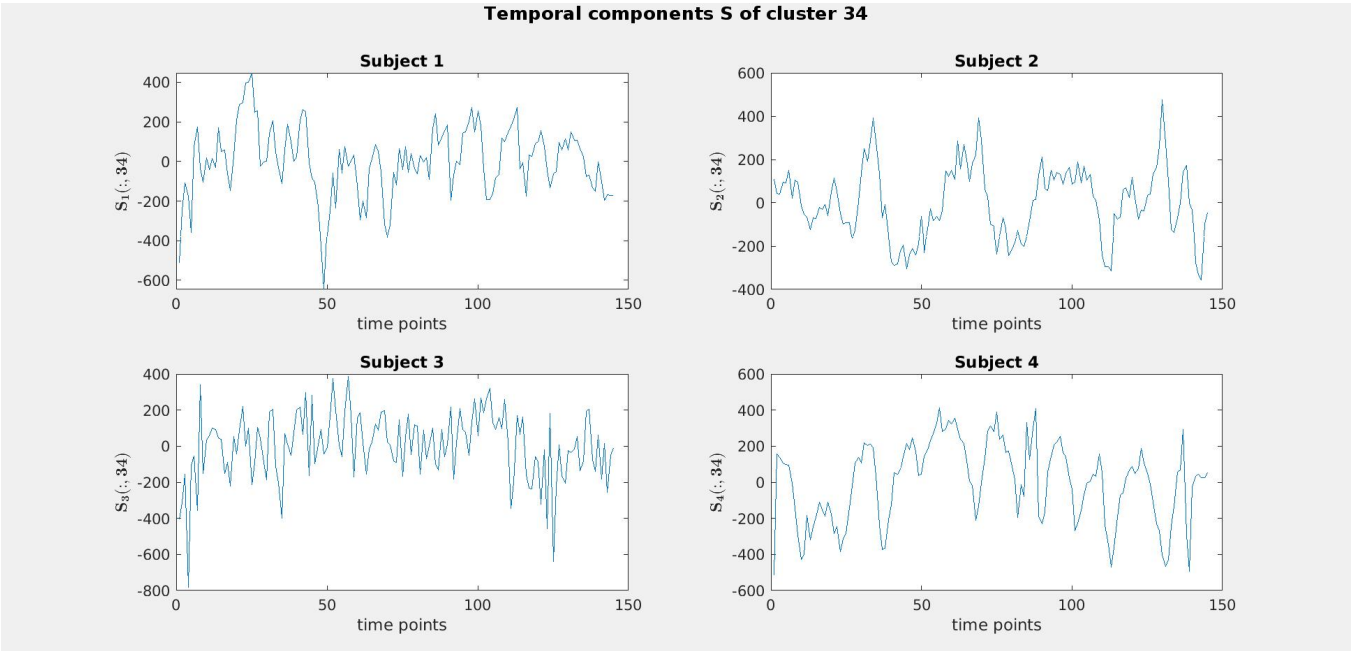


Figure B.706: Waveform of temporal components $\mathbf{S}_k(:, 34)$ for $k = 1, \ldots, 4$ and R=42.

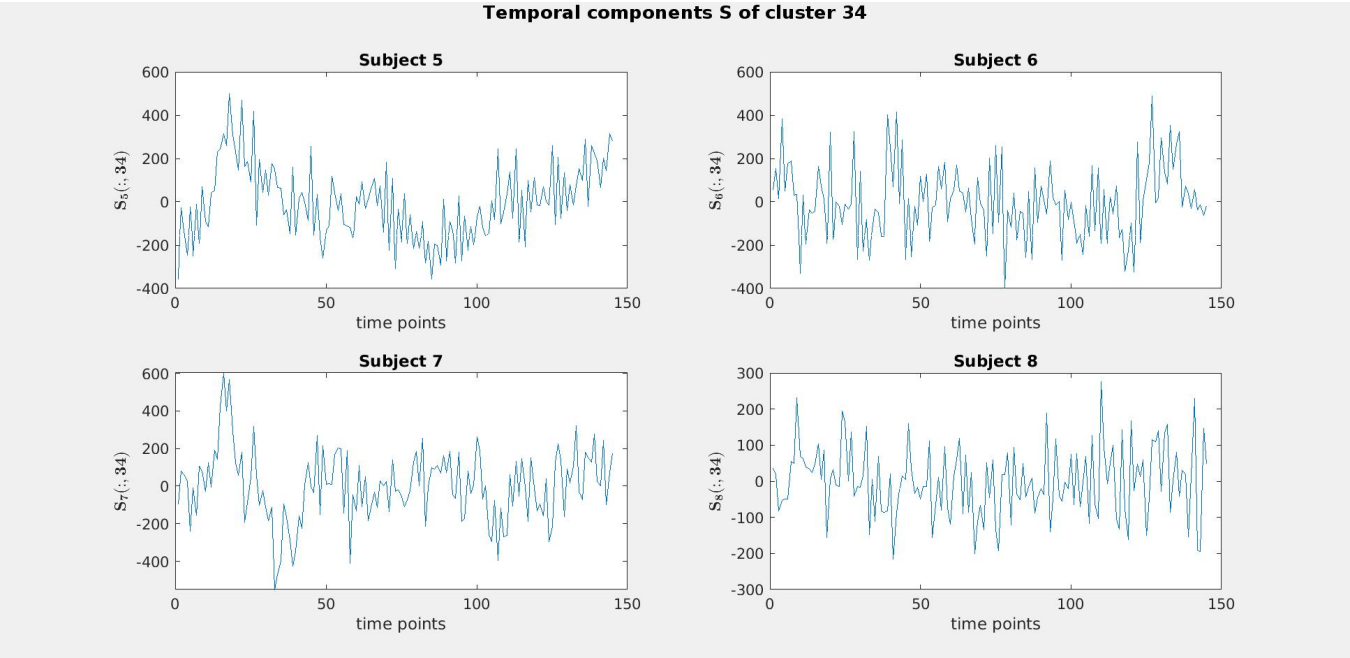


Figure B.707: Waveform of temporal components $\mathbf{S}_k(:, 34)$ for $k = 5, \ldots, 8$ and R=42.

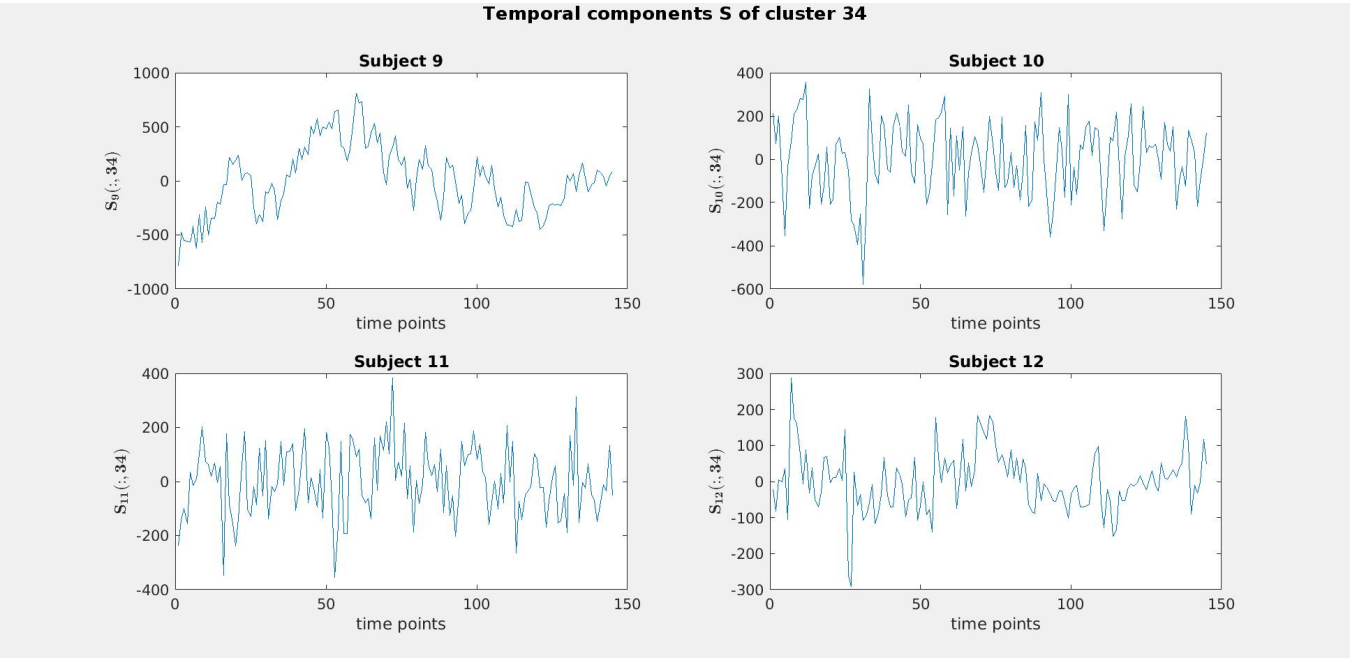


Figure B.708: Waveform of temporal components $\mathbf{S}_k(:, 34)$ for $k = 9, \ldots, 12$ and R=42.

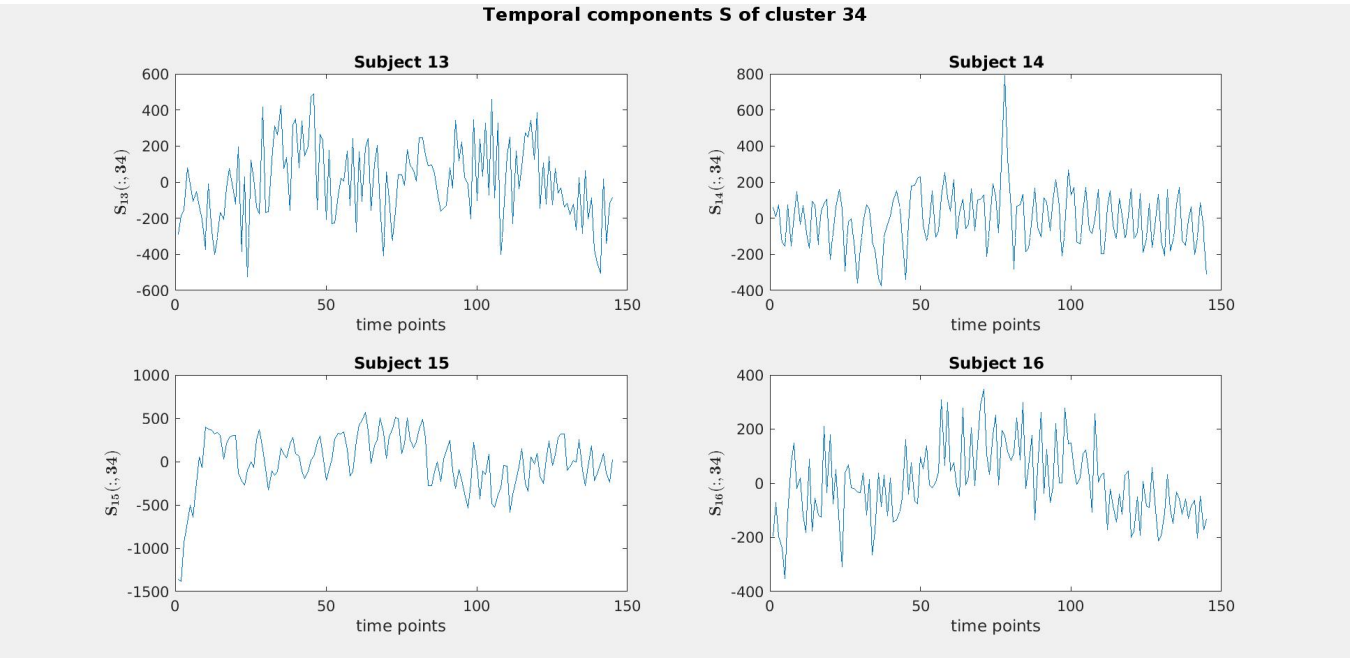


Figure B.709: Waveform of temporal components $\mathbf{S}_k(:, 34)$ for $k = 13, \ldots, 16$ and R=42.

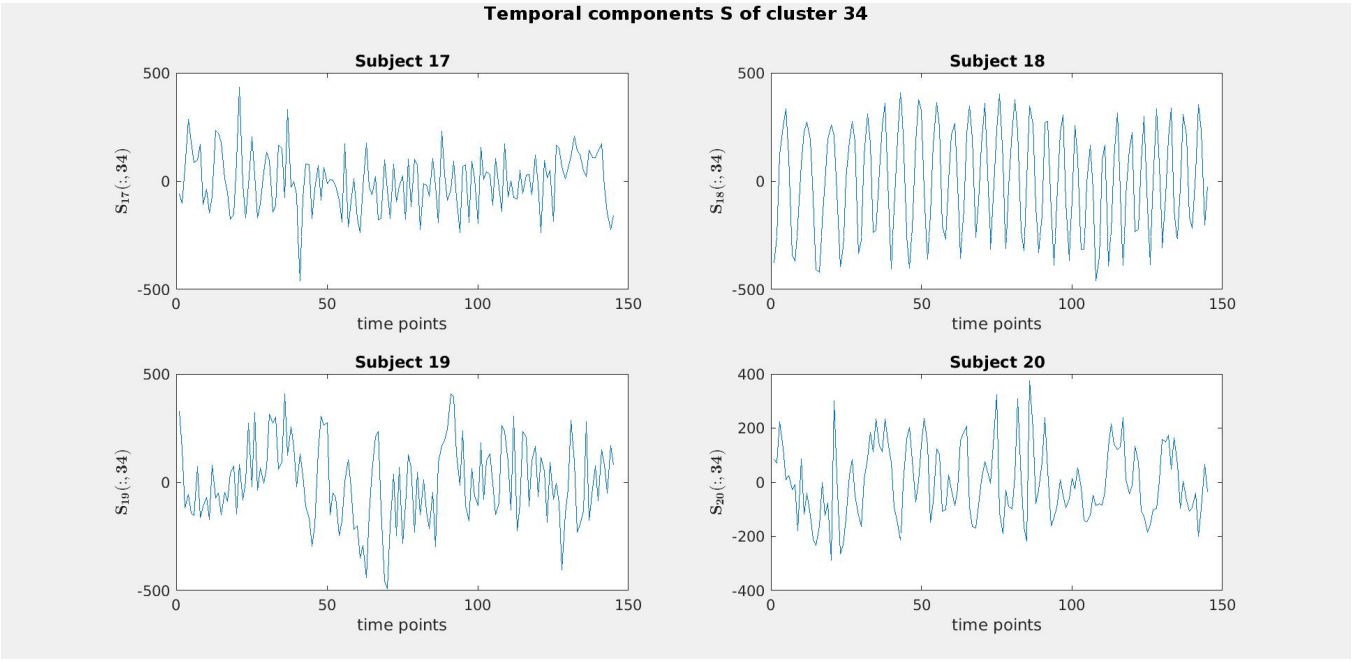


Figure B.710: Waveform of temporal components $\mathbf{S}_k(:, 34)$ for $k = 17, \ldots, 20$ and R=42.

