Channel Estimation For Large MIMO Systems Under Hardware Impairments

Konstantinos Papadopoulos

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

April 2016

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

Contents

List of Figures

List of Acronyms

CSI	Channel state information
CRB	Cramer-Rao bound
EVM	Error Vector Magnitude
FDD	Frequency Division Duplex
I/Q	In-phase Quadrature
LAS	Likelihood Ascent Search
LOS	Line of Sight
LS	Least Squares
LMMSE	Linear Minimum Mean Square Error
MF	Matched Filter
MIMO	Multiple Input Multiple Output
OFDM	Orthogonal Frequency-Division Multiplexing
PDF	Probability Density Function
RF	Radio Frequency
RTS	Reactive Tabu Search
SNR	Signal-to-Noise ratio
TLS	Total Least Squares
TDD	Time Division Duplex
ULA	Uniform Linear Array
WL	Widely-Linear
ZF	Zero Forcing

Abstract

In this thesis, we study the problem of channel estimation under hardware impairments for Large Multiple Input Multiple Output(MIMO) systems. Large MIMO are wireless communication systems with tens to hundreds of antennas at the base station which offer numerous advantages over conventional MIMO, such as improved performance and energy efficiency [9]. Large MIMO systems are often built with low-cost components, which may lead to hardware imperfections and cause distortion of the base station's and the mobile station's signal. In order to perform channel estimation, we develop an accurate system model taking into consideration the distortion caused by hardware impairments. For this system model, we extend the Linear Minimum Mean Square Error (LMMSE) estimator for Large MIMO systems proposed in [5] for a multi-user system. The proposed LMMSE estimator considers the distortion at both ends of the communication link and achieves better performance in terms of relative estimation error per antenna over Signal-to-Noise ratio (SNR) compared to estimators used for conventional MIMO systems, such as the LMMSE and the Least Squares (LS) estimator. Furthermore, the Cramer-Rao bound (CRB) for the channel estimation of the system is calculated.

Sommaire

Dans cette thèse, nous étudions le problème d'estimation des canaux en cas de dèfaillances matèrielles sur les systémes Large MIMO. Les LMIMO sont équipés de dizaines à centaines d'antennes sur la station de base, offrant de nombreux avantages par rapport aux systèmes MIMO traditionnels, tels que de meilleures performances et une efficacité énergétique accrue. Les grands systèmes MIMO sont souvent fabriqués à partir de composants à faible cout, ce qui peut conduire à des imperfections matérielles et se solder par du bruit de distorsion au niveau de la station de base et des utilisateurs. Afin d'élaborer un modèle plus fidèle pour ce type de système, nous prenons en considération le bruit causé par les défaillances matérielles durant le processus d'estimation des canaux.

Pour ce modèle de système, nous étendons l'estimateur à erreur quadratique moyenne minimale linéaire (LMMSE) aux systèmes LMIMO; cet estimateur est proposé ici pour un système multi-utilisateurs. L'estimateur LMMSE présenté prend en compte le bruit de distorsion aux deux extrémités et offre de meilleures performances en termes d'erreur d'estimation relative du rapport signal-bruit (SNR) par antenne comparé aux estimateurs utilisés pour les systèmes MIMO classiques, tels que les estimateurs LMMSE et à méthode des moindres carrés (LS). En outre, un calcul est effectué de la borne Cramer-Rao (CRB) du système.

Acknowledgments

First and foremost, I would like to express my sincerest gratitude to my advisor Professor Ioannis Psaromiligkos. I thank him for offering me the chance to be part of his research group, for his honest advises in difficult times, and for providing me with any kind of support during my so far studies at McGill. I would also like to thank my lab mates for their help, inspiration and good company. In addition I would also like to thank Ivan who initially worked with me on the subject of Large MIMO systems. Furthermore, I am grateful to my friends for their support. Of course, I would also like to thank Aikaterini Dimogiorgi for supporting me in every decision I've made and always being there for me. Finally, I thank my family whose lifelong love and support makes everything seem possible and means the world to me.

Chapter 1

Introduction

Over the last years, there has been an increasing interest on systems with large antenna arrays known as Massive or Large MIMO systems. Large MIMO systems include a base station with tens to hundreds of antennas serving a significantly lower number of users compared to the number of antennas. The main idea behind these systems is to extend conventional MIMO on a greater scale in order to achieve better performance mainly in terms of energy and spectrum efficiency [9].

Although Large MIMO systems have a lot of advantages there is a number of open problems that have attracted research interest over the last years. A notable difference between conventional and Large MIMO is the quality of the hardware. Large MIMO systems often use low quality components due to the large number of antennas which are particularly prone to impairments. These impairments may cause in-phase/quadrature (I/Q) imbalance, may limit the capacity of the system and possibly cause inaccuracies on the estimation of the channel by creating distortion at both ends of the communication link.

Channel estimation has always been a prominent research field in communication systems. Channel state information (CSI) is essential in order to use a communication link. There are several research efforts on that field for conventional as well as Large MIMO systems. However, the fact that the hardware is often assumed to be perfect may lead to an inaccurate system model for Large MIMO.

The objective of this thesis is to analyze the impact of hardware impairments in channel estimation for a Large MIMO system. Thus, the system model used assumes distortion caused by hardware impairments on both the base station and the users. We use as basis the LMMSE proposed in [5] in order to present a LMMSE estimator considering hardware impairments for a multi-user Large MIMO system. In order to evaluate the performance of the proposed estimator we compare it with the Least Squares (LS) and the Total Least Squares (TLS) estimators as well as the LMMSE used for conventional MIMO systems and we calculate the Cramer-Rao bound (CRB) on the channel estimation of the system.

This thesis is organized as follows. In Chapter 2, we include an overview on Large MIMO systems presenting their main advantages and challenges. Furthermore, we provide a theoretical background on channel estimation and on the CRB. In Chapter 3, the proposed LMMSE estimator is presented and compared to the impairment-ignoring LMMSE, the LS and the TLS estimators for a large MIMO system. Chapter 4 focuses on calculating the CRB of the system presented in Chapter 3. Finally, Chapter 5 includes a summary of the thesis and possible future problems.

Chapter 2

Background

2.1 Large MIMO systems

MIMO are systems with multiple antennas at both ends of the communication link. MIMO systems take advantage of the multi-path propagation and with the use of space-time signal processing they achieve better quality or data rate compared to single antenna systems [21]. As a result, MIMO technology has been popular in wireless communication systems over the last years. They have been an essential element of communication standards including the latest, LTE-Advanced.

MIMO systems consist of a base station communicating with a number of users in a cell environment. These systems are described as Multi-User MIMO or MU-MIMO. Due to the use of multiple antennas at the base station, they manage to achieve more degrees of freedom compared to single antenna systems [9]. These degrees of freedom occur from the fact that there are more antennas at the base station than the number of users served concurrently. As a result, MU-MIMO systems offer significant advantages, such as increased data rate and enhanced reliability. The latest MU-MIMO systems use 8-10 antennas at the base station [24]. The reasons that most of the systems are limited to that number of antennas are hardware complexity, energy cost as well as the physical space needed to accommodate the antennas.

Over the last few years, there has been an extensive research on systems with a large number of antennas at the base station. Such systems are called Large or Massive MIMO systems. The main idea behind these systems is to use those extra antenna elements in order to direct the signal more precisely to the receivers either in a line of sight environment or in a rich scattering one [24].

Large MIMO systems have been proven to have advantages over conventional MIMO systems [9], mainly in terms of improved performance and energy efficiency and they can be the basis for the development of future networks. Working in a larger scale comes with a number of new challenges as well as a number of traditional research questions that have to be revisited, some of them being channel estimation, signal detection and radio frequency (RF) chain management.

2.1.1 Challenges

In this section the main difficulties in developing Large MIMO systems will be presented. The system model that will be used includes a base station with N_r antennas serving N_t terminals with one antenna each.

The first problem that was noticed in Large MIMO was the communication link. In conventional MIMO systems, a duplex communication link is established which can either be frequency division duplex (FDD) or time division duplex (TDD). In FDD transmission, two frequency bands separated by a guard band are used, one for the uplink and one for the downlink. On the other hand, in TDD, there is one frequency band for both links and the base station and the terminals alternate their transmission over time.

In conventional MIMO systems, for a pilot based method, the terminals send pilots to the base station, based on which the base station estimates the channel for the uplink. For TDD, since the uplink and the downlink share the same bandwidth the channel is reciprocal, which means that the downlink channel is the reverse of the channel estimated for the uplink and as such it can be easily calculated by the base station without the use of extra pilots. On the other hand, for FDD, the downlink channel has to be estimated. This creates two problems. Firstly, optimal downlink pilots have to be orthogonal between the antennas, which means that the amount of time-frequency resources scales with the number of the antennas at the base station. Secondly, the amount of channel responses that each user has to calculate also scales with the number of antennas at the base station. Hence, both in term of resources and complexity, FDD is more difficult in Large MIMO compared to conventional MIMO. Despite the challenges, there have been research efforts on implementing FDD for Large MIMO. The main proposed solution suggests to map the highly correlated antennas at the base station to a single value representing that group of antennas [15] having a system model with significantly less antennas at the base station. As a result, the effective channel matrix will be smaller and savings in the downlink training can be achieved.

Channel estimation in TDD transmission is easily performed since the terminals send orthogonal pilots and the base station performs the estimation for the uplink. The base station also estimates the downlink channel due to channel reciprocity [19], since both ends operate on the same frequency band. However, the communication link also consists of the antennas, radio-frequency (RF) chains and other transceivers' hardware. Hardware on both ends may cause the transmitted and received signals to have phase and amplitude differences, which creates different channel realizations at the uplink and the downlink. In order to solve that problem, [31] proposes calibration of the hardware as a solution. Hardware calibration can be performed with the use of an external reference in order to compensate for the differences mentioned previously [28]. Researchers in [31] propose the use of relative calibration which means that as long as each base station antenna's channel measurement deviates from the real one by the same multiplicative factor, channel reciprocity holds.

At this point, it is important to specify that the cell that the base station is working in, is surrounded by other $N_c - 1$ cells. In order to perform the channel estimation in TDD using orthogonal pilots we would need $N_c \times N_t$ pilot symbols for each user, a number which becomes very large as N_c grows. If the orthogonal pilots are less than $N_t \times N_c$, the pilots from one cell will have to be reused in another, a phenomenon called *pilot contamination*. In pilot contamination, the base station overhears the pilot transmission of a terminal from another cell which uses the same pilots as one within its cell. Consequently, interference from that terminal occurs during the channel estimation. Pilot contamination can occur for conventional MIMO as well but is not as common since the number of terminals that can be served simultaneously is significantly less compared to Large MIMO. Many research groups have worked on pilot contamination as it is a prominent problem in Large MIMO. A number of solutions have been proposed which vary from new precoding techniques to non-linear estimation methods [12].

