TECHNICAL UNIVERSITY OF CRETE
ELECTRICAL AND COMPUTER ENGINEERING DEPARTMENT
TELECOMMUNICATIONS DIVISION

Practical Channel Estimation for Reconfigurable Reflection Surfaces in Next Generation Wireless Networks

by

Georgios Kotridis

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

ELECTRICAL AND COMPUTER ENGINEERING

October 15, 2021

THESIS COMMITTEE

Professor Aggelos Bletsas, *Thesis Supervisor*
Professor George N. Karystinos
Associate Professor Vasileios Samoladas

Contents

List of Figures

Abstract

After the introduction of the Multiple Input Multiple Output and mmWave technologies to the communication world, great research interest has been recently attracted to find new ways of meeting the need for higher data rates that are both energy efficient and have low hardware cost. Reconfigurable Intelligent Surface (RIS) is a hardware technology composed of passive, software controlled metamaterials with reconfigurable scattering properties, meaning that the right adjustments can lead to constructive addition of EM waves at the receiver's end. Due to the passive nature of the elements there can be no signal processing of the incoming signals between the transmitter and the receiver, leading to challenging channel estimation. This work first studies prior art in the literature regarding channel estimation for MIMO systems using the zero-mean error Linear Minimum Mean Square Estimator (LMMSE), where the estimator is an affine transformation of the received signal. It offers in-depth derivation of both the estimator and the MSE formulas, which can be applied to any MIMO system including the RIS case for Rician fading, proven first theoretically and then followed by simulation results. We then use the estimation of all the channel coefficients as input-data to a novel algorithm which computes the optimal element-configuration in $\mathcal{O}(M\log(M))$ time, where M is the number of RIS elements. Simulations indicate that after the algorithm's application, the average power improvement reaches 12 dB gain for $M = 6162$, while the impact of error for the same value of M may lead to a gain loss in the order of 3 dB.

Chapter 1

Introduction

Wireless communications have become essential to our everyday lives. The overgrowing need for bigger data transfer and the ubiquitous accessibility to the Internet, call for higher and robust data transmission rates. Despite the remarkable efforts in upgrading both the hardware and the software domain of wireless networks by introducing various cutting-edge technologies such as Mulitple Input Mulitple Output (MIMO) and mmWave communications in the current 5-th Generation (5G), the problems of hardware cost and energy consumption still remain unresolved.

An overlooked resource throughout the literature that can be exploited to overcome these issues is the environment between the transceiver link, which plays a passive role to data transmission. Reconfigurable Intelligent Surface (RIS) is a novel hardware technology that comes to fill this gap by real-time, software-defined control of the wireless environment to improve communication performance. RIS is generally composed of engineered materials which have scattering properties and can be reconfigured. In this work we simulate a RIS assisted wireless network by using a total of M passive, ultra low-cost, wireless and batteryless Radio Frequency Identification (RFID) tags as elements with K possible states, with each state corresponding to a different phase-shift and if properly adjusted, they can lead to a constructive signal at the receiver.

The lack of active components on a RIS makes it less energy consuming than a traditional Amplify-and-Forward (AF) relay transceiver while keeping the hardware cost to a minimum. However, the question that naturally arises is, will this translate to a lower gain at the receiver's end? If the gains turn out to be comparable, what are the prerequisites for a sufficient Channel State Information (CSI) considering a large number of elements? Is there an algorithm we can use to find the optimal element-state configuration which maximizes the received signal's strength before the environment changes and how will a non-ideal CSI impact the performance of the algorithm?

Previous research on RIS assumes mainly perfect CSI. To tackle the problem of fast channel estimation we propose the Linear Minimum Mean Square Error (LMMSE) estimator with zero-mean channel estimation error, which is an affine function of the received signal's measurements. We show by simulations that for static environments the estimator responds well assuming Rice fading with a strong Line of Sights (LoS) deterministic component. We then use the estimated channel coefficients as the input of a novel algorithm, which can compute the optimal RIS configuration with $\mathcal{O}(M\log(M))$ complexity instead of an exponential $\mathcal{O}(K^M)$ that the exhaustive approach requires. Finally, simulations indicate that an insufficient channel estimation may lead to non-negligible gain loss reaching up to 3dB.

Notation: $\mathbf{0}_N$ denotes the all-zeros vector. The phase of complex number z is denoted as $\angle z$, while $\Re\{z\}$ denotes the real part of z. The distribution of a proper complex Gaussian $N \times 1$ vector $\boldsymbol{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$ is denoted by $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{\Sigma}) \triangleq \frac{1}{\pi^N \det(\Sigma)} \mathrm{e}^{-(\boldsymbol{x}-\boldsymbol{\mu})^H \mathbf{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$; the special case of a circularly symmetric complex Gaussian $N \times 1$ vector corresponds by definition to $\mathcal{CN}(\mathbf{0}_N, \mathbf{\Sigma})$; expectation of function $\mathrm{g}(\cdot)$ of random variable x is denoted by $\mathbb{E}[\mathrm{g}(x)]$. Finally, the traces of a matrix is denoted by $\mathrm{tr}(\cdot)$ while the Kronecker Product between two matrices is denoted by the symbol $\otimes$.

Chapter 2

Prior Art

2.1 Work in [1]

2.1.1 LMMSE estimator

We begin by defining the system model as:

$$X = \sqrt{\frac{\rho}{M}}\, S\, H + V \tag{2.1}$$

where $X \in \mathbb{C}^{T\times N}$ is the received symbols matrix, $S \in \mathbb{C}^{T\times M}$ is the transmitted symbols matrix, $H \in \mathbb{C}^{M\times N}$ is the channel matrix and $V \in \mathbb{C}^{T\times N}$ is the thermal noise matrix. Dimensions M and N refer to the number of transmit and receive antennas respectively, while T is the channel coherence time interval, counted in symbol periods, within which the channel coefficients remain invariant.

Given that we can only transmit T symbols in total before the channel coefficients' values change, we choose a fraction of them, say T_τ, to be pilot symbols in order to find the estimation of matrix H, while the rest of them are devoted to data transmission.
We henceforth define the pilot system model as:

$$X_\tau = c\, S_\tau\, H + V_\tau \tag{2.2}$$

where $c = \sqrt{\frac{\rho_\tau}{M}}$ is a known constant, $X_\tau \in \mathbb{C}^{T_\tau\times N}$ is the received pilot symbol-signal, $S_\tau \in \mathbb{C}^{T_\tau\times M}$ is the transmitted pilot symbol-signal and $V_\tau \in \mathbb{C}^{T_\tau\times N}$. The dimension T_τ now stands for the pilot-symbols time interval.

We can simplify the process of estimating H, without loss of generality, by assuming that H_{ij}, V_{kl} i.i.d. $\sim \mathcal{CN}\,(0,1)$ for any $i \in \{1,...,M\}$, $j \in \{1,...,N\}$, $k \in \{1,...,T_\tau\}$, $l \in \{1,...,N\}$, while also assuming that $\operatorname{tr}\left(S_\tau S_\tau^{\mathrm{H}}\right) = MT_\tau$ for normalization purposes.

A few fundamental questions naturally arise. How many of the available symbols should be devoted to training and how many to data? On the one hand, increasing the pilot symbols leads to a more accurate estimation of the coefficients, while on the other hand it leaves less room for communication. In which way does the choice of S_τ affect the derivation of the MMSE? The rest of the section is devoted to answering the latter in a supplementary way to [1].

We wish to find $\hat{H}$ such that $\mathbb{E}[\|H - \hat{H}\|_F^2]$ is minimized, when $\hat{H}$ is a linear (affine) function of the measurements, i.e. $\hat{H} = AX_\tau + B$, where $A \in \mathbb{C}^{M\times T_\tau}$, $B \in \mathbb{C}^{M\times N}$. Subsequently, finding the right A and B is equivalent to estimating $\hat{H}$, since X_τ is known to the receiver.

For the purpose of simplifying the calculations that follow, we start by estimating an arbitrary column of H, say $H_{:,i} = \mathbf{h}$, $i \in \{1, ..., N\}$. By modifying accordingly the rest of the quantities in (2), we get the following system model:

$$\mathbf{x} = c\, S_\tau\, \mathbf{h} + \mathbf{v} \tag{2.3}$$

where $\mathbf{x} \in \mathbb{C}^{T_\tau\times 1}$, $\mathbf{h} \in \mathbb{C}^{M\times 1}$, $\mathbf{v} \in \mathbb{C}^{T_\tau\times 1}$.

Hence, we now try to find $\hat{\mathbf{h}}_{\text{LMMSE}} = \underset{\hat{\mathbf{h}}=A\,\mathbf{x}+b}{\operatorname{argmin}}\{\mathbb{E}[\|\mathbf{h} - \hat{\mathbf{h}}\|^2]\}$. A and $\mathbf{b}$ can be computed in closed form as (detailed proof can be found in the Appendix):

$$A = C_{\mathbf{hx}}\, C_{\mathbf{x}}^{-1} \tag{2.4}$$

$$\mathbf{b} = \mu_{\mathbf{h}} - A\,\mu_{\mathbf{x}} \tag{2.5}$$

where $C_{\mathbf{hx}} = \mathbb{E}[(\mathbf{h} - \mu_{\mathbf{h}})(\mathbf{x} - \mu_{\mathbf{x}})^{\mathrm{H}}]$ is the cross-covariance matrix between $\mathbf{h}$ and $\mathbf{x}$ and $C_{\mathbf{x}} = \mathbb{E}[(\mathbf{x} - \mu_{\mathbf{x}})(\mathbf{x} - \mu_{\mathbf{x}})^{\mathrm{H}}]$ is the covariance matrix of $\mathbf{x}$.

Due to the fact that the entries of H and V are independent (and as a consequence uncorrelated), we obtain:

$$C_{\mathbf{hv}} = C_{\mathbf{vh}}^T = \mathbf{0}_{M\times T_\tau} \tag{2.6}$$

while for the mean of $\mathbf{x}$, we derive:

$$\mu_{\mathbf{x}} = \mathbb{E}[\mathbf{x}] = c\, S_\tau \mathbb{E}[\mathbf{h}] + \mathbb{E}[\mathbf{v}] = c\, S_\tau\, \mu_{\mathbf{h}} + \mu_{\mathbf{v}} = \mathbf{0}_{T_\tau\times 1} \tag{2.7}$$

Having the relations (2.6) and (2.7) in mind, $C_{\mathbf{x}}$ and $C_{\mathbf{hx}}$ can be expressed as follows:

$$
\begin{aligned}
C_{\mathbf{x}} &= \mathbb{E}[(\mathbf{x} - \mu_{\mathbf{x}})\,(\mathbf{x} - \mu_{\mathbf{x}})^{\mathrm{H}}] \\
&= \mathbb{E}[\mathbf{x}\,\mathbf{x}^{\mathrm{H}}] \\
&= \mathbb{E}[(c\,S_\tau\,\mathbf{h} + \mathbf{v})\left(c\,\mathbf{h}^{\mathrm{H}}\,S_\tau^{\mathrm{H}} + \mathbf{v}^{\mathrm{H}}\right)] \\
&= \mathbb{E}[c^2\,S_\tau\,\mathbf{h}\,\mathbf{h}^{\mathrm{H}} S_\tau^{\mathrm{H}} + c\,S_\tau\,\mathbf{h}\,\mathbf{v}^{\mathrm{H}} + c\,\mathbf{v}\,\mathbf{h}^{\mathrm{H}} S_\tau^{\mathrm{H}} + \mathbf{v}\,\mathbf{v}^{H}] \\
&= c^2\,S_\tau\,\mathbb{E}[\mathbf{h}\,\mathbf{h}^{\mathrm{H}}]\,S_\tau^{\mathrm{H}} + c\,S_\tau\,\mathbb{E}[\mathbf{h}\,\mathbf{v}^{\mathrm{H}}] + c\,\mathbb{E}[\mathbf{v}\,\mathbf{h}^{\mathrm{H}}]\,S_\tau + C_{\mathbf{v}} \\
&= c^2\,S_\tau\,C_{\mathbf{h}}\,S_\tau^{\mathrm{H}} + c\,S_\tau\,C_{\mathbf{hv}} + c\,C_{\mathbf{vh}}\,S_\tau + C_{\mathbf{v}} \\
&= c^2\,S_\tau\,\mathbf{I_M}\,S_\tau^{\mathrm{H}} + \mathbf{0}_{T_\tau \times T_\tau} + \mathbf{0}_{T_\tau \times T_\tau} + \mathbf{I}_{T_\tau} \\
&= \frac{\rho_\tau}{M}\,S_\tau\,S_\tau^{\mathrm{H}} + \mathbf{I}_{T_\tau}
\end{aligned} \tag{2.8}
$$

$$
\begin{aligned}
C_{\mathbf{hx}} &= \mathbb{E}[(\mathbf{h} - \mu_{\mathbf{h}})\,(\mathbf{x} - \mu_{\mathbf{x}})^{\mathrm{H}}] \\
&= \mathbb{E}[\mathbf{h}\,\mathbf{x}^{\mathrm{H}}] \\
&= \mathbb{E}[\mathbf{h}\left(c\,\mathbf{h}^{\mathrm{H}} S_\tau^{\mathrm{H}} + \mathbf{v}^{\mathrm{H}}\right)] \\
&= \mathbb{E}[c\,\mathbf{h}\,\mathbf{h}^{\mathrm{H}} S_\tau^{\mathrm{H}} + \mathbf{h}\,\mathbf{v}^{\mathrm{H}}] \\
&= c\,C_{\mathbf{h}}\,S_\tau^{\mathrm{H}} + C_{\mathbf{hv}} \\
&= c\,\mathbf{I_M}\,S_\tau^{\mathrm{H}} + \mathbf{0}_{M \times T_\tau} \\
&= \sqrt{\frac{\rho_\tau}{M}}\,S_\tau^{\mathrm{H}}
\end{aligned} \tag{2.9}
$$

Substituting (2.8) and (2.9) to A and $\mathbf{b}$ yields:

$$
A = C_{\mathbf{hx}}\,C_{\mathbf{x}}^{-1} = \sqrt{\frac{\rho_\tau}{M}}\,S_\tau^{H}\left(\frac{\rho_\tau}{M}\,S_\tau\,S_\tau^{\mathrm{H}} + \mathbf{I}_{T_\tau}\right)^{-1} \tag{2.10}
$$

Matrix $\frac{\rho_\tau}{M} S_\tau\,S_\tau^{\mathrm{H}}$ is Positive Semi Definite (PSD) since $\mathbf{x}^{\mathrm{H}} S_\tau S_\tau^{\mathrm{H}} \mathbf{x} = \| S_\tau^{\mathrm{H}}\,\mathbf{x}\|^2 \geq 0$ for any $\mathbf{x} \in \mathbb{C}^{T_\tau}$. Adding $\mathbf{I}_{T_\tau}$ to the PSD matrix, the sum is Positive Definite (PD), which means that the final matrix is also non singular, i.e. invertible.

Due to (2.7) we obtain:

$$
\mathbf{b} = \mu_{\mathbf{h}} - A\,\mu_{\mathbf{x}} = \mathbf{0}_{T_\tau \times 1} \tag{2.11}
$$

Hence the LMMSE estimate of $\mathbf{h}$ can now be computed:

$$\begin{aligned}
\hat{\mathbf{h}} &= A\,\mathbf{x} + \mathbf{b} \\
&= \sqrt{\frac{\rho_\tau}{M}}\, S_\tau^{\mathrm{H}} \left(\frac{\rho_\tau}{M}\, S_\tau\, S_\tau^{\mathrm{H}} + \mathbf{I}_{T_\tau}\right)^{-1} \mathbf{x} \\
&= \sqrt{\frac{\rho_\tau}{M}} \left(\frac{\rho_\tau}{M} \left(S_\tau\, S_\tau^{\mathrm{H}} + \frac{M}{\rho_\tau}\mathbf{I}_{T_\tau}\right)\right)^{-1} \mathbf{x} \\
&= \sqrt{\frac{\rho_\tau}{M}}\, \frac{M}{\rho_\tau} \left(S_\tau\, S_\tau^{\mathrm{H}} + \frac{M}{\rho_\tau}\mathbf{I}_{T_\tau}\right)^{-1} \mathbf{x} \\
&= \sqrt{\frac{M}{\rho_\tau}} S_\tau^{\mathrm{H}} \left(S_\tau\, S_\tau^{\mathrm{H}} + \frac{M}{\rho_\tau}\mathbf{I}_{T_\tau}\right)^{-1} X_{:,i}
\end{aligned} \tag{2.12}$$

Following the same series of steps, one can compute the LMMSE estimate for every column of H. Thus, the LMMSE estimate of H can be expressed by simply replacing $X_{:,i}$ with the whole matrix X_τ:

$$\hat{H} = \sqrt{\frac{M}{\rho_\tau}}\, S_\tau^{\mathrm{H}} \left(S_\tau\, S_\tau^{\mathrm{H}} + \frac{M}{\rho_\tau}\mathbf{I}_\tau\right)^{-1} X_\tau \tag{2.13}$$

In order to conclude to the same relation as the one in [1], we make use of the Push Through Identity:

$$U(\mathbf{I} + VU)^{-1} = (UV + \mathbf{I})^{-1} U \tag{2.14}$$

After applying it to (2.13), we obtain:

$$\hat{H} = \sqrt{\frac{M}{\rho_\tau}} \left(S_\tau^{\mathrm{H}}\, S_\tau + \frac{M}{\rho_\tau}\mathbf{I}_M\right)^{-1} S_\tau^{\mathrm{H}}\, X_\tau \tag{2.15}$$

Matrix $S_\tau^{\mathrm{H}}\, S_\tau + \frac{M}{\rho_\tau}\mathbf{I}_M$ is also invertible, since $\mathbf{x}^{\mathrm{H}}\, S_\tau^{\mathrm{H}}\, S_\tau\, \mathbf{x} = \|S_\tau\, \mathbf{x}\|^2 \geq 0$ is PSD and $\frac{M}{\rho_\tau}\mathbf{I}_M$ is PD, which leads to the sum of them being PD.