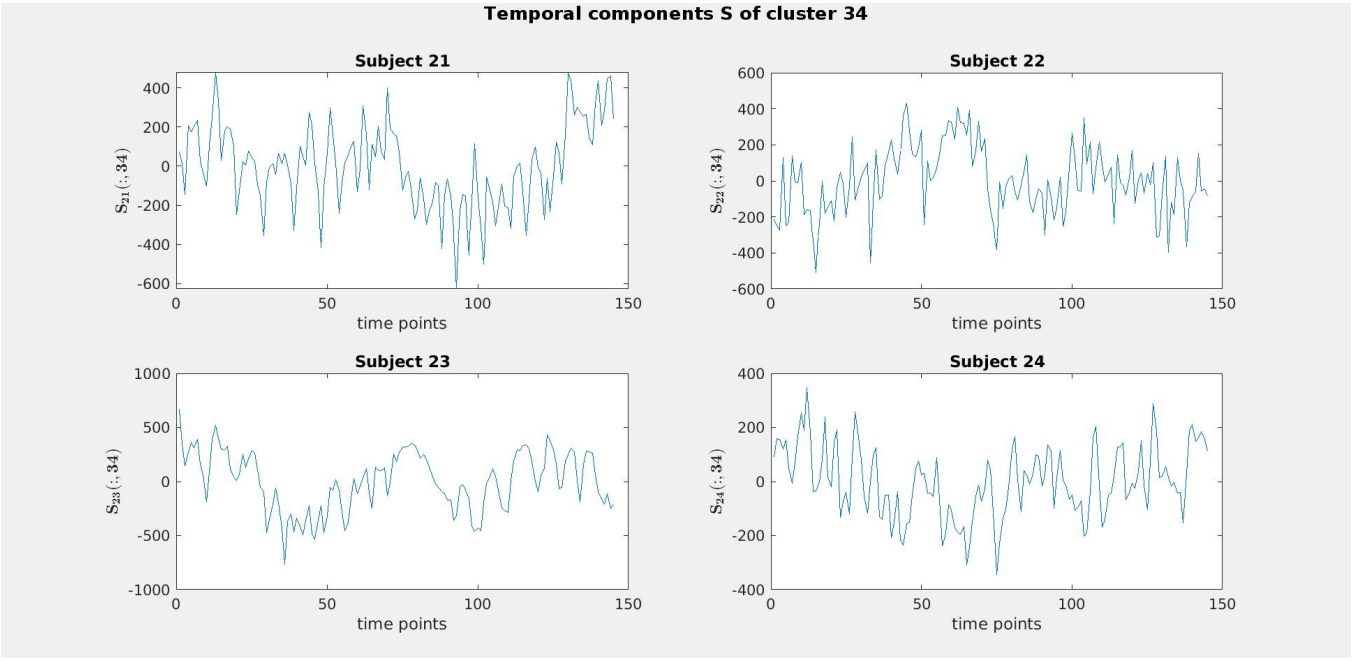


Figure B.711: Waveform of temporal components $\mathbf{S}_k(:, 134)$ for $k = 21, \ldots, 24$ and R=42.

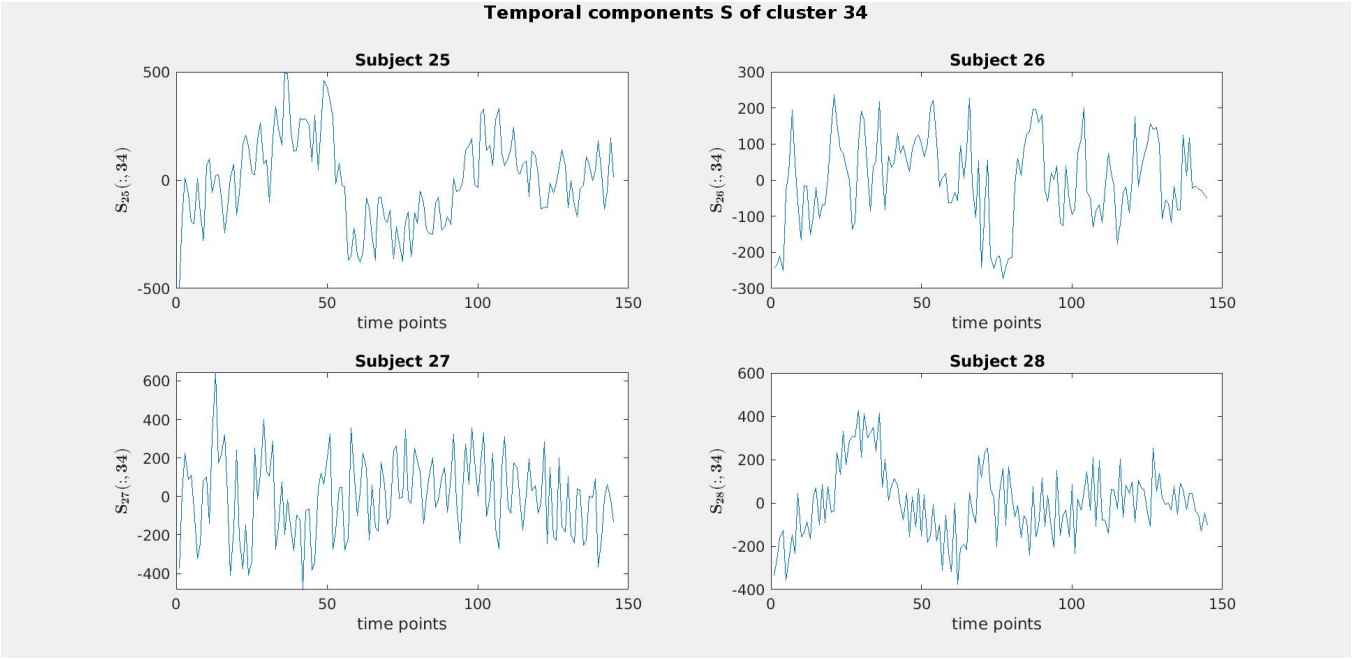


Figure B.712: Waveform of temporal components $\mathbf{S}_k(:, 34)$ for $k = 25, \ldots, 28$ and R=42.

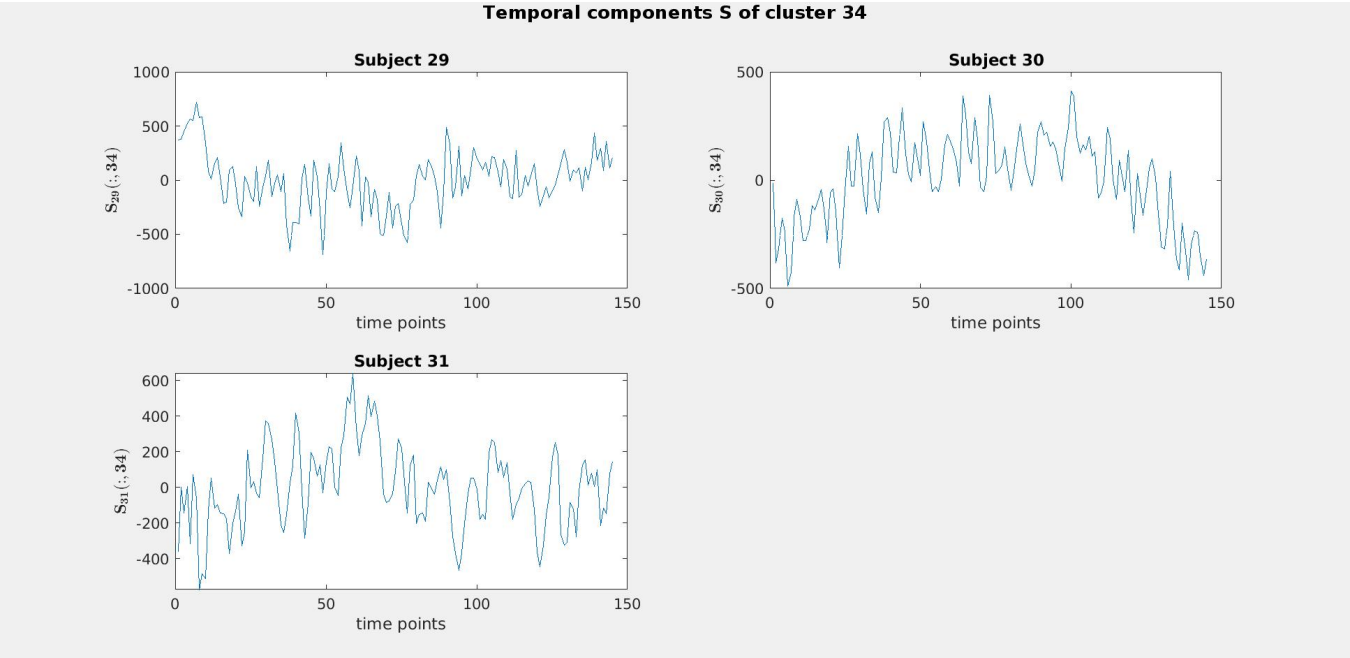


Figure B.713: Waveform of temporal components $\mathbf{S}_k(:, 34)$ for $k = 29, \ldots, 31$ and R=42.

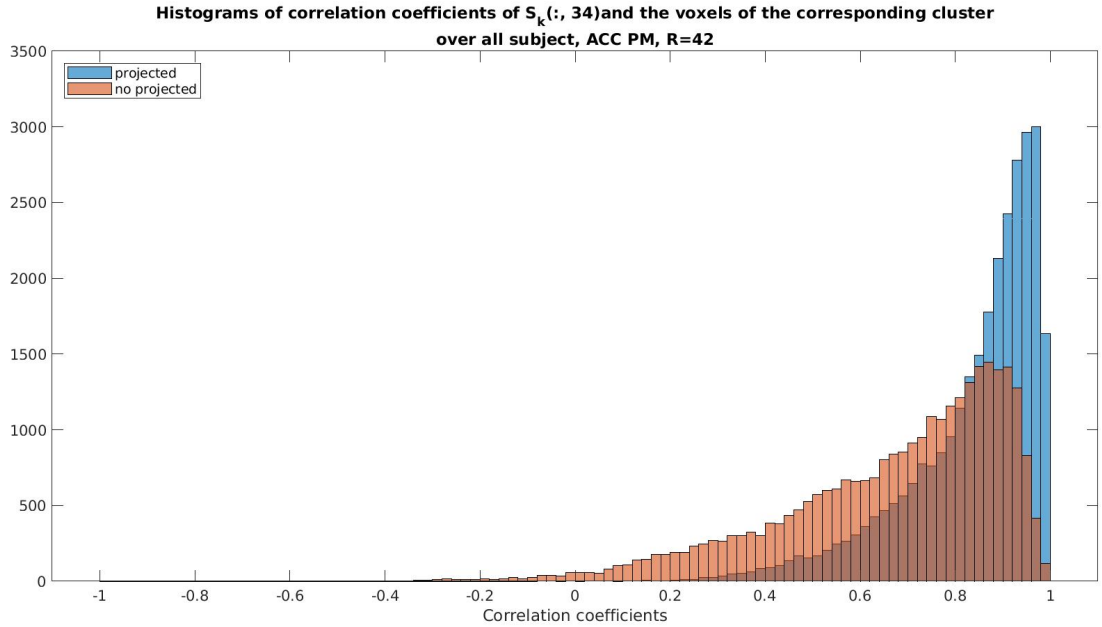


Figure B.714: For the column 34 and subject k, $k = 1, \ldots, K$, we compute the correlation coefficients between $\mathbf{S}_k(:, 34)$ and the time series of voxels which belong to this cluster. We gather the correlation coefficients over all subjects and illustrate the histograms of the correlation coefficients for initial (red color) and projected data (blue color).

B.5 Correlation coefficients of temporal components S_k, $R = 27$

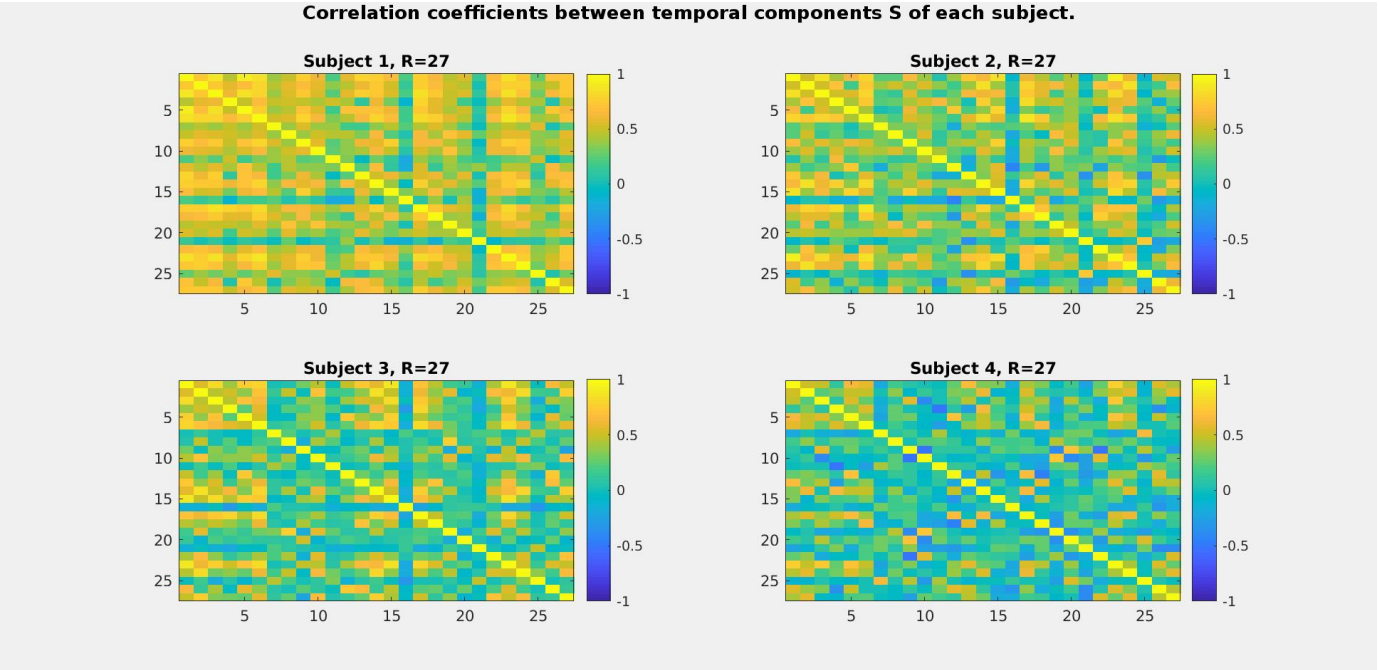


Figure B.715: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 1, \ldots, 4$ and $R = 27$.

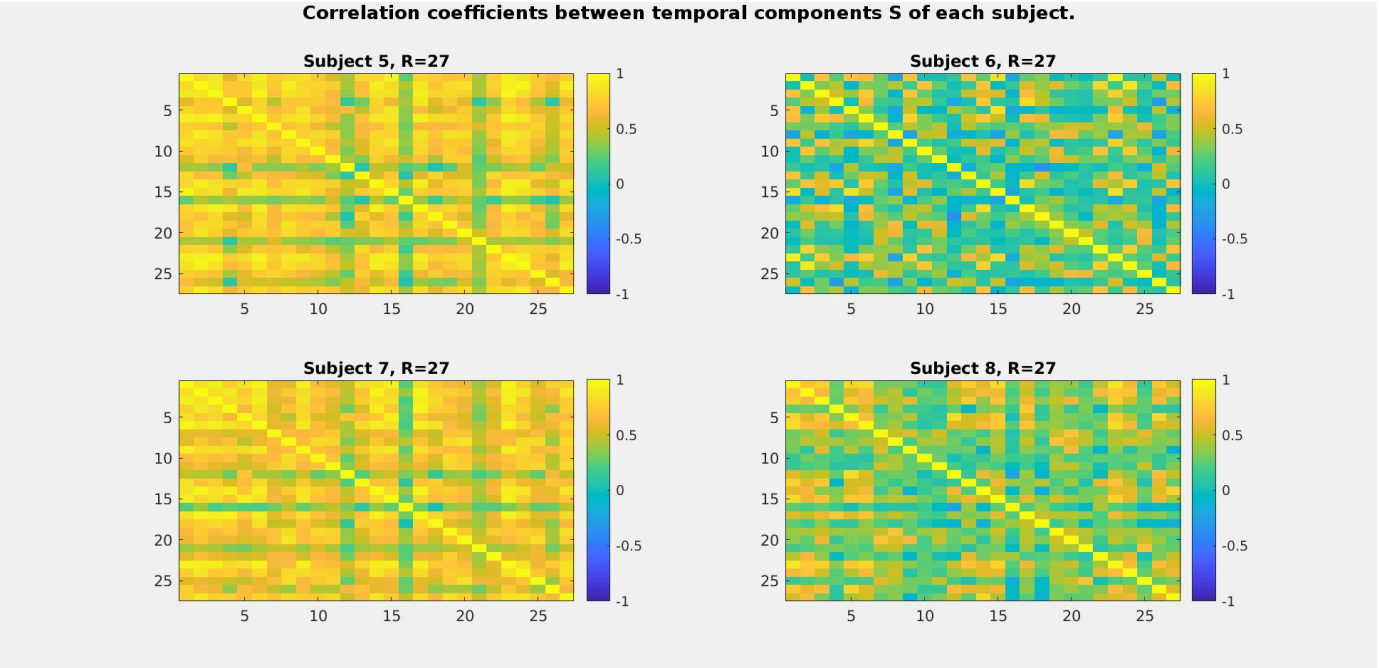


Figure B.716: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 5, \ldots, 8$ and $R = 27$.