2.1.2 Advantages

As mentioned before, large MIMO offers numerous advantages over conventional MIMO. Some of the most important ones are:

Energy efficiency

In conventional MIMO, in a non line of sight (LOS) environment the field strength is usually focused by the transmit antennas to a geographical point where the transmitted signals add constructively. In [24], it is shown that it is possible with large antenna arrays to focus the electromagnetic field to the point where the receiver is with great precision. As a result, there are no unnecessary power peaks in the general area around the user, something that can happen in a conventional MIMO system. Hence, Large MIMO are allowed to have multiple users in a small area, since they will not cause distortion to each other.

No added computational complexity

When the idea of Large MIMO was introduced, it was initially claimed that by increasing the number of antenna elements the resource allocation and signal processing would be significantly more complicated. Nevertheless, that is not the case. Resource allocation includes the process of assigning time and frequency resources to the users. Finding the best sub-carriers for each user can be challenging in conventional MIMO. Small-scale fading causes channel variations and varies at the order of milliseconds creating the need for accurate and fast resource allocation algorithms. However, in Large MIMO systems the channel variations mainly depend on large-scale fading in the time domain, which varies significantly slower than the small-scale fading, rendering the resource allocation techniques used in conventional MIMO unnecessary [6]. The researchers of [6] also proved that the complexity of signal processing is indeed increased but only scales linearly with the number of antennas at the base station.

Resistance to fading / interference

Fading is a limiting factor in wireless transmissions especially in high scattering environments. In conventional MIMO, a fading dip occurs when the signal travels through multiple paths that add destructively and as a result the signal's strength becomes significantly lower. In Large MIMO, this problem is less common since the probability that the signals from all the antennas add destructively becomes lower as the number of antennas grows [9]. In the case of interference, the excess degrees of freedom can be used to cancel signals from external sources, like users from other cells.

Inexpensive or Impaired Hardware

While in conventional MIMO impaired hardware may lead to problems [9], in large MIMO systems, they usually do not. That occurs since one defective antenna among 200 is less of a problem compared to one impaired element among 4 or 8 antennas. Moreover, instead of having a 40 Watt amplifier per antenna [30], it is more cost efficient to use more, smaller, even in the order of milli-watts, amplifiers. As mentioned previously, the use of a large number of small antennas enables focusing of the electromagnetic field to a specific point with great precision.

2.1.3 Research Questions

As described thus far, Large MIMO offer new advantages as well as new challenges compared to conventional MIMO. While unveiling new problems is one aspect, there is a number of traditional questions, for which the solutions proposed for conventional MIMO do not directly apply to large systems. Channel estimation under hardware impairments, which is analyzed in Chapter 3, is among them. Other examples of research questions that have to be revisited are described below.

Detection

Detection is one of the most vital parts of a communication process. Achieving a near optimal detection with low complexity can be difficult as the number of antennas grow, since algorithms commonly used in conventional MIMO do not suit large MIMO. Detectors which can achieve near optimum performance in conventional MIMO, such as the sphere decoder, have high complexity in large arrays, while the ones with low complexity, such as zero-forcing (ZF), achieve relatively low performance [29]. As a result, algorithms which scale well with the antenna array's size are needed. It has been found that heuristic detectors scale well as the number of antennas grow while having low complexity [7]. The main two examples are the Likelihood Ascent Search (LAS) and the Reactive Tabu Search (RTS) algorithms.

In order to present these methods, we first present the system model, given by $Y = HX + N$ where $H \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix, $X \in \mathbb{C}^{N_t}$ is the transmitted signal, $Y \in \mathbb{C}^{N_r}$ is the received signal and $N \in \mathbb{C}^{N_r}$ is the complex Gaussian noise. The number of antennas at the transmitter and the receiver are N_t and N_r, respectively, and are both in the order of tens to hundreds.

Both the LAS and the RTS algorithm begin with an initial solution vector $X_{(0)} = BY$, where $B \in \mathbb{C}^{N_t \times N_r}$ can be a matched filter (MF) or a zero-forcing (ZF) filter. A neighborhood of this vector is defined arbitrarily, for example it can include all the vectors that differ from $X_{(0)}$ in one entry. The detector computes the value of the maximum likelihood cost function, $||Y - HX_{(i)}||^2$, of the current solution vector $X_{(0)}$ as well as that of each neighboring vector. The LAS algorithm defines as $X_{(1)}$ the neighbor with the best cost function, and finds its own neighborhood. The algorithm repeats the previous steps until there is no neighbor who achieves better performance in terms of maximum likelihood cost than the current solution. The RTS is similar to the LAS but continues for a pre-defined number of iterations. This happens in order to avoid being trapped in the first local minimum it visits. Having stored all the solutions that it visited, the RTS chooses the best one in terms of maximum likelihood cost.

These algorithms have overall low complexity with the most complex operations being the calculation of $X_{(0)}$ which is in the magnitude of $O(N_t N_r)$ and the search part being $O(N_t)$, both calculated in [7]. Thus, the overall complexity is $O(N_t N_r)$ per transmitted QAM or PAM symbol. As for its performance, it improves as the number of antennas increases for the same number of users [27].

There have been proposed some variations of the LAS and the RTS detectors with the main ones coming from [8] where it is suggested using random initial vectors instead of using MF or ZF in order to lower the complexity. Another proposed idea includes the use of multiple instances of LAS or RTS algorithms simultaneously in order to have a larger variety of possible solutions [16].

Radio Frequency (RF) chains

The management of the radio frequency chains in a conventional MIMO system has been a subject of interest, especially in systems with $8-12$ antennas, due to the fact that RF chains are very costly it terms of power and hardware. In order to mitigate this cost, antenna selection has been proposed. It is common to have a system with N_t and N_r antennas but only L_t and L_r RF chains at the transmitter and the receiver, respectively, with $L_t < N_t$ and $L_r < N_r$. In [25], it is suggested that antenna selection should be performed in order to achieve the maximum possible signal-to-noise ratio (SNR). Hence, the L_t and L_r antennas with the highest channel gain are chosen by exhaustive search.

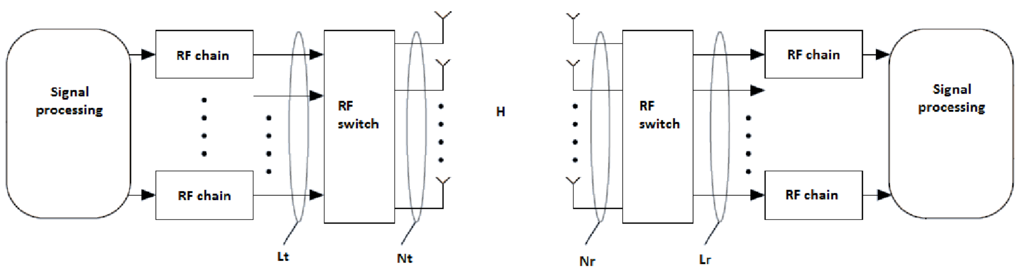


Figure 2.1: System with L_t RF chains and N_t antennas at the transmitter and L_r RF chains and N_r antennas at the receiver.

In MIMO systems, the circuit power consumption scales linearly with the amount of RF chains used and it can become a significant problem for large antenna arrays. Hence, the antenna selection process can be even more necessary compared to conventional MIMO systems. In the specific case of having a system with a base station with many antennas and a user with one, [22] proves that the optimal configuration with respect to ergodic capacity is to use half of the existing antennas at the base station when we have perfect CSI. In more general cases, it is stated that the optimal RF chain management is achieved by maximizing the average rate $C = \max_{\phi\in\Phi}\max_{\psi\in\Psi} E_H[C(H_{\phi\psi})]$ where $C(H_{\phi\psi})$ is the capacity of the channel $H_{\phi\psi}$. The channel $H_{\phi\psi}$ defines the link between the set ϕ of L_t antennas at the transmitter and the set ψ of L_r antennas at the receiver, where Φ and Ψ include all the possible sets of antenna combinations at the transmitter and the receiver, respectively.

In [18], it is proposed to use a single RF chain configuration at the transmitter. The proposed transmitter operates with orthogonal frequency-division multiplexing (OFDM) and achieves high power efficiency with the use of clipping techniques, in order to limit the total transmitted power.

As stated previously, the hardware used in Large MIMO is usually inexpensive with the exception of RF chains. This happens because communication systems are more sensitive to RF chain impairments than other hardware elements, such as amplifiers or filters. More specifically, non-ideal RF chains can cause I/Q imbalance which distorts the transmitted signal. I/Q imbalance transforms a circular signal, whose probability density function (PDF) is invariant to rotations, to non-circular. As a result, it degrades the receiver's performance since many detection algorithms rely on the circularity of the received signal [14]. More specifically, I/Q imbalance distorts the received signal as, $y_{imb} = K_1 y + K_2 y^*$, where y is the ideal received signal and y^* its complex conjugate, K_1 and K_2 are the I/Q imbalance coefficients. To address this challenge, [2] presents a system which uses widely-linear (WL) beamforming. This method includes a signal transformation at the receiver in the form of $y_{WL} = w_1^H y_{imb} + w_2^H y_{imb}^*$ where the imbalanced received signal and its complex conjugate are multiplied by weights, w_1 and w_2 respectively and then added up, in order to eliminate the I/Q imbalance coefficients. The result is $y_{WL} = (w_1^H K_1 y + w_2^H K_2^* y) + (w_1^H K_2 y^* + w_2^H K_1^* y^*)$, where the first term of the sum is the wanted received signal and the second is the unwanted interference which can be eliminated with the use of the proper weight coefficients.

2.2 Channel Estimation

CSI is essential in order to establish a communication link. Channel estimation is the procedure with which we obtain CSI. Many common methods use a set of known pilots sent from one end of the communication link to the other in order to acquire knowledge of the channel. Well known pilot-based methods, used in MIMO, are the least squares (LS) and the linear minimum mean square error (LMMSE) estimators [4]. In order to present these estimators, the system model will be expressed as

$$Y = HX + N \tag{2.1}$$

where X is the $N_t \times T$ transmitted pilots with N_t being the number of transmit antennas and T the pilot length with $T \geq N_t$ in order to have more equations than unknowns. The received signal, Y, is a $N_r \times T$ matrix, $H \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix and $N \in \mathbb{C}^{N_r \times T}$ is the Gaussian noise matrix with 0 mean and $\sigma_N^2 I$ variance, with I being the identity matrix.