2.1.2 LMMSE

We will now try to find the value of the LMMSE. Firstly, we define the covariance matrices in a way that will simplify the computation's procedure:

$$R_{X_\tau} \triangleq \mathbb{E}[\mathrm{vec}(X_\tau)\, \mathrm{vec}(X_\tau)^{\mathrm{H}}] \tag{2.16}$$

$$R_{HX_\tau} \triangleq \mathbb{E}[\mathrm{vec}(H)\, \mathrm{vec}(X_\tau)^{\mathrm{H}}] \tag{2.17}$$

$$R_{X_\tau H} \triangleq \mathbb{E}[\mathrm{vec}(X_\tau)\,\mathrm{vec}(H)^{\mathrm{H}}] \tag{2.18}$$

$$R_{\tilde{H}} \triangleq \mathbb{E}[\mathrm{vec}(\tilde{H})\,\mathrm{vec}(\tilde{H})^{\mathrm{H}}] = R_H - R_{HX_\tau} R_{X_\tau}^{-1} R_{X_\tau H} \tag{2.19}$$

where $\tilde{H} = H - \hat{H}$ is the zero-mean channel estimation error. Consequently, the LMMSE can be written as:

$$\mathsf{LMMSE} = \sigma_{\tilde{H}}^2 = \frac{1}{NM}\mathrm{tr}\left(R_{\tilde{H}}\right) \tag{2.20}$$

Our goal is to minimize the linear MSE with respect to the training symbols matrix S_τ. To achieve that, we will first write the aforementioned relations in a more analytical way, while assuming that:

$$\mathrm{vec}(\mathrm{mtx}) = \mathrm{mtx}_v \tag{2.21}$$

to simplify the notation.

Moreover, due to the ubiquity of the Kronecker Product's application to the derivation that follows, we first list some fundamental properties/identities [3] for the reader's convenience:

$$A \otimes 0 = 0 \otimes A = 0 \tag{2.22}$$

$$\mathrm{vec}\left(AB\right) = \left(\mathbf{I}_{\mathrm{size}(B,2)} \otimes A\right) \mathrm{vec}\left(B\right) \tag{2.23}$$

$$\left(A \otimes B\right)^{-1} = A^{-1} \otimes B^{-1} \tag{2.24}$$

$$\left(A \otimes B\right)^{\mathrm{H}} = A^{\mathrm{H}} \otimes B^{\mathrm{H}} \tag{2.25}$$

The Woodbury Matrix Identity:

$$\mathbf{I} - V\left(\mathbf{I} + UV\right)^{-1} U = \left(\mathbf{I} + VU\right)^{-1} \tag{2.26}$$

The vectorized covariance matrices can be written as:

$$
\begin{aligned}
R_{HX_\tau} &= \mathbb{E}\left[\mathrm{vec}(H)\mathrm{vec}(X_\tau)^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[H_v\left(c\,S_\tau\,H + V_\tau\right)_v^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[H_v\left(\left(\mathbf{I}_N \otimes c\,S_\tau\right)H_v\right)^{\mathrm{H}} + H_v V_{\tau_v}^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[H_v H_v^{\mathrm{H}}\left(\mathbf{I}_N \otimes c\,S_\tau\right)^{\mathrm{H}}\right] + \mathbf{E}\left[H_v V_{\tau_v}^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[H_v H_v^{\mathrm{H}}\right]\left(\mathbf{I}_N \otimes c\,S_\tau^{\mathrm{H}}\right) + \mathbf{0}_{(MN)\times(T_\tau N)} \\
&= \mathbf{I}_{MN}\left(\mathbf{I}_N \otimes c\,S_\tau^{\mathrm{H}}\right) \\
&= c\left(\mathbf{I}_N \otimes S_\tau^{\mathrm{H}}\right) \\
&= \sqrt{\frac{\rho_\tau}{M}}\left(\mathbf{I}_N \otimes S_\tau^{\mathrm{H}}\right)
\end{aligned} \tag{2.27}
$$

$$
\begin{aligned}
R_{X_\tau} &= \mathbb{E}\left[\mathrm{vec}(X_\tau)\mathrm{vec}(X_\tau)^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[\left(c\,S_\tau H + V_\tau\right)_v\left(c\,S_\tau H + V_\tau\right)_v^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[\left(\left(\mathbf{I}_N \otimes c\,S_\tau\right)H_v + V_{\tau_v}\right)\left(\left(\mathbf{I}_N \otimes c\,S_\tau\right)H_v + V_{\tau_v}\right)^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[\left(\left(\mathbf{I}_N \otimes c\,S_\tau\right)H_v + V_{\tau_v}\right)\left(H_v^{\mathrm{H}}\left(\mathbf{I}_N \otimes c\,S_\tau^{\mathrm{H}}\right) + V_{\tau_v}^{\mathrm{H}}\right)\right] \\
&= \mathbb{E}\left[\left(\mathbf{I}_N \otimes c\,S_\tau\right)H_v H_v^{\mathrm{H}}\left(\mathbf{I}_N \otimes c\,S_\tau^{\mathrm{H}}\right) + \left(\mathbf{I}_N \otimes c\,S_\tau\right)H_v V_{\tau_v}^{\mathrm{H}} + V_{\tau_v}H_v^{\mathrm{H}}\left(\mathbf{I}_N \otimes c\,S_\tau^{\mathrm{H}}\right) + V_{\tau_v}V_{\tau_v}^{\mathrm{H}}\right] \\
&= \left(\mathbf{I}_N \otimes c\,S_\tau\right)\mathbb{E}\left[H_v H_v^{\mathrm{H}}\right]\left(\mathbf{I}_N \otimes c\,S_\tau^{\mathrm{H}}\right) + \left(\mathbf{I}_N \otimes c\,S_\tau\right)\mathbb{E}\left[H_v V_{\tau_v}^{\mathrm{H}}\right] + \\
&\quad + \mathbb{E}\left[V_{\tau_v}H_v^{\mathrm{H}}\right]\left(\mathbf{I}_N \otimes c\,S_\tau^{\mathrm{H}}\right) + \mathbb{E}\left[V_{\tau_v}V_{\tau_v}^{\mathrm{H}}\right] \\
&= \left(\mathbf{I}_N \otimes c\,S_\tau\right)\mathbf{I}_{MN}\left(\mathbf{I}_N \otimes c\,S_\tau^{\mathrm{H}}\right) + \mathbf{0}_{(T_\tau N)\times(T_\tau N)} + \mathbf{0}_{(T_\tau N)\times(T_\tau N)} + \mathbf{I}_{\tau N} \\
&= c^2\left(\mathbf{I}_N \otimes S_\tau\right)\left(\mathbf{I}_N \otimes S_\tau^{\mathrm{H}}\right) + \mathbf{I}_{\tau N} \\
&= \frac{\rho_\tau}{M}\left(\mathbf{I}_N \otimes S_\tau\right)\left(\mathbf{I}_N \otimes S_\tau^{\mathrm{H}}\right) + \mathbf{I}_\tau \otimes \mathbf{I}_N
\end{aligned} \tag{2.28}
$$

$$
\begin{aligned}
R_{X_\tau H} &= \mathbb{E}\left[\mathrm{vec}(X_\tau)\mathrm{vec}(H)^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[\left(c\,S_\tau H + V_\tau\right)_v H_v^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[\left(\left(\mathbf{I}_N \otimes c\,S_\tau\right)H_v H_v^{\mathrm{H}}\right) + V_{\tau_v}H_v^{\mathrm{H}}\right] \\
&= \left(\mathbf{I}_N \otimes c\,S_\tau\right)\mathbb{E}\left[H_v H_v^{\mathrm{H}}\right] + \mathbb{E}\left[V_{\tau_v}H_v^{\mathrm{H}}\right] \\
&= c\left(\mathbf{I}_N \otimes S_\tau\right)\mathbf{I}_{MN} + \mathbf{0}_{(T_\tau N)\times(MN)} \\
&= \sqrt{\frac{\rho_\tau}{M}}\left(\mathbf{I}_N \otimes S_\tau\right)
\end{aligned} \tag{2.29}
$$

$$
\begin{aligned}
R_{\tilde{H}} &= R_H - R_{HX_\tau} R_{X_\tau}^{-1} R_{X_\tau H} \\
&= \mathbf{I}_M \otimes \mathbf{I}_N - \sqrt{\frac{\rho_\tau}{M}} \left(\mathbf{I}_N \otimes S_\tau^{\mathrm{H}}\right) \left(\frac{\rho_\tau}{M} \left(\mathbf{I}_N \otimes S_\tau\right) \left(\mathbf{I}_N \otimes S_\tau^{\mathrm{H}}\right) + \mathbf{I}_\tau \otimes \mathbf{I}_N\right)^{-1} \sqrt{\frac{\rho_\tau}{M}} \left(\mathbf{I}_N \otimes S_\tau\right) \\
&= \left(\frac{\rho_\tau}{M} \left(\mathbf{I}_N \otimes S_\tau^{\mathrm{H}}\right) \left(\mathbf{I}_N \otimes S_\tau\right) + \left(\mathbf{I}_M \otimes \mathbf{I}_N\right)\right)^{-1} \\
&= \left(\frac{\rho_\tau}{M} \left(\mathbf{I}_N \otimes S_\tau^{\mathrm{H}}\right) \left(\mathbf{I}_N \otimes S_\tau\right) + \left(\mathbf{I}_N \otimes \mathbf{I}_M\right)\right)^{-1} \\
&= \left(\frac{\rho_\tau}{M} \left(\mathbf{I}_N \mathbf{I}_N\right) \otimes \left(S_\tau^{\mathrm{H}} S_\tau\right) + \left(\mathbf{I}_N \otimes \mathbf{I}_M\right)\right)^{-1} \\
&= \left(\mathbf{I}_N \otimes \left(\frac{\rho_\tau}{M} S_\tau^{\mathrm{H}} S_\tau + \mathbf{I}_M\right)\right)^{-1} \\
&= \mathbf{I}_N \otimes \left(\frac{\rho_\tau}{M} S_\tau^{\mathrm{H}} S_\tau + \mathbf{I}_M\right)^{-1} \qquad (2.30)
\end{aligned}
$$

Let us define a new matrix as $W = \left(\frac{\rho_\tau}{M} S_\tau^{\mathrm{H}} S_\tau + \mathbf{I}_M\right)$. As we have already discussed, this matrix is PD and hence its eigenvalues are positive. Having also in mind that every matrix can be uniquely described by its Eigenvalue Decomposition: $W = Q \Lambda_{W_i} Q^{-1}$ and that $\mathrm{tr}(W) = \mathrm{tr}(\Lambda_{W_i}) = \sum_{i=1}^{M} \lambda_{W_i}$, where Λ_{W_i} is a diagonal matrix with the eigenvalues of W (λ_{W_i}) as its non-zero elements, we can derive the following for W^{-1}:

$$
\begin{aligned}
\mathrm{tr}(W^{-1}) &= \mathrm{tr}((Q \Lambda_{W_i} Q^{-1})^{-1}) \\
&= \mathrm{tr}(Q^{-1} (Q \Lambda_{W_i})^{-1}) \\
&= \mathrm{tr}(Q \Lambda_{W_i}^{-1} Q^{-1}) \\
&= \sum_{i=1}^{M} \frac{1}{\lambda_{W_i}} \qquad (2.31)
\end{aligned}
$$

After making use of the property: $\Lambda_{A+\mathbf{I}} = \Lambda_A + \Lambda_{\mathbf{I}}$, equation (2.31) becomes:

$$
\mathrm{tr}(W^{-1}) = \sum_{i=1}^{M} \frac{1}{1 + \frac{\rho_\tau}{M} \lambda_{(S^{\mathrm{H}} S)_i}} \qquad (2.32)
$$

where $\lambda_{(S^{\mathrm{H}} S)_i}$ is the i-th eigenvalue of the product $S_\tau^{\mathrm{H}} S_\tau$ under the constraint: $\sum_{i=1}^{M} \lambda_i \leq M T_\tau$.

As it was previously mentioned, our goal is to minimize the MSE with respect to S_τ. Looking at equation (2.20), this translates to minimizing the equation (2.32) with respect to the eigenvalues of $S_\tau^{\mathrm{H}} S_\tau$. This optimization problem (considering also the constraints) is solved by choosing $\lambda_1 = \lambda_2 = ... = \lambda_M = T_\tau$ [1], which leads to:

$$S_\tau^{\mathrm{H}} S_\tau = T_\tau \mathbf{I}_M \tag{2.33}$$

Substituting equation (2.33) to (2.30) we obtain:

$$\begin{aligned} R_{\tilde{H}} &= \mathbf{I}_N \otimes \left(\frac{\rho_\tau}{M} S_\tau^{\mathrm{H}} S_\tau + \mathbf{I}_M \right)^{-1} \\ &= \mathbf{I}_N \otimes \left(\frac{\rho_\tau}{M} T_\tau \mathbf{I}_M + \mathbf{I}_M \right)^{-1} \\ &= \mathbf{I}_N \otimes \left(\mathbf{I}_M \left(\frac{\rho_\tau}{M} T_\tau + 1 \right) \right)^{-1} \\ &= \mathbf{I}_N \otimes \mathbf{I}_M \left(\frac{1}{\frac{\rho_\tau}{M} T_\tau + 1} \right) \end{aligned} \tag{2.34}$$

while the MMSE from (2.20) becomes:

$$\begin{aligned} \mathrm{MSE} &= \frac{1}{NM} \mathrm{tr} \left(R_{\tilde{H}} \right) \\ &= \frac{1}{NM} \mathrm{tr} \left(\mathbf{I}_N \otimes \mathbf{I}_M \left(\frac{1}{\frac{\rho_\tau}{M} T_\tau + 1} \right) \right) \\ &= \left(\frac{1}{NM} \right) \left(\frac{NM}{\frac{\rho_\tau}{M} T_\tau + 1} \right) \\ &= \frac{1}{1 + \frac{\rho_\tau}{M} T_\tau} \end{aligned} \tag{2.35}$$

2.2 Work in [2]

Lozano's system differs from Hassibi's, in that we now have to model a cluster within the system. Since the out-of-cluster signals do not have a fixed power value, they cannot be simply merged with the noise term.

We begin by defining the system model in a way that will include the out-of-cluster interference, to better describe a wireless network:

$$y_n = \sqrt{\mathsf{SINR}_n} \sum_{k=1}^{K} \sqrt{g_{nk}}\, h_{nk}\, x_k + z'_n \tag{2.36}$$

where $n \in \{1, ..., N\}, k \in \{1, ..., K\}$, $x_k \in \mathbb{C}$ is the transmitted symbol's signal from k-th antenna with unit variance, g_{nk} is the normalized channel power gain, SINR_n[1] is the Signal-to-Interference-plus-Noise ratio of n-th receiver and $h_{nk}, z'_n \sim \mathcal{CN}(0, 1)$ are the normalized channel coefficient from transmitter k to receiver n and the thermal noise at the receiver's end respectively. It follows that $y_n \in \mathbb{C}$ is the received symbol's signal at n-th receiver's antenna.

While at first glance this model may seem more complicated than the one from the previous section, in actuality not only is it far simpler to analyze (since we now deal with scalar values instead of vectors and matrices), but we can also modify it to resemble the form of the latter. Thus, the proofs that follow will not deviate conceptually far from what has already been derived.

We assume to have L symbols available before the channel coefficients' values change and we select αL of them to be pilot symbols in order to estimate the channel, where $\alpha \in (0, 1)$. Since there are K transmitters, only $T = \frac{\alpha L}{K}$ pilot symbols are needed for each channel coefficient h_{nk}. Naturally, the inequality $\alpha L \geq K$ is satisfied to ensure that at least one pilot symbol will be used for the estimation of each h_{nk}.

If we shift our focus to an arbitrary tranceiver pair, say n-th receiver and k-th transmitter, and incorporate the pilot symbols to the system model, we obtain the following modification of (2.36):

$$\mathbf{y} = c\, h\, \mathbf{x} + \mathbf{z} \tag{2.37}$$

where $\mathbf{y}, \mathbf{x}, \mathbf{z} \in \mathbb{C}^T$, $h = h_{nk}$ and $c = \sqrt{\mathsf{SINR}_n\, g_{nk}}$.

One can notice that the pilot-symbol model is closely related to the one we used to describe Hassibi's system, but will the LMMSE estimator and the LMMSE relations reflect this similarity? The answer lies in the derivation below.

[1] $\frac{1}{\mathsf{SINR}_n} = \frac{1}{\mathrm{SIR}_n} + \frac{1}{\mathrm{SNR}_n}$ where SIR_n is the Signal-to-Interference ratio.

2.2.1 LMMSE estimator

Following the steps of the previous section, we wish to find an estimation of h ($\hat{h}$) such that $\mathbb{E}[|h-\hat{h}|^2]$ is minimized, when $\hat{h}$ is a linear (affine) function of the measurements, i.e. $\hat{h} = \mathbf{a}^{\mathrm{H}}\,\mathbf{y} + b$, where $\mathbf{a} \in \mathbb{C}^T$, $b \in \mathbb{C}$. Subsequently, finding the right $\mathbf{a}$ and b is equivalent to finding $\hat{h}$, since $\mathbf{y}$ is known to the receiver.
According to [4], [5], $\mathbf{a}$ and b can be computed in closed form as follows:

$$\mathbf{a} = C_{\mathbf{y}}^{-1} C_{\mathbf{y}h} \tag{2.38}$$

and

$$b = \mu_h - \mathbf{a}^{\mathrm{H}} \mu_{\mathbf{y}} \tag{2.39}$$

where $C_{\mathbf{y}h} = \mathbb{E}[(\mathbf{y}-\mu_{\mathbf{y}})(h-\mu_h)^*]$ is the cross covariance matrix between $\mathbf{y}$ and h, while $C_{\mathbf{y}} = \mathbb{E}[(\mathbf{y}-\mu_{\mathbf{y}})(\mathbf{y}-\mu_{\mathbf{y}})^{\mathrm{H}}]$ is the covariance matrix of $\mathbf{y}$.