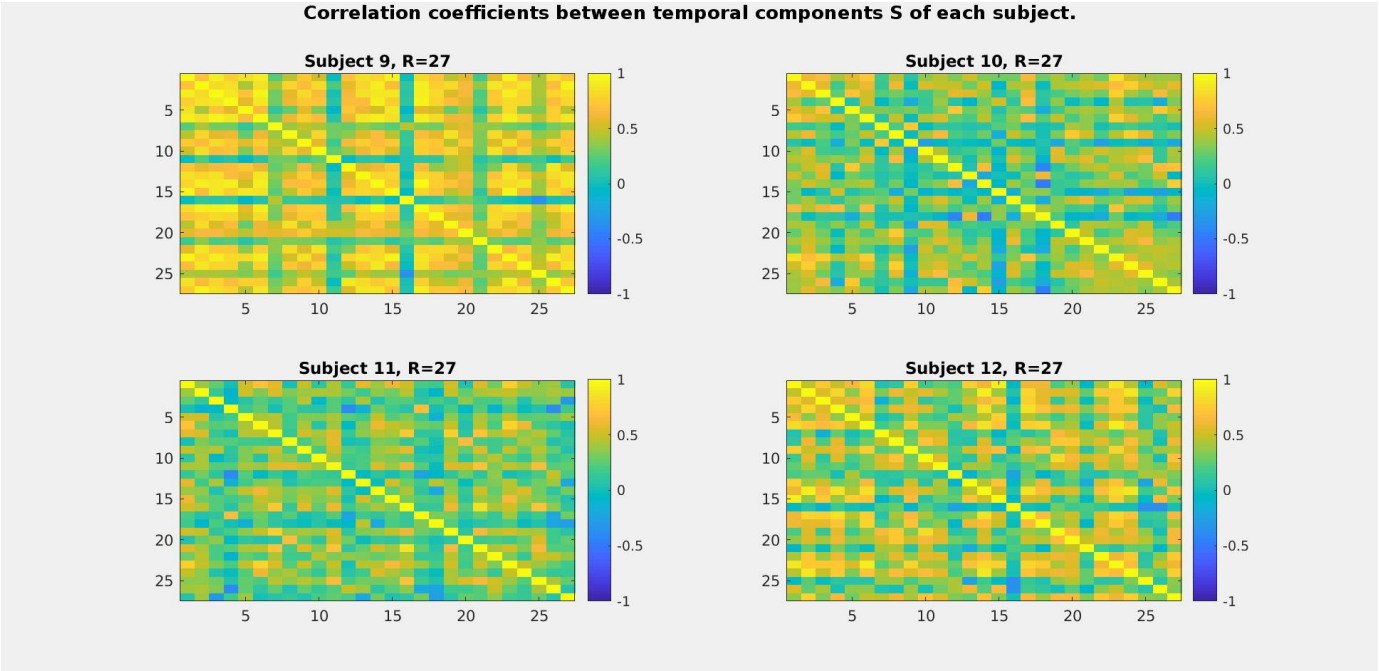


Figure B.717: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 9, \ldots, 12$ and $R = 27$.

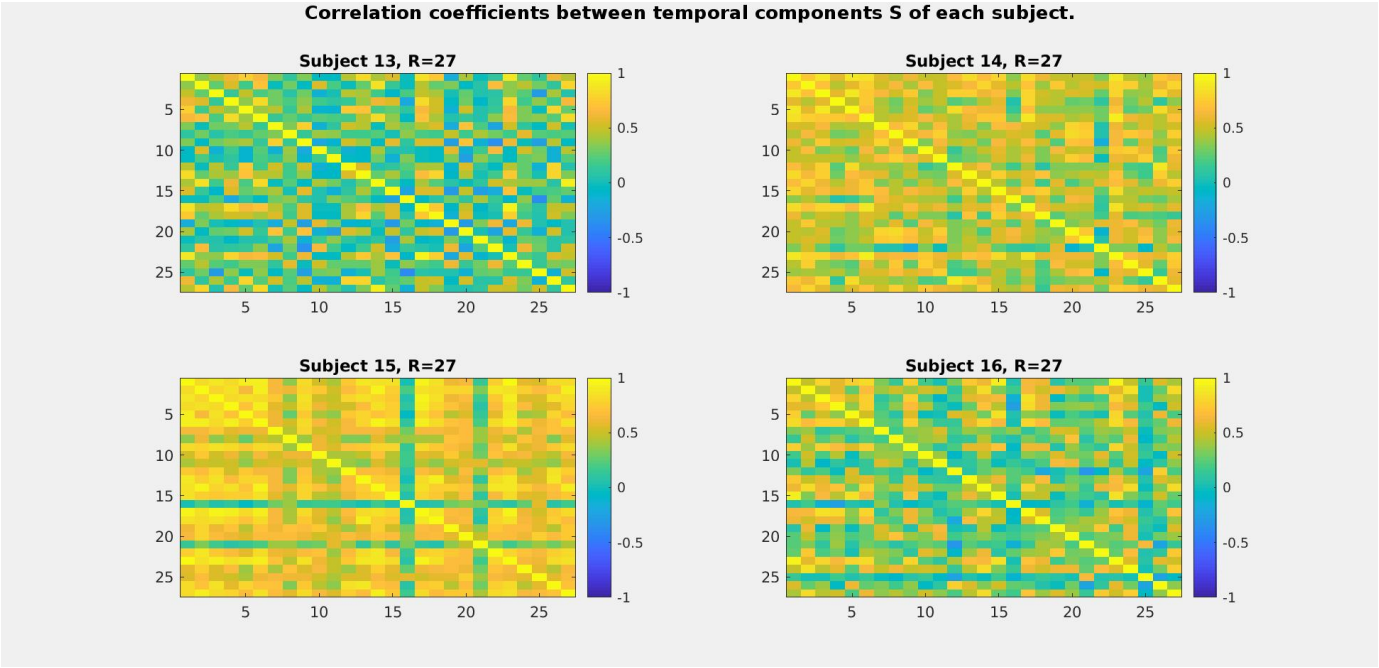


Figure B.718: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 13, \ldots, 16$ and $R = 27$.

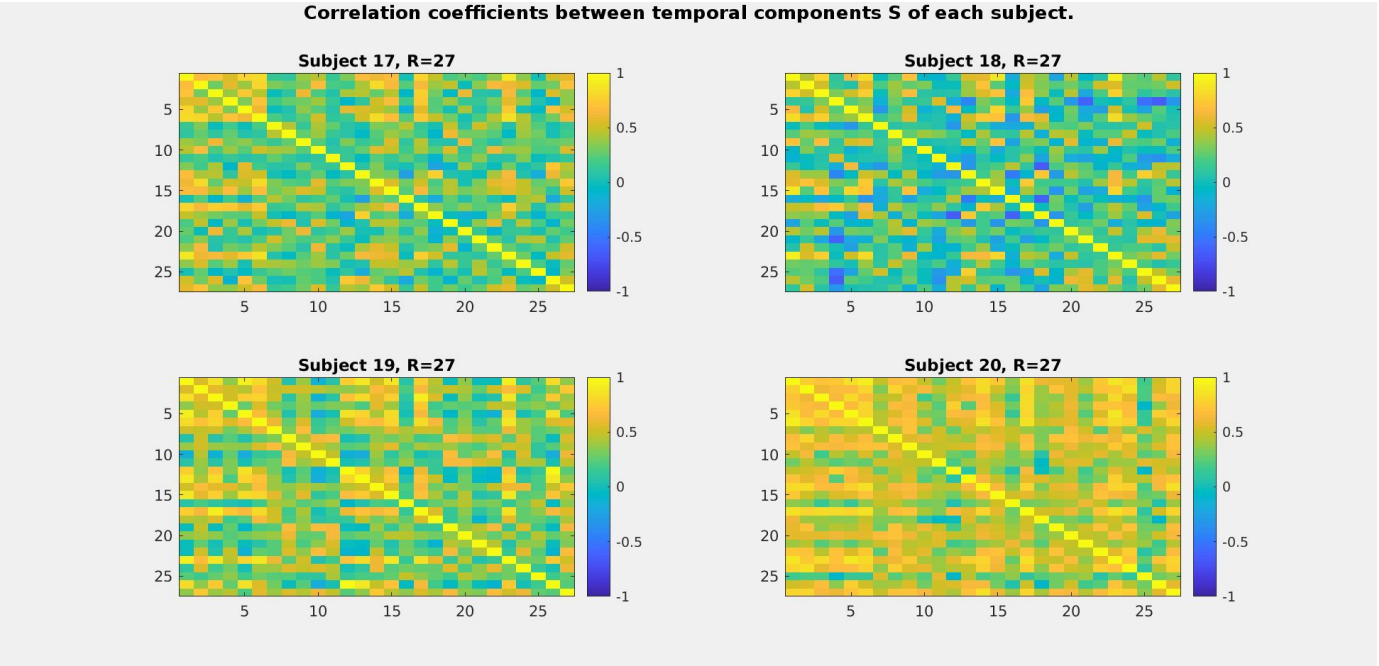


Figure B.719: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 17, \ldots, 20$ and $R = 27$.

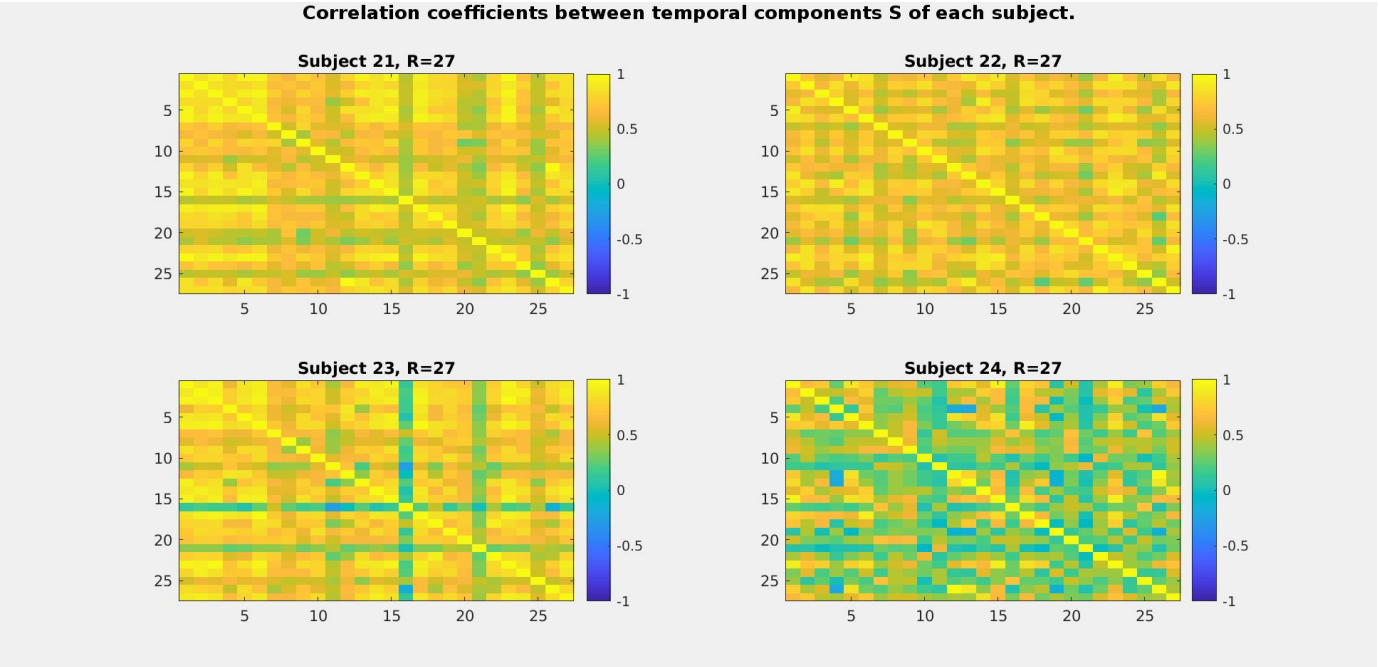


Figure B.720: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 21, \ldots, 24$ and $R = 27$.

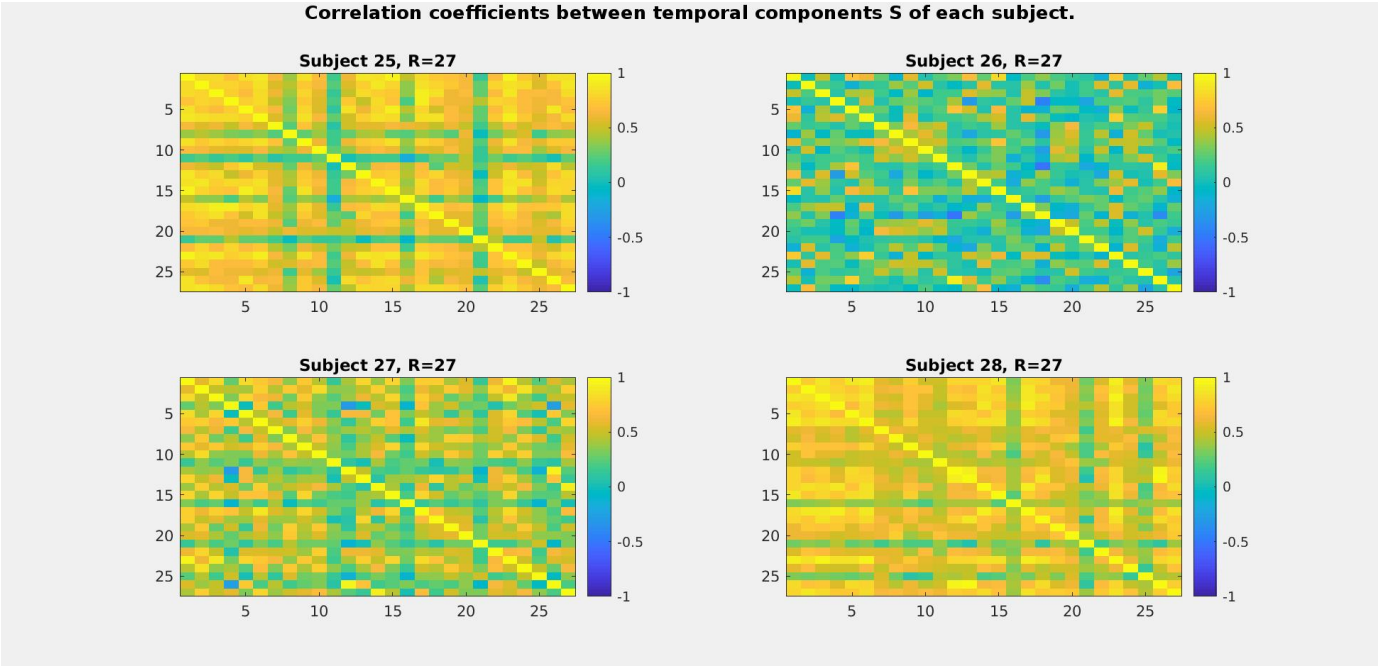


Figure B.721: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 25, \ldots, 28$ and $R = 27$.