2.2.1 Least Squares (LS)

The LS estimator is a method with low complexity which does not require any knowledge about the channel statistics. The LS estimator is solving the problem:

$$\min_{\hat{H}_{LS}} ||Y - \hat{H}_{LS}X||^2 \tag{2.2}$$

The result of which is :

$$\hat{H}_{LS} = YX^{\dagger} \tag{2.3}$$

where $X^{\dagger}$ is the pseudo-inverse of X:

$$X^{\dagger} = X^H(XX^H)^{-1} \tag{2.4}$$

In order to find the optimal set of pilots we solve [4]:

$$\min_{X} E[||H - \hat{H}_{LS}||_F^2] \text{ subject to } ||X||^2 = P_X \tag{2.5}$$

where $E[\cdot]$ denotes expectation over N and P_X is the given power constant of the pilots. Using (2.1) and (2.3) we get:

$$\begin{aligned} E[||H - \hat{H}_{LS}||_F^2] &= E[||NX^{\dagger}||_F^2] \\ &= \sigma_N^2 Itr(XX^H)^{-1} \end{aligned} \tag{2.6}$$

where $tr((XX^H)^{-1})$ is the trace of the matrix $(XX^H)^{-1}$. Hence, the optimization problem in (2.5) can be modified as:

$$\min_{X} tr(XX^H)^{-1} \text{ subject to } tr(XX^H) = P_X \tag{2.7}$$

For a training sequence to be optimal with respect to (2.7) it has to satisfy the given equation as proven in [3]:

$$XX^H = \frac{P_X}{T} I \tag{2.8}$$

where I is the $N_t \times N_t$ identity matrix. XX^H is Hermitian and positive semi-definite, thus, any training matrix with orthogonal rows with norm $\sqrt{\frac{P_X}{T}}$ can be a solution to (2.7).

2.2.2 Total Least Squares (TLS)

TLS is an extension of the LS estimator with the main difference being that there are perturbations at the transmitted signal as well.

In order to present TLS the system model in (2.1) is re-written in vector form stacking the rows into one column as:

$$y = (X + V)h + n \tag{2.9}$$

where X is the $N_rT \times N_rN_t$ transmitted pilots with $T \geq N_t$; V is the $N_rT \times N_rN_t$ distortion at the user-end; y is the $N_rT \times 1$ received signal vector; $h \in \mathbb{C}^{N_rN_t \times 1}$ is the channel vector and $n \in \mathbb{C}^{N_rT \times 1}$ is a Gaussian noise vector.

Initially we want to minimize the perturbations, which are the noise variables at the transmitted and the receive signal:

$$\min_{\hat{V},\hat{n}} ||[\hat{V} \quad \hat{n}]||_F \text{ subject to } y - \hat{n} \in \text{Range}(X + \hat{V}) \tag{2.10}$$

where $||\cdot||_F$ denotes the Frobenius norm. Having the values of $\hat{V}$ and $\hat{n}$ we solve an equivalent problem to the LS one:

$$\min_{\hat{h}_{TLS}} ||(X + \hat{V})\hat{h}_{TLS} - (y - \hat{n})||^2 \tag{2.11}$$

In order to solve (2.11) we define $A = [X \quad y]$ and $U^T AV = \Sigma$ its singular value decomposition (SVD). The $N_r N_t + 1$ singular values of A are $\sigma_1, \sigma_2, \ldots, \sigma_{N_r N_t}, \sigma_{N_r N_t+1}$, with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{N_r N_t} \geq \sigma_{N_r N_t+1}$. The TLS estimator, $\hat{h}_{TLS}$, is equal to [10]:

$$\hat{h}_{TLS} = (X^H X - \sigma_{N_r N_t+1} I)^{-1} X^H y \tag{2.12}$$

In order to have a TLS solution the matrix, $(X^H X - \sigma_{N_r N_t+1} I)$ has to be non-singular. According to [10], that occurs when $\sigma_{N_r N_t+1} > \sigma_{N_r N_t}$ which ensures that the last eigenvalue has multiplicity one. In the case of $\sigma_{N_r N_t+1} = 0$ the results of the LS and the TLS estimators are identical.

2.2.3 Linear Minimum Mean Square Estimator (LMMSE)

Although the LS estimator is a well known method, it does not always provide the best results. When the channel statistics are known it is preferable to use the linear MMSE estimator since it achieves better performance [13]. This estimator can be expressed in the following form [4]:

$$\hat{H}_{LMMSE} = YA \tag{2.13}$$

In order to minimize the MSE we must find the optimal value of A:

$$\arg\min_{A} E[||H - \hat{H}_{LMMSE}||_F^2] \tag{2.14}$$

The estimation error becomes:

$$\begin{aligned} e &= E[||H - \hat{H}_{LMMSE}||_F^2] \\ &= tr(R_H) - tr(R_H X A) - tr(A^H X^H R_H) + tr(A^H(X^H R_H X + \sigma_N^2)A) \end{aligned} \tag{2.15}$$

where $R_H = E[(H - E(H))(H - E(H))^H]$ is the covariance of the channel and $R_Y = E[(Y - E(Y))(Y - E(Y))^H]$ the covariance matrix of the received signal. The optimal value of A is given by solving $\nabla_A e = 0$, to obtain:

$$A_o = (X^H R_H X + \sigma_N^2 I)^{-1} X^H R_H \tag{2.16}$$

Thus, the LMMSE estimator from (2.11) becomes:

$$\hat{H}_{LMMSE} = Y(X^H R_H X + \sigma_N^2 I)^{-1} X^H R_H \tag{2.17}$$

2.3 Cramer-Rao Bound

In estimation theory, the Cramer-Rao bound (CRB) or Cramer-Rao lower bound (CRLB) expresses a lower bound on the variance of estimators. The bound is primarily used for unbiased estimators but can be used for biased ones with known bias as well with the proper modifications. The covariance matrix of a vector of unbiased estimators $\hat{\vartheta} = [\hat{\vartheta}_1, \hat{\vartheta}_2, \ldots, \hat{\vartheta}_N]^T$ of a parameter column vector $\vartheta = [\vartheta_1, \vartheta_2, \ldots, \vartheta_N]^T$ is lower bounded by the inverse of the Fisher information matrix [23]:

$$\text{cov}(\hat{\vartheta}) \geq I(\vartheta)^{-1} \tag{2.18}$$

where the matrix inequality $A \geq B$ means that the matrix $A - B$ is positive semidefinite [23]. Since the trace of a semidefinite matrix is greater or equal to 0, (2.18) becomes:

$$\text{tr}(\text{cov}(\hat{\vartheta})) \geq \text{tr}(I(\vartheta)^{-1}) \tag{2.19}$$

The Fisher information is the answer to the question how much information can a vector of data $x = [x_1, x_2, \ldots, x_N]$ provide about an unknown parameter. We consider a vector of unknown parameters ϑ and x is the measurements vector with a known PDF $p(x; \vartheta)$, the Fisher information measures the amount of information about ϑ that can be collected by observing x and is a $N \times N$ matrix whose $(i, j)_{th}$ entry is given by:

$$I_{ij}(\vartheta) = E[\frac{\partial}{\partial \vartheta_i} \text{log} p(x; \vartheta) \frac{\partial}{\partial \vartheta_j} \text{log} p(x; \vartheta)] = -E[\frac{\partial^2}{\partial \vartheta_i \partial \vartheta_j} \text{log} p(x; \vartheta)] \tag{2.20}$$

where $\log p(x;\vartheta)$ is the natural logarithm of the pdf of x and $E[\cdot]$ denotes the expectation over x. The CRB can be considered as a benchmark in order to compare unbiased estimators. The ones that achieve it are called minimum variance unbiased estimators (MVUE).

Chapter 3

Channel Estimation for a Large MIMO System under Hardware Impairments

As mentioned in Chapter 2, Large MIMO systems offer numerous advantages over conventional MIMO systems. There has been extensive research on various topics with one of the most important ones being channel estimation.

One of the major differences between conventional and Large MIMO systems is the level of hardware impairments. The transceivers on both ends of the communication link consist of various components, such as amplifiers, filters etc. Since Large MIMO use inexpensive hardware, there can be imperfections on any of these. These impairments may cause I/Q-imbalance, limit the capacity of the system and cause inaccuracies on the estimation of the channel by creating distortion at both ends of the communication link [9]. In order to characterize the level of distortion, [28] uses the error vector magnitude (EVM) which is defined as $EVM = \frac{E[||\tilde{d}-d||^2]}{E[||d||^2]}$, where d is the ideal transmitted signal and $\tilde{d}$ is the distorted one. The latest LTE standards have strict requirements and use transceivers with low EVM, usually below 0.08 [5]. However, Large MIMO can use hardware with more relaxed constraints and EVM in the interval [0.05, 0.15].

In this chapter, the impact of hardware impairments at the base station and the users, on the process of channel estimation, will be studied building on the work of [5] and extending it to multi-user systems. The nature of these impairments can vary and it is very difficult to accurately model the distortion from each impaired component [28]. Researchers in [11] have conducted experiments and measurements in order to model the noise from impaired hardware. Although there are multiple sources of distortion it is shown that modeling the noise as Gaussian accurately reflects the real-world residual impairments. More specifically, the total noise caused at the base station and the users is modeled as Gaussian noise with 0 mean and standard deviation defined by the EVM in the interval [0.05, 0.15]. This model is adopted for this thesis, as it captures the main characteristics of impaired hardware and is experimentally verified.

3.1 Channel estimation for a single-user Large MIMO system

In this section, we examine a system with a base station with N_r antennas and a single antenna user. We explain in detail the LMMSE estimator proposed in [5] as it will be the basis for the proposed LMMSE estimator for multi-user systems in Section 3.2. In order to acquire CSI using TDD transmission, the user sends pilots to the base station. With these pilots the base station performs the channel estimation for the uplink. Usually in

Large MIMO systems, the users have only knowledge of the statistics of the channel [20]. Another method includes the base station usong the uplink pilots to estimate the downlink channel, due to channel reciprocity, and send feedback to the users. Researchers in [15] have proposed that the users can perform the estimation too, by mapping multiple antennas of the base station into one value in order to have a system model with less antennas at the base station.

The uplink system model is given by:

$$y = h(x + i_{UE}) + i_{BS} + n \tag{3.1}$$

where $x \in \mathbb{C}$ is the deterministic pilot signal; h is the $N_r \times 1$ Gaussian channel matrix with 0 mean and covariance matrix $R = E[hh^H]$. The additive Gaussian noise, $n \in \mathbb{C}^{N_r \times 1}$, has zero mean and covariance $\sigma_n^2 I$. The noise from the hardware impairments at the user-end is $i_{UE} \in \mathbb{C}^{1\times 1} \sim \mathcal{CN}(0, \sigma_{UE}^2)$ and $i_{BS} \in \mathbb{C}^{N_r \times 1} \sim \mathcal{CN}(0, \sigma_{BS}^2)$ is the noise from the hardware impairments at the base station.