As for the covariance matrix of the channel coefficient and the noise vector, since h and the entries of $\mathbf{z}$ are independent, they are also uncorrelated, which leads to:

$$C_{h\mathbf{z}} = C_{\mathbf{z}h}^T = \mathbf{0}_{T\times 1} \tag{2.40}$$

while for the mean of $\mathbf{y}$, we derive:

$$\mu_{\mathbf{y}} = \mathbb{E}[\mathbf{y}] = c\,\mathbf{x}\,\mathbb{E}[h] + \mathbb{E}[\mathbf{z}] = c\,\mathbf{x}\,\mu_h + \mu_{\mathbf{z}} = \mathbf{0}_{W\times 1} \tag{2.41}$$

Having the equations (2.40) and (2.41) in mind, $C_{\mathbf{y}}$ and $C_{\mathbf{y}h}$ can be calculated as follows:

$$
\begin{aligned}
C_{\mathbf{y}} &= \mathbb{E}[(\mathbf{y}-\mu_{\mathbf{y}})\,(\mathbf{y}-\mu_{\mathbf{y}})^{\mathrm{H}}] \\
&= \mathbb{E}[\mathbf{y}\,\mathbf{y}^{\mathrm{H}}] \\
&= \mathbb{E}[(c\,h\,\mathbf{x}+\mathbf{z})\left(c\,\mathbf{x}^{\mathrm{H}}h^{*}+\mathbf{z}^{\mathrm{H}}\right)] \\
&= \mathbb{E}[c^{2}h\,h^{*}\mathbf{x}\,\mathbf{x}^{\mathrm{H}}+c\,\mathbf{x}\,h\,\mathbf{z}^{\mathrm{H}}+\mathbf{z}\,h^{*}c\,\mathbf{x}^{\mathrm{H}}+\mathbf{z}\,\mathbf{z}^{\mathrm{H}}] \\
&= c^{2}\,\mathbb{E}[h\,h^{*}]\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}}+c\,\mathbf{x}\,\mathbb{E}[h\,\mathbf{z}^{\mathrm{H}}]+c\,\mathbb{E}[\mathbf{z}\,h^{*}]\,\mathbf{x}^{\mathrm{H}}+\mathbb{E}[\mathbf{z}\,\mathbf{z}^{\mathrm{H}}] \\
&= c^{2}\,\mathbb{E}[|h|^{2}]\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}}+c\,\mathbf{x}C_{h\mathbf{z}}+c\,C_{\mathbf{z}h}\,\mathbf{x}^{\mathrm{H}}+C_{\mathbf{z}} \\
&= c^{2}\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}}+\mathbf{0}_{T\times T}+\mathbf{0}_{T\times T}+\mathbf{I}_{T} \\
&= \mathsf{SINR}_{n}\,g_{nk}\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}}+\mathbf{I}_{T}
\end{aligned} \tag{2.42}
$$

$$
\begin{aligned}
C_{\mathbf{y}h} &= \mathbb{E}[(\mathbf{y}-\mu_{\mathbf{y}})\,(h-\mu_{h})^{\mathrm{H}}] \\
&= \mathbb{E}[\mathbf{y}\,h^{*}] \\
&= \mathbb{E}[(c\,h\,\mathbf{x}+\mathbf{z})\,h^{*}] \\
&= \mathbb{E}[c\,h\,h^{*}\mathbf{x}+\mathbf{z}\,h^{*}] \\
&= c\,\mathbb{E}[h\,h^{*}]\,\mathbf{x}+\mathbb{E}[\mathbf{z}\,h^{*}] \\
&= c\,\mathbb{E}[|h|^{2}]\,\mathbf{x}+C_{\mathbf{z}h} \\
&= c\,\mathbf{x}+\mathbf{0}_{T\times 1} \\
&= \sqrt{\mathsf{SINR}_{n}\,g_{nk}}\,\mathbf{x}
\end{aligned} \tag{2.43}
$$

Substituting (2.42) and (2.43) to $\mathbf{a}$ leads to:

$$
\mathbf{a} = C_{\mathbf{y}}^{-1}C_{\mathbf{y}h} = \left(\mathsf{SINR}_{n}\,g_{nk}\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}}+\mathbf{I}_{T}\right)^{-1}\sqrt{\mathsf{SINR}_{n}\,g_{nk}}\,\mathbf{x} \tag{2.44}
$$

and as for b:

$$
b = \mu_{h}-\mathbf{a}^{\mathrm{H}}\,\mu_{\mathbf{y}} = 0-\mathbf{a}^{\mathrm{H}}\,\mathbf{0}_{T\times 1} = 0 \tag{2.45}
$$

Hence the LMMSE estimate of h can now be calculated as follows:

$$
\begin{aligned}
\hat{h} &= \mathbf{a}^{\mathrm{H}}\,\mathbf{y} + b \\
&= \left(\left(\mathsf{SINR}_n\, g_{nk}\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}} + \mathbf{I}_T\right)^{-1}\sqrt{\mathsf{SINR}_n\, g_{nk}}\,\mathbf{x}\right)^{\mathrm{H}}\mathbf{y} \\
&= \sqrt{\mathsf{SINR}_n\, g_{nk}}\,\mathbf{x}^{\mathrm{H}}\left(\left(\mathsf{SINR}_n\, g_{nk}\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}} + \mathbf{I}_T\right)^{-1}\right)^{\mathrm{H}}\mathbf{y} \\
&= \sqrt{\mathsf{SINR}_n\, g_{nk}}\,\mathbf{x}^{\mathrm{H}}\left(\left(\mathsf{SINR}_n\, g_{nk}\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}} + \mathbf{I}_T\right)^{\mathrm{H}}\right)^{-1}\mathbf{y} \\
&= \sqrt{\mathsf{SINR}_n\, g_{nk}}\,\mathbf{x}^{\mathrm{H}}\left(\mathsf{SINR}_n\, g_{nk}\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}} + \mathbf{I}_T\right)^{-1}\mathbf{y} \\
&\overset{(*)}{=} \sqrt{\mathsf{SINR}_n\, g_{nk}}\,\left(\mathsf{SINR}_n\, g_{nk}\,\mathbf{x}^{\mathrm{H}}\,\mathbf{x} + 1\right)^{-1}\mathbf{x}^{\mathrm{H}}\,\mathbf{y} \\
&\overset{(**)}{=} \frac{\sqrt{\mathsf{SINR}_n\, g_{nk}}\,\mathbf{x}^{\mathrm{H}}\,\mathbf{y}}{1 + T\,\mathsf{SINR}_n\, g_{nk}}
\end{aligned}
\tag{2.46}
$$

where at $(*)$ we used the Woodbury Matrix Identity Eq. (2.26) and at $(**)$ we used the optimal training symbols based on Hassibi's derivation.

2.2.2 LMMSE

We define the zero-mean channel estimation error as $\tilde{h} = h - \hat{h}$. Subsequently, the LMMSE is defined as:

$$
\mathrm{LMMSE} = \sigma^2_{\tilde{h}} = \sigma^2_h - C_{h\mathbf{y}} C_{\mathbf{y}}^{-1} C_{\mathbf{y}h} \tag{2.47}
$$

As also stated in the previous section, our goal is to minimize the linear MSE. To do that, we first need to compute the cross-covariance matrix $C_{h\mathbf{y}}$. Since the training symbols is also a variable $\mathbf{x}$ to the problem, we can further reduce the MSE.

$$
\begin{aligned}
C_{h\mathbf{y}} &= \mathbb{E}[(h - \mu_h)\,(\mathbf{y} - \mu_{\mathbf{y}})^{\mathrm{H}}] \\
&= \mathbb{E}[h\mathbf{y}^{\mathrm{H}}] \\
&= \mathbb{E}[h\left(c\,\mathbf{x}^{\mathrm{H}} h^* + \mathbf{z}^{\mathrm{H}}\right)] \\
&= \mathbb{E}[c\,h\,h^*\mathbf{x}^{\mathrm{H}} + h\mathbf{z}^{\mathrm{H}}] \\
&= c\,\mathbb{E}[h\,h^*]\mathbf{x}^{\mathrm{H}} + \mathbb{E}[h\mathbf{z}^{\mathrm{H}}] \\
&= c\,\mathbb{E}[|h|^2]\mathbf{x}^{\mathrm{H}} + C_{\mathbf{z}h} \\
&= c\,\mathbf{x}^{\mathrm{H}} + \mathbf{0}_{1\times T} \\
&= \sqrt{\mathsf{SINR}_n\, g_{nk}}\,\mathbf{x}^{\mathrm{H}}
\end{aligned}
\tag{2.48}
$$

We can now calculate the LMMSE:

$$\begin{aligned}
\sigma_{\tilde{h}}^2 &= \sigma_h^2 - C_{h\mathbf{y}} C_{\mathbf{y}}^{-1} C_{\mathbf{y}h} \\
&= 1 - c\,\mathbf{x}^{\mathrm{H}} \left(c^2\,\mathbf{x}\,\mathbf{x}^{\mathrm{H}} + \mathbf{I}_T\right)^{-1} c\mathbf{x} \\
&\overset{(*)}{=} \left(c\,\mathbf{x}^{\mathrm{H}} c\,\mathbf{x} + 1\right)^{-1} \\
&= \frac{1}{1 + c^2\,\mathbf{x}^{\mathrm{H}}\,\mathbf{x}} \\
&\overset{(**)}{=} \frac{1}{1 + \mathsf{SINR}_n\, g_{nk}\, T} \\
&= \frac{1}{1 + \mathsf{SINR}_n\, g_{nk}\, \frac{\alpha L}{K}}
\end{aligned} \tag{2.49}$$

where at $(*)$ we used the Woodbury Matrix Identity Eq. (2.26) and at $(**)$ we used the optimal training symbols based on Hassibi's derivation.
One can notice that the LMMSE estimate of h as well as the LMMSE Eq. (2.49) are highly correlated to the ones from the previous section, i.e. Eq. (2.35).

Chapter 3

RIS Channel Estimator

3.1 System Model

3.1.1 Channel Model

A source-destination link is assisted by an array of M tags/RIS elements. The following *large-scale* channel path-loss model is adopted [6]:

$$\mathsf{L}_{\mathrm{X}} \propto \left(\frac{\lambda}{4\pi d_0^{\mathrm{X}}}\right)^2 \left(\frac{d_0^{\mathrm{X}}}{d_{\mathrm{X}}}\right)^{v_{\mathrm{X}}}, \tag{3.1}$$

where $\mathrm{X} \in \{\mathrm{SD}, \mathrm{ST}_m, \mathrm{T}_m\mathrm{D}\}$ denotes the source-to-destination, source-to-tag m and tag m-to-destination, respectively; λ is the carrier wavelength, d_0^{X} is a reference distance, v_{X} is the path-loss exponent and d_{X} is the distance for link X.

Flat fading is assumed; complex channel coefficient h_{SD}, h_{ST_m} and $h_{\mathrm{T}_m\mathrm{D}}$ denotes the baseband channel coefficients for the source-destination, source-tag and tag-reader link, respectively. Due to strong line-of-sight (LoS) signals present in the problem, *small-scale* Rice flat fading channel model [6] is mainly adopted:[1]

$$h_{\mathrm{T}_m\mathrm{D}} \sim \mathcal{CN}\left(\sqrt{\frac{\kappa_{\mathrm{T}_m\mathrm{D}}}{\kappa_{\mathrm{T}_m\mathrm{D}}+1}}\,\sigma_{\mathrm{T}_m\mathrm{D}}, \frac{\sigma^2_{\mathrm{T}_m\mathrm{D}}}{\kappa_{\mathrm{T}_m\mathrm{D}}+1}\right), \tag{3.2}$$

where $h_{\mathrm{T}_m\mathrm{D}} \triangleq |h_{\mathrm{T}_m\mathrm{D}}|\, e^{-j\phi_{\mathrm{T}_m\mathrm{D}}}$, $\kappa_{\mathrm{T}_m\mathrm{D}}$ is the power ratio between the deterministic LoS component and the scattering components and $\mathbb{E}[|h_{\mathrm{T}_m\mathrm{D}}|^2] = \sigma^2_{h_{\mathrm{T}_m\mathrm{D}}}$ is the average power of the scattering components. For link budget normalization purposes, $\sigma^2_{h_{\mathrm{T}_m\mathrm{D}}} = 1$ will be also assumed (other values could be easily accommodated into the large-scale, average coefficients). Similar notation and assumptions hold for h_{ST_m}, $m \in \{1, 2, \ldots, M\}$ and h_{SD}. It is noted that for $\kappa = 0$, Rice is simplified to Rayleigh fading.

Quasi-static block fading is assumed, i.e., the channel remains constant for L_c (source-destination link) symbols and changes independently between

---

[1] The complex channel is the superposition of $\sqrt{\frac{\kappa_{\mathrm{T}_m\mathrm{D}}}{\kappa_{\mathrm{T}_m\mathrm{D}}+1}}\,\sigma_{\mathrm{T}_m\mathrm{D}}\, e^{j\theta} + \mathcal{CN}\left(0, \frac{\sigma^2_{\mathrm{T}_m\mathrm{D}}}{\kappa_{\mathrm{T}_m\mathrm{D}}+1}\right)$ with $\theta \sim \mathcal{U}[0, 2\pi)$.

channel coherence time periods. Channel coefficients h_{SD}, $\{h_{\mathrm{ST}_m}\}$, $\{h_{\mathrm{T}_m\mathrm{D}}\}$, $m \in \{1, 2, \ldots, M\}$ are assumed independent in the numerical results. Furthermore, the following notation is also adopted:

$$h_m = h_{\mathrm{ST}_m} h_{\mathrm{T}_m\mathrm{D}} = |h_{\mathrm{ST}_m} h_{\mathrm{T}_m\mathrm{D}}| \, e^{-j\phi_m}, m \in \{1, 2, \ldots, M\},$$
$$h_0 = h_{\mathrm{SD}}, m = 0, \tag{3.3}$$

3.1.2 Signal Model

The baseband equivalent of the source message $m(t)$ is given by:

$$c(t) = \sqrt{2P}\, m(t) \tag{3.4}$$

where $\mathbb{E}[|m(t)|^2] = 1$. Different normalization could be incorporated into the large-scale coefficients. The baseband complex equivalent of the scattered waveform from tag m is given by [7]:

$$u_m(t) = \sqrt{\eta\, \mathsf{L}_{\mathrm{ST}_m}}\, \left[A_s - \Gamma_m(t)\right]\, h_{\mathrm{ST}_m}\, c(t), \tag{3.5}$$

$$\Gamma_m(t) \in \{\Gamma_1, \Gamma_2, \ldots, \Gamma_K\}, \tag{3.6}$$

where $\Gamma_m(t)$ stands for the modified (complex) reflection coefficient for tag m, assuming that the tag can terminate its antenna between K loads and η models the (limited) tag power scattering efficiency. It is noted that for passive (amplification-free) tags, $|\Gamma_k| \leq 1$, while for commercial RFID tags, $K = 2$. Parameter A_s stands for the load-independent *structural mode* that solely depends on tag's antenna [8], commonly overlooked in the literature; $A_s = 0$ only for *minimum scattering* antennas, i.e., antennas that do not reflect anything when terminated at open (i.e., infinite) load.

The received demodulated complex baseband signal at the destination is given by the superposition of the source and all tags' backscattered signals propagated through wireless channels h_{SD} and $\{h_{\mathrm{T}_m\mathrm{D}}\}$, respectively:

$$\begin{aligned} y(t) &= \sqrt{\mathsf{L}_{\mathrm{SD}}}\, h_{\mathrm{SD}}\, c(t) + \sum_{m=1}^{M} \sqrt{\mathsf{L}_{\mathrm{T}_m\mathrm{D}}}\, h_{\mathrm{T}_m\mathrm{D}}\, u_m(t) + n(t) \\ &= \sqrt{\mathsf{L}_{\mathrm{SD}}}\, h_{\mathrm{SD}}\, c(t) + n(t) \\ &\quad + \sum_{m=1}^{M} \sqrt{\eta\, \mathsf{L}_{\mathrm{ST}_m} \mathsf{L}_{\mathrm{T}_m\mathrm{D}}}\, h_{\mathrm{ST}_m}\, h_{\mathrm{T}_m\mathrm{D}}\, \left[A_s - \Gamma_m(t)\right]\, c(t) \\ &= \sqrt{2P} \left[\sqrt{g_0}\, h_0 + \sum_{m=1}^{M} \sqrt{g_m}\, h_m\, \mathcal{Y}_m(t) \right] m(t) + n(t), \end{aligned} \tag{3.7}$$

where $n(t)$ is the thermal noise at the receiver and

$$g_0 = \mathsf{L}_{\mathrm{SD}}, \tag{3.8}$$

$$g_m = \eta\, \mathsf{L}_{\mathrm{ST}_m} \mathsf{L}_{\mathrm{T}_m\mathrm{D}}\, \mathbb{E}\left[|A_s - \Gamma_m(t)|^2\right], \tag{3.9}$$

$$\mathcal{Y}_m(t) = \frac{A_s - \Gamma_m(t)}{\sqrt{\mathbb{E}\left[|A_s - \Gamma_m(t)|^2\right]}}, \tag{3.10}$$

$$y_m\left[\Gamma_m(t)\right] \triangleq \sqrt{g_m}\, h_m\, \mathcal{Y}_m(t). \tag{3.11}$$

Notice that $\mathbb{E}[|h_0|^2] = \mathbb{E}[|h_m|^2] = \mathbb{E}[|\mathcal{Y}_m|^2] = 1$ since $\mathbb{E}[|h_m|^2] = \mathbb{E}[|h_{\mathrm{ST}_m}|^2]\,\mathbb{E}[|h_{\mathrm{T}_m\mathrm{D}}|^2]$, due to the followed assumptions. It is also noted that $\mathbb{E}\left[|A_s - \Gamma_m(t)|^2\right] = (1/K)\sum_{k=1}^{K} |A_s - \Gamma_k|^2$.

The above model is valid when coupling among the tags is negligible, i.e., the tags are separated by distance at least equal to $\lambda/2$. Additive thermal noise $n(t)$ is modelled by a complex, circularly symmetric, additive Gaussian noise process with $\mathbb{E}[|n(t)|^2] = N_0 B$, where B stands for receiver's bandwidth.[2]

3.2 LMMSE estimate

3.2.1 Rayleigh Fading

Let us first try to find the possible values of the term $\mathcal{Y}_m(t)$. Assume that $\kappa_{\mathrm{SD}} = \kappa_{\mathrm{ST_m}} = \kappa_{\mathrm{T}_m\mathrm{D}} = 0$, which leads to h_{SD}, h_{ST_m}, $h_{\mathrm{T}_m\mathrm{D}}$ being modeled by Rayleigh distribution. If we also consider that $\sigma = 1$ for every possible link, then (3.2) yields: h_{SD}, h_{ST_m} , $h_{\mathrm{T}_m\mathrm{D}} \sim \mathcal{CN}(0, 1)$.

Suppose now that $K = 2$ and $\Gamma_1 = 0$, $\Gamma_2 = A_s$ are the possible load states.

[2] $N_0 = k_b T_\theta$, where k_b and T_θ are the Boltzmann constant and receiver temperature, respectively.

Thus, term $\mathcal{Y}_m(t)$ becomes:

$$\begin{aligned}\mathcal{Y}_m(t) &= \frac{A_s - \Gamma_m(t)}{\sqrt{\mathbb{E}\left[|A_s - \Gamma_m(t)|^2\right]}}\\ &= \begin{cases} \dfrac{A_s - 0}{\sqrt{\frac{|A_s-0|^2+|A_s-A_s|^2}{1+1}}} &, \quad \Gamma_m = 0\\ \dfrac{A_s - A_s}{\sqrt{\frac{|A_s-0|^2+|A_s-A_s|^2}{1+1}}} &, \quad \Gamma_m = A_s\end{cases}\\ &= \begin{cases} \sqrt{2}\,\dfrac{A_s}{|A_s|} &, \quad \Gamma_m = 0\\ 0 &, \quad \Gamma_m = A_s\end{cases}\end{aligned}$$

Regarding the pilot symbols for the estimation of the various h_m, $m \in \{0, 1, ..., M\}$, we assume to have L symbol periods within which the coefficients' values remain invariant. Following the steps of the previous sections, we choose a portion of them, in order to estimate the channel coefficients, while the rest of them will be devoted to data transmission. Specifically, we presume that the total number of pilot symbols available is $N_{\text{tot}} = \alpha L$ where $\alpha \in (0, 1)$, which will be used to estimate all $M + 1$ coefficients. This leads to $N_{\text{tr}} = \frac{T_{\text{tot}}}{M+1} = \frac{\alpha L}{M+1}$ [3] of them to be used for the estimation of each h_m.