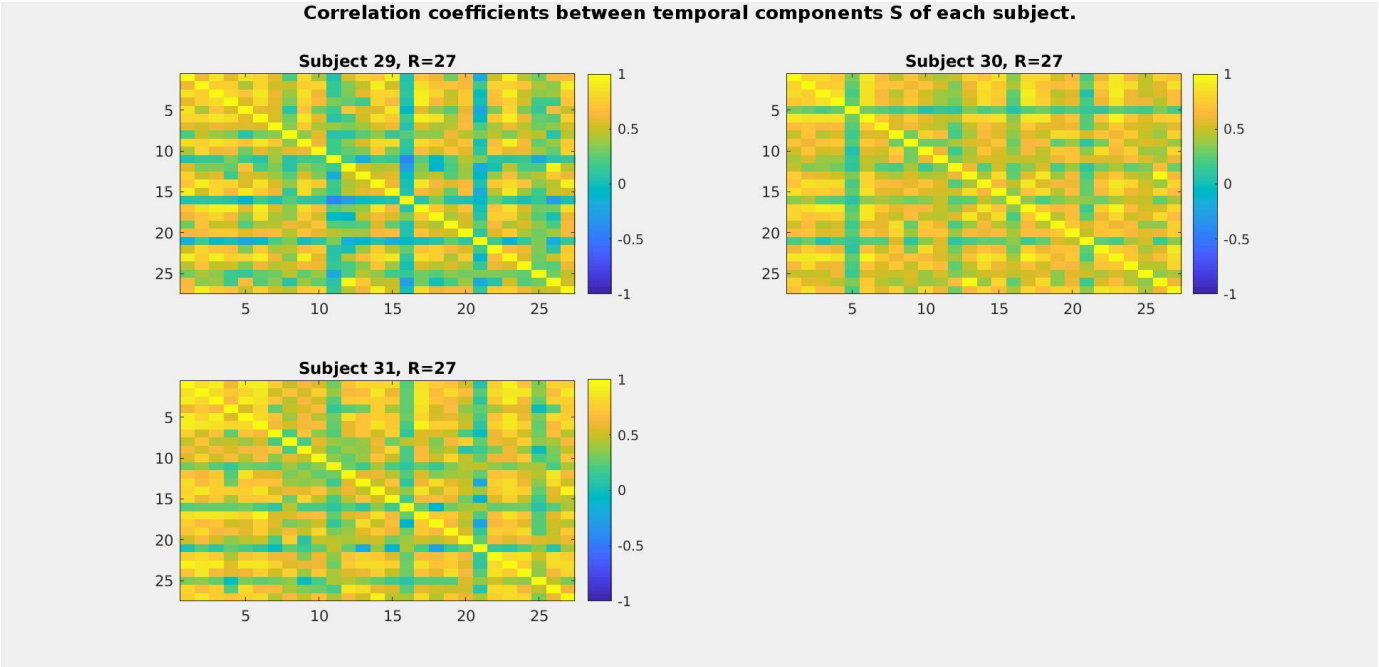


Figure B.722: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 29, \ldots, 31$ and $R = 27$.

B.6 Correlation coefficients of temporal components S_k, $R = 42$

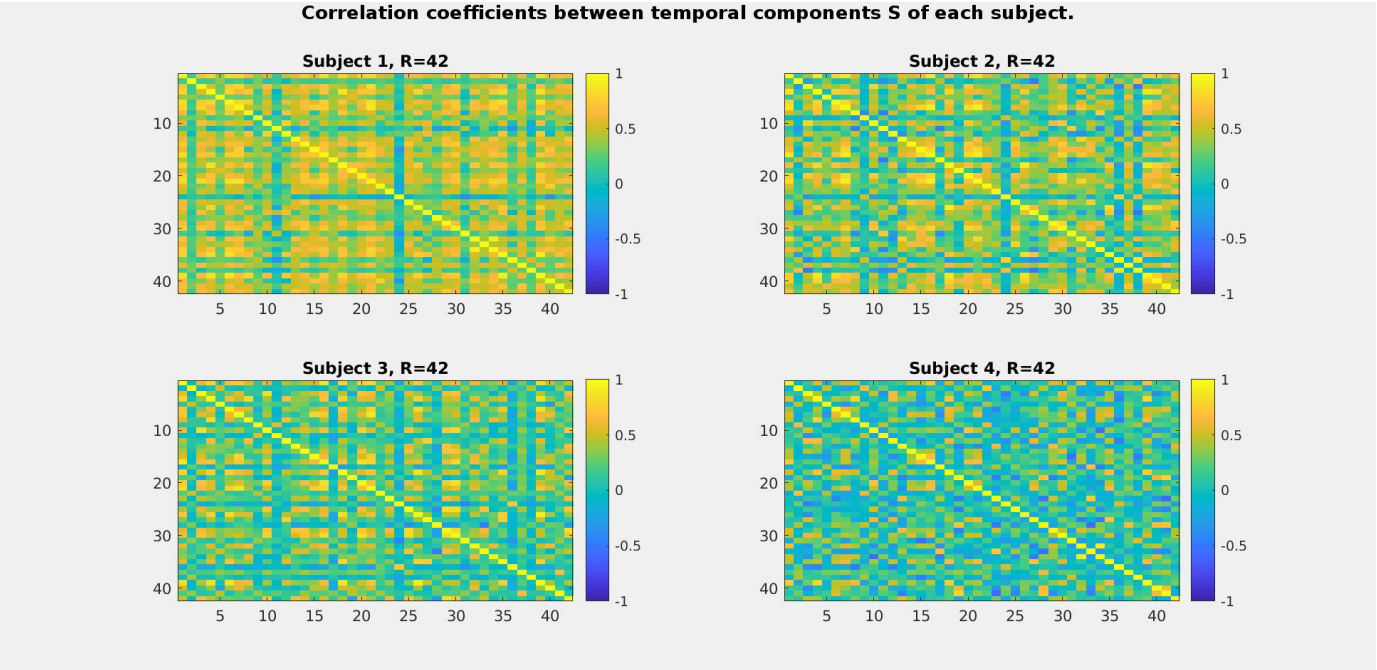


Figure B.723: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 1, \ldots, 4$ and $R = 42$.

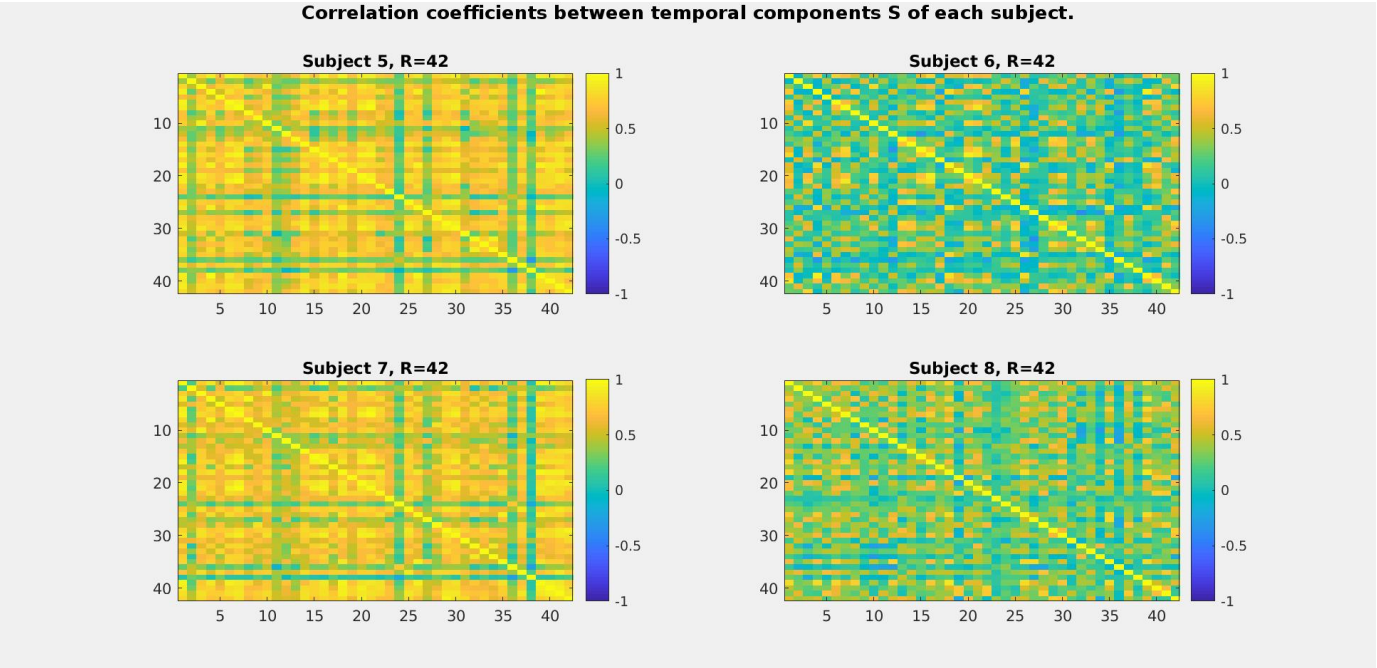


Figure B.724: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 5, \ldots, 8$ and $R = 42$.

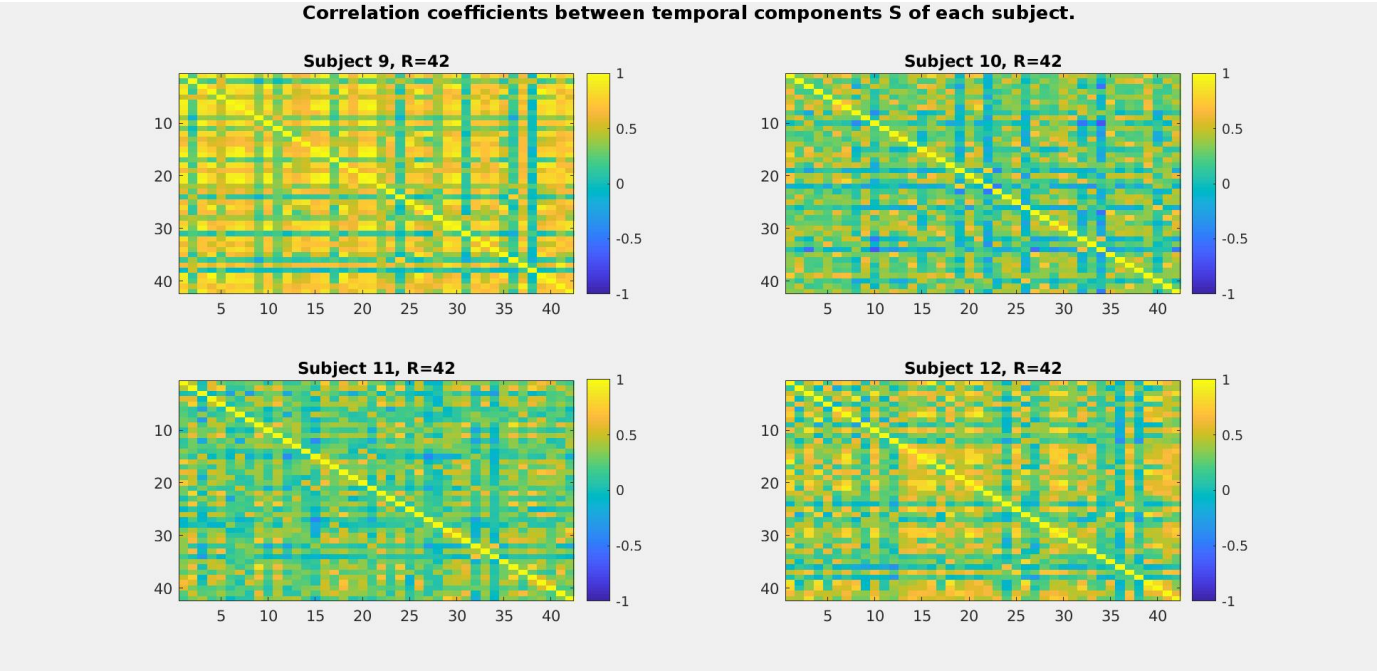


Figure B.725: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 9, \ldots, 12$ and $R = 42$.

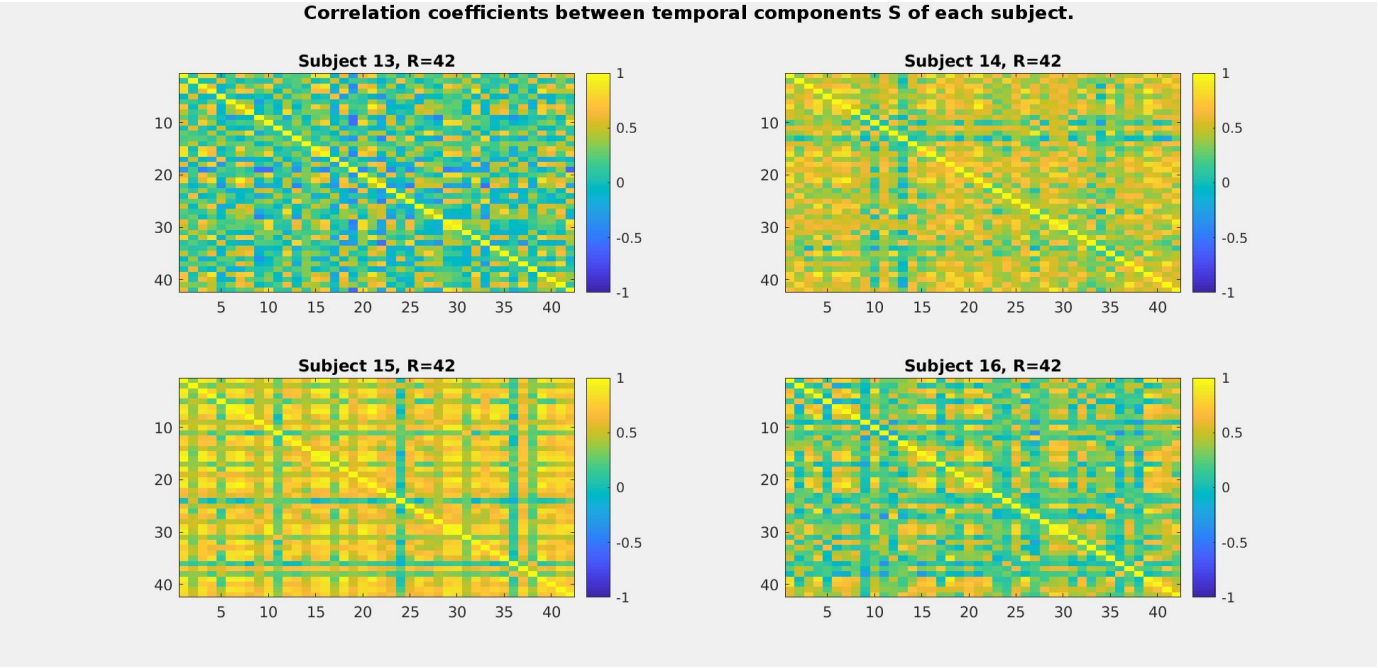


Figure B.726: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 13, \ldots, 16$ and $R = 42$.

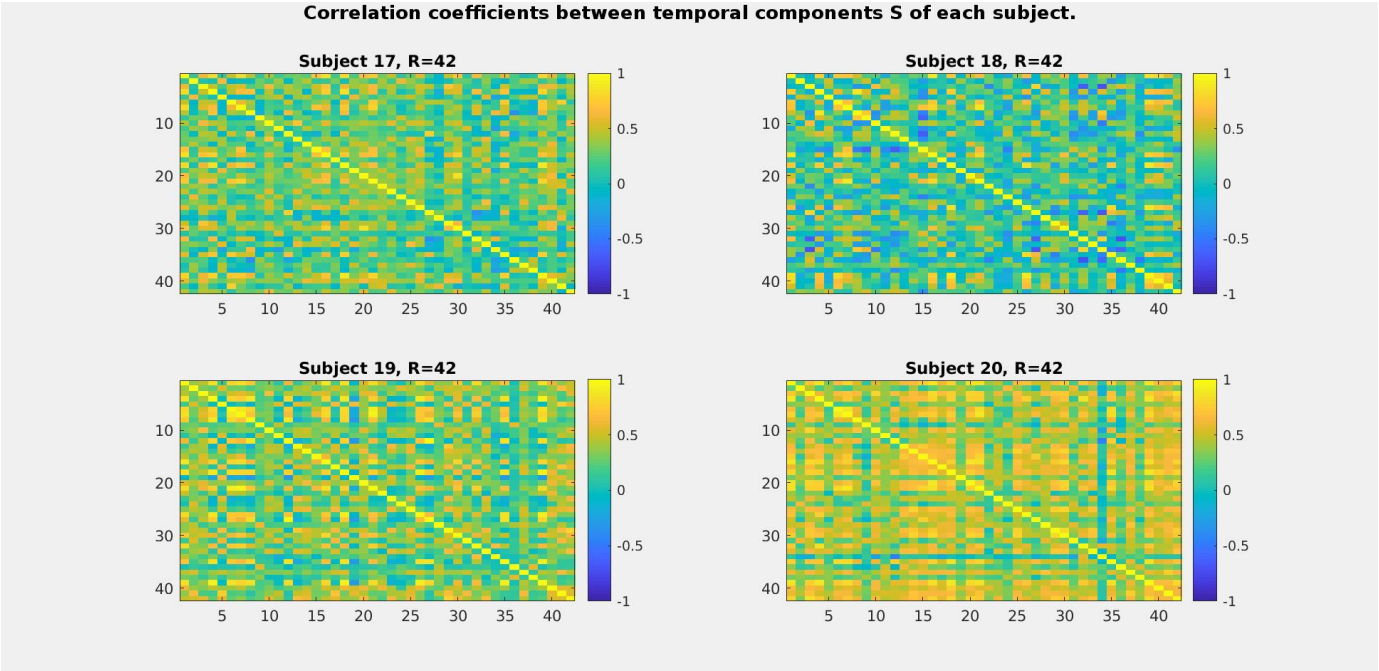


Figure B.727: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 17, \ldots, 20$ and $R = 42$.