The variance of the distortion variables is given by [5]:

$$\sigma_{UE}^2 = \kappa_{UE}\, p \tag{3.2}$$

$$\sigma_{BS}^2 = \kappa_{BS}\, p\, diag(|h_1|^2, \dots, |h_{N_r}|^2) \tag{3.3}$$

where κ_{UE} and κ_{BS} characterize the level of impairments at the user-end and the base station, respectively. The transmitted signal's power is given by $p = |x|^2$ and $\text{diag}(|h_1|^2, \dots)$ is the $N_r \times N_r$ diagonal matrix with $|h_i|^2$ values at the diagonal.

The conventional LMMSE estimator is given in Chapter 2:

$$\hat{h}_{cLMMSE} = x^* R(pR + \sigma_n^2 I)^{-1} y \tag{3.4}$$

However, for the system presented in (3.1) we have to take into consideration the interference caused by hardware impairments for the LMMSE estimator in order to achieve better performance. Hence, the value of $R_y = E[yy^H]$, is:

$$\begin{aligned} R_y &= E[(h(x+i_{UE}) + i_{BS} + n)(h(x+i_{UE}) + i_{BS} + n)^H] \\ &= E[h(x+i_{UE})(h(x+i_{UE}))^H] + E[hi_{BS}^H] + E[i_{BS}h^H] + E[i_{BS}i_{BS}^H] + E[nn^H] \\ &= E[hxx^H h^H + hi_{UE}i_{UE}^H h^H] + E[hi_{BS}^H] + E[i_{BS}h^H] + E[i_{BS}i_{BS}^H] + E[nn^H] \\ &= p(1+\kappa_{UE})R_h + E[hi_{BS}^H] + E[i_{BS}h^H] + p\kappa_{BS}R_{diag} + \sigma_n^2 I \end{aligned} \tag{3.5}$$

where the terms $h(x+i_{UE})n^H$, $i_{BS}n^H$, $h(x+i_{UE})^H n$ and $i_{BS}^H n$ are equal to 0 since they are products of uncorrelated random variables and R_{diag} is the $N_r \times N_r$ diagonal matrix with values $|R_{11}|, \ldots, |R_{N_rN_r}|$. From (3.3), the covariance matrix of σ_{BS}^2 is equal to $\kappa_{BS}pdiag(|h_1|^2, \ldots, |h_{N_r}|^2)$. However, the values of the channel vector h are unknown. Therefore, the values of the channel covariance R_{diag} are used for the proposed estimator since the channel statistics are known.

The distortion at the base station i_{BS} is rewritten as $z\sigma_{BS}$ with $z \sim \mathcal{CN}(0,1)$ since i_{BS}

is Gaussian. Hence, (3.7) becomes:

$$\begin{aligned}
R_y &= E[(h(x+i_{UE})+i_{BS}+n)(h(x+i_{UE})+i_{BS}+n)^H] \\
&= p(1+\kappa_{UE})R_h + E[h\sigma_{BS}z^H] + E[\sigma_{BS}zh^H] + p\kappa_{BS}R_{diag} + \sigma_N I \\
&= p(1+\kappa_{UE})R_h + p\kappa_{BS}R_{diag} + \sigma_n^2 I
\end{aligned}$$

since h and z are uncorrelated with zero mean.

The proposed LMMSE estimator in [5], taking into consideration the hardware impairments becomes from (3.4):

$$\hat{h}_{pLMMSE} = x^* R(p(1+\kappa_{UE})R_h + p\kappa_{BS}R_{diag} + \sigma_n^2 I)^{-1} y \tag{3.6}$$

The MSE of which is:

$$MSE = E[||\hat{h}_{pLMMSE} - h||_2^2] = tr(C) \tag{3.7}$$

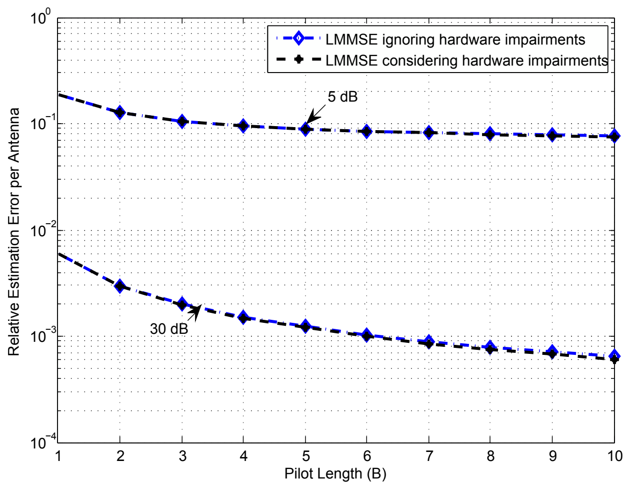
where C is the error covariance matrix calculated as follows:

$$\begin{aligned}
C &= E[(\hat{h}_{pLMMSE} - h)(\hat{h}_{pLMMSE} - h)^H] \\
&= E[(x^* R R_y^{-1} y - h)(x^* R R_y^{-1} y - h)^H] \\
&= E[x^* R R_y^{-1} y y^H (R_y^{-1})^H R^H x^T] - E[x^* R R_y^{-1} y h^H] - E[h y^H (R_y^{-1})^H R^H x^T] + E[hh^H] \\
&= pR(R_y^{-1})^H R^H - pRR_y^{-1}R - pR(R_y^{-1})^H R^H + R \\
&= R - pRR_y^{-1}R
\end{aligned} \tag{3.8}$$

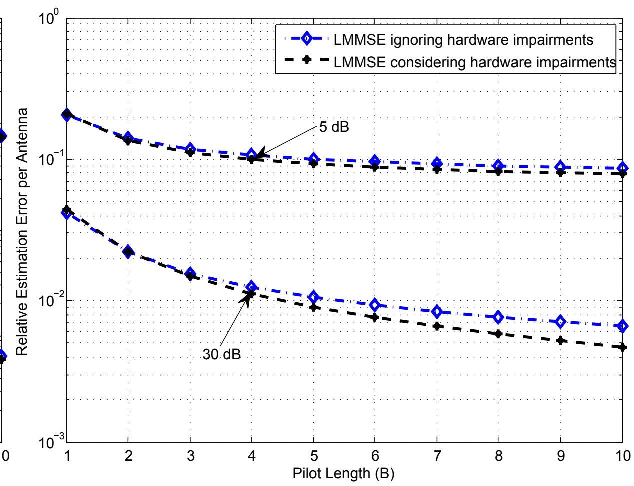
Thus far we have presented the estimator using one pilot. Having T pilots can be translated into performing T separate LMMSE estimators and averaging over them [5].

$$\hat{h}_{avg} = \frac{1}{T}\sum_{i=1}^{T}\hat{h_i} \tag{3.9}$$

The performance of the LMMSE estimator taking into consideration the hardware impairments compared to the traditional LMMSE estimator as the pilot length varies from 1 to 10 is shown below.



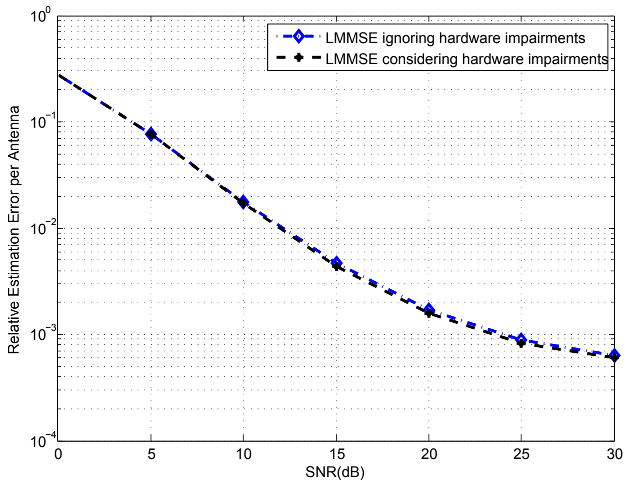
(a) $\kappa_{BS} = \kappa_{UE} = 0.05^2$



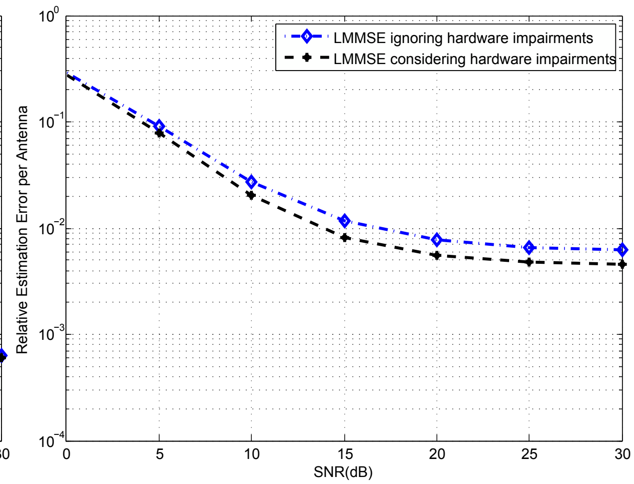
(b) $\kappa_{BS} = \kappa_{UE} = 0.15^2$

Figure 3.1: Estimation error per antenna for conventional LMSSE and LMMSE considering hardware impairments over pilot length for SNR equal to $5dB$ and $30dB$, where SNR is defined as ratio of the signal over the AWGN noise. Single-user system with 50 antennas at the base station.

Figure 3.1 presents the estimation error per antenna which is defined as $\frac{E[||h-\hat{h}||^2]}{N_r}$ over 1000 Monte-Carlo realizations of the channel, the distortion on both ends and the Gaussian noise.



(a) $\kappa_{BS} = \kappa_{UE} = 0.05^2$



(b) $\kappa_{BS} = \kappa_{UE} = 0.15^2$

Figure 3.2: Estimation error per antenna for conventional LMMSE and LMMSE considering hardware impairments estimators over SNR. Single-user system with 50 antennas at the base station.

In Figure 3.2 above, the conventional LMMSE estimator is compared to the LMMSE estimator taking into consideration hardware impairments as SNR varies from $0dB$ to $30dB$ over 1000 Monte-Carlo realizations of the channel, the noise on both ends and the Gaussian noise in a system with 50 antennas at the base station and one single antenna user. The covariance matrix, R of the channel is a symmetric toeplitz matrix with value 1 at the diagonal and correlation factor $r = 0.7$ such as $R_{(i,j)} = r^{|i-j|}$. This model describes a uniform linear array (ULA) where the correlation factor determines the eigenvalue spread in R [20]. The level of the impairments is chosen to be 0.05^2 as the lower bound where there is a difference in the performance of the two LMMSE estimators and 0.15^2 as the upper bound of the parameter defined in [17].