Our goal is to find the right $\hat{h}_m$ which will minimize $\mathbb{E}[|h_m - \hat{h}_m|^2]$, when $\hat{h}_m$ is a linear (affine) function of the received signal's measurements, i.e. $\hat{h}_m = \mathbf{a}_m^{\text{H}}\,\mathbf{y} + b_m$, where $\mathbf{a}_m \in \mathbb{C}^{N_{\text{tr}}\times 1}$, $b_m \in \mathbb{C}$, $\forall\, m \in \{0, 1, 2, \ldots, M\}$. Subsequently, finding the right $\mathbf{a}_m$ and b_m is equivalent to finding $\hat{h}_m$, since $\mathbf{y}$ is known to the receiver. As it is shown in previous sections, $\mathbf{a}_m$ and b_m can be computed in closed form as follows:

$$\mathbf{a}_m = C_{\mathbf{y}}^{-1} C_{\mathbf{y}h_m} \tag{3.12}$$

and

$$b_m = \mu_{h_m} - \mathbf{a}^{\text{H}} \mu_{\mathbf{y}} \tag{3.13}$$

where $\mu_{\mathbf{y}} \triangleq \mathbb{E}[\mathbf{y}]$, $C_{\mathbf{y}} \triangleq \mathbb{E}[(\mathbf{y} - \mu_{\mathbf{y}})(\mathbf{y} - \mu_{\mathbf{y}})^{\text{H}}]$ is the covariance matrix of $\mathbf{y}$, $\mu_{h_m} \triangleq \mathbb{E}[h_m]$ and $C_{\mathbf{y}h_m} \triangleq \mathbb{E}[(\mathbf{y} - \mu_{\mathbf{y}})(h_m - \mu_{h_m})^*]$ is the cross-covariance vector between $\mathbf{y}$ and h_m.

[3] α is selected in such way so that $N_{\text{tr}}$is a positive integer.

Direct Link Estimation

In order to estimate all RIS coefficients, estimation of h_0 is a preceding necessity. We presume all tags are terminated to a load such that $\Gamma_m(t) = A_s$ for all RIS elements ($m \in \{1, 2, \dots, M\}$). Hence, (3.7) becomes:

$$y(t) = \sqrt{2Pg_0}\, h_0\, m(t) + n(t) \tag{3.14}$$

We define the SNR at the receiver as:

$$\mathsf{SNR}_0 \triangleq \frac{2Pg_0}{N_0 B} \tag{3.15}$$

Thus, (3.14) can be modified to be:

$$\bar{y}_0(t) = \sqrt{\mathsf{SNR}_0}\, h_0\, m(t) + \bar{n}(t) \tag{3.16}$$

where $\bar{n}(t) \sim \mathcal{CN}(0, 1)$ is the normalized thermal noise.

We can further simplify equation (3.16), while also incorporating the pilot symbols to it:

$$\bar{\mathbf{y}}_0 = c_0\, h_0\, \mathbf{m} + \bar{\mathbf{n}} \tag{3.17}$$

where $c_0 = \sqrt{\mathsf{SNR}_0}$ and $\bar{\mathbf{y}}_0, \mathbf{m} \in \mathbb{C}^{N_{\mathrm{tr}}}$ and $\bar{\mathbf{n}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_{\mathrm{tr}}})$.

One can easily notice the similarities between the RIS and the Lozano model. We therefore expect that the derivation of both the LMMSE estimator and the LMMSE will be a straightforward procedure. However, this might not be the case regarding the Ricean fading which will be analyzed right after this section.

Now that we have disentangled the system model, we proceed to first find the estimator for the direct link channel, i.e. $\hat{h}_0$ where $\hat{h}_0 = \mathbf{a}_0^{\mathrm{H}}\, \bar{\mathbf{y}}_0 + b_0$. To achieve that we introduce some fundamental properties that derive from the dependency relations of the system's parameters.

Since h_0 and the entries of $\bar{\mathbf{n}}$ are independent, they are also uncorrelated, which leads to:

$$C_{\bar{\mathbf{n}}h_0} = C_{h_0\bar{\mathbf{n}}}^{\mathrm{T}} = \mathbf{0}_{N_{\mathrm{tr}}\times 1} \tag{3.18}$$

while for the mean of $\bar{\mathbf{y}}_0$, we derive:

$$\mu_{\bar{\mathbf{y}}_0} = \mathbb{E}[\bar{\mathbf{y}}_0] = c_0\, \mathbf{m}\, \mathbb{E}[h_0] + \mathbb{E}[\bar{\mathbf{n}}] = c_0\, \mathbf{m}\, \mu_{h_0} + \mu_{\bar{\mathbf{n}}} = \mathbf{0}_{N_{\mathrm{tr}}\times 1} \tag{3.19}$$

Having the equations (3.18) and (3.19) in mind, $C_{\bar{\mathbf{y}}_0}$ and $C_{\bar{\mathbf{y}}_0 h_0}$ can be calculated as follows:

$$\begin{aligned}
C_{\bar{\mathbf{y}}_0} &= \mathbb{E}[(\bar{\mathbf{y}}_0 - \mu_{\bar{\mathbf{y}}_0})\,(\bar{\mathbf{y}}_0 - \mu_{\bar{\mathbf{y}}_0})^{\mathrm{H}}] \\
&= \mathbb{E}[\bar{\mathbf{y}}_0 \bar{\mathbf{y}}_0^{\mathrm{H}}] \\
&= \mathbb{E}[(c_0 h_0 \mathbf{m} + \bar{\mathbf{n}})\left(c_0 \mathbf{m}^{\mathrm{H}} h_0^* + \bar{\mathbf{n}}^{\mathrm{H}}\right)] \\
&= \mathbb{E}[c_0^2 h_0 h_0^* \mathbf{m}\mathbf{m}^{\mathrm{H}} + c_0 \mathbf{m} h_0 \bar{\mathbf{n}}^{\mathrm{H}} + \bar{\mathbf{n}} h_0^* c_0 \mathbf{m}^{\mathrm{H}} + \bar{\mathbf{n}}\bar{\mathbf{n}}^{\mathrm{H}}] \\
&= c_0^2 \mathbb{E}[h_0 h_0^*]\mathbf{m}\mathbf{m}^{\mathrm{H}} + c_0 \mathbf{m}\mathbb{E}[h_0 \bar{\mathbf{n}}^{\mathrm{H}}] + c_0 \mathbb{E}[\bar{\mathbf{n}} h_0^*]\mathbf{m}^{\mathrm{H}} + \mathbb{E}[\bar{\mathbf{n}}\bar{\mathbf{n}}^{\mathrm{H}}] \\
&= c_0^2 \mathbb{E}[|h_0|^2]\mathbf{m}\mathbf{m}^{\mathrm{H}} + c_0 \mathbf{m} C_{h_0 \bar{\mathbf{n}}} + c_0 C_{\bar{\mathbf{n}} h_0} \mathbf{m}^{\mathrm{H}} + C_{\bar{\mathbf{n}}} \\
&= c_0^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{0}_{N_{\mathrm{tr}} \times N_{\mathrm{tr}}} + \mathbf{0}_{N_{\mathrm{tr}} \times N_{\mathrm{tr}}} + \mathbf{I}_{N_{\mathrm{tr}}} \\
&= \mathsf{SNR}_0\, \mathbf{m}\, \mathbf{m}^{\mathrm{H}} + \mathbf{I}_{N_{\mathrm{tr}}}
\end{aligned} \tag{3.20}$$

$$\begin{aligned}
C_{\bar{\mathbf{y}}_0 h_0} &= \mathbb{E}[(\bar{\mathbf{y}}_0 - \mu_{\bar{\mathbf{y}}_0})\,(h_0 - \mu_{h_0})^*] \\
&= \mathbb{E}[\bar{\mathbf{y}}_0 h_0^*] \\
&= \mathbb{E}[(c_0 h_0 \mathbf{m} + \bar{\mathbf{n}})\, h_0^*] \\
&= \mathbb{E}[c_0 h_0 h_0^* \mathbf{m} + \bar{\mathbf{n}} h_0^*] \\
&= c_0 \mathbb{E}[h_0 h_0^*]\mathbf{m} + \mathbb{E}[\bar{\mathbf{n}} h_0^*] \\
&= c_0 \mathbb{E}[|h_0|^2]\mathbf{m} + C_{\bar{\mathbf{n}} h_0} \\
&= c_0 \mathbf{m} + \mathbf{0}_{N_{\mathrm{tr}} \times 1} \\
&= \sqrt{\mathsf{SNR}_0}\, \mathbf{m}
\end{aligned} \tag{3.21}$$

Substituting the last relations to $\mathbf{a}_0$, we derive:

$$\mathbf{a}_0 = C_{\bar{\mathbf{y}}_0}^{-1} C_{\bar{\mathbf{y}}_0 h_0} = \left(\mathsf{SNR}_0 \mathbf{m}\, \mathbf{m}^{\mathrm{H}} + \mathbf{I}_{N_{\mathrm{tr}}}\right)^{-1} \sqrt{\mathsf{SNR}_0}\, \mathbf{m} \tag{3.22}$$

and as for b:

$$b_0 = \mu_{h_0} - \mathbf{a}_0^{\mathrm{H}}\, \mu_{\bar{\mathbf{y}}_0} = 0 - \mathbf{a}_0^{\mathrm{H}}\, \mathbf{0}_{N_{\mathrm{tr}} \times 1} = 0 \tag{3.23}$$

Thus, the LMMSE estimate of h_0 can now be calculated:

$$
\begin{aligned}
\hat{h}_0 &= \mathbf{a}_0^{\mathrm{H}} \bar{\mathbf{y}}_0 + b_0 \\
&= \left(\left(\mathsf{SNR}_0 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{I}_{N_{\mathrm{tr}}} \right)^{-1} \sqrt{\mathrm{SNR}_0}\, \mathbf{m} \right)^{\mathrm{H}} \bar{\mathbf{y}}_0 \\
&= \sqrt{\mathsf{SNR}_0}\, \mathbf{m}^{\mathrm{H}} \left(\left(\mathsf{SNR}_0 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{I}_{N_{\mathrm{tr}}} \right)^{-1} \right)^{\mathrm{H}} \bar{\mathbf{y}}_0 \\
&= \sqrt{\mathsf{SNR}_0}\, \mathbf{m}^{\mathrm{H}} \left(\mathsf{SNR}_0 \left(\mathbf{m}\mathbf{m}^{\mathrm{H}} + \frac{1}{\mathsf{SNR}_0} \mathbf{I}_{N_{\mathrm{tr}}} \right)^{H} \right)^{-1} \bar{\mathbf{y}}_0 \\
&= \frac{1}{\sqrt{\mathsf{SNR}_0}}\, \mathbf{m}^{\mathrm{H}} \left(\mathbf{m}\mathbf{m}^{\mathrm{H}} + \frac{1}{\mathsf{SNR}_0} \mathbf{I}_{N_{\mathrm{tr}}} \right)^{-1} \bar{\mathbf{y}}_0 \\
&\overset{(*)}{=} \frac{1}{\sqrt{\mathsf{SNR}_0}} \left(\mathbf{m}^{\mathrm{H}}\mathbf{m} + \frac{1}{\mathsf{SNR}_0} \right)^{-1} \mathbf{m}^{\mathrm{H}} \bar{\mathbf{y}}_0 \\
&\overset{(**)}{=} \frac{1}{\sqrt{\mathsf{SNR}_0}} \left(N_{\mathrm{tr}} + \frac{1}{\mathsf{SNR}_0} \right)^{-1} \mathbf{m}^{\mathrm{H}} \bar{\mathbf{y}}_0 \\
&= \frac{1}{\sqrt{\mathsf{SNR}_0} \left(N_{\mathrm{tr}} + \frac{1}{\mathsf{SNR}_0} \right)} \mathbf{m}^{\mathrm{H}} \bar{\mathbf{y}}_0 \\
&= \frac{\sqrt{\mathsf{SNR}_0}\, \mathbf{m}^{\mathrm{H}} \bar{\mathbf{y}}_0}{1 + \mathsf{SNR}_0 N_{\mathrm{tr}}}
\end{aligned}
\tag{3.24}
$$

where at $(*)$ we used the Push Through Identity and at $(**)$ the optimal training symbols' property from the Hassibi derivation (2.33) when there is only 1 transmitter antenna.

We define the zero-mean channel estimation error as $\tilde{h}_0 = h_0 - \hat{h}_0$. Subsequently, the LMMSE is defined as:

$$
\mathsf{LMMSE} = \sigma^2_{\tilde{h}_0} = \sigma^2_{h_0} - C_{h_0\mathbf{y}} C_{\mathbf{y}}^{-1} C_{\mathbf{y}h_0} \tag{3.25}
$$

As previously stated, our goal is to minimize the linear MSE. To do that, we first need to compute the cross-covariance matrix $C_{h_0\mathbf{y}}$, while also taking into account the optimal training symbols vector's property which will further reduce the MSE.

$$
\begin{aligned}
C_{h_0\mathbf{y}} &= \mathbb{E}[(h_0 - \mu_{h_0})\,(\mathbf{y} - \mu_{\mathbf{y}})^{\mathrm{H}}] \\
&= \mathbb{E}[h_0 \mathbf{y}^{\mathrm{H}}] \\
&= \mathbb{E}[h_0\left(c_0 \mathbf{m}^{\mathrm{H}} h_0^* + \bar{\mathbf{n}}^{\mathrm{H}}\right)] \\
&= \mathbb{E}[c_0 h_0 h_0^* \mathbf{m}^{\mathrm{H}} + h_0 \bar{\mathbf{n}}^{\mathrm{H}}] \\
&= c_0 \mathbb{E}[h_0 h_0^*] \mathbf{m}^{\mathrm{H}} + \mathbb{E}[h_0 \bar{\mathbf{n}}^{\mathrm{H}}] \\
&= c_0 \mathbb{E}[|h_0|^2] \mathbf{m}^{\mathrm{H}} + C_{\bar{\mathbf{n}} h_0} \\
&= c_0 \mathbf{m}^{\mathrm{H}} + \mathbf{0}_{1\times N_{\mathrm{tr}}} \\
&= \sqrt{\mathsf{SNR}_0}\,\mathbf{m}^{\mathrm{H}}
\end{aligned} \tag{3.26}
$$

We are now in the position to calculate the LMMSE as:

$$
\begin{aligned}
\sigma^2_{\tilde{h}_0} &= \sigma^2_{h_0} - C_{h_0\mathbf{y}} C_{\mathbf{y}}^{-1} C_{\mathbf{y}h_0} \\
&= 1 - c_0 \mathbf{m}^{\mathrm{H}} \left(c_0^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{I}_{N_{\mathrm{tr}}}\right)^{-1} c_0 \mathbf{m} \\
&\overset{(*)}{=} \left(c_0 \mathbf{m}^{\mathrm{H}} c_0 \mathbf{m} + 1\right)^{-1} \\
&= \frac{1}{1 + c_0^2 \mathbf{m}^{\mathrm{H}} \mathbf{m}} \\
&= \frac{1}{1 + N_{\mathrm{tr}}\,\mathsf{SNR}_0}
\end{aligned} \tag{3.27}
$$

where at $(*)$ we used the Woodbury Matrix Identity (2.26). We notice that both the LMMSE estimate and the LMMSE are closely related to the ones we derived in the Lozano case (2.49).

Compound Link Estimation

As for the estimation of the RIS channel coefficients the procedure is almost identical. We assume that every RIS element is terminated at open load Γ_2, except the one we are interested in estimating, say h_m. Considering also the fact that the estimation of g_0 and h_0 has preceded, the signal model is modified as follows:

$$
y_{\mathrm{init}}(t) = \left[\sqrt{2Pg_0}\,h_0 + \sqrt{2}\,\frac{A_s}{|A_s|}\sqrt{2Pg_m}\,h_m\right] m(t) + n(t) \tag{3.28}
$$

where the factor of $\sqrt{2}\,\frac{A_s}{|A_s|}$ is included due to the term's $\mathcal{Y}_m(t)$ value for

$\Gamma_m = 0$.

After incorporating the pilot symbols to the model, while also applying the normalization with respect to noise power, we obtain:

$$\bar{\mathbf{y}}_{\text{init}} = \left[\sqrt{\mathsf{SNR}_0}\, h_0 + \sqrt{2}\,\frac{A_s}{|A_s|}\,\sqrt{\mathsf{SNR}_m}\, h_m\right]\mathbf{m} + \bar{\mathbf{n}} \tag{3.29}$$

which can be further simplified to be the same equation as in the case of the direct-link channel:

$$\bar{\mathbf{y}} = \bar{\mathbf{y}}_{\text{init}} - \sqrt{\mathsf{SNR}_0}\,\hat{h}_0\,\mathbf{m} = c_m\, h_m\,\mathbf{m} + \bar{\mathbf{n}} \tag{3.30}$$

where $c_m = \frac{A_s}{|A_s|}\sqrt{2\mathsf{SNR}_m}$ with $\mathsf{SNR}_m \triangleq \frac{2Pg_m}{N_0 B}$, (we presume perfect h_0 estimation, with $\hat{h}_0 = h_0$), $\mathbf{m} \in \mathbb{C}^N_{\text{tr}}$ and $\bar{\mathbf{n}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_{\text{tr}}})$. We notice that $\mathbf{y}$ has, again, zero-mean which leads to the same estimator and error relations as in the prior model, with the only difference being the constant term c_m.

Hence, following the same steps as in the case of h_0, we can easily derive that:

$$\hat{h}_m = \frac{\sqrt{2\,\mathsf{SNR}_m}\,\frac{A_s}{|A_s|}\,\mathbf{m}^{\mathrm{H}}\mathbf{y}}{1 + 2\,N_{\text{tr}}\,\mathsf{SNR}_m} \tag{3.31}$$

and

$$\sigma^2_{\tilde{h}_m} = \frac{1}{1 + 2\,N_{\text{tr}}\,\mathsf{SNR}_m} \tag{3.32}$$

3.2.2 Rice Fading

As previously mentioned, we expect that the derivation of the Rice Fading case will deviate to an extent from the one of the Rayleigh case, especially regarding the estimation of the compound coefficients. To be exact, we presume that $\kappa_{\text{SD}} = \kappa_{\text{ST}_\text{m}} = \kappa_{\text{T}_m\text{D}} = \kappa \neq 0$, which leads to h_{SD}, h_{ST_m}, $h_{\text{T}_m\text{D}}$ being modeled by Rice distribution. Recall that $\sigma = 1$, which translates to:

$$|\mu_{h_i}| = \sqrt{\frac{\kappa}{\kappa + 1}} \tag{3.33}$$

with $i \in \{\text{SD}, \text{ST}_\text{m}, \text{T}_\text{m}\text{D}\}$. One can easily deduct that $\mu_{h_{\text{SD}}} \neq \mu_{h_m}$ ($h_m = h_{\text{ST}_\text{m}}\, h_{\text{T}_\text{m}\text{D}}$), which will be shown shortly.