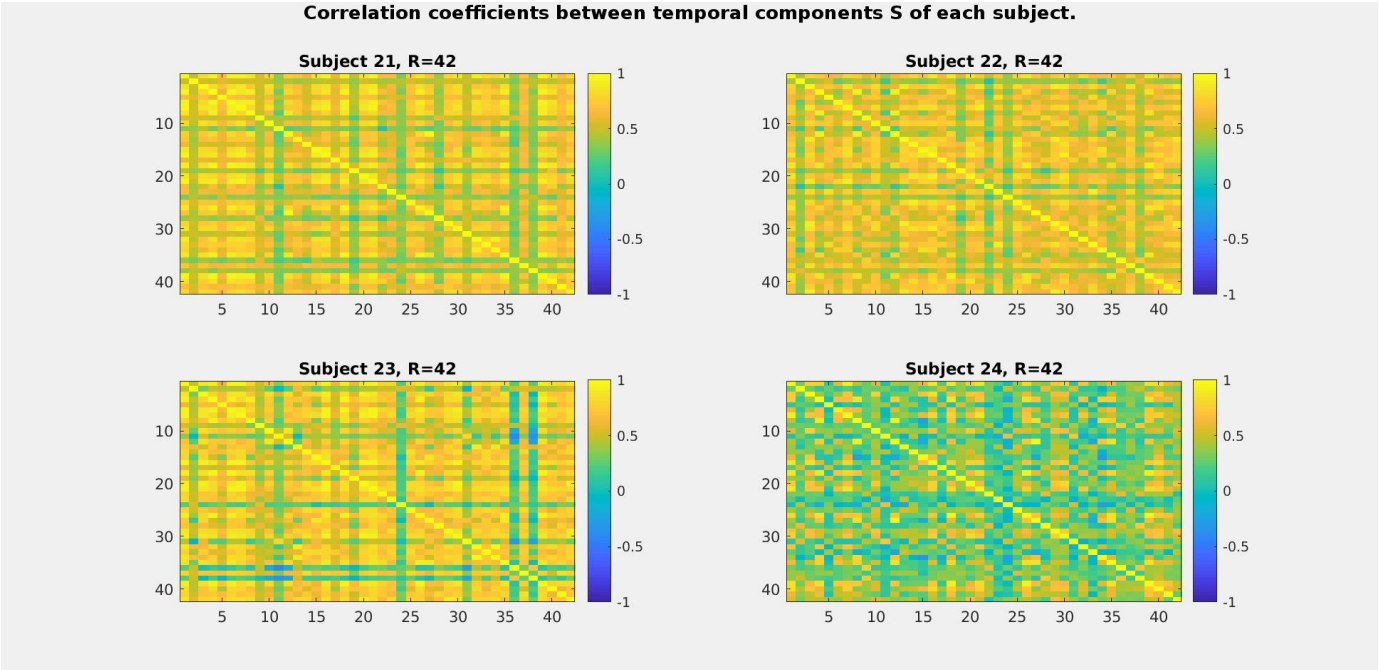


Figure B.728: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 21, \ldots, 24$ and $R = 42$.

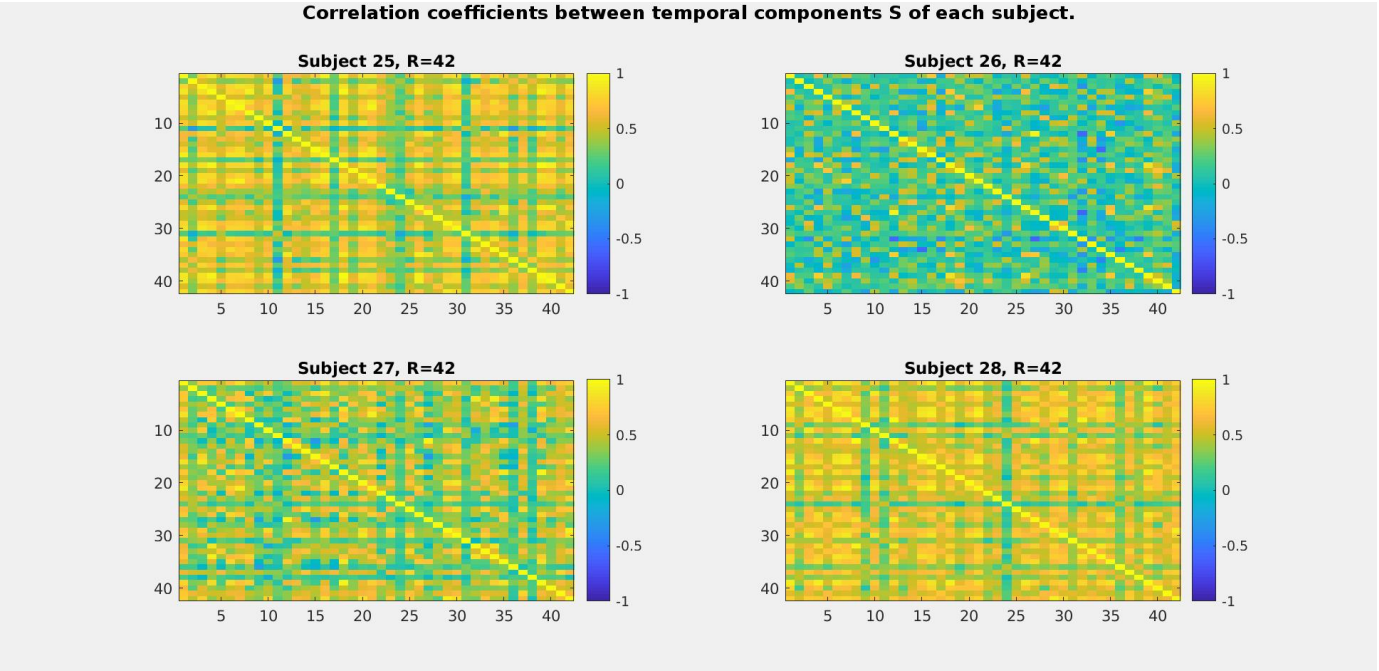


Figure B.729: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 25, \ldots, 28$ and $R = 42$.

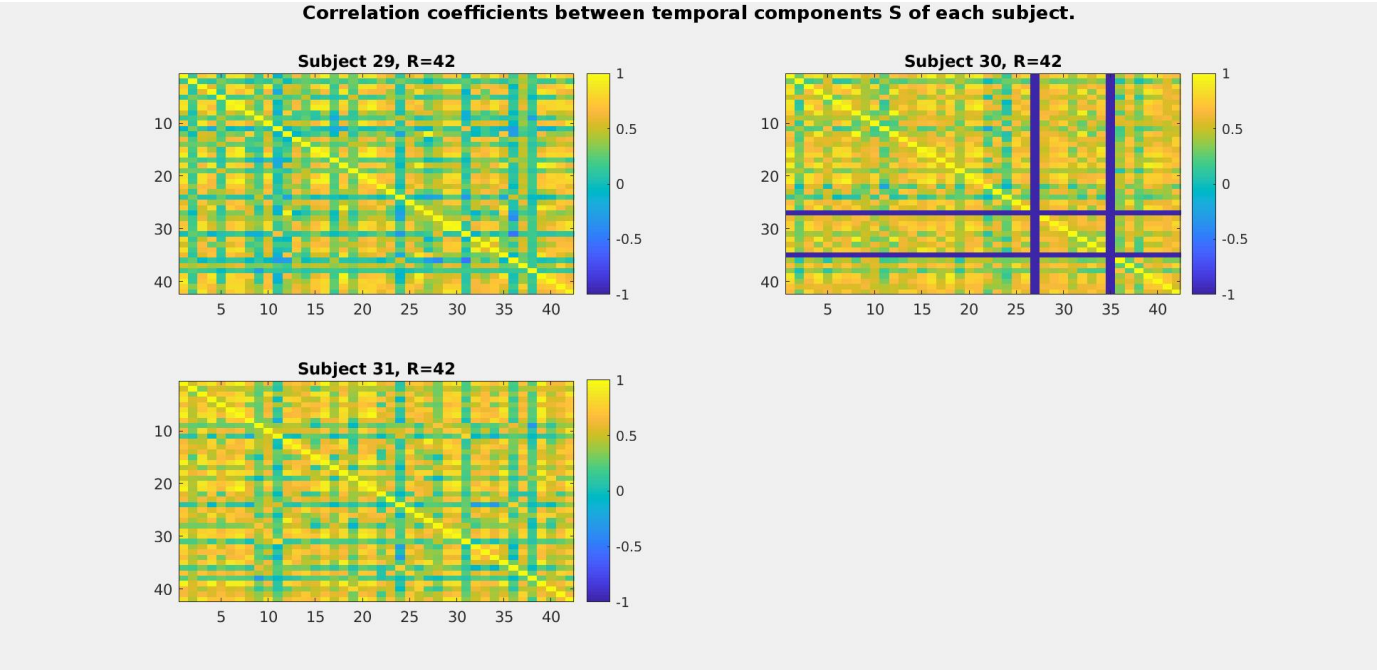


Figure B.730: Correlation coefficients between temporal components $\mathbf{S}k$, for $k = 29, \ldots, 31$ and $R = 42$.

B.7 Comparison of estimated A matrices for common subspace dimension 27 and 42

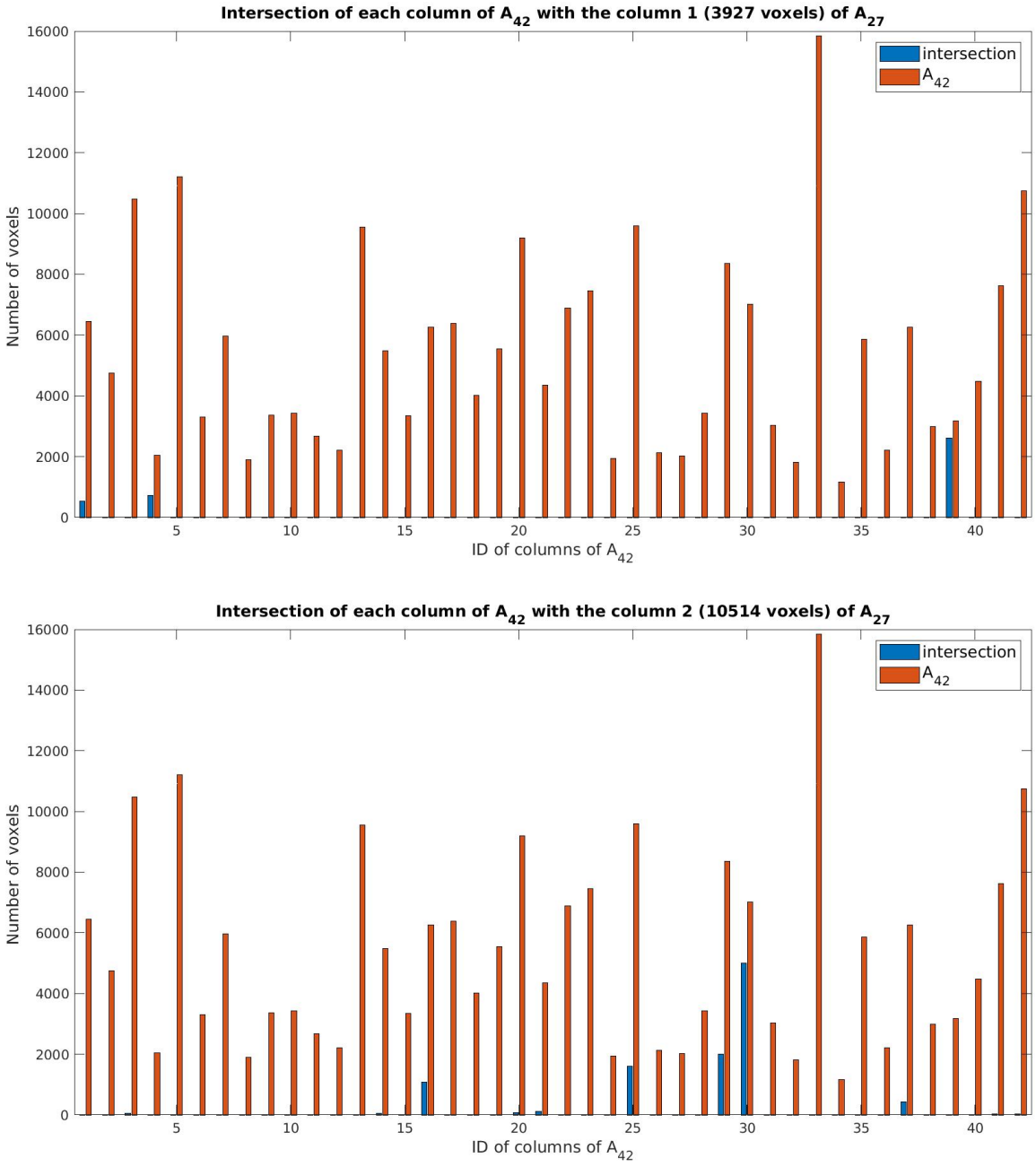


Figure B.731: For the column r, for $r = 1, 2$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

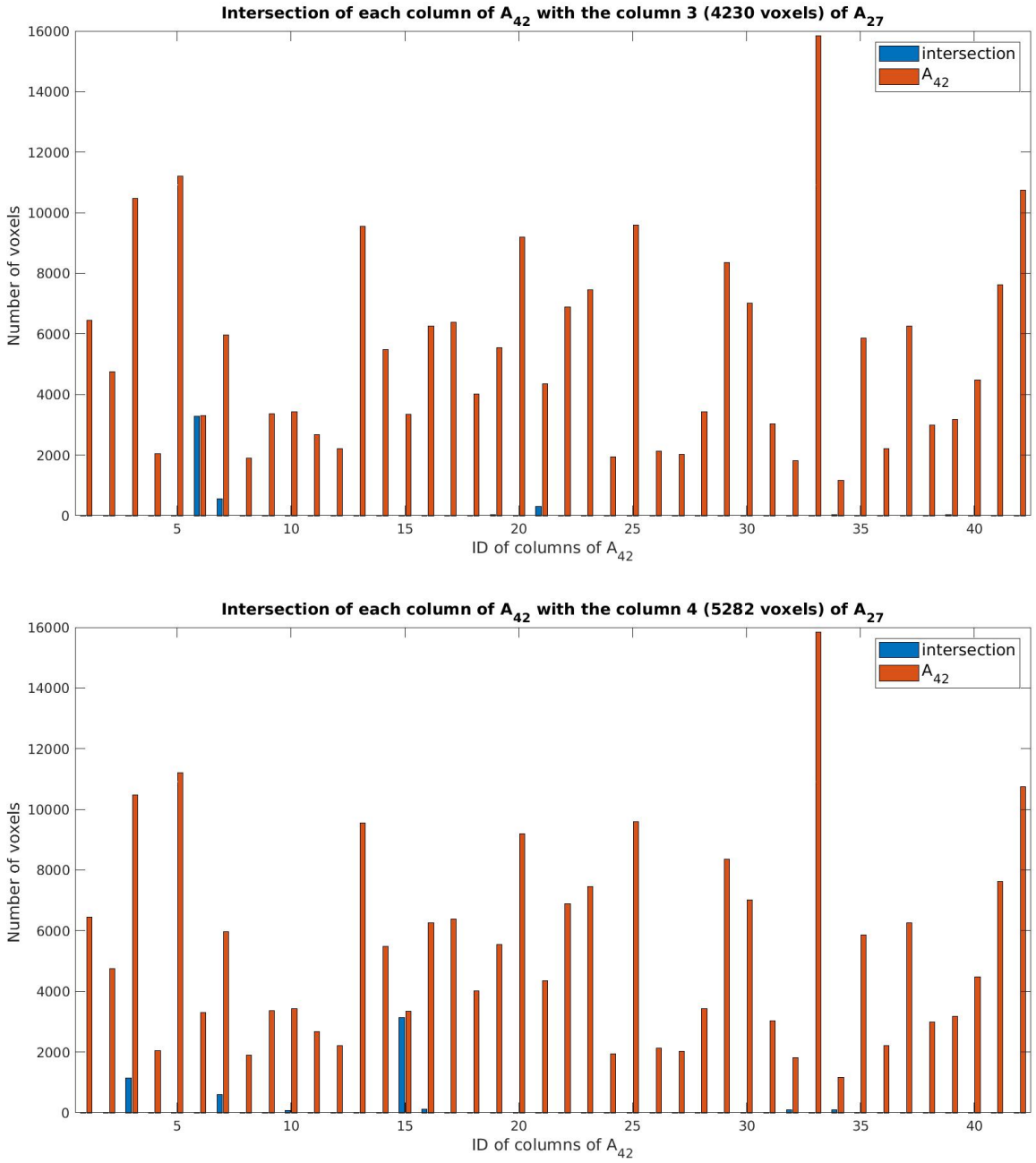


Figure B.732: For the column r, for $r = 3,\ 4$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