3.2 Channel estimation for a multi-user Large MIMO system

In order to extend the model to a multi-user system, a Large MIMO system is considered with N_r antennas at the base station and N_t users with one antenna each. The least squares (LS) and the total least squares (TLS) estimators for a multi-user system will be presented. In order to perform the pilot-based channel estimation at the uplink, the following system model is considered:

$$Y = H(X + I_{UE}) + I_{BS} + N \tag{3.10}$$

where $X \in \mathbb{C}^{N_t \times T}$ is the deterministic pilot signal; T is the length of the training sequence and is set to be equal or larger than N_t in order to have more equations than unknowns. The average power of the transmitted signal is p. The Gaussian channel H is a $N_r \times N_t$ matrix. The additive Gaussian noise $N \in \mathbb{C}^{N_r \times T}$ has zero mean and covariance matrix $\sigma_N^2 I$. The distortion that occurs at the user-end, $I_{UE} \in \mathbb{C}^{N_t \times T}$, consists of N_t independent Gaussian random variables with zero mean and covariance matrix Σ_{UE}^2 and $I_{BS} \in \mathbb{C}^{N_r \times T} \sim \mathcal{CN}(0, \Sigma_{BS}^2)$ is the distortion that occurs at the base station.

The channel matrix H can be written as:

$$H = R_r^{1/2} H_w R_t^{1/2} \tag{3.11}$$

with $R_r \in \mathbb{C}^{N_r \times N_r}$ and $R_t \in \mathbb{C}^{N_t \times N_t}$ being respectively the transmit and receive spatial correlation matrices given by $R_{(i,j)_r} = r_r^{|i-j|}$ and $R_{(i,j)_t} = r_t^{|i-j|}$, respectively; where r_r and r_t

are the receive and transmit correlation factors. Moreover, $H_w \in \mathbb{C}^{N_r \times N_t}$ has i.i.d. complex Gaussian entries with zero mean and 0.5 variance [1]. Hence, the channel correlation matrix is a symmetric toeplitz given by:

$$R = R_t \otimes R_r \tag{3.12}$$

where $\otimes$ is the Kronecker product.

It will be helpful to rewrite (3.10) as follows:

$$vec(Y) = (\bar{X} + \bar{I}_{UE})vec(H) + vec(I_{BS}) + vec(N) \tag{3.13}$$

where $vec(\cdot)$ represents the column vectorization of the matrix argument. The transmitted signal, $\bar{X}$ is a $N_rT \times N_rN_t$ Toeplitz matrix with first row equal to $[x_{1,1}\underbrace{0\ldots0}_{N_r}\ldots x_{N_t,1}\underbrace{0\ldots0}_{N_r}]$ and first column equal to $[x_{1,1}\underbrace{0\ldots0}_{N_r}\ldots x_{1,T}\underbrace{0\ldots0}_{N_r}]$, with $x_{i,j} = X_{(i,j)}$. The user-end noise matrix, $\bar{I}_{UE}$, has the same form with first row equal to $[i_{UE_{1,1}}\underbrace{0\ldots0}_{N_r}\ldots i_{UE_{N_t,1}}\underbrace{0\ldots0}_{N_r}]$ and first column equal to $[i_{UE_{1,1}}\underbrace{0\ldots0}_{N_r}\ldots i_{UE_{1,T}}\underbrace{0\ldots0}_{N_r}]$, with $i_{UE_{i,j}} = I_{UE_{(i,j)}}$.

The covariance of the distortion of each user, i_{UEi} is given by a $T \times T$ matrix:

$$\Sigma_{UE}^2 = p \begin{bmatrix} \kappa_{UE} & q_{UE} & \dots & q_{UE} \\ q_{UE} & \kappa_{UE} & \dots & q_{UE} \\ \dots & \dots & \ddots & \dots \\ q_{UE} & \dots & q_{UE} & \kappa_{UE} \end{bmatrix} \tag{3.14}$$

where q_{UE} is the correlation between two noise samples from the same user, $q_{UE} = E[i_{UE_{ij}} i^H_{UE_{ij'}}]$ with $j \neq j'$, and is in the interval $[0,\ \kappa_{UE}]$. As mentioned before, the source of the noise can vary. A user with one source of distortion will have highly correlated samples, while a user with multiple impaired components will have lower values of q_{UE}. We assume that all the users will have the same hardware and as a result will develop noise with the same covariance matrix Σ_{UE}^2.

The covariance of the distortion at the base station, $vec(I_{BS})$, is similarly to the single-user case given by:

$$\Sigma_{BS}^2 = \kappa_{BS}\, p\, diag(|h_1|^2, \dots, |h_{N_t N_r}|^2) \tag{3.15}$$

where the matrix $diag(|h_1|^2, \dots, |h_{N_t N_r}|^2)$ is the $N_r N_t \times N_r N_t$ diagonal matrix including the elements of the channel $vec(H)$ along its diagonal. The values of κ_{UE} and κ_{BS} determine the degree of the hardware impairments at the user-end and the base station, respectively, and are usually in the range $[0.05^2,\ 0.15^2]$ [17].

The LS and TLS estimators from Chapter 2 are modified accordingly for the system model in (3.15). The LS estimator of the MIMO channel $vec(H)$ is given by [1]:

$$vec(\hat{H}_{LS}) = \bar{X}^{\dagger} vec(Y) \tag{3.16}$$

The TLS estimator of $vec(H)$ is [4]:

$$\hat{H}_{TLS} = (\bar{X}^H \bar{X} - \sigma_{NN_t+1} I)^{-1} \bar{X}^H vec(Y) \tag{3.17}$$

where $\sigma_{N_r N_t+1}$ is the last singular value of the matrix $[\bar{X} \;\; vec(Y)]$. The TLS solution exists as long as $\sigma^{\bar{X}}_{N_r N_t} > \sigma_{N_r N_t+1}$ as explained in Chapter 2.

The conventional linear MMSE estimation of the MIMO channel H is given by [1]:

$$vec(\hat{H}_{LMMSE}) = R_{HY} R_Y^{-1} vec(Y) \tag{3.18}$$

where

$$R_{HY} = E[vec(H) vec(Y)^H] \tag{3.19}$$

$$R_Y = E[vec(Y) vec(Y)^H] \tag{3.20}$$

Equation (3.21) can be expanded as:

$$\begin{aligned}
R_{HY} &= E[vec(H)(vec(H)^H(\bar{X}^H + \bar{I}_{UE}^H) + vec(I_{BS})^H + vec(N)^H)] \\
&= E[vec(H)vec(H)^H\bar{X}^H] + E[vec(H)vec(H)^H\bar{I}_{UE}^H] \\
&+ E[vec(H)vec(I_{BS})^H] + E[vec(H)vec(N)^H)] \\
&= R\bar{X}^H
\end{aligned} \tag{3.21}$$

Furthermore, equation (3.22) can be extended as:

$$\begin{aligned}
R_Y &= E[(\bar{X}vec(H) + vec(N))(vec(H)^H\bar{X}^H + vec(N)^H)] \\
&= E[\bar{X}vec(H)vec(H)^H\bar{X}^H] + E[vec(N)vec(N)^H] \\
&= \bar{X}R\bar{X}^H + \sigma_N^2 I
\end{aligned} \tag{3.22}$$

Combining (3.23) and (3.24), (3.20) becomes:

$$vec(\hat{H}_{cLMMSE}) = R\bar{X}^H(\bar{X}R\bar{X}^H + \sigma_N^2 I)^{-1}vec(Y) \tag{3.23}$$

The proposed linear MMSE estimation of $vec(H)$ is also based on (3.20). For the proposed LMMSE estimator $vec(\hat{H}_{pLMMSE})$, equation (3.21) can be expanded as:

$$\begin{aligned} R_{HY} &= E[vec(H)(vec(H)^H\bar{X}^H + vec(N)^H)] \\ &= E[vec(H)vec(H)^H\bar{X}^H] + E[vec(H)vec(H)^H\bar{I}_{UE}^H] \\ &+ E[vec(H)vec(I_{BS})^H] + E[vec(H)vec(N)^H)] \\ &= R\bar{X}^H \end{aligned} \tag{3.24}$$

Moreover, equation (3.22) can be extended as:

$$\begin{aligned} R_Y &= E[((\bar{X} + \bar{I}_{UE})vec(H) + vec(I_{BS}) + vec(N)) \\ &(vec(H)^H(\bar{X}^H + \bar{I}_{UE}^H) + vec(I_{BS})^H + vec(N)^H)] \\ &= E[\bar{X}vec(H)vec(H)^H\bar{X}^H] + E[\bar{I}_{UE}vec(H)vec(H)^H\bar{I}_{UE}^H] \\ &+ E[vec(I_{BS})vec(I_{BS})^H] + E[vec(N)vec(N)^H] \\ &= \bar{X}R\bar{X}^H + E[\bar{I}_{UE}vec(H)vec(H)^H\bar{I}_{UE}^H] + \kappa_{BS}pdiag(R) + \sigma_N^2 I \end{aligned} \tag{3.25}$$

Similarly to the single-user case i_{BS} can be rewritten as $\Sigma_{BS}^2 z$, with z a vector with $\mathcal{CN}(0,1)$ where z is uncorrelated with the channel H in (3.26) and (3.27).