Following the same signal model as in the Rayleigh Fading case (3.17), we firstly need to find the optimal estimation of the direct link channel, $\hat{h}_0 = \hat{h}_{\text{SD}}$, such that $\mathbb{E}[|h_0 - \hat{h}_0|^2]$ is minimized, when $\hat{h}_0$ is a linear (affine) function of the measurements, i.e. $\hat{h}_0 = \mathbf{a}^{\text{H}}\mathbf{y} + b$, where $\mathbf{a} \in \mathbb{C}^{T\times 1}$, $b \in \mathbb{C}$. Finding the right $\mathbf{a}$ and b is equivalent to finding $\hat{h}_0$, since $\mathbf{y}$ is known at the receiver's end. We will then assume that h_0 is known and proceed to estimate each of the RIS elements' compound channel coefficients.
As it is to be expected by now, the closed form equations for $\mathbf{a}$ and b are:

$$\mathbf{a}_j = C_{\mathbf{y}}^{-1} C_{\mathbf{y}h_j} \tag{3.34}$$

and

$$b_j = \mu_{h_j} - \mathbf{a}_j^{\text{H}} \mu_{\mathbf{y}} \tag{3.35}$$

where $j \in [0, 1, ..., M]$.
where $C_{\mathbf{y}h_j} = \mathbb{E}[(\mathbf{y} - \mu_{\mathbf{y}})(h_j - \mu_h)^{\text{H}}]$ is the cross-covariance vector between $\mathbf{y}$ and h_j, while $C_{\mathbf{y}} = \mathbb{E}[(\mathbf{y} - \mu_{\mathbf{y}})(\mathbf{y} - \mu_{\mathbf{y}})^{\text{H}}]$ is the covariance matrix of $\mathbf{y}$.

Since h_j and the entries of $\bar{\mathbf{n}}$ are independent, they are also uncorrelated, which leads to:

$$C_{\bar{\mathbf{n}}h_j} = C_{h_j\bar{\mathbf{n}}}^{T} = \mathbf{0}_{T\times 1} \tag{3.36}$$

while for the mean of $\mathbf{y}$, we derive:

$$\mu_{\mathbf{y}} = \mathbb{E}[\mathbf{y}] = c\mathbf{m}\mathbb{E}[h_j] + \mathbb{E}[\bar{\mathbf{n}}] = c\mathbf{m}\mu_{h_j} + \mu_{\bar{\mathbf{n}}} = c\mu_{h_j}\mathbf{m} \tag{3.37}$$

Lastly, the variance of h_j is:

$$\sigma_{h_j}^2 = \mathbb{E}\left[|h_j|^2\right] - \left|\mathbb{E}\left[h_j\right]\right|^2 = 1 - |\mu_{h_j}|^2 \tag{3.38}$$

Direct Link Estimation

Having the equations (3.36) and (3.37) (3.38) in mind, $C_{\mathbf{y}}$ and $C_{\mathbf{y}h_0}$ can be calculated as follows:

$$\begin{aligned}
C_{\mathbf{y}} &= \mathbb{E}\left[(\mathbf{y}-\mu_{\mathbf{y}})(\mathbf{y}-\mu_{\mathbf{y}})^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[(c_0 h_0 \mathbf{m} + \bar{\mathbf{n}} - \mu_{\mathbf{y}})\left(c_0 \mathbf{m}^{\mathrm{H}} h_0^* + \bar{\mathbf{n}}^{\mathrm{H}} - \mu_{\mathbf{y}}^{\mathrm{H}}\right)\right] \\
&= \mathbb{E}[c_0^2 h_0 h_0^* \mathbf{m}\mathbf{m}^{\mathrm{H}} + c_0 h_0 \mathbf{m}\bar{\mathbf{n}}^{\mathrm{H}} - c_0 h_0 \mathbf{m}\mu_{\mathbf{y}}^{\mathrm{H}} + c_0 \bar{\mathbf{n}} h_0^* \mathbf{m}^{\mathrm{H}} + \\
&\quad + \bar{\mathbf{n}}\bar{\mathbf{n}}^{\mathrm{H}} - \bar{\mathbf{n}}\mu_{\mathbf{y}}^{\mathrm{H}} - c_0 \mu_{\mathbf{y}} h_0^* \mathbf{m}^{\mathrm{H}} - \mu_{\mathbf{y}}\bar{\mathbf{n}}^{\mathrm{H}} + \mu_{\mathbf{y}}\mu_{\mathbf{y}}^{\mathrm{H}}] \\
&= c_0^2 \mathbb{E}[h_0 h_0^*]\mathbf{m}\mathbf{m}^{\mathrm{H}} + c_0 \mathbf{m}\mathbb{E}[h_0 \bar{\mathbf{n}}^{\mathrm{H}}] - c_0 \mathbb{E}[h_0]\mathbf{m}\mu_{\mathbf{y}}^{\mathrm{H}} + c_0 \mathbb{E}[\bar{\mathbf{n}} h_0^*]\mathbf{m}^{\mathrm{H}} + \\
&\quad + \mathbb{E}[\bar{\mathbf{n}}\bar{\mathbf{n}}^{\mathrm{H}}] - \mathbb{E}[\bar{\mathbf{n}}]\mu_{\mathbf{y}}^{\mathrm{H}} - c_0 \mu_{\mathbf{y}}\mathbb{E}[h_0^*]\mathbf{m}^{\mathrm{H}} - \mu_{\mathbf{y}}\mathbb{E}[\bar{\mathbf{n}}^{\mathrm{H}}] + \mu_{\mathbf{y}}\mu_{\mathbf{y}}^{\mathrm{H}} \\
&= c_0^2 \mathbb{E}[|h_0|^2]\mathbf{m}\mathbf{m}^{\mathrm{H}} + c_0 \mathbf{m} C_{h_0\bar{\mathbf{n}}} - c_0 \mu_{h_0}\mathbf{m}\mu_{\mathbf{y}}^{\mathrm{H}} + c_0 \mathbf{0}_{T\times 1}\mathbf{m}^{\mathrm{H}} + C_{\bar{\mathbf{n}}} - \\
&\quad - \mathbf{0}_{T\times 1}\mu_{\mathbf{y}}^{\mathrm{H}} - c_0 \mu_{h_0}^* \mu_{\mathbf{y}}\mathbf{m}^{\mathrm{H}} - c_0 \mu_{\mathbf{y}} \mathbf{0}_{1\times T} + \mu_{\mathbf{y}}\mu_{\mathbf{y}}^{\mathrm{H}} \\
&= c_0^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{0}_{T\times T} - c_0 \mu_{h_0}\mathbf{m}\mu_y^{\mathrm{H}} + \mathbf{0}_{T\times T} + \mathbf{I}_T - \\
&\quad - \mathbf{0}_{T\times T} - c_0 \mu_{h_0}^* \mu_y \mathbf{m}^{\mathrm{H}} - \mathbf{0}_{T\times T} + \mu_y \mu_y^{\mathrm{H}} \\
&= c_0^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} - c_0 \mu_{h_0}\mathbf{m}\mu_y^{\mathrm{H}} + \mathbf{I}_T - c_0 \mu_{h_0}^* \mu_y \mathbf{m}^{\mathrm{H}} + c_0 \mu_{h_0}\mathbf{m}\mu_y^{\mathrm{H}} \\
&= c_0^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{I}_T - c_0^2 \mu_{h_0}\mu_{h_0}^* \mathbf{m}\mathbf{m}^{\mathrm{H}} \\
&= c_0^2 \mathbf{m}\mathbf{m}^{\mathrm{H}}\left(1 - |\mu_{h_0}|^2\right) + \mathbf{I}_T \\
&= c_0^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} \sigma_{h_0}^2 + \mathbf{I}_T \\
&= \mathrm{SNR}_0 \sigma_{h_0}^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{I}_T
\end{aligned} \tag{3.39}$$

$$\begin{aligned}
C_{\mathbf{y}h_0} &= \mathbb{E}\left[(\mathbf{y}-\mu_{\mathbf{y}})(h_0-\mu_{h_0})^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[(c_0 h_0 \mathbf{m} + \bar{\mathbf{n}} - \mu_{\mathbf{y}})\left(h_0^* - \mu_{h_0}^*\right)\right] \\
&= \mathbb{E}\left[c h_0 h_0^* \mathbf{m} - c h_0 \mathbf{m}\mu_{h_0}^* + \bar{\mathbf{n}} h_0^* - \bar{\mathbf{n}}\mu_{h_0}^* - \mu_{\mathbf{y}} h_0^* + \mu_{\mathbf{y}}\mu_{h_0}^*\right] \\
&= c_0 \mathbb{E}[h_0 h_0^*]\mathbf{m} - c_0 \mathbb{E}[h_0]\mu_{h_0}^* \mathbf{m} + \mathbb{E}[\bar{\mathbf{n}} h_0^*] - \mathbb{E}[\bar{\mathbf{n}}]\mu_{h_0}^* - \mu_{\mathbf{y}}\mathbb{E}[h_0^*] + \mu_{\mathbf{y}}\mu_{h_0}^* \\
&= c_0 \mathbb{E}[|h_0|^2]\mathbf{m} - c_0 |\mu_{h_0}|^2 \mathbf{m} + C_{\bar{\mathbf{n}}h_0} - \mathbf{0}_{T\times 1} - \mu_{\mathbf{y}}\mu_{h_0}^* + c_0 \mu_{h_0}\mathbf{m}\mu_{h_0}^* \\
&= c_0 \mathbf{m} - c_0 |\mu_{h_0}|^2 \mathbf{m} - c_0 \mu_{h_0}\mathbf{m}\mu_{h_0}^* + c_0 |\mu_{h_0}|^2 \mathbf{m} \\
&= c_0 \mathbf{m}\left(1 - |\mu_{h_0}|^2\right) \\
&= c_0 \mathbf{m}\sigma_{h_0}^2 \\
&= \sqrt{\mathrm{SNR}_0}\, \sigma_{h_0}^2 \mathbf{m}
\end{aligned} \tag{3.40}$$

Substituting these two quantities to **a** and b, we obtain:

$$\begin{aligned}
\mathbf{a}_0 &= C_{\mathbf{y}}^{-1} C_{\mathbf{y}h_0} \\
&= \left(c_0^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} \sigma_{h_0}^2 + \mathbf{I}_T\right)^{-1} c_0 \mathbf{m} \sigma_{h_0}^2 \\
&\overset{(*)}{=} c_0 \mathbf{m} \left(c_0^2 \mathbf{m}^{\mathrm{H}} \mathbf{m} \sigma_{h_0}^2 + 1\right)^{-1} \sigma_{h_0}^2 \\
&= \frac{c_0 \sigma_{h_0}^2}{c_0^2 \sigma_{h_0}^2 \mathbf{m}^{\mathrm{H}} \mathbf{m} + 1} \mathbf{m} \\
&\overset{(**)}{=} \frac{c_0 \sigma_{h_0}^2}{c_0^2 \sigma_{h_0}^2 T + 1} \mathbf{m} \\
&= \frac{\sqrt{\mathrm{SNR}_0}\, \sigma_{h_0}^2}{T\, \mathrm{SNR}_0\, \sigma_{h_0}^2 + 1} \mathbf{m}
\end{aligned} \tag{3.41}$$

and

$$\begin{aligned}
b_0 &= \mu_{h_0} - \mathbf{a}_0^{\mathrm{H}} \mu_{\mathbf{y}} \\
&= \mu_{h_0} - \left(\frac{c_0 \sigma_{h_0}^2}{c_0^2 \sigma_{h_0}^2 T + 1} \mathbf{m}\right)^{\mathrm{H}} c_0 \mathbf{m} \mu_{h_0} \\
&= \mu_{h_0} - \frac{c_0^2 \sigma_{h_0}^2 \mu_{h_0}}{c_0^2 \sigma_{h_0}^2 T + 1} \mathbf{m}^{\mathrm{H}} \mathbf{m} \\
&\overset{(**)}{=} \frac{\mu_{h_0}(c_0^2 \sigma_{h_0}^2 T + 1) - c_0^2 \sigma_{h_0}^2 \mu_{h_0} T}{c_0^2 \sigma_{h_0}^2 T + 1} \\
&= \frac{\mu_{h_0}}{c_0^2 \sigma_{h_0}^2 T + 1} \\
&= \frac{\mu_{h_0}}{T\, \mathrm{SNR}_0\, \sigma_{h_0}^2 + 1}
\end{aligned} \tag{3.42}$$

where at $(*)$ we used the Push Through Identity and at $(**)$ the optimal training symbols' property from Hassibi's derivation (2.33), when there is only 1 transmitter antenna.

Thus, the LMMSE estimate of h_0 can now be calculated:

$$\begin{aligned}\hat{h}_0 &= \mathbf{a}_0^{\mathrm{H}}\mathbf{y} + b_0 \\ &= \frac{\sqrt{\mathrm{SNR}_0}\,\sigma_{h_0}^2}{T\,\mathrm{SNR}_0\,\sigma_{h_0}^2+1}\mathbf{m}^{\mathrm{H}}\mathbf{y} + \frac{\mu_{h_0}}{T\,\mathrm{SNR}_0\,\sigma_{h_0}^2+1} \\ &= \frac{\sqrt{\mathrm{SNR}_0}\,\sigma_{h_0}^2\,\mathbf{m}^{\mathrm{H}}\mathbf{y} + \mu_{h_0}}{T\,\mathrm{SNR}_0\,\sigma_{h_0}^2+1}\end{aligned} \tag{3.43}$$

We define the zero-mean channel estimation error as $\tilde{h}_0 = h_0 - \hat{h}_0$. Subsequently, the LMMSE is defined as:

$$\mathrm{LMMSE} = \sigma_{\tilde{h}_0}^2 = \sigma_{h_0}^2 - C_{h_0\mathbf{y}}C_{\mathbf{y}}^{-1}C_{\mathbf{y}h_0} \tag{3.44}$$

As we have already stated, our goal is to minimize the linear MSE. To achieve this, we first need to compute the cross-covariance vector $C_{h_0\mathbf{y}}$, while also taking into account the optimal training symbols vector, which will further reduce the MSE.

$$\begin{aligned}C_{h_0\mathbf{y}} &= \mathbb{E}\left[(h_0-\mu_{h_0})\,(\mathbf{y}-\mu_{\mathbf{y}})^{\mathrm{H}}\right] \\ &= \mathbb{E}\left[(h_0-\mu_{h_0})\left(c_0h_0^*\mathbf{m}^{\mathrm{H}} + \bar{\mathbf{n}}^{\mathrm{H}} - \mu_{\mathbf{y}}^{\mathrm{H}}\right)\right] \\ &= \mathbb{E}\left[c_0h_0h_0^*\mathbf{m}^{\mathrm{H}} + h_0\bar{\mathbf{n}}^{\mathrm{H}} - h_0\mu_{\mathbf{y}}^{\mathrm{H}} - c_0h_0^*\mu_{h_0}\mathbf{m}^{\mathrm{H}} - \mu_{h_0}\bar{\mathbf{n}}^{\mathrm{H}} + \mu_{h_0}\mu_y^{\mathrm{H}}\right] \\ &= c_0\mathbb{E}[|h_0|^2]\mathbf{m}^{\mathrm{H}} + \mathbb{E}[h_0\bar{\mathbf{n}}^{\mathrm{H}}] - \mathbb{E}[h_0]\mu_{\mathbf{y}}^{\mathrm{H}} - c_0\mathbb{E}[h_0^*]\mu_{h_0}\mathbf{m}^{\mathrm{H}} - \mu_{h_0}\mathbb{E}[\bar{\mathbf{n}}^{\mathrm{H}}] + \mu_{h_0}\mu_y^{\mathrm{H}} \\ &= c_0\mathbf{m}^{\mathrm{H}} + \mathbf{0}_{1\times T} - \mu_{h_0}\mu_{\mathbf{y}}^{\mathrm{H}} - c_0\mu_{h_0}\mu_{h_0}^*\mathbf{m}^{\mathrm{H}} - \mathbf{0}_{1\times T} + \mu_{h_0}\mu_y^{\mathrm{H}} \\ &= c_0\mathbf{m}^{\mathrm{H}} - c_0|\mu_{h_0}|^2\mathbf{m}^{\mathrm{H}} \\ &= c_0\mathbf{m}^{\mathrm{H}}\left(1-|\mu_{h_0}|^2\right) \\ &= c_0\mathbf{m}^{\mathrm{H}}\sigma_{h_0}^2 \\ &= \sqrt{\mathrm{SNR}_0}\,\mathbf{m}^{\mathrm{H}}\,\sigma_{h_0}^2\end{aligned} \tag{3.45}$$

We can now calculate the MMSE:

$$\begin{aligned}
\sigma^2_{\tilde{h}_0} &= \sigma^2_{h_0} - C_{h_0\mathbf{y}} C_{\mathbf{y}}^{-1} C_{\mathbf{y}h_0} \\
&= \sigma^2_{h_0} - \left(c_0 \mathbf{m}^{\mathrm{H}} \sigma^2_{h_0}\right) \left(\frac{c_0 \sigma^2_{h_0}}{c_0^2 \sigma^2_{h_0} T + 1} \mathbf{m} \right) \\
&= \sigma^2_{h_0} - \frac{c_0^2 \sigma^2_{h_0} \sigma^2_{h_0}}{c_0^2 \sigma^2_{h_0} T + 1} \mathbf{m}^{\mathrm{H}} \mathbf{m} \\
&= \frac{\sigma^2_{h_0} (c_0^2 \sigma^2_{h_0} T + 1) - c_0^2 \sigma^2_{h_0} \sigma^2_{h_0} T}{c_0^2 \sigma^2_{h_0} T + 1} \\
&= \frac{\sigma^2_{h_0}}{c_0^2 \sigma^2_{h_0} T + 1} \\
&= \frac{\sigma^2_{h_0}}{T \, \mathrm{SNR}_0 \, \sigma^2_{h_0} + 1}
\end{aligned} \tag{3.46}$$

Compound Link Estimation

The procedure regarding the estimation of the RIS compound channel estimation is differentiated to an extent. We begin by assuming that every RIS element is terminated at open load Γ_2, except the one we are interested in, say $h_m = h_{\mathrm{ST_m}} h_{\mathrm{T_mD}}$. Considering also the fact that the estimation of g_0 and h_0 has preceded, the signal model is modified to be:

$$y_{\mathrm{init}}(t) = \left[\sqrt{2Pg_0}\, h_0 + \sqrt{2}\, \sqrt{2Pg_m}\, h_m \right] m(t) + n(t) \tag{3.47}$$

where the factor of $\sqrt{2}$ is included due to the way $\mathcal{Y}_m(t)$ is defined.