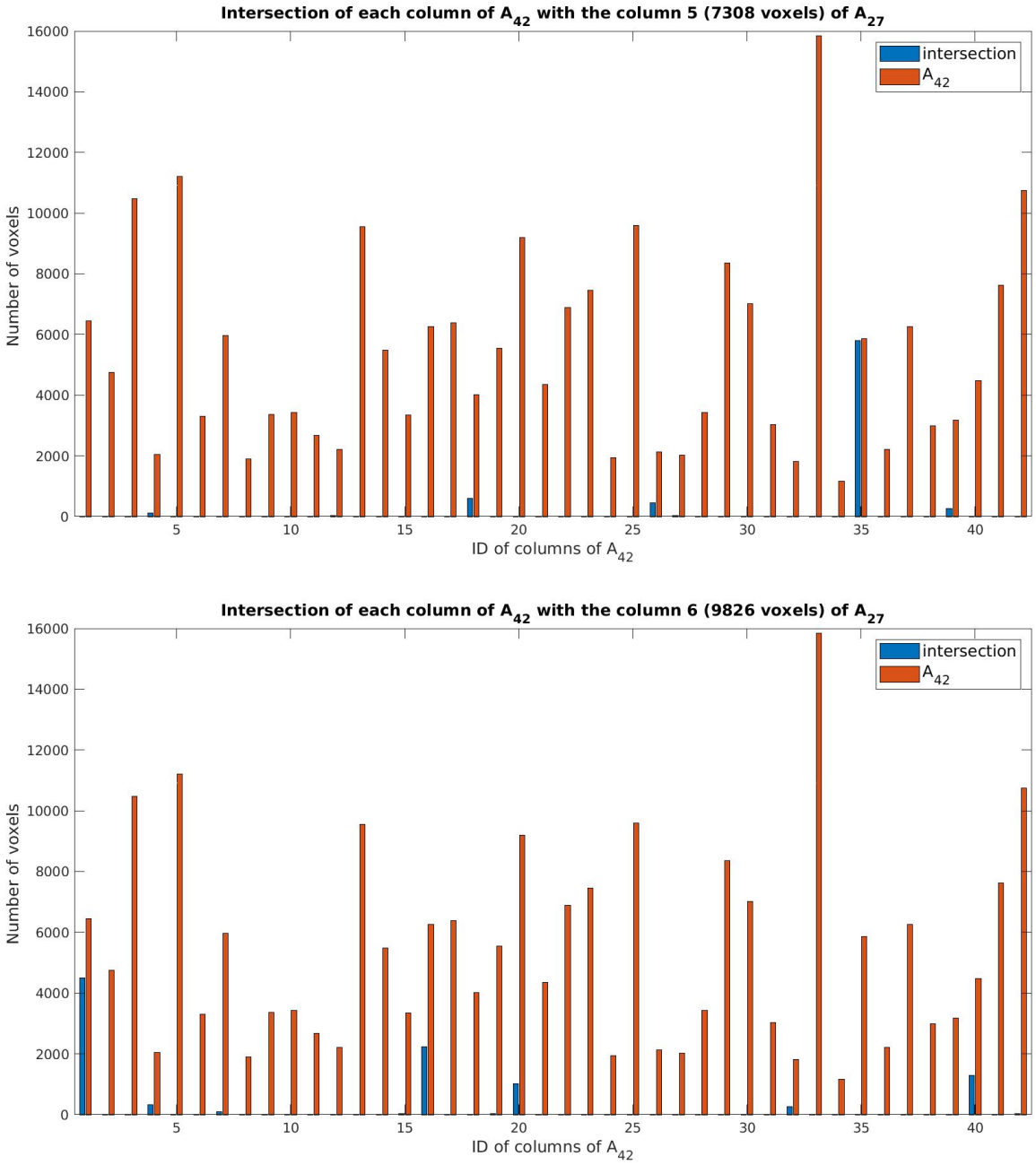


Figure B.733: For the column r, for $r = 5,\ 6$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

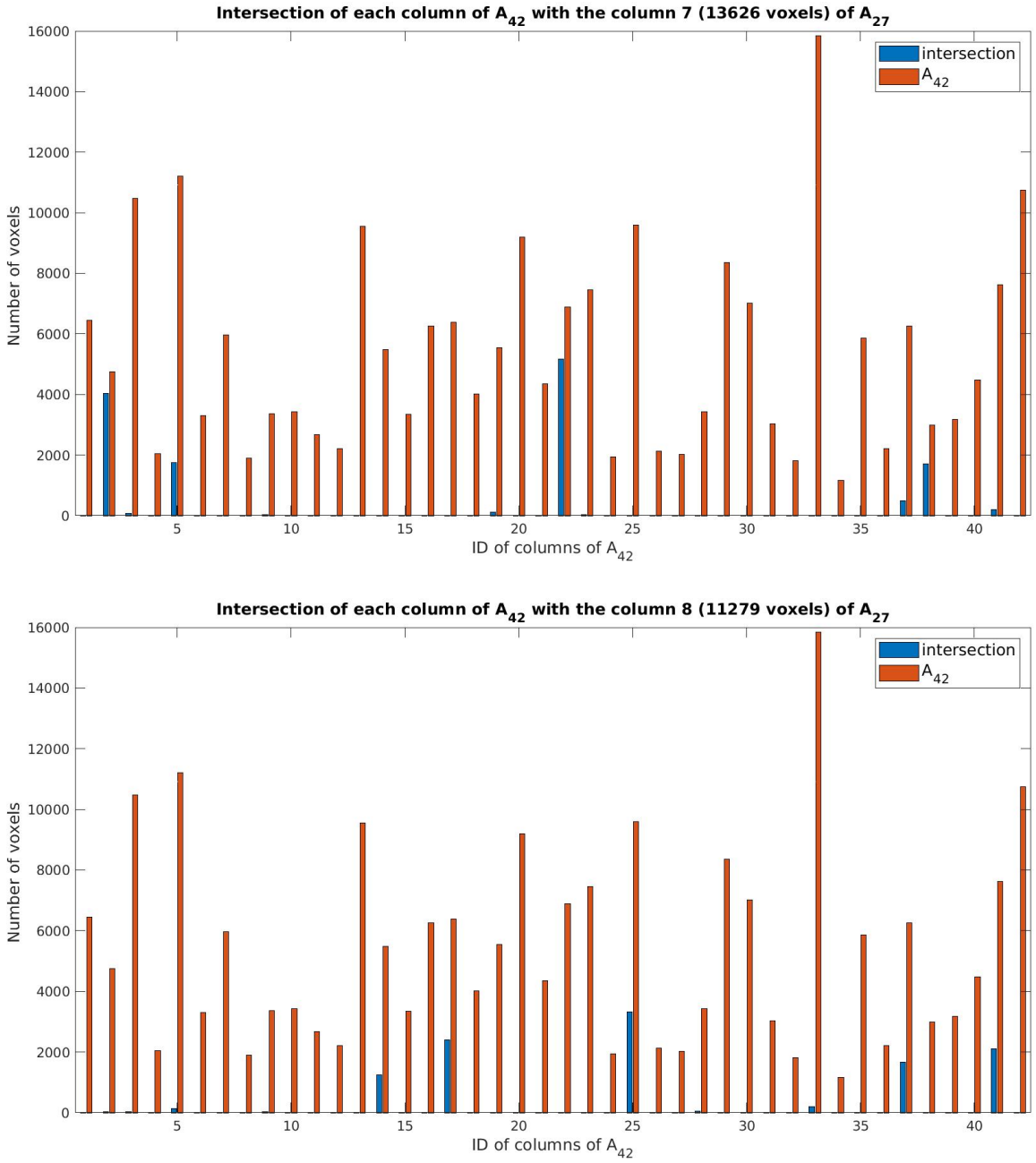


Figure B.734: For the column r, for $r = 7, 8$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

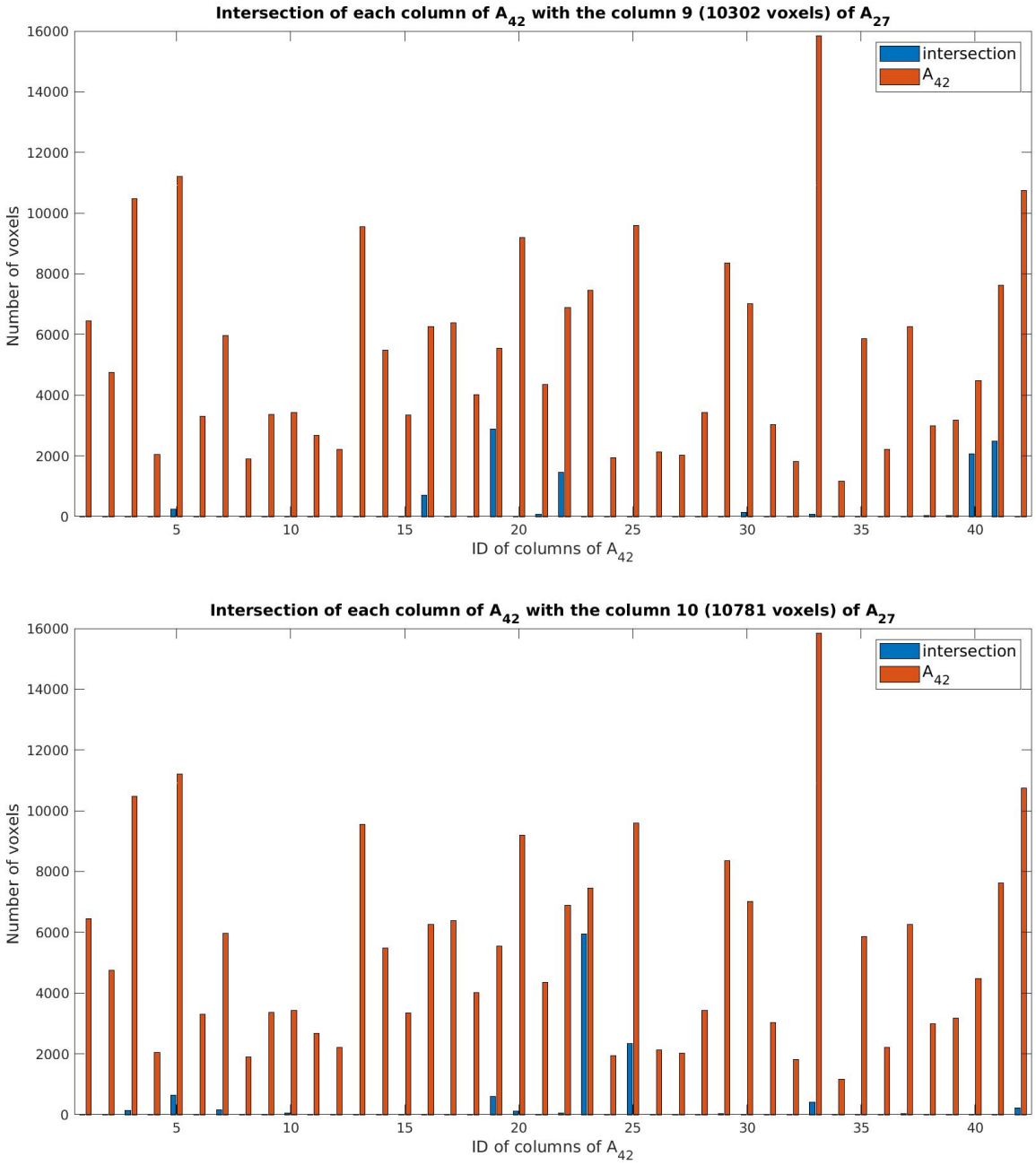


Figure B.735: For the column r, for $r = 9,\ 10$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

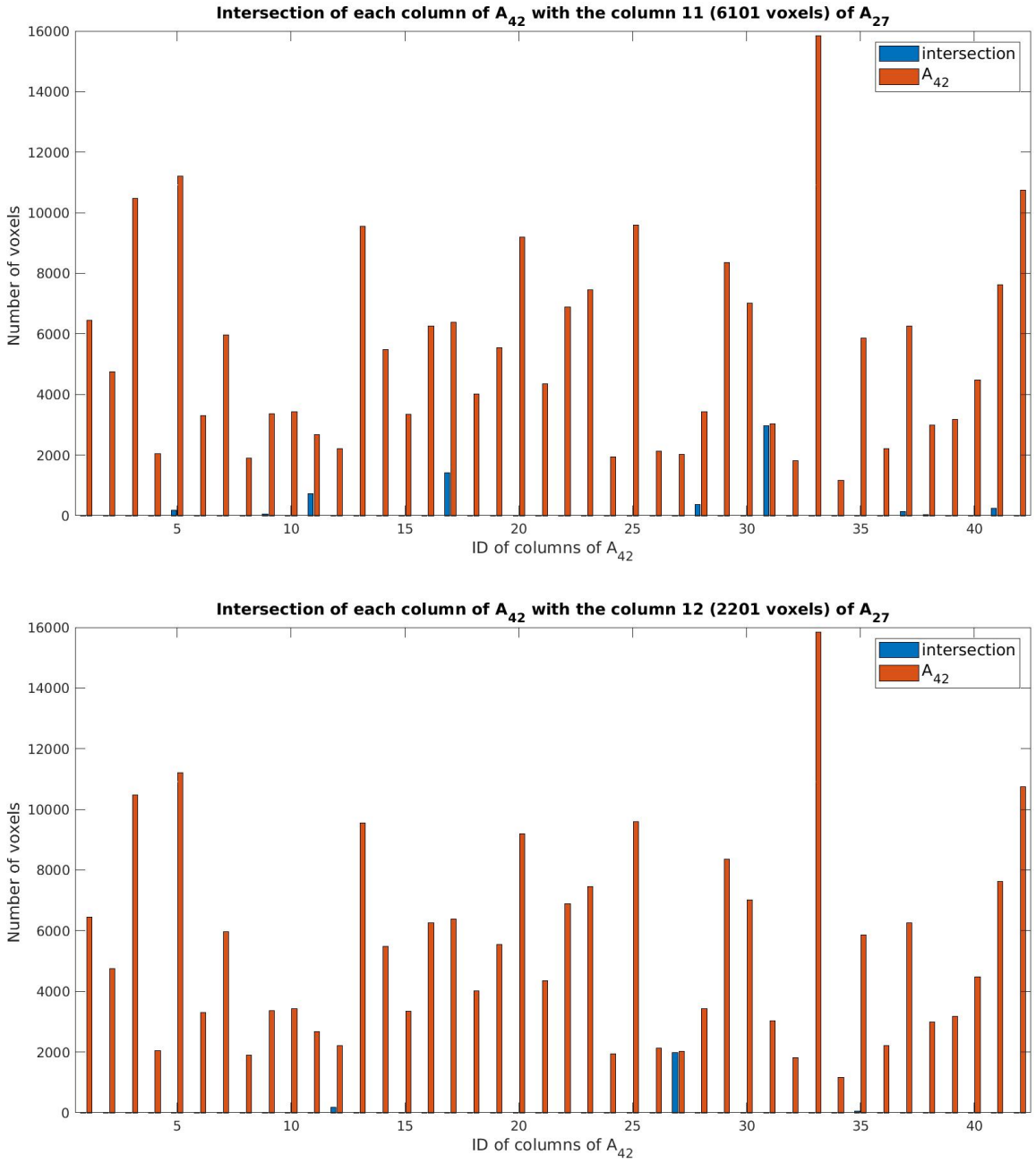


Figure B.736: For the column r, for $r = 11,\ 12$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