The term $E[\bar{I}_{UE}vec(H)vec(H)^H\bar{I}_{UE}^H]$ can be rewritten as $E[\bar{I}_{UE}vec(H)(\bar{I}_{UE}vec(H))^H]$. Setting $A = \bar{I}_{UE}vec(H)$ we get:

$$A = \begin{bmatrix} i_{UE_{11}}h_{11} + i_{UE_{21}}h_{12} + \cdots + i_{UE_{N_t1}}h_{1N_t} \\ i_{UE_{11}}h_{21} + i_{UE_{21}}h_{22} + \cdots + i_{UE_{N_t1}}h_{2N_t} \\ \vdots \\ i_{UE_{11}}h_{N_r1} + i_{UE_{21}}h_{N_r2} + \cdots + i_{UE_{N_t1}}h_{N_rN_t} \\ i_{UE_{12}}h_{11} + i_{UE_{22}}h_{12} + \cdots + i_{UE_{N_t2}}h_{1N_t} \\ \vdots \\ i_{UE_{1T}}h_{11} + i_{UE_{2T}}h_{12} + \cdots + i_{UE_{N_tT}}h_{1N_t} \\ \vdots \\ i_{UE_{1T}}h_{N_r1} + i_{UE_{2T}}h_{N_r2} + \cdots + i_{UE_{N_tT}}h_{N_rN_t} \end{bmatrix}$$

The $(i,j)_{th}$ element of $E[AA^H]$ is given by multiplying the i_{th} element of A with the j_{th} element of A . For example the $(1,1)$ element of $E[AA^H]$ is:

$$\begin{aligned} E[AA^H]_{(1,1)} &= E[i_{UE_{11}}h_{11}(i_{UE_{11}}h_{11})^H + \ldots \\ &+ i_{UE_{N_t1}}h_{1N_t}(i_{UE_{(N_t-1)1}}h_{1(N_t-1)})^H + i_{UE_{N_t1}}h_{1N_t}(i_{UE_{N_t1}}h_{1N_t})^H] \\ &= E[i_{UE_{11}}h_{11}(i_{UE_{11}}h_{11})^H] + \cdots + E[i_{UE_{N_t1}}h_{N_t1}(i_{UE_{N_t1}}h_{N_t1})^H] \\ &= c\kappa_{UE}R_{(1,1)} + c\kappa_{UE}R_{(2,2)} + \cdots + c\kappa_{UE}R_{(N_t,N_t)} \\ &= N_t c\kappa_{UE}R_{(1,1)} \end{aligned} \tag{3.26}$$

Since $i_{UE_{11}}$ and $i_{UE_{21}}$ denote noise from different users and as a result are uncorrelated. Furthermore, R is a symmetric toeplitz matrix with $R_{(j,j)} = R_{(1,1)}$.

Similarly, $E[AA^H]_{(2,1)} = N_t c \kappa_{UE} R_{(2,1)}$.

However, after taking the product of the first and the N_{th} element of A we obtain according to (3.16):

$$\begin{aligned} E[AA^H]_{(N,1)} &= E[(i_{UE_{11}} h_{11} + \cdots + i_{UE_{N_t 1}} h_{1N_t})(i_{UE_{12}} h_{11} + i_{UE_{22}} h_{12} + \cdots + i_{UE_{N_t 2}} h_{1N_t})^H] \\ &= E[i_{UE_{11}} h_{11} (i_{UE_{12}} h_{11})^H] + \cdots + E[i_{UE_{N_t 1}} h_{N_t 1} (i_{UE_{N_t 2}} h_{N_t 1})^H] \\ &= N_t c q_{UE} R_{11} \end{aligned} \tag{3.27}$$

Hence, the $TN_r \times TN_r$ matrix $R_A = \mathbb{E}[\bar{I}_{UE} vec(H) vec(H)^H \bar{I}_{UE}^H]$ is given by:

$$R_A = U_{UE} \otimes R_r \tag{3.28}$$

where U_{UE} is the covariance of the noise at each user and is the $T \times T$ matrix given in (3.16) and R_r is the $N_r \times N_r$ receive spatial correlation of the channel matrix.

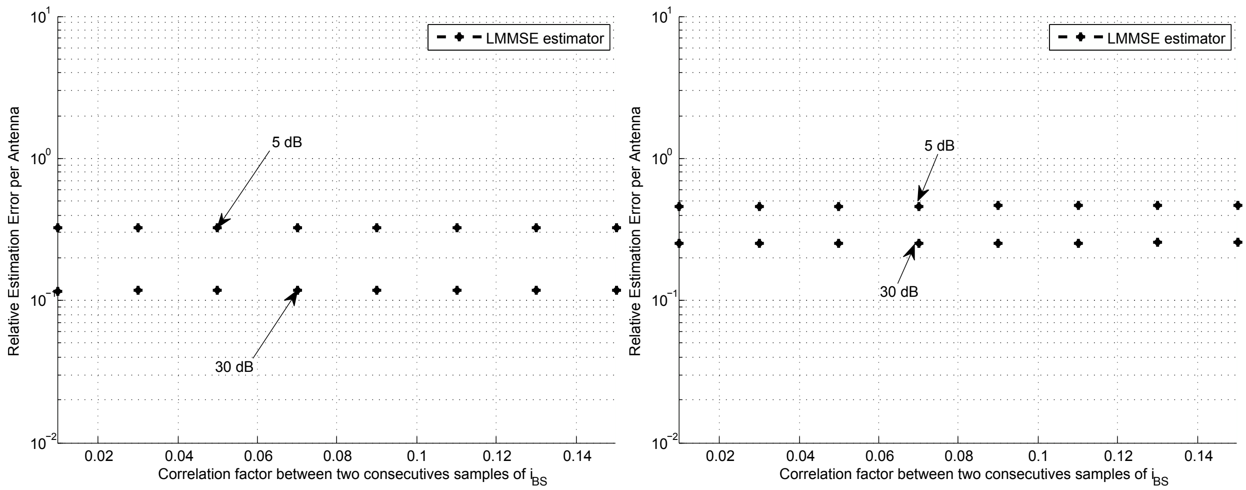
The covariance matrix R_A has the form:

$$R_A \triangleq E[AA^H] = N_t c \begin{bmatrix} \kappa_{UE} R_{(1,1)} & \dots & \kappa_{UE} R_{(1,N_r)} & \dots & q_{UE} R_{(1,1)} & \dots & q_{UE} R_{(1,N_r)} \\ \dots & \ddots & \dots & \ddots & \dots & \ddots & \dots \\ \kappa_{UE} R_{(N_r,1)} & \dots & \kappa_{UE} R_{(N_r,N_r)} & \dots & q_{UE} R_{(N_r,1)} & \dots & q_{UE} R_{(N_r,N_r)} \\ \dots & \ddots & \dots & \ddots & \dots & \ddots & \dots \\ q_{UE} R_{(N_r,1)} & \dots & q_{UE} R_{(N_r,N_r)} & \dots & \kappa_{UE} R_{(N_r,1)} & \dots & \kappa_{UE} R_{(N_r,N_r)} \end{bmatrix}$$

Combining (3.26) and (3.27) we get the proposed LMMSE estimator:

$$vec(\hat{H}_{pLMMSE}) = R\bar{X}^H (\bar{X} R \bar{X}^H + R_A + p\kappa_{BS} diag(R) + \sigma_N^2 I_{NN_R})^{-1} vec(Y) \tag{3.29}$$

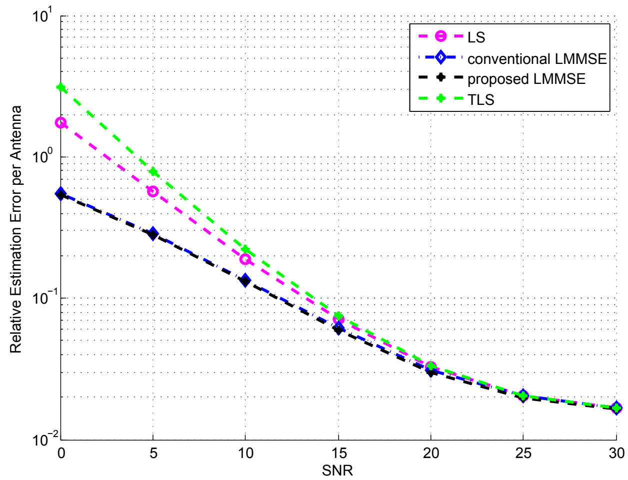
As mentioned before, the correlation, q_{UE}, of the distortion i_{UE} of each user, depends on the nature of the impairments that caused the distortion. Figure 3.3 shows the performance of the proposed LMMSE estimator as q_{UE} varies from 0.01 to 0.15. It is shown that the value of q_{UE} does not have an impact on the performance of the proposed LMMSE estimator for neither low nor high SNR values. The simulation was performed over 1000 Monte-Carlo realizations of the channel, the noise on both ends and the Gaussian noise.



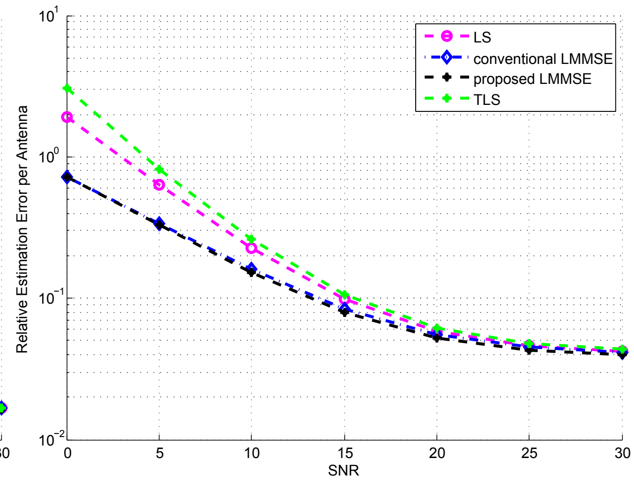
(a) System with 5 users (b) System with 15 users

Figure 3.3: Estimation error per antenna for proposed LMMSE estimator over corellation of the noise on each user for SNR equal to $5dB$ and $30dB$. Multi-user system with 50 antennas at the base station.

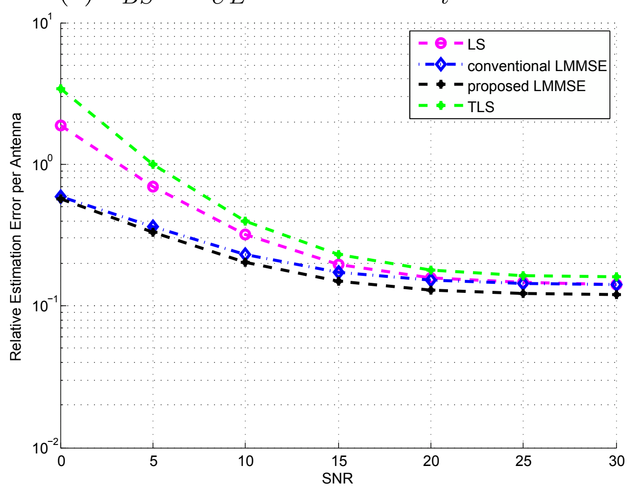
In Figure 3.4 below, the proposed LMMSE estimator is compared to the LS, the TLS and the conventional LMMSE estimator as SNR varies from 0 to $30dB$ in a system with 50 antennas at the base station and multiple users. The level of the impairments at both ends is chosen to be 0.05^2 and 0.15^2. The simulation was performed over 1000 Monte-Carlo realizations of the channel and the noise samples.