The same pilot-symbols model as in the Rayleigh case is adopted:

$$\mathbf{y}_{\mathrm{init}} = \left[\sqrt{\mathrm{SNR}_0}\, h_0 + \sqrt{2\mathrm{SNR}_m}\, h_m \right] \mathbf{m} + \bar{\mathbf{n}} \tag{3.48}$$

which can be further simplified to be the same equation as the one we used for the direct-link channel:

$$\mathbf{y} = c_m \, h_m \, \mathbf{m} + \bar{\mathbf{n}} \tag{3.49}$$

where $c_m = \sqrt{2\mathrm{SNR}_m}$ with $\mathrm{SNR}_m = \frac{2Pg_m}{N_0 B}$, $\mathbf{y} = \mathbf{y}_{\mathrm{init}} - \sqrt{\mathrm{SNR}_0}\, \hat{h}_0 \mathbf{m} \in \mathbb{C}^T$

(perfect h_0 estimation is assumed, with $\hat{h}_0 = h_0$) and $\mathbf{m}, \bar{\mathbf{n}} \in \mathbb{C}^T$. In contrast to the Rayleigh case, we now notice that $\mathbf{y}$ not only has a non-zero mean, but it also differs from the one of h_0, equation (3.33), since we now focus on the compound channel coefficient h_m.

More specifically, $|\mu_{h_m}|$ and μ_y can be calculated as follows:

$$|\mu_{h_m}| = |\mathbb{E}\left[h_{\mathrm{ST_m}} h_{\mathrm{T_mD}}\right]| = |\mu_{h_{\mathrm{ST_m}}}||\mu_{h_{\mathrm{T_mD}}}| = \frac{\kappa}{\kappa+1} \tag{3.50}$$

and

$$\mu_{\mathbf{y}} = \mathbb{E}[\mathbf{y}] = c_m\,\mathbf{m}\,\mathbb{E}[h_m] + \mathbb{E}[\bar{\mathbf{n}}] = c_m\,\mathbf{m}\,\mu_{h_m} + \mu_{\bar{\mathbf{n}}} = c_m\,\mu_{h_m}\mathbf{m} \tag{3.51}$$

with the variance of the compound channel being:

$$\sigma^2_{h_m} = \mathbb{E}\left[|h_m|^2\right] - |\mathbb{E}\left[h_m\right]|^2 = 1 - |\mu_{h_m}|^2 = 1 - \left(\frac{\kappa}{\kappa+1}\right)^2 = \frac{2\kappa+1}{(\kappa+1)^2} \tag{3.52}$$

Having the equations (3.36) and (3.51) (3.52) in mind, $C_{\mathbf{y}}$ and $C_{\mathbf{y}h_m}$ can be calculated as follows:

$$\begin{aligned}
C_{\mathbf{y}} &= \mathbb{E}\left[(\mathbf{y}-\mu_{\mathbf{y}})\,(\mathbf{y}-\mu_{\mathbf{y}})^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[(c_m h_m \mathbf{m} + \bar{\mathbf{n}} - \mu_{\mathbf{y}})\left(c_m \mathbf{m}^{\mathrm{H}} h_m^* + \bar{\mathbf{n}}^{\mathrm{H}} - \mu_{\mathbf{y}}^{\mathrm{H}}\right)\right] \\
&= \mathbb{E}[c_m^2 h_m h_m^* \mathbf{m}\mathbf{m}^{\mathrm{H}} + c_m h_m \mathbf{m}\bar{\mathbf{n}}^{\mathrm{H}} - c_m h_m \mathbf{m}\mu_{\mathbf{y}}^{\mathrm{H}} + c_m \bar{\mathbf{n}} h_m^* \mathbf{m}^{\mathrm{H}} + \\
&\qquad + \bar{\mathbf{n}}\bar{\mathbf{n}}^{\mathrm{H}} - \bar{\mathbf{n}}\mu_{\mathbf{y}}^{\mathrm{H}} - c_m \mu_{\mathbf{y}} h_m^* \mathbf{m}^{\mathrm{H}} - \mu_{\mathbf{y}}\bar{\mathbf{n}}^{\mathrm{H}} + \mu_{\mathbf{y}}\mu_{\mathbf{y}}^{\mathrm{H}}] \\
&= c_m^2 \mathbb{E}[h_m h_m^*]\mathbf{m}\mathbf{m}^{\mathrm{H}} + c_m \mathbf{m}\mathbb{E}[h_m \bar{\mathbf{n}}^{\mathrm{H}}] - c_m \mathbb{E}[h_m]\mathbf{m}\mu_{\mathbf{y}}^{\mathrm{H}} + c_m \mathbb{E}[\bar{\mathbf{n}} h_m^*]\mathbf{m}^{\mathrm{H}} + \\
&\qquad + \mathbb{E}[\bar{\mathbf{n}}\bar{\mathbf{n}}^{\mathrm{H}}] - \mathbb{E}[\bar{\mathbf{n}}]\mu_{\mathbf{y}}^{\mathrm{H}} - c_m \mu_{\mathbf{y}}\mathbb{E}[h_m^*]\mathbf{m}^{\mathrm{H}} - \mu_{\mathbf{y}}\mathbb{E}[\bar{\mathbf{n}}^{\mathrm{H}}] + \mu_{\mathbf{y}}\mu_{\mathbf{y}}^{\mathrm{H}} \\
&= c_m^2 \mathbb{E}[|h_m|^2]\mathbf{m}\mathbf{m}^{\mathrm{H}} + c_m \mathbf{m} C_{h_m\bar{\mathbf{n}}} - c_m \mu_{h_m}\mathbf{m}\mu_{\mathbf{y}}^{\mathrm{H}} + c_m \mathbf{0}_{T\times 1}\mathbf{m}^{\mathrm{H}} + C_{\bar{\mathbf{n}}} - \\
&\qquad - \mathbf{0}_{T\times 1}\mu_{\mathbf{y}}^{\mathrm{H}} - c_m \mu_{h_m}^* \mu_{\mathbf{y}}\mathbf{m}^{\mathrm{H}} - c_m \mu_{\mathbf{y}}\mathbf{0}_{1\times T} + \mu_{\mathbf{y}}\mu_{\mathbf{y}}^{\mathrm{H}} \\
&= c_m^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{0}_{T\times T} - c_m \mu_{h_m}\mathbf{m}\mu_y^{\mathrm{H}} + \mathbf{0}_{T\times T} + \mathbf{I}_T - \\
&\qquad - \mathbf{0}_{T\times T} - c_m \mu_{h_m}^* \mu_y \mathbf{m}^{\mathrm{H}} - \mathbf{0}_{T\times T} + \mu_y \mu_y^{\mathrm{H}} \\
&= c_m^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} - c_m \mu_{h_m}\mathbf{m}\mu_y^{\mathrm{H}} + \mathbf{I}_T - c_m \mu_{h_m}^* \mu_y \mathbf{m}^{\mathrm{H}} + c_m \mu_{h_m}\mathbf{m}\mu_y^{\mathrm{H}} \\
&= c_m^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{I}_T - c_m^2 \mu_{h_m}\mu_{h_m}^* \mathbf{m}\mathbf{m}^{\mathrm{H}} \\
&= c_m^2 \mathbf{m}\mathbf{m}^{\mathrm{H}}\left(1 - |\mu_{h_m}|^2\right) + \mathbf{I}_T \\
&= 2\,\mathrm{SNR}_m \sigma^2_{h_m}\mathbf{m}\mathbf{m}^{\mathrm{H}} + \mathbf{I}_T
\end{aligned} \tag{3.53}$$

$$\begin{aligned}
C_{\mathbf{y}h_m} &= \mathbb{E}\left[(\mathbf{y}-\mu_{\mathbf{y}})\left(h_m-\mu_{h_m}\right)^{\mathrm{H}}\right] \\
&= \mathbb{E}\left[(c_m h_m \mathbf{m} + \bar{\mathbf{n}} - \mu_{\mathbf{y}})\left(h_m^* - \mu_{h_m}^*\right)\right] \\
&= \mathbb{E}\left[c_m h_m h_m^* \mathbf{m} - c_m h \mathbf{m} \mu_{h_m}^* + \bar{\mathbf{n}} h_m^* - \bar{\mathbf{n}} \mu_{h_m}^* - \mu_{\mathbf{y}} h_m^* + \mu_{\mathbf{y}} \mu_{h_m}^*\right] \\
&= c_m \mathbb{E}[h_m h_m^*]\mathbf{m} - c_m \mathbb{E}[h_m]\mu_{h_m}^* \mathbf{m} + \mathbb{E}[\bar{\mathbf{n}} h_m^*] - \mathbb{E}[\bar{\mathbf{n}}]\mu_{h_m}^* - \mu_{\mathbf{y}} \mathbb{E}[h_m^*] + \mu_{\mathbf{y}} \mu_{h_m}^* \\
&= c_m \mathbb{E}[|h_m|^2]\mathbf{m} - c_m |\mu_{h_m}|^2 \mathbf{m} + C_{\bar{\mathbf{n}} h_m} - \mathbf{0}_{T\times 1} - \mu_{\mathbf{y}} \mu_{h_m}^* + c_m \mu_{h_m} \mathbf{m} \mu_{h_m}^* \\
&= c_m \mathbf{m} - c_m |\mu_{h_m}|^2 \mathbf{m} - c_m \mu_{h_m} \mathbf{m} \mu_{h_m}^* + c_m |\mu_{h_m}|^2 \mathbf{m} \\
&= c_m \mathbf{m}\left(1 - |\mu_{h_m}|^2\right) \\
&= \sqrt{2\,\mathrm{SNR}_m}\, \sigma_{h_m}^2 \mathbf{m}
\end{aligned} \tag{3.54}$$

Substituting these two quantities to $\mathbf{a}$, we derive:

$$\begin{aligned}
\mathbf{a}_m &= C_{\mathbf{y}}^{-1} C_{\mathbf{y}h_m} \\
&= \left(c_m^2 \mathbf{m}\mathbf{m}^{\mathrm{H}} \sigma_{h_m}^2 + \mathbf{I}_T\right)^{-1} c_m \mathbf{m} \sigma_{h_m}^2 \\
&\overset{(*)}{=} c_m \mathbf{m}\left(c_m^2 \mathbf{m}^{\mathrm{H}} \mathbf{m} \sigma_{h_m}^2 + 1\right)^{-1} \sigma_{h_m}^2 \\
&= \frac{c_m \sigma_{h_m}^2 \mathbf{m}}{c_m^2 \sigma_{h_m}^2 \mathbf{m}^{\mathrm{H}} \mathbf{m} + 1} \\
&\overset{(**)}{=} \frac{c_m \sigma_{h_m}^2}{c_m^2 \sigma_{h_m}^2 T + 1} \mathbf{m} \\
&= \frac{\sqrt{2\,\mathrm{SNR}_m}\, \sigma_{h_m}^2}{2\,T\,\mathrm{SNR}_m \sigma_{h_m}^2 + 1} \mathbf{m}
\end{aligned} \tag{3.55}$$

and as for b:

$$\begin{aligned}
b_m &= \mu_{h_m} - \mathbf{a}_m^{\mathrm{H}} \mu_{\mathbf{y}} \\
&= \mu_{h_m} - \left(\frac{c_m \sigma_{h_m}^2}{c_m^2 \sigma_{h_m}^2 T + 1} \mathbf{m}\right)^{\mathrm{H}} c_m \mu_{h_m} \mathbf{m} \\
&= \mu_{h_m} - \frac{c_m^2 \sigma_{h_m}^2 \mu_{h_m}}{c_m^2 \sigma_{h_m}^2 T + 1} \mathbf{m}^{\mathrm{H}} \mathbf{m} \\
&\overset{(**)}{=} \frac{\mu_{h_m}(c_m^2 \sigma_{h_m}^2 T + 1) - c_m^2 \sigma_{h_m}^2 T \mu_{h_m}}{c_m^2 \sigma_{h_m}^2 T + 1} \\
&= \frac{\mu_{h_m}}{c_m^2 \sigma_{h_m}^2 T + 1} \\
&= \frac{\mu_{h_m}}{2\,T\,\mathrm{SNR}_m \sigma_{h_m}^2 + 1}
\end{aligned} \tag{3.56}$$

where at (∗) we used the Push Through Identity and at (∗∗) the optimal training symbols' property from the Hassibi derivation (2.33), for 1 transmitter antenna in our case.

Hence, the LMMSE estimate of h can now be calculated:

$$\begin{aligned}\hat{h}_m &= \mathbf{a}_m^{\mathrm{H}}\mathbf{y} + b_m \\ &= \frac{c_m \sigma_{h_m}^2}{2T\,\mathrm{SNR}_m \sigma_{h_m}^2 + 1}\mathbf{m}^{\mathrm{H}}\mathbf{y} + \frac{\mu_{h_m}}{2T\,\mathrm{SNR}_m \sigma_{h_m}^2 + 1} \\ &= \frac{\sqrt{2\mathrm{SNR}_m}\,\sigma_{h_m}^2 \mathbf{m}^{\mathrm{H}}\mathbf{y} + \mu_{h_m}}{2\,T\,\mathrm{SNR}_m \sigma_{h_m}^2 + 1}\end{aligned} \tag{3.57}$$

As also stated in the previous section, our goal is to minimize the MMSE with respect to the training symbols' vector $\mathbf{m}$. As a last measure, in order to find the MMSE, we need to calculate the cross-covariance matrix $C_{h\mathbf{y}}$. It is by now evident that $C_{h\mathbf{y}}$ is just the hermitian of, the already calculated matrix, $C_{\mathbf{y}h}$. Invoking the equation (3.54), we get:

$$C_{h_m\mathbf{y}} = \sqrt{2\mathrm{SNR}_m}\,\sigma_{h_m}^2 \mathbf{m}^{\mathrm{H}} \tag{3.58}$$

We can now calculate the MMSE for the channel coefficient of the m-th RIS element:

$$\begin{aligned}\sigma_{\tilde{h}_m}^2 &= \sigma_{h_m}^2 - C_{h_m\mathbf{y}} C_{\mathbf{y}}^{-1} C_{\mathbf{y}h_m} \\ &= \sigma_{h_m}^2 - \left(c_m \sigma_{h_m}^2 \mathbf{m}^{\mathrm{H}}\right)\left(\frac{c_m \sigma_{h_m}^2}{c_m^2 \sigma_{h_m}^2 T + 1}\mathbf{m}\right) \\ &= \sigma_{h_m}^2 - \left(\frac{c_m^2 \sigma_{h_m}^2 \sigma_{h_m}^2}{c_m^2 \sigma_{h_m}^2 T + 1}\right)\mathbf{m}^{\mathrm{H}}\mathbf{m} \\ &= \frac{\sigma_{h_m}^2 (c_m^2 \sigma_{h_m}^2 T + 1) - c_m^2 \sigma_{h_m}^2 \sigma_{h_m}^2 T}{c_m^2 \sigma_{h_m}^2 T + 1} \\ &= \frac{\sigma_{h_m}^2}{c_m^2 \sigma_{h_m}^2 T + 1} \\ &= \frac{\sigma_{h_m}^2}{2\,T\,\mathrm{SNR}_m \sigma_{h_m}^2 + 1}\end{aligned} \tag{3.59}$$

A proof of concept which corroborates the preceding results by applying the formulas in practice can be found in [9].

Chapter 4

RIS Application

4.1 Optimal Gain Algorithm

Since the introduction of RIS to the communication world, there has been an overgrowing concern on whether there exists an algorithm which will optimally select the configuration that maximizes the signal's power at the receiver's end. If it does exist, can this algorithm respond before the environment changes and the problem's parameters are altered?

According to Eq. (3.7), the following instantaneous power maximization problem is formulated:

$$\max_{\{\mathcal{Y}_m(t)\}} \left| \sqrt{g_0}\, h_0 + \sum_{m=1}^{M} \sqrt{g_m}\, h_m\, \mathcal{Y}_m(t) \right|^2 2P \tag{4.1}$$

$$\max_{\{\mathcal{Y}_m(t)\}} \left| \underbrace{\sqrt{g_0}\, h_0}_{y_0} + \sum_{m=1}^{M} \sqrt{g_m}\, h_m\, \mathcal{Y}_m(t) \right| \tag{4.2}$$

$$= \max_{\{\Gamma_m(t)\}} \left| y_0 + \sum_{m=1}^{M} y_m \left[\Gamma_{m(t)}\right] \right|, \tag{4.3}$$

If we were to find the solution via brute force search, we would need to examine K^M load configurations since there are M elements with K possible load states, i.e. $\Gamma_m(t) \in \{\Gamma_1, \Gamma_2, \ldots, \Gamma_K\}$.