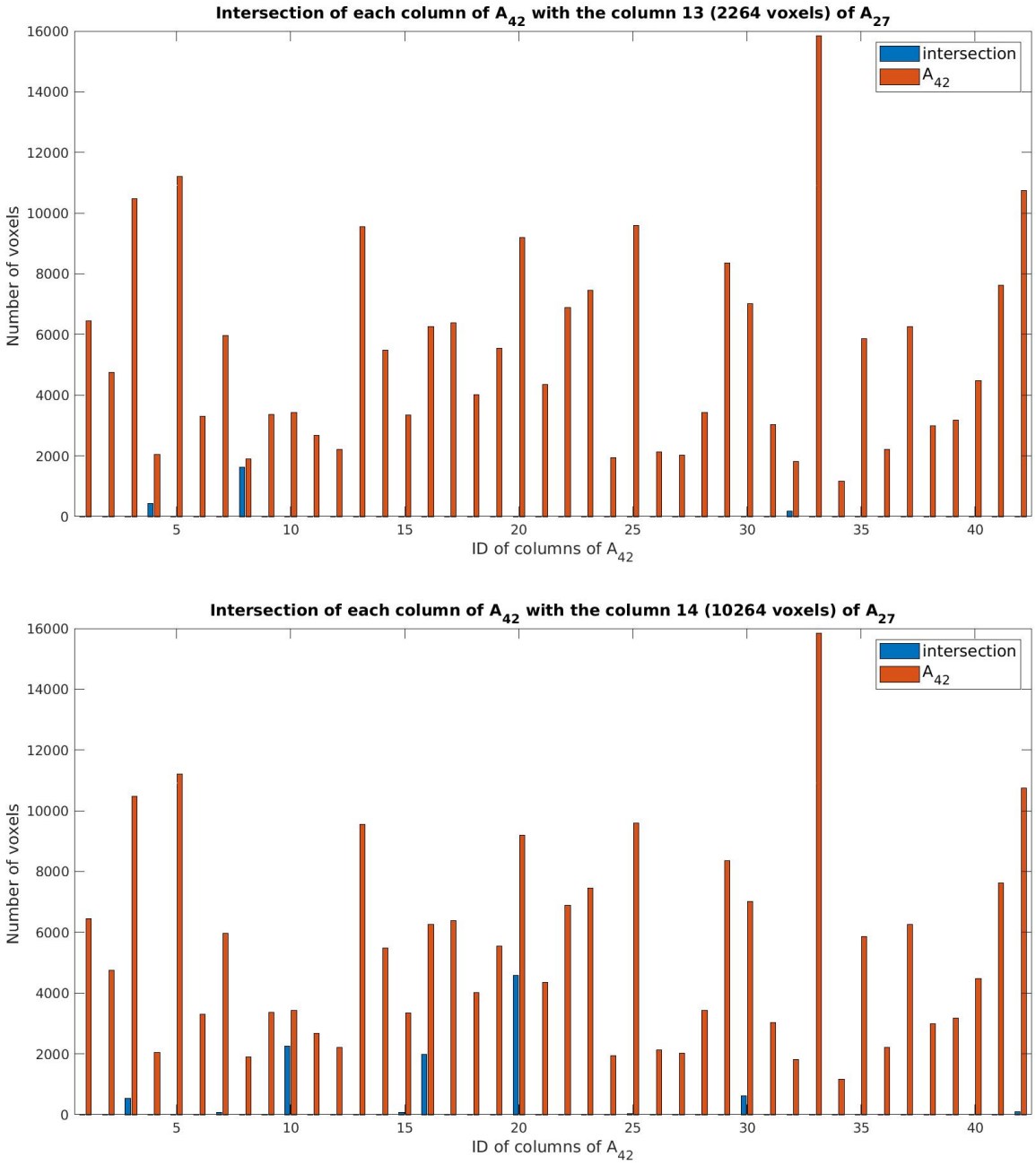


Figure B.737: For the column r, for $r = 13,\ 14$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

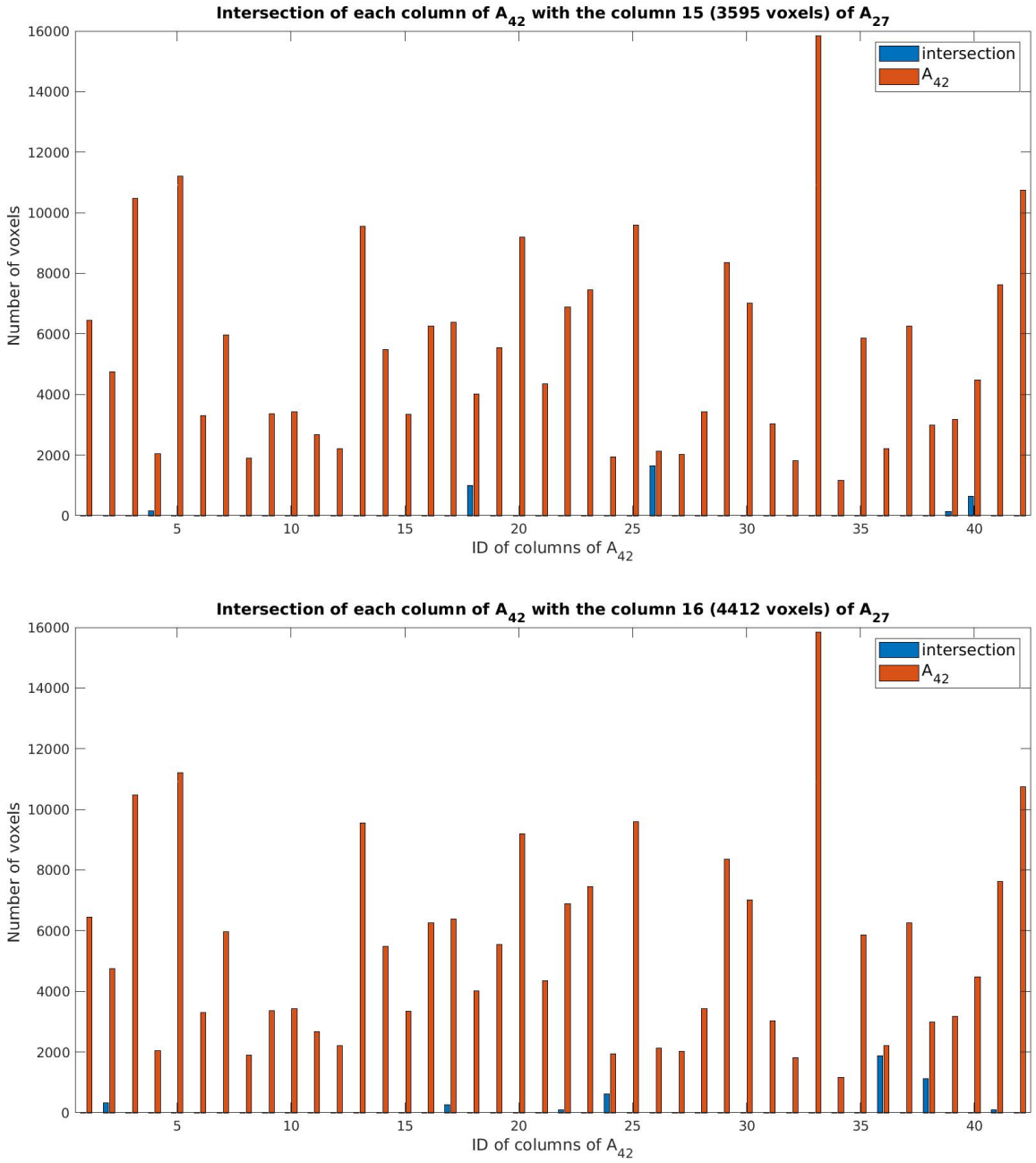


Figure B.738: For the column r, for $r = 15,\ 16$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

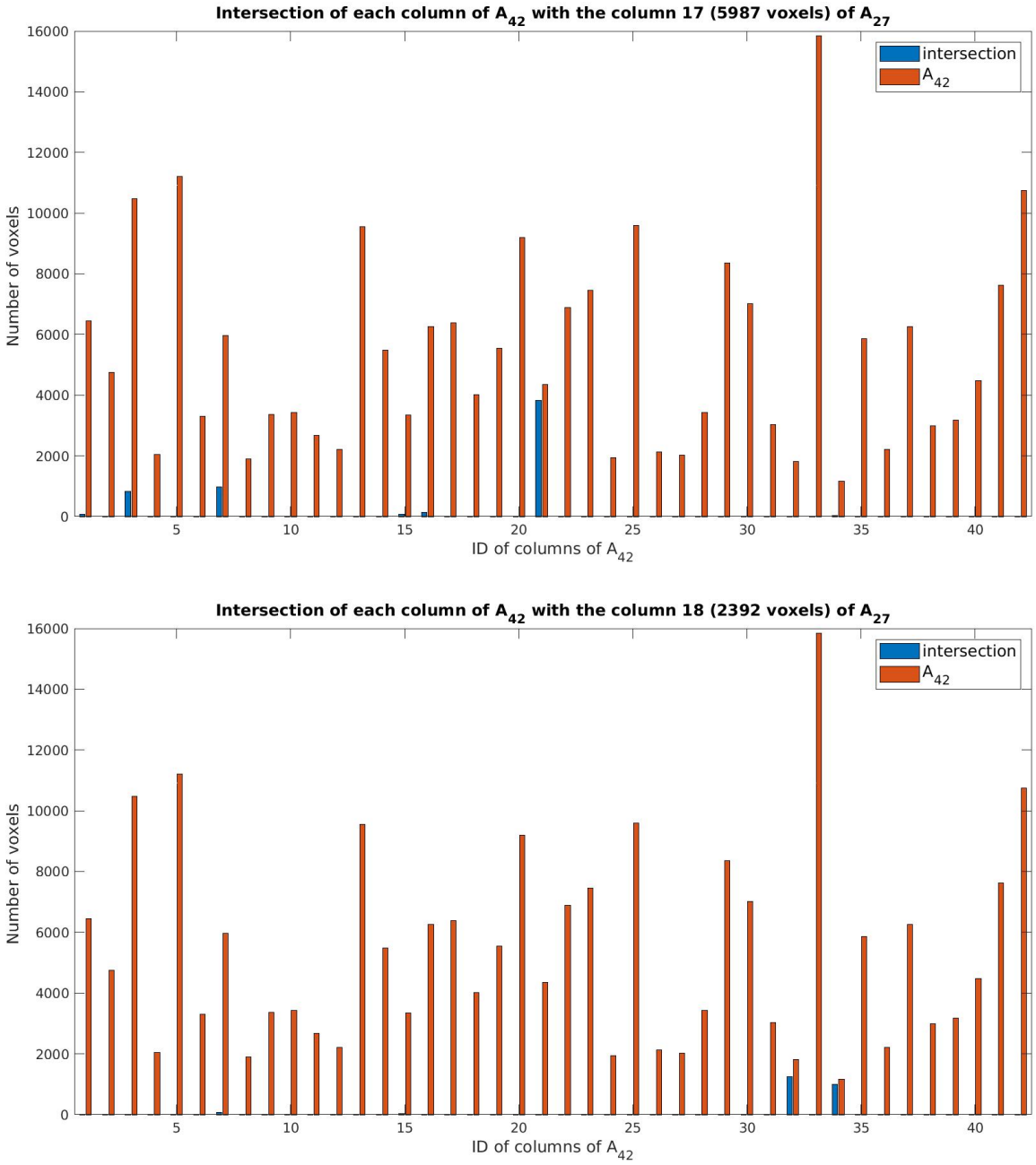


Figure B.739: For the column r, for $r = 17,\ 18$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

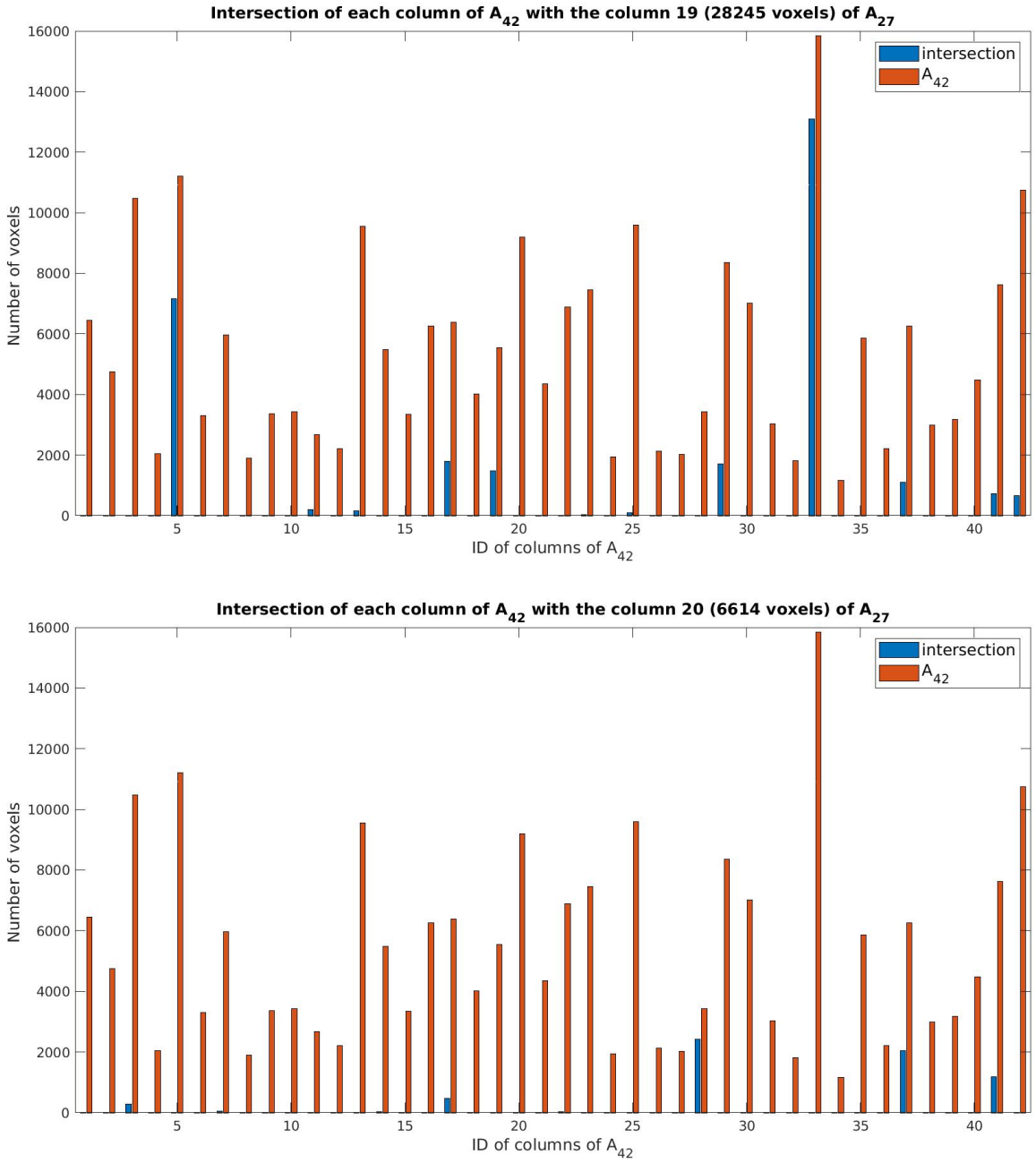


Figure B.740: For the column r, for $r = 19,\ 20$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