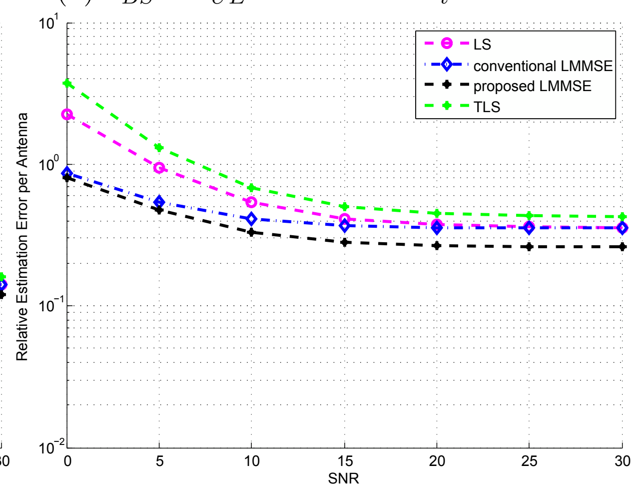
(a) $\kappa_{BS} = \kappa_{UE} = 0.05^2$ and $N_t = 5$ users



(b) $\kappa_{BS} = \kappa_{UE} = 0.05^2$ and $N_t = 15$ users

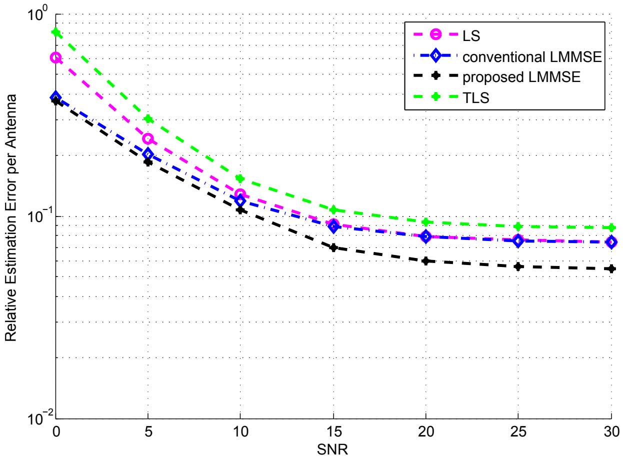


(c) $\kappa_{BS} = \kappa_{UE} = 0.15^2$ and $N_t = 5$ users

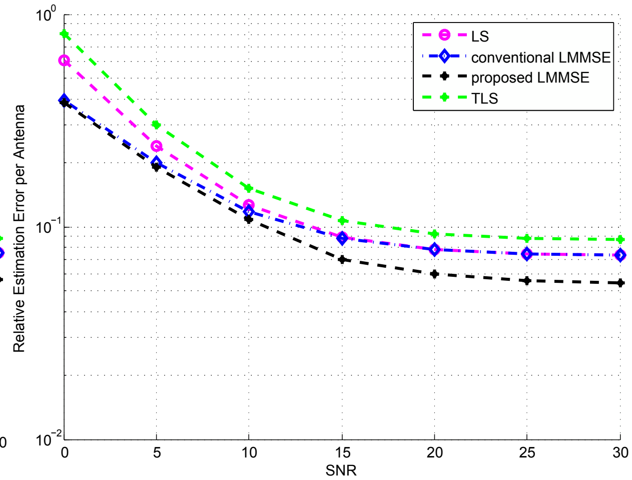


(d) $\kappa_{BS} = \kappa_{UE} = 0.15^2$ and $N_t = 15$ users

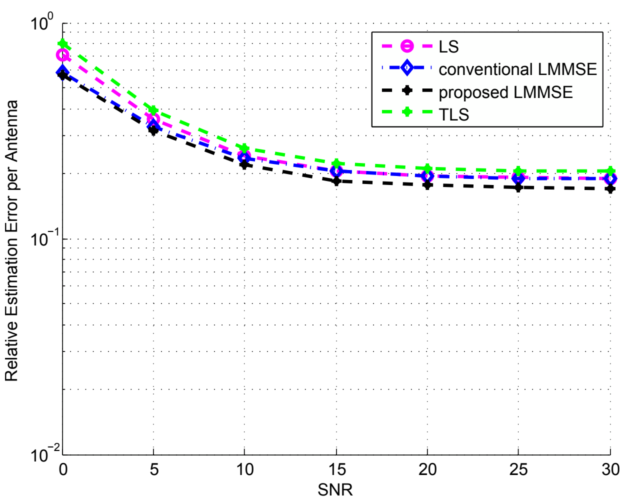
Figure 3.4: Estimation error per antenna for LS, TLS, conventional and proposed LMMSE estimators over SNR. Multi-user system with 50 antennas at the base station.



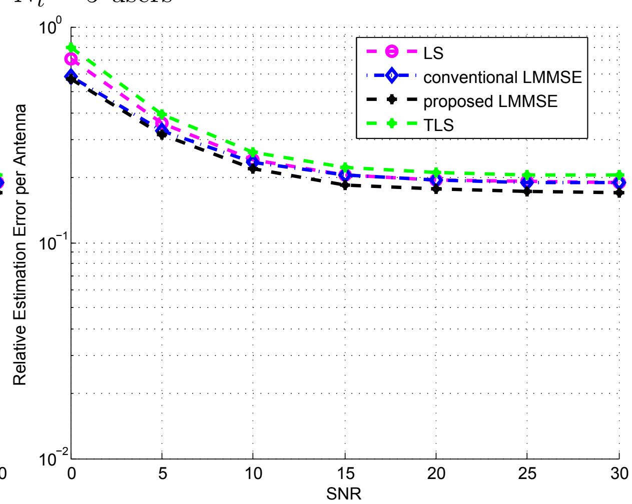
(a) $N_r = 100$ antennas at the base station and $N_t = 5$ users



(b) $N_r = 200$ antennas at the base station and $N_t = 5$ users



(c) $N_r = 100$ antennas at the base station and $N_t = 15$ users



(d) $N_r = 200$ antennas at the base station and $N_t = 15$ users

Figure 3.5: Estimation error per antenna for LS, TLS, conventional and proposed LMMSE estimators over SNR with level of impairments being $\kappa_{BS} = \kappa_{UE} = 0.15^2$.

Chapter 4

Cramer-Rao Lower Bound for a System with Hardware Impairments

As stated in Chapter 2, the covariance matrix of every vector of unbiased estimators $\hat{\vartheta} = [\hat{\vartheta}_1, \hat{\vartheta}_2, \ldots, \hat{\vartheta}_N]^T$ of a parameter column vector $\vartheta = [\vartheta_1, \vartheta_2, \ldots, \vartheta_N]^T$ is lower bounded by CRB which is defined as the inverse of the Fisher information matrix [23]:

$$cov(\hat{\vartheta}) \geq I(\vartheta)^{-1} \tag{4.1}$$

The $(i, j)_{th}$ entry of I is given by:

$$I_{ij}(\vartheta) = E[\frac{d}{d\vartheta_i}\text{log}p(x;\vartheta)\frac{d}{d\vartheta_j}\text{log}p(x;\vartheta)] = -E[\frac{d^2}{d\vartheta_i\vartheta_j}\text{log}p(x;\vartheta)] \tag{4.2}$$

where $\text{log}p(x;\vartheta)$ is the natural logarithm of the pdf of x, $E[\cdot]$ denotes expectation over x.

4.1 Cramer-Rao bound for a single-user Large MIMO system

The system explored has N_r antennas at the base station and 1 user with 1 antenna element. The modulation is QPSK or QAM and the system model used is:

$$Y = h(x + i_{UE}) + I_{BS} + N \tag{4.3}$$

where $x \in \mathbb{C}^{1\times T}$ is the deterministic pilot signal with length T; h is the $N_r \times 1$ deterministic channel matrix. The additive Gaussian noise, $N \in \mathbb{C}^{N_r\times T}$, has zero mean and covariance matrix $\sigma_N^2 I$. The distortion that occurs at the user-end is $i_{UE} \in \mathbb{C}^{1\times T} \sim \mathcal{CN}(0, \sigma_{UE}^2)$ and $i_{BS} \in \mathbb{C}^{N_r\times T} \sim \mathcal{CN}(0, \sigma_{BS}^2)$ is the distortion that occurs at the base station. The signal at the j_{th} antenna is:

$$y_j = h_j(x + i_{UE}) + i_{BS_j} + n_j \tag{4.4}$$

where $h_j = a_j + \mathrm{i}b_j$ is the channel between the j_{th} antenna and the single user and $x = c + \mathrm{i}d$ is the pilot signal. The mean and the variance of y_j are $(a_j c - b_j d) + \mathrm{i}(a_j d + b_j c)$ and $\sigma_j^2 = |h_j|^2\sigma_{UE}^2 + \sigma_{BS}^2 + \sigma_N^2$, respectively. Hence, the 2×2 Fisher information matrix of the signal at the j_{th} antenna is given by:

$$I(a_j, b_j) = \begin{bmatrix} -E[(\frac{d^2 p(y;a_j,b_j)}{da_j\, da_j})] - E[(\frac{d^2 p(y;a_j,b_j)}{da_j\, db_j})] \\ -E[(\frac{d^2 p(y;a_j,b_j)}{db_j\, da_j})] - E[(\frac{d^2 p(y;a_j,b_j)}{db_j\, db_j})] \end{bmatrix} \tag{4.5}$$

In order to calculate the CRB it is preferable to use two real Gaussian variables y_{r_j} and y_{im_j}, instead of y_j, with means $a_j c - b_j d$ and $a_j d + b_j c$, respectively, and the same variance as y_j. Since we use a deterministic channel when calculating the CRB the only random variables are the distortion samples i_{UE} and I_{BS} and the noise N. Hence, the received signals $y_1, y_2 \dots y_{N_r}$ are independent. The PDF of y is:

$$p(y_1, y_2, ..., y_{N_r}) = \prod_{j=1}^{N_r} \prod_{i=1}^{T} \frac{1}{\sqrt{2\pi\sigma_j^2}} exp(-\sum_{i=1}^{T} \frac{(y_{r_{ji}} - (a_j c_i - b_j d_i))^2 + (y_{im_{ji}} - (a_j d_i + b_j c_i))^2}{2\sigma_j^2}) \tag{4.6}$$

The natural logarithm of the probability function is:

$$\text{log}p(y_1, y_2, ..., y_{N_r}) = \text{log}p(y_1) + \text{log}p(y_2) + \cdots + \text{log}(y_{N_r}) \tag{4.7}$$

Taking the derivative of (4.7) with respect to a_j we get:

$$\begin{aligned} \frac{d \log p(y_1, y_2, ..., y_{N_r})}{da_j} = &- \frac{T}{2} \frac{2a_j \sigma_{UE}^2}{\sigma_j^2} \\ &+ \sum_{i=1}^{T} \frac{c_i(y_{r_{ji}} - (a_j c_i - b_j d_i))}{\sigma_j^2} \\ &+ \sum_{i=1}^{T} \frac{a\sigma_{UE}^2 (y_{r_{ji}} - (a_j c_i - b_j d_i))^2}{(\sigma_j^2)^2} \\ &+ \sum_{i=1}^{T} \frac{d_i(y_{im_{ji}} - (b_j c_i + a_j d_i))}{\sigma_j^2} \\ &+ \sum_{i=1}^{T} \frac{a_j \sigma_{UE}^2 (y_{im_{ji}} - (b_j c_i + a_j d_i))^2}{(\sigma_j^2)^2} \end{aligned} \tag{4.8}$$

Taking the derivative of (4.7) with respect to b_j we get:

$$\begin{aligned}
\frac{d \log p(y_1, y_2, ..., y_{N_r})}{db_j} = &- \frac{T}{2}\frac{2b_j\sigma_{UE}^2}{\sigma_j^2} \\
&- \sum_{i=1}^{T} \frac{d_i(y_{r_{ji}} - (a_j c_i - b_j d_i))}{\sigma_j^2} \\
&+ \sum_{i=1}^{T} \frac{b\sigma_{UE}^2(y_{r_{ji}} - (a_j c_i - b_j d_i))^2}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{c_i(y_{im_{ji}} - (b_j c_i + a_j d_i))}{\sigma_j^2} \\
&+ \sum_{i=1}^{T} \frac{b_j\sigma_{UE}^2(y_{im_{ji}} - (b_j c_i + a_j d_i))^2}{(\sigma_j^2)^2}
\end{aligned} \tag{4.9}$$

In order to calculate $I_{11}(h_j)$ from (4.5) we must calculate the second derivative of (4.8) with respect to a_j:

$$\begin{aligned}\frac{d^2 \log p(y_1, y_2, ..., y_{N_r})}{da_j\, da_j} =& \frac{T(2a_j^2(\sigma_{UE}^2)^2 - 2\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \\ &+ \sum_{i=1}^{T} \frac{-c_i^2\sigma_j^2 - 2a_j\sigma_{UE}^2 c_i(y_{r_{ji}} - (a_jc_i - b_jd_i))}{(\sigma_j^2)^2} \\ &+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2(y_{r_{ji}} - (a_jc_i - b_jd_i))^2 - a_j^2\sigma_{UE}^2 c_i(y_{r_{ji}} - (a_jc_i - b_jd_i))}{(\sigma_j^2)^2} \\ &- \sum_{i=1}^{T} \frac{4a_j^2(\sigma_{UE}^2)^2\sigma_j^2(y_{r_{ji}} - (a_jc_i - b_jd_i))^2}{(\sigma_j^2)^4} \\ &+ \sum_{i=1}^{T} \frac{-d_i^2\sigma_j^2 - 2a_j\sigma_{UE}^2 d_i(y_{im_{ji}} - (b_jc_i + a_jd_i))}{(\sigma_j^2)^2} \\ &+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2(y_{im_{ji}} - (a_jd_i + b_jc_i))^2 - a_j^2\sigma_{UE}^2 c_i(y_{im_{ji}} - (a_jd_i + b_jc_i))}{(\sigma_j^2)^2} \\ &- \sum_{i=1}^{T} \frac{4a_j^2(\sigma_{UE}^2)^2\sigma_j^2(y_{im_{ji}} - (a_jd_i + b_jc_i))^2}{(\sigma_j^2)^4} \qquad (4.10)\end{aligned}$$

We take the expected value over y:

$$
\begin{aligned}
E[\frac{d^2 \log p(y_1, ..., y_{N_r})}{da_j da_j}] =& \frac{T(2a_j^2(\sigma_{UE}^2)^2 - 2\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{-c_i^2\sigma_j^2 - 2a_j\sigma_{UE}^2 c_i(E[y_{r_{ji}}] - (a_j c_i - b_j d_i))}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2[(E[y_{r_{ji}}] - (a_j c_i - b_j d_i))^2 - a_j^2 c_i(E[y_{r_{ji}}] - (a_j c_i - b_j d_i))]}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{4a_j^2(\sigma_{UE}^2)^2\sigma_j^2(E[y_{r_{ji}}] - (a_j c_i - b_j d_i))^2}{(\sigma_j^2)^4} \\
&+ \sum_{i=1}^{T} \frac{-d_i^2\sigma_j^2 - 2a_j\sigma_{UE}^2 d_i(E[y_{im_{ji}}] - (b_j c_i + a_j d_i))}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{(E[y_{im_{ji}}] - (a_j d_i + b_j c_i))^2 - a_j^2 c_i(E[y_{im_{ji}}] - (a_j d_i + b_j c_i))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{4a_j^2(\sigma_{UE}^2)^2\sigma_j^2(E[y_{im_{ji}}] - (a_j d_i + b_j c_i))^2}{(\sigma_j^2)^4} \\
=& - \sum_{i=1}^{T} \frac{c_i^2 + d_i^2}{\sigma_j^2} + \frac{T(2a_j^2(\sigma_{UE}^2)^2 - 2\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \qquad (4.11)
\end{aligned}
$$

Similarly, in order to calculate $I_{21}(h_j)$ from (4.5) we must calculate the second derivative of (4.8) with respect to b_j:

$$\begin{aligned}
\frac{d^2 \log p(y_1, y_2, ..., y_{N_r})}{da_j\, db_j} =& \frac{2a_j b_j T(\sigma_{UE}^2)^2}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{c_i d_i \sigma_j^2 - 2a_{jm} c_{mi} \sigma_{UE}^2 (y_{r_{ji}} - (a_j c_i - b_j d_i))}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{2a_j d_i (y_{r_{ji}} - (a_j c_i - b_j d_i))}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{4a_j b_j (\sigma_{UE}^2)^2 \sigma_j^2 (y_{r_{ji}} - (a_j c_i - b_j d_i))^2}{(\sigma_j^2)^4} \\
&- \sum_{i=1}^{T} \frac{c_i d_i \sigma_j^2 - 2a_j d_i \sigma_{UE}^2 (y_{im_{ji}} - (a_j d_i + b_j c_i))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{2a_j c_i (y_{im_{ji}} - (a_j d_i + b_j c_i))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{4a_j b_j (\sigma_{UE}^2)^2 \sigma_j^2 (y_{im_{ji}} - (a_j d_i + b_j c_i))^2}{(\sigma_j^2)^4}
\end{aligned} \tag{4.12}$$

The expectation of which is:

$$\begin{aligned}
E[\frac{d^2 \log p(y_1, y_2, ..., y_{N_r})}{da_j db_j}] =& \frac{2a_j b_j T(\sigma_{UE}^2)^2}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{c_i d_i \sigma_j^2 - 2a_{jm} c_{mi} \sigma_{UE}^2 (E[y_{r_{ji}}] - (a_j c_i - b_j d_i))}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{2a_j d_i (E[y_{r_{ji}}] - (a_j c_i - b_j d_i))}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{4a_j b_j (\sigma_{UE}^2)^2 \sigma_j^2 (E[y_{r_{ji}}] - (a_j c_i - b_j d_i))^2}{(\sigma_j^2)^4} \\
&- \sum_{i=1}^{T} \frac{c_i d_i \sigma_j^2 - 2a_j d_i \sigma_{UE}^2 (E[y_{im_{ji}}] - (a_j d_i + b_j c_i))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{2a_j c_i (E[y_{im_{ji}}] - (a_j d_i + b_j c_i))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{4a_j b_j (\sigma_{UE}^2)^2 \sigma_j^2 (E[y_{im_{ji}}] - (a_j d_i + b_j c_i))^2}{(\sigma_j^2)^4} \\
=& \frac{2a_j b_j (\sigma_{UE}^2)^2 T}{(\sigma_j^2)^2} \qquad (4.13)
\end{aligned}$$

According to Schwartz's theorem [26], since the second partial derivatives of a logarithmic function are continuous, its partial derivations are commutative. Thus, we have

$E[\frac{d^2 \log p(y_1,y_2,...,y_{N_r})}{db da}] = E[\frac{d^2 \log p(y_1,y_2,...,y_N)}{da db}]$. Differentiating (4.9) with respect to b_j we get:

$$\begin{aligned}
\frac{d^2 \log p(y_1, y_2, ..., y_{N_r})}{db_j db_j} =& \frac{T(2b_j^2(\sigma_{UE}^2)^2 - 2\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{-d_i^2\sigma_j^2 - 2a_j\sigma_{UE}^2 c_i(y_{r_{ji}} - (a_jc_i - b_jd_i))}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2(y_{r_{ji}} - (a_jc_i - b_jd_i))^2 - b_j^2\sigma_{UE}^2 d_i(y_{r_{ji}} - (a_jc_i - b_jd_i))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{4b_j^2(\sigma_{UE}^2)^2\sigma_j^2(y_{r_{ji}} - (a_jc_i - b_jd_i))^2}{(\sigma_j^2)^4} \\
&+ \sum_{i=1}^{T} \frac{-c_i^2\sigma_j^2 - 2b_j\sigma_{UE}^2 d_i(y_{im_{ji}} - (b_jc_i + a_jd_i))}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2(y_{im_{ji}} - (a_jd_i + b_jc_i))^2 - a_j^2 d_i(y_{im_{ji}} - (a_jd_i + b_jc_i))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{4b_j^2(\sigma_{UE}^2)^2\sigma_j^2(y_{im_{ji}} - (a_jd_i + b_jc_i))^2}{(\sigma_j^2)^4}
\end{aligned} \tag{4.14}$$

Similarly to (4.11), taking the expected value over y gives:

$$E[\frac{d^2 \log p(y_1, y_2, ..., y_{N_r})}{db_j db_j}] = -\sum_{i=1}^{T} \frac{c_i^2 + d_i^2}{\sigma_j^2} + \frac{T(2b_j^2(\sigma_{UE}^2)^2 - 2\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \tag{4.15}$$

Hence the information matrix of the signal at the j_{th} antenna is given by:

$$I(a_j, b_j) = \begin{bmatrix} \sum_{i=1}^{T} \frac{|x|_i^2}{\sigma_j^2} - \frac{T(2a_j^2(\sigma_{UE}^2)^2 - \sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} & -\frac{2a_j b_j (\sigma_{UE}^2)^2 T}{(\sigma_j^2)^2} \\ -\frac{2a_j b_j (\sigma_{UE}^2)^2 T}{(\sigma_j^2)^2} & \sum_{i=1}^{T} \frac{|x|_i^2}{\sigma_