The problem is similar to noncoherent sequence detection of orthogonally-modulated sequences, solved with log-linear complexity in [10]. The trick is to introduce an auxiliary scalar variable ϕ into the problem of Eq. (4.3):

$$\begin{aligned} &\max_{\{\Gamma_m(t)\}} \max_{\phi\in[0,2\pi)} \Re\left\{ e^{-j\phi} \left(y_0 + \sum_{m=1}^{M} y_m \left[\Gamma_{m(t)}\right] \right) \right\} = \\ &\max_{\phi\in[0,2\pi)} \max_{\{\Gamma_m(t)\}} \left(\Re\left\{e^{-j\phi} y_0\right\} + \sum_{m=1}^{M} \Re\left\{ e^{-j\phi} y_m \left[\Gamma_{m(t)}\right] \right\} \right) \end{aligned} \tag{4.4}$$

4.1.1 $K = 2$ Loads

For a given point $\phi \in [0, 2\pi)$, the innermost maximization in Eq. (4.4) is separable for each $\Gamma_m(t)$ and hence, splits into independent maximizations for any $m = 1, 2, \ldots, M$:

$$\begin{aligned}\hat{\Gamma}_m(t) &= \underset{\Gamma_m(t)\in\{\Gamma_1,\Gamma_2\}}{\arg\max}\ \Re\left\{e^{-j\phi} y_m\left[\Gamma_m(t)\right]\right\} \\ &\Leftrightarrow \Re\left\{e^{-j\phi} y_m\left[\Gamma_1\right]\right\} \underset{\hat{\Gamma}_m(t)=\Gamma_2}{\overset{\hat{\Gamma}_m(t)=\Gamma_1}{\gtrless}} \Re\left\{e^{-j\phi} y_m\left[\Gamma_2\right]\right\} \\ &\Leftrightarrow \Re\left\{e^{-j\phi}\left(y_m\left[\Gamma_1\right] - y_m\left[\Gamma_2\right]\right)\right\} \underset{\hat{\Gamma}_m(t)=\Gamma_2}{\overset{\hat{\Gamma}_m(t)=\Gamma_1}{\gtrless}} 0 \\ &\Leftrightarrow \cos\left(\phi - \underline{/y_m\left[\Gamma_1\right] - y_m\left[\Gamma_2\right]}\right) \underset{\hat{\Gamma}_m(t)=\Gamma_2}{\overset{\hat{\Gamma}_m(t)=\Gamma_1}{\gtrless}} 0 \end{aligned} \tag{4.5}$$

Given the relation in Eq. (4.4), the optimal load sequence $\hat{\boldsymbol{\Gamma}}^{\text{opt}}$ can be found by varying ϕ from 0 to 2π. It is further noticed that, as ϕ scans $[0, 2\pi)$, the decision $\hat{\Gamma}_m(t)$ changes, according to Eq. (4.5), only when:

$$\begin{aligned}&\cos\left(\phi - \underline{/y_m\left[\Gamma_1\right] - y_m\left[\Gamma_2\right]}\right) = 0 \\ \Leftrightarrow\ &\phi = \underbrace{\pm\frac{\pi}{2} + \underline{/y_m\left[\Gamma_1\right] - y_m\left[\Gamma_2\right]} \pmod{2\pi}}_{\phi_m^{(1)},\phi_m^{(2)}}.\end{aligned} \tag{4.6}$$

Therefore, the sequence $\hat{\boldsymbol{\Gamma}} = \left[\hat{\Gamma}_1(t), \hat{\Gamma}_2(t), \ldots, \hat{\Gamma}_M(t)\right]^T$ changes only at $\left(\phi_1^{(1)}, \phi_1^{(2)}, \phi_2^{(1)}, \phi_2^{(2)}, \cdots, \phi_M^{(1)}, \phi_M^{(2)}\right)$. For the remaining part of this section, we assume that the above $2M$ points are distinct and nonzero, i.e., $\phi_m^{(j)} \neq \phi_l^{(k)}$ and $\phi_m^{(j)} \neq 0$, for any $j, k, \in \{1, 2\}$ and $m, l \in \{1, 2, \ldots, M\}$ with $m \neq l$. There is a case where the above assumption does not hold, examined in [10]. If the above points are sorted in ascending order, i.e.,

$$\begin{aligned}&(\theta_1, \theta_2, \cdots, \theta_{2M}) = \\ &= \mathsf{sort}\left(\phi_1^{(1)}, \phi_1^{(2)}, \phi_2^{(1)}, \phi_2^{(2)}, \cdots, \phi_M^{(1)}, \phi_M^{(2)}\right),\end{aligned} \tag{4.7}$$

then the decision $\hat{\boldsymbol{\Gamma}}$ will remain constant in each one of the $2M + 1$ intervals $(\theta_i, \theta_{i+1}), i \in \{0, 1, \ldots, 2M\}$, with $\theta_0 = 0$ and $\theta_{2M+1} = 2\pi$. The goal is the identification of the $2M + 1$ sequences that correspond to these intervals,[1]

[1] It can be shown that the sequence at $[0, \theta_1)$ is the same with the sequence at $[\theta_{2M}, 2\pi)$ and thus, $2M$ intervals/sequences should be identified.

one of which gives the optimal $\hat{\mathbf{\Gamma}}^{\text{opt}}$, i.e., the one that offers the maximum power; thus, the quality of each sequence is calculated with the norm metric of Eq. (4.3), which explicitly includes the direct channel h_0. Based on the above, the sorting operation in Eq. (4.7) is dominant in terms of computational cost, which is $\mathcal{O}(M \log M)$ and not 2^M.

4.1.2 $K > 2$ Loads

The method described above can be generalized to $K > 2$ loads, i.e., $\Gamma_m(t)$ belongs in $\{\Gamma_1, \Gamma_2, \ldots, \Gamma_K\}$. The solution is given by selecting the largest value of $\Re\left\{\mathrm{e}^{-j\phi} y_m[\Gamma_k]\right\}$ among all $k \in \{1, 2, \ldots K\}$, which results in testing $2M \times (K-1)$ changes of ϕ and as a result, same number of sequence changes and not $2M \times \binom{K}{2}$, as one would expect; the rest of the steps are exactly the same as in $K = 2$. Formal proof and details can be found in [10], omitted due to space constraints. Notice that the norm metric for the quality of each sequence must include h_0. The complexity of the algorithm is again $\mathcal{O}(M \log M)$ for $M > K$ and not K^M.

Notice that the change in loads' number also impacts the possible values of the term $\mathcal{Y}_m(t)$ which can be calculated as:

$$\begin{aligned}
\mathcal{Y}_m(t) &= \frac{A_s - \Gamma_m(t)}{\sqrt{\mathbb{E}\left[|A_s - \Gamma_m(t)|^2\right]}} \\
&= \begin{cases} \dfrac{A_s - \Gamma_k}{\sqrt{(1/K)\sum_{k=1}^{K} |A_s - \Gamma_k|^2}} &, \quad \Gamma_m = \Gamma_k \\ \dfrac{A_s - A_s}{\sqrt{(1/K)\sum_{k=1}^{K} |A_s - \Gamma_k|^2}} &, \quad \Gamma_m = A_s \end{cases} \\
&= \begin{cases} \dfrac{\sqrt{K}\,(A_s - \Gamma_k)}{\sqrt{\sum_{k=1}^{K} |A_s - \Gamma_k|^2}} &, \quad \Gamma_m = \Gamma_k \\ 0 &, \quad \Gamma_m = A_s \end{cases}
\end{aligned}$$

This change, however, does not alter the behaviour or the efficiency of the LMMSE estimator, since the results are modified solely by this constant term.

Chapter 5

Numerical Results

In this section, we study the behavior of MMSE as a function of the number of RIS elements (M), while the percentage the of pilot symbols (α) used remains fixed to 1% (Figs 5.1, 5.2). Notice that $\alpha\, L_c > M$ throughout the course of the experiment. We then plot the MMSE as a function of α assuming $M = 4096$ in order to investigate the effect of the large scale case on the system (Figs 5.3, 5.4). Both Rayleigh and Rician fading were examined.

Regarding all the figures that follow bellow, we consider $\kappa_{\mathrm{SD}} = \kappa_{\mathrm{ST_m}} = \kappa_{\mathrm{T}_m\mathrm{D}} = 0$ for the Rayleigh case, $\kappa_{\mathrm{SD}} = \kappa_{\mathrm{ST_m}} = \kappa_{\mathrm{T}_m\mathrm{D}} = 15$ for Rician fading, $\eta = 10\%$, $d_0^{\mathrm{X}} = 3$m, $d_{\mathrm{RIS-SD}} = 8$ m, $f_2 = 870$ MHz, $B = 48$ MHz, and 10 dB relative end-2-end antenna gain for the backscattered links compared to direct link, assuming that the source and destination antennas point towards the RIS (in order to assist its operation). Channel coherence time in number of symbols is set to $L_c = 24 \times 10^5$, corresponding to 100 ms and SD link is set at 48 Mbps using QPSK modulation. The Normalized MSE (NMSE) is defined as: $\mathrm{NMSE} = \mathbb{E}\left[\frac{|\hat{h}-h|^2}{|h|^2}\right]$. It is lastly worth mentioning that all presented illustrations are a product of averaged results after 10^5 realizations of the channels' coefficients h_i, $i \in \{\mathrm{ST_m}, \mathrm{T_mD}, \mathrm{SD}\}$, while assuming that the angles of both the LoS and the compound links' channel coefficients remain invariant throughout the course of the simulation.

We specifically set $d_{\mathrm{SD}} = 30$ m, $v_{\mathrm{X}} = 4$ and $P = 0$ dBm for Figures 5.1 and 5.3, which lead to $\mathrm{SNR}_0 = 21$ dB for the LoS link. Figs. 5.2, 5.4 and Figs. 5.5, 5.6 are offered with $d_{\mathrm{SD}} = 15$ m, $v_{\mathrm{X}} = 3$ where $P = 20$ dBm, which yields $\mathrm{SNR}_m = -20$ dB concerning the m-th element's received signal for the first two Figs, while P is varied between 0 and 20 dBm for the latter Figs.

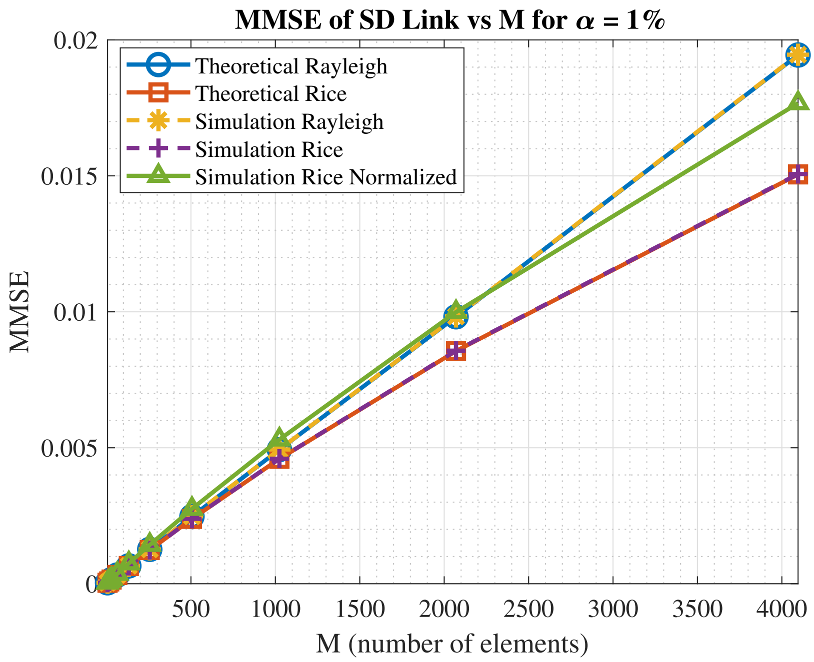

Figure 5.1: MMSE of LoS Link vs the Number of RIS Elements for fixed Training Symbols Value

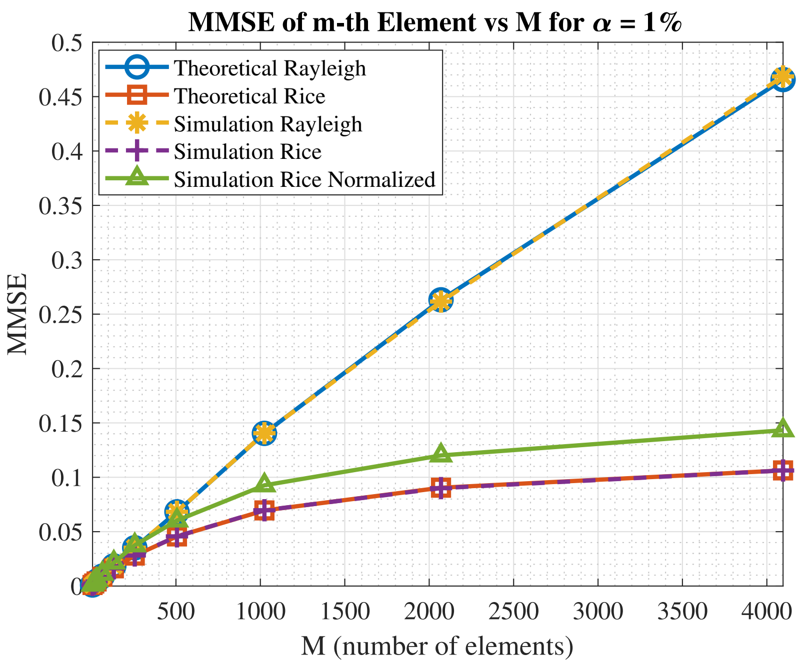

Figure 5.2: MMSE of m-th Tag's Path vs the Number of RIS Elements for a Fixed Training Symbols Value

The achievable MMSE performance in terms of the number of RIS elements is depicted in Figures 5.1 and 5.2. Focusing first on the LoS path, we notice that the simulation completely matches the equations derived in previous sections. As it was expected, an increase of M leads to an increase of the MMSE, due to the shortening of the training symbols' number corresponding to the estimation of each h_i and hence to h_{SD}. The Rice case outperforms the Rayleigh one as a result of the addition of a much more powerful and deterministic LoS component, which leads to a significantly smaller variance with respect to the values of each h_i. Lastly, the Normalized value of the MMSE is close to the real value due to the norm of each h_i being close to 1 in the average case scenario.

The same hold regarding the estimation of a random tag's channel coefficient in Fig. 5.2, only now the value of the MMSE is considerably greater due to the decrease in signal strength. Recall that although h_m is of normalized power, it is also the compound coefficient meaning that each g_m incorporates both the Source to Tag and Tag to Destination large-scale losses, a metric analogous to SNR_m. Additionally, the decrease in power leads to a much larger gap between the Rayleigh and the Rice case.

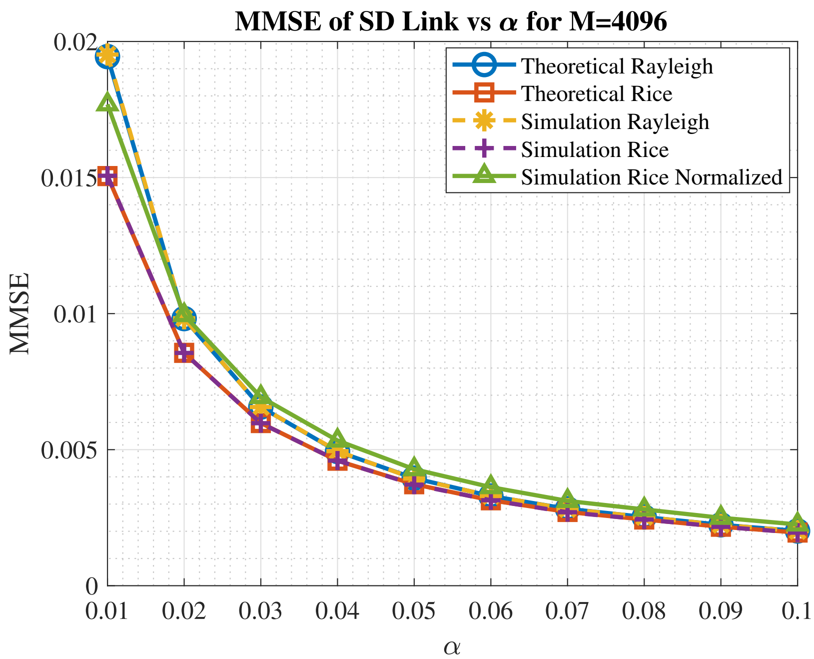

Figure 5.3: MMSE of LoS Link vs the Number of Training Symbols for a Fixed RIS Elements number

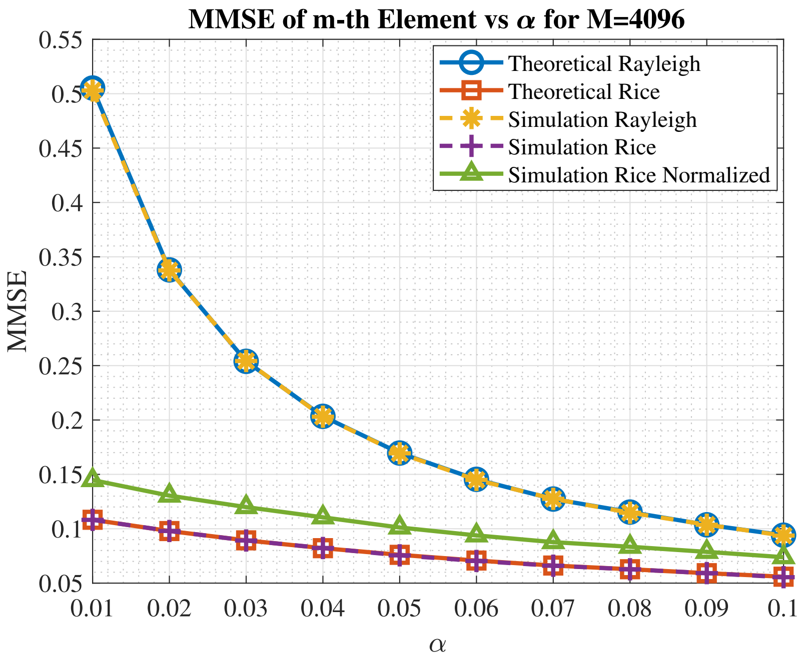

Figure 5.4: MMSE of m-th Tag's Path vs the Number of RIS Elements for a fixed training symbols value

Figures 5.3 and 5.4 illustrate how the MMSE behaves with respect to the number of pilot symbols α. For both the LoS path estimation as well as the estimation of the m-th tag's coefficient it is evident that as α increases, the MMSE is reduced. Moreover at high training symbols number, the product $\text{SNR}_0 T$ increases to a value with order of magnitude much greater than κ (recall that $\sigma_0^2 = \frac{1}{\kappa+1}$). Hence, by observing the LoS link figure we notice that at high α values the Rice case overlaps with the Rayleigh one. On the other hand, considering more realistic values for α leads to a gap between the Rice and the Rayleigh that is certainly non-negligible, especially regarding the estimation of a tag's coefficient as we see in Fig. 5.4.

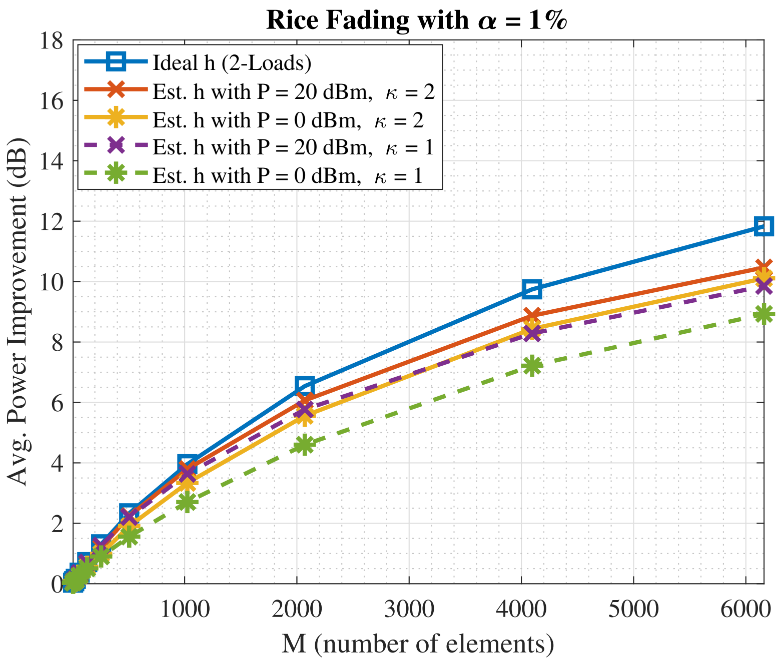

Figure 5.5: Impact of Channel Estimation Error to Power Improvement, $\alpha = 1\%$

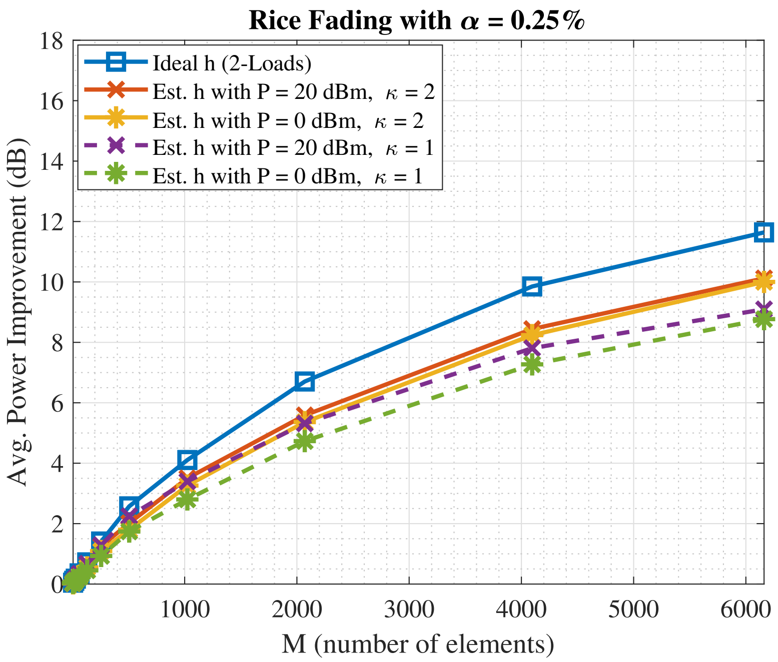

Figure 5.6: Impact of Channel Estimation Error to Power Improvement, $\alpha = 0.25\%$

Finally, Figs. 5.5 and 5.6 quantify the impact of channel estimation error $\tilde{h}_m$ to the optimal-gain algorithm's performance assuming Rice fading. The optimal configuration is found based on the estimated channels $\{\hat{h}_m\}$, while the power of the received signal is computed using the true channel coefficients. Channel coherence time in number of symbols is set to $L_c = 24 \times 10^5$, corresponding to channel coherence time of 100 ms and SD link at 48 Mbps with QPSK modulation. In these plots α is fixed at 1% (0.25% for Fig. 5.6) while the transmitted signal's power and parameter κ are varied. It is found that the estimation error for large number of elements (M) has an important impact to the algorithm's performance, as the gap between the ideal case and the estimation reaches 2 dB. The effect is even more evident for smaller values of P, which in combination with a weaker LoS component $(\kappa = 1)$, yields a 3 dB difference.

Chapter 6

Conclusion/Future Work

In this work we discussed the channel estimation of RIS assisted wireless network system and RIS' contribution to the received signal's average power gain after the application of the Optimal Gain Algorithm using the estimated channels as input.

In the first chapter we focused on the prior work in channel estimation considering MIMO systems using both the matrix, vector and scalar versions of the LMMSE estimator formulas, with an analytical derivation included, corroborating the results. The second chapter offers an analytical system (channel and signal) model which includes both the small and the large-scale fading as well as specific parameters relevant to reflection (or backscatter) radio, such as antenna structural mode and reflection efficiency. We then applied the LMMSE estimator comparing the MSE of Rice fading to the one of Rayleigh fading. Finally, we examined the contribution of the RIS to the system after the application of the Optimal Gain Algorithm taking into account the impact of the channel estimation error the the algorithm's performance.

We conclude that the Rice fading is more suited to describe the channel model than Rayleigh fading in the RIS assisted scenario, since it takes into account the Line of Sight link which (in the general case) has a power of orders of magnitude greater than multipath's power. This results in a smaller MSE and, consequently, to a superior channel estimation, minimizing the gain-loss of the algorithm. However, in both fading cases the gain-loss is non negligible, especially considering how sensitive the system is to the number of training symbols used for estimation .

Although the gains obtained from the algorithm are impressive, we have assumed a static environment, which leaves room for a sufficient estimation of the constant terms used in the LMMSE formulas and, therefore, a sufficient channel estimation. However, this is not a feasible approach in mobile scenarios where fast varying channel terms make an appearance and the channel coherence time shortens. This case requires the solution of the, far more taxing, joint estimation problem, which includes the estimation of both the slow and the fast varying terms.

Chapter 7

Appendix

7.1 Orthogonality of the LMMSE estimator

From MIT notes

7.2 Orthogonality of the LMMSE estimator (vector case)

In this section we will try to prove the two orthogonality principles [11], which are a prerequisite for the derivation of the Wiener-Hopf equations. Let us consider the same pilot-symbol model as in Chapter 1, that is:

$$\mathbf{x} = cS_\tau \mathbf{h} + \mathbf{v} \tag{7.1}$$

where $\mathbf{x} \in \mathbb{C}^{T_\tau \times 1}$, $\mathbf{h} \in \mathbb{C}^{M \times 1}$, $\mathbf{v} \in \mathbb{C}^{T_\tau \times 1}$.

7.2.1 Orthogonality Principle 1

Suppose that the set of all the affine estimators of $\mathbf{x}$ is defined as:

$$\mathcal{A} := \{\hat{\mathbf{h}} \in \mathbb{C}^{M \times 1} : \hat{\mathbf{h}} = A\mathbf{x} + \mathbf{b}\} \tag{7.2}$$

The first principle states that $\hat{\mathbf{h}}_{\text{L}} = \mathbf{h}_{\text{LMMSE}}$ is the LMMSE estimator of $\mathbf{h}$, if and only if

$$\mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\text{L}})\mathbf{z}^{\text{H}}] = 0 \tag{7.3}$$

where $\mathbf{z}$ is any affine transformation of $\mathbf{x}$, i.e. $\mathbf{z} \in \mathcal{A}$.

Proof (=>).
Let us first assume that $\hat{\mathbf{h}}_{\text{L}} \in \mathcal{A}$ satisfies (7.3), but is not the LMMSE estimator, i.e. there exists a $\hat{\mathbf{h}}_0 \neq \hat{\mathbf{h}}_{\text{L}}$: $\hat{\mathbf{h}}_0 \in \mathcal{A}$ and

$$\mathbb{E}[\|\mathbf{h} - \hat{\mathbf{h}}_0\|^2] \leq \mathbb{E}[\|\mathbf{h} - \hat{\mathbf{h}}_{\text{L}}\|^2] \tag{7.4}$$

We can rewrite the first half of the inequality as:

$$\begin{aligned}\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_0\|^2] &= \mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L}+\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0\|^2] \\ &= \mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L}\|^2]+\mathbb{E}[\|\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0\|^2]+ \\ &+2\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L})^\mathrm{H}(\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0)] \end{aligned} \tag{7.5}$$

Considering that $\hat{\mathbf{h}}_\mathrm{L}\in\mathcal{A}$ and $\hat{\mathbf{h}}_0\in\mathcal{A}$, it follows a fortiori that $(\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0)\in\mathcal{A}$ and combined with (7.3) we obtain:

$$\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L})(\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0)^\mathrm{H}]=0 \tag{7.6}$$

Recalling that trace is a linear operator, it is evident that for any complex vectors $\mathbf{x}_1$, $\mathbf{x}_2$:

$$\mathbb{E}[\mathbf{x}_2^\mathrm{H}\mathbf{x}_1]=\mathbb{E}[\mathrm{trace}(\mathbf{x}_2^\mathrm{H}\mathbf{x}_1)]=\mathbb{E}[\mathrm{trace}(\mathbf{x}_1\mathbf{x}_2^\mathrm{H})]=\mathrm{trace}(\mathbb{E}[\mathbf{x}_1\mathbf{x}_2^\mathrm{H}]) \tag{7.7}$$

This result, along with (7.6), implies that $2\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L})^\mathrm{H}(\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0)]=\mathrm{trace}(\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L})(\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0)^\mathrm{H}])=0$.
Consequently, (7.5) can be written as:

$$\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_0\|^2]=\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L}\|^2]+\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_0\|^2] \tag{7.8}$$

Substituting (7.8) to (7.4) we obtain:

$$\begin{aligned}&\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_0\|^2]\le\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L}\|^2] \\ &\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L}\|^2]+\mathbb{E}[\|\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0\|^2]\le\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L}\|^2] \\ &\mathbb{E}[\|\hat{\mathbf{h}}_\mathrm{L}-\hat{\mathbf{h}}_0\|^2]\le 0\end{aligned} \tag{7.9}$$

which is only true if $\hat{\mathbf{h}}_0=\hat{\mathbf{h}}_\mathrm{L}$.

Proof (<=).
Let us now assume that $\hat{\mathbf{h}}_\mathrm{L}$ is the LMMSE estimator but does not satisfy (7.3), i.e. there exists a $\mathbf{z}_0\in\mathcal{A}$ for which $\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L})\mathbf{z}_0^\mathrm{H}]\neq 0$.

We introduce another linear estimator of the form: $\hat{\mathbf{h}}_0=\hat{\mathbf{h}}_\mathrm{L}+C\mathbf{z}_0$. Then the MSE is modified to be $\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_0\|^2]=\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L}-C\mathbf{z}_0\|^2]$. After differentiating with respect to C and setting the result to 0, we obtain:

$$C_\mathrm{min}=\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_\mathrm{L})\mathbf{z}^\mathrm{H}]\mathbb{E}[\mathbf{z}_0\mathbf{z}_0^\mathrm{H}]^{-1} \tag{7.10}$$

Hence, we conclude that the optimal $\hat{\mathbf{h}}_0$ is: $\hat{\mathbf{h}}_0 = \hat{\mathbf{h}}_{\mathrm{L}} + C_{\min}\mathbf{z}_0$.

We continue by simplifying the new MSE in order to compare it to the initial one:

$$\begin{aligned}\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_0\|^2] &= \mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}} - C_{\min}\mathbf{z}_0\|^2] \\ &= \mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}}\|^2] + \mathbb{E}[\mathbf{z}_0^{\mathrm{H}} C_{\min}^{\mathrm{H}} C_{\min}\mathbf{z}_0] - \\ &- 2\mathbb{E}[\mathbf{z}_0^{\mathrm{H}} C_{\min}^{\mathrm{H}}(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})]\end{aligned} \tag{7.11}$$

After applying (7.7) to the second term of (7.12) we obtain:

$$\begin{aligned}\mathbb{E}[\mathbf{z}_0^{\mathrm{H}} C_{\min}^{\mathrm{H}} C_{\min}\mathbf{z}_0] &= \mathrm{trace}(\mathbb{E}[C_{\min}\mathbf{z}_0\mathbf{z}_0^{\mathrm{H}} C_{\min}^{\mathrm{H}}]) \\ &= \mathrm{trace}(C_{\min}\mathbb{E}[\mathbf{z}_0\mathbf{z}_0^{\mathrm{H}}] C_{\min}^{\mathrm{H}}) \\ &= \mathrm{trace}(\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}]\mathbb{E}[\mathbf{z}_0\mathbf{z}_0^{\mathrm{H}}]^{-1} \\ &\qquad \mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}]^{\mathrm{H}})\end{aligned} \tag{7.12}$$

Following the same procedure regarding the third term of (7.12), we derive:

$$\begin{aligned}\mathbb{E}[\mathbf{z}_0^{\mathrm{H}} C_{\min}^{\mathrm{H}}(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})] &= \mathrm{trace}(\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}\mathbf{z}_0^{\mathrm{H}} C_{\min}^{\mathrm{H}}]) \\ &= \mathrm{trace}(\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}\mathbf{z}_0^{\mathrm{H}}] C_{\min}^{\mathrm{H}}) \\ &= \mathrm{trace}(\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}]\mathbb{E}[\mathbf{z}_0\mathbf{z}_0^{\mathrm{H}}]^{-1} \\ &\qquad \mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}]^{\mathrm{H}})\end{aligned} \tag{7.13}$$

Substituting the last two equations in (7.12) leads to:

$$\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_0\|^2] = \mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}}\|^2] - \mathrm{trace}(\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}]\mathbb{E}[\mathbf{z}_0\mathbf{z}_0^{\mathrm{H}}]^{-1}\mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}]^{\mathrm{H}}) \tag{7.14}$$

Since the trace of a PSD matrix is a non-negative scalar value, we can directly compare the two MSEs: $\mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_0\|^2] \leq \mathbb{E}[\|\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}}\|^2]$,a result which contradicts to the statement of $\hat{\mathbf{h}}_{\mathrm{L}}$ being the LMMSE estimator.
Thus the proof is concluded.

7.2.2 Orthogonality Principle 2

The second principle states that $\hat{\mathbf{h}}_{\mathrm{L}} \in \mathcal{A}$ is the LMMSE estimator if and only if:

$$\mathbb{E}[\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}}] = 0 \quad \text{and} \quad \mathbb{E}[(\mathbf{h}-\hat{\mathbf{h}}_{\mathrm{L}})\mathbf{x}^{\mathrm{H}}] = 0 \tag{7.15}$$

Both the necessity and the sufficiency can be proven straightforwardly from Orthogonality Principle 1.

Proof (=>).
Suppose that $\hat{\mathbf{h}}_{\mathrm{L}}$ is the LMMSE estimator. Setting $A = 0$ and $\mathbf{b} = \mathbf{0}$ to $\mathbf{z}$, it is derived from (7.3) that: $\mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}] = \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})] = 0$.
Setting, therefore, $A = \mathbf{I}$ and $\mathbf{b} = \mathbf{0}$ to $\mathbf{z}$ leads to: $\mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}] = \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})\mathbf{x}^{\mathrm{H}}] = 0$.

Proof (<=).
Now assume that $\mathbb{E}[\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}}] = 0 \quad \text{and} \quad \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})\mathbf{x}^{\mathrm{H}}] = 0$.
Substituting $\mathbf{z} = A\mathbf{x} + \mathbf{b}$ to $\mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}]$ gives us:

$$
\begin{aligned}
\mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})\mathbf{z}^{\mathrm{H}}] &= \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})(\mathbf{x}^{\mathrm{H}}A^{\mathrm{H}} + \mathbf{b}^{\mathrm{H}})] \\
&= \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})\mathbf{x}^{\mathrm{H}}]A^{\mathrm{H}} + \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})\mathbf{b}^{\mathrm{H}})] \\
&\overset{(7.15)}{=} \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}}_{\mathrm{L}})(\mathbf{x}^{\mathrm{H}}A^{\mathrm{H}} + \mathbf{b}^{\mathrm{H}})] \\
&= \mathbf{0}_{M\times M} + \mathbf{0}_{M\times M} = \mathbf{0}_{M\times M}
\end{aligned} \tag{7.16}
$$

Thus $\hat{\mathbf{h}}_{\mathrm{L}}$ is the LMMSE estimator and the proof is completed.

7.2.3 Wiener-Hopf equations

As a by-product of the aforementioned derivation, we can now proceed to the task of finding closed-form expressions for the optimal A and $\mathbf{b}$.
It follows directly from (7.3), that:

$$\mathbf{b} = \mathbb{E}[\mathbf{h} - A\mathbf{x}] = \mathbb{E}[\mathbf{h}] - A\mathbb{E}[\mathbf{x}] \tag{7.17}$$

As for the case of A, we make use of (7.15):

$$\mathbb{E}[(\mathbf{h} - A\mathbf{x} - \mathbf{b})\mathbf{x}^{\mathrm{H}}] \overset{(7.17)}{=} \mathbb{E}[(\mathbf{h} - \mathbb{E}[\mathbf{h}] - A(\mathbf{x} - \mathbb{E}[\mathbf{x}]))\,\mathbf{x}^{\mathrm{H}}] = 0 \tag{7.18}$$

From which we obtain:

$$\mathbb{E}[(\mathbf{h} - \mathbb{E}[\mathbf{h}])\mathbf{x}^{\mathrm{H}}] = A\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])\mathbf{x}^{\mathrm{H}}] \tag{7.19}$$

Thus, since $\mathbb{E}[\mathbf{x}] = \mathbb{E}[\mathbf{h}] = 0$ (shown in Chapter 1), we get:

$$A = C_{\mathbf{hx}}C_{\mathbf{x}}^{-1} \tag{7.20}$$

Bibliography

[1] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[2] A. Lozano, R. W. Heath, and J. G. Andrews, "Fundamental Limits of Cooperation," *IEEE Trans. Inform. Theory*, vol. 59, no. 9, pp. 5213–5226, Sep. 2013.

[3] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra.* SIAM, 2001.

[4] S. M. Kay, *Fundamentals of statistical signal processing. [Volume I]. , Estimation theory.* Upper Saddle River (N.J.): Prentice Hall, 1993.

[5] S. Trampitsch, "Complex-valued data estimation: Second-order statistics and widely linear estimators," Ph.D. dissertation, 04 2013.

[6] A. Goldsmith, *Wireless Communications.* New York, NY, USA: Cambridge University Press, 2005.

[7] J. Kimionis, A. Bletsas, and J. N. Sahalos, "Increased Range Bistatic Scatter Radio," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 1091–1104, Mar. 2014.

[8] A. Bletsas, A. G. Dimitriou, and J. N. Sahalos, "Improving Backscatter Radio Tag Efficiency," *IEEE Trans. Microwave Theory Tech.*, vol. 58, no. 6, pp. 1502–1509, Jun. 2010.

[9] I. Vardakis, G. Kotridis, S. Peppas, K. Skyvalakis, G. Vougioukas, and A. Bletsas, "Intelligently Wireless Batteryless RF-Powered Reconfigurable Surface," in *IEEE GLOBECOM, accepted*, Madrid, Spain, Dec. 2021.

[10] P. N. Alevizos, Y. Fountzoulas, G. N. Karystinos, and A. Bletsas, "Log-Linear-Complexity GLRT-Optimal Noncoherent Sequence Detection for Orthogonal and RFID-Oriented Modulations," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1600–1612, Apr. 2016.

[11] Namrata Vaswani, "Jointly Gaussian random variables, MMSE and linear MMSE estimation," last visited on 30/8/2021. [Online]. Available: https://www.ece.iastate.edu/~namrata/EE527_Spring12/jointGauss_MMSE_linearMMSE.pdf