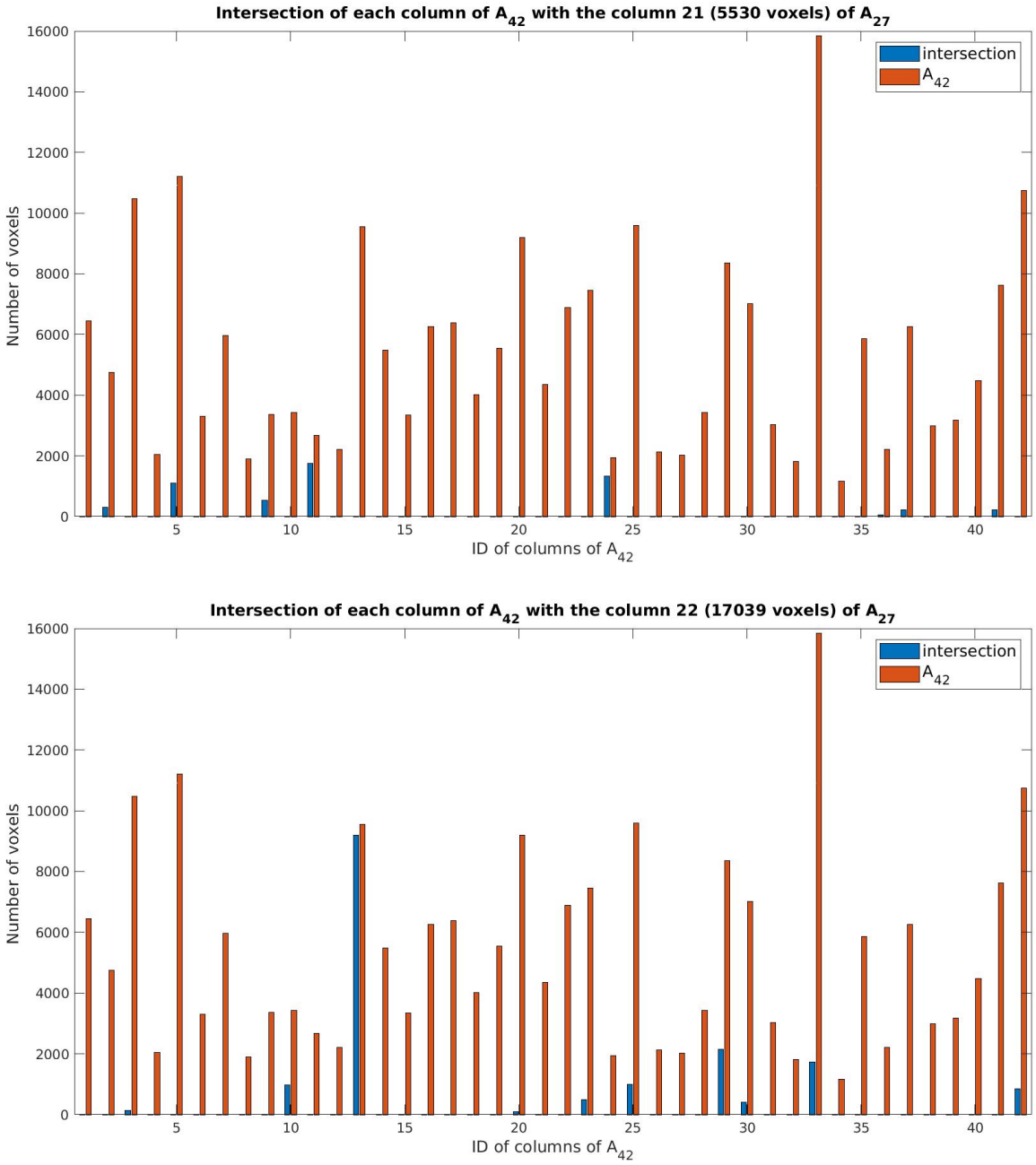


Figure B.741: For the column r, for $r = 21,\ 22$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

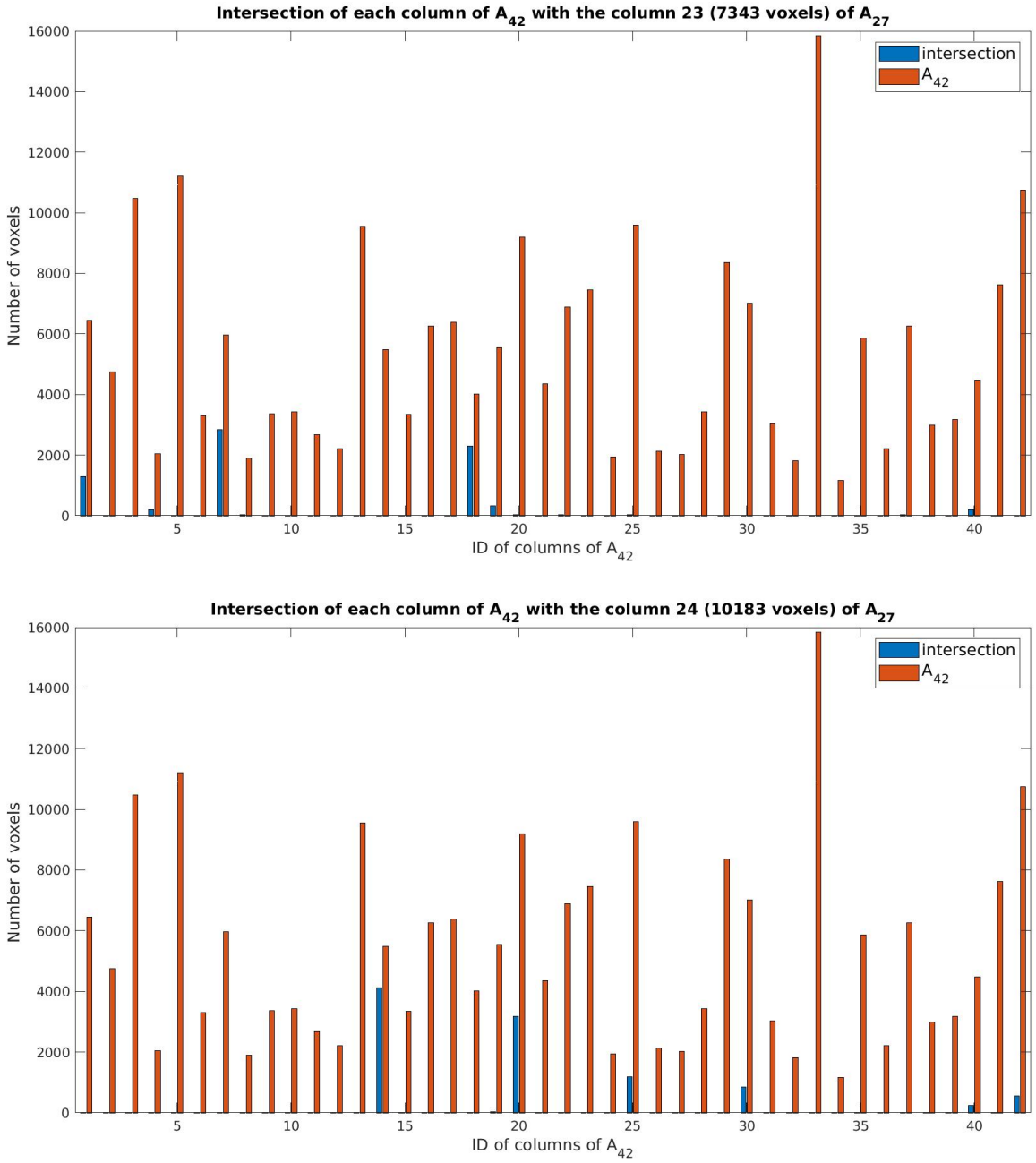


Figure B.742: For the column r, for $r = 23,\ 24$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

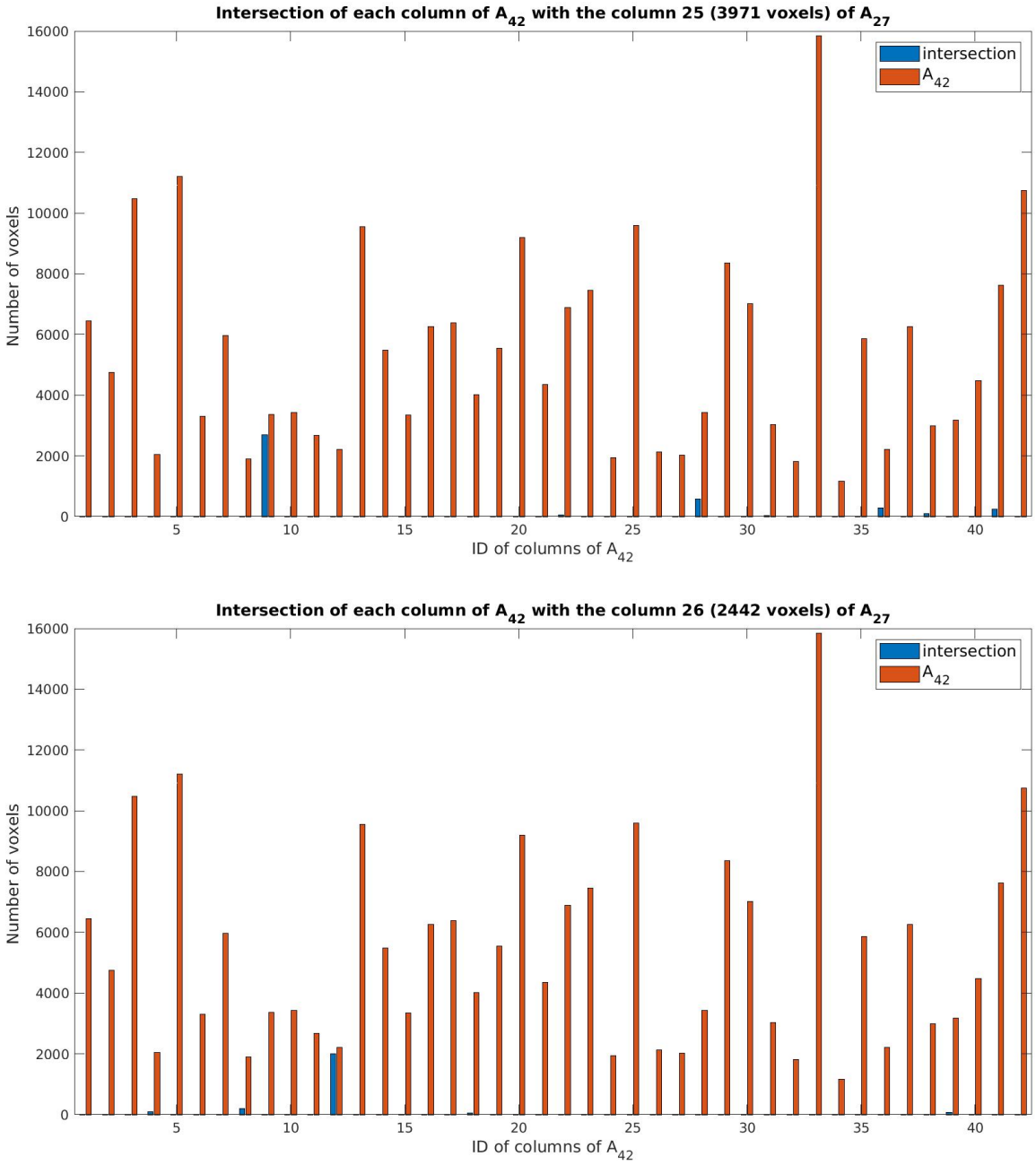


Figure B.743: For the column r, for $r = 25,\ 26$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

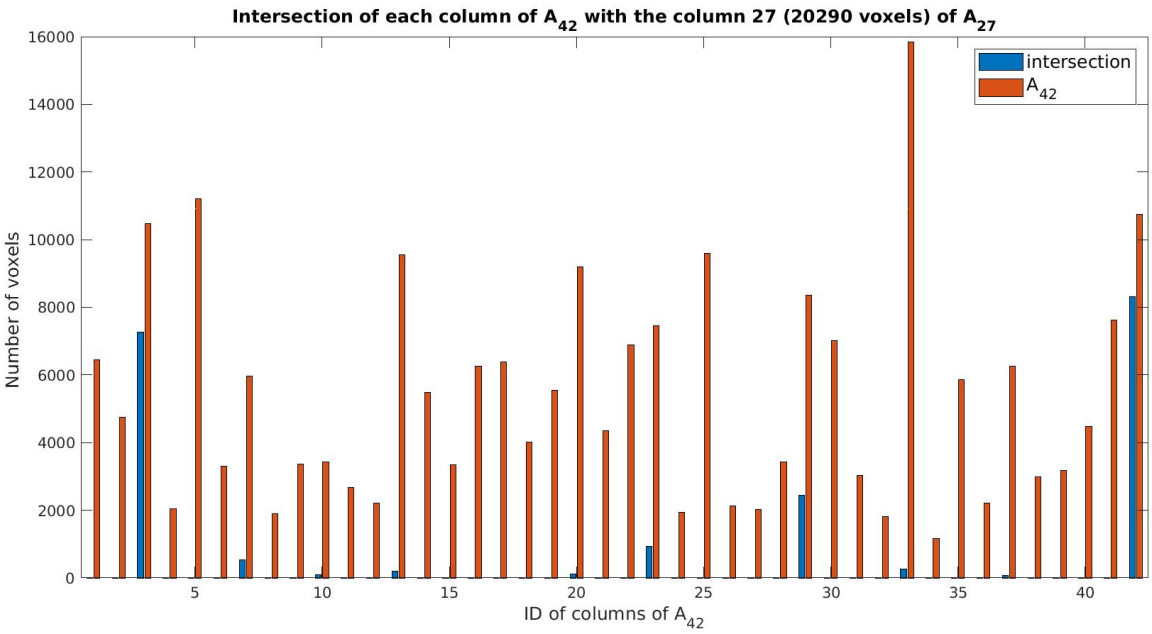


Figure B.744: For the column r, for $r = 27$, of the $\mathbf{A}_{27}^{est}$, we illustrate the size of the intersection between the non-zero elements of column r and the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (blue bar). We also illustrate the number of the non-zero elements of each column of the $\mathbf{A}_{42}^{est}$ (red bar).

Bibliography

[1] P. A. Karakasis, A. P. Liavas, N. D. Sidiropoulos, P. G. Simos, and E. Papadaki, "Multi-subject resting-state fmri data analysis via generalized canonical correlation analysis," in *2020 28th European Signal Processing Conference (EUSIPCO)*, 2021, pp. 1040–1044.

[2] B. Thirion, G. Varoquaux, E. Dohmatob, and J.-B. Poline, "Which fmri clustering gives good brain parcellations?" *Frontiers in Neuroscience*, vol. 8, p. 167, 2014. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2014.00167

[3] B. Yeo, F. Krienen, J. Sepulcre, M. Sabuncu, D. Lashkari, M. Hollinshead, J. Roffman, J. Smoller, L. Zllei, J. Polimeni, B. Fischl, H. Liu, and R. Buckner, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *J Neurophysiol*, 2011.

[4] F. G. Ashby, "An introduction to fmri," in *An introduction to model-based cognitive neuroscience*. Springer, 2015, pp. 99–112.

[5] I. Bankman, *Handbook of Medical Image Processing andAnalysis*. SanDiego, CA, USA: Academic Press, 2008.

[6] K. FRISTON, J. ASHBURNER, S. KIEBEL, T. NICHOLS, and W. PENNY, *Statistical Parametric Mapping The Analysis of Functional Brain Images*. Academic Press, 2007.

[7] J. Ashburner and K. J. Friston, "Nonlinear spatial normalization using basis functions," *Human brain mapping*, vol. 7, no. 4, pp. 254–266, 1999.

[8] A. COLLIGNON, F. MAES, D. DELAERE, D. VANDERMEULEN, P. SUETENS, and G. MARCHAL, "Automated multi-modality image registration based on information theory."

[9] R. Frackowiak, K. Friston, C. Frith, R. Dolan, C. Price, S. Zeki, J. Ashburner, and W. Penny, *Human Brain Function*, 2nd ed. Academic Press, 2003. [Online]. Available: http://www.fil.ion.ucl.ac.uk/spm/doc/books/hbf2/

[10] B. D. Haan and C. Rorden, "An introduction to functional mri," 2004.

[11] P. Thevenaz, T. Blu, and M. Unser, "Interpolation revisited [medical images application]," *IEEE Transactions on Medical Imaging*, vol. 19, no. 7, pp. 739–758, 2000.

[12] P. Horst, "Generalized canonical correlations and their applications to experimental data." *Journal of clinical psychology*, vol. 17, pp. 331–47, 1961.

[13] J. D. Carroll, "Generalization of canonical correlation analysis to three or more sets of variables," *Proceedings of the 76th annual convention of the American Psychological Association*, vol. 3, p. 227228, 1968.

[14] J. R. KETTENRING, "Canonical analysis of several sets of variables," *Biometrika*, vol. 58, no. 3, pp. 433–451, 12 1971. [Online]. Available: https://doi.org/10.1093/biomet/58.3.433

[15] G. W. Stewart and J. guang Sun, *Matrix Perturbation Theory*. Academic Press, 1990.

[16] B. Li, G. Zhou, and A. Cichocki, "Two efficient algorithms for approximately orthogonal nonnegative matrix factorization," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 843–846, 2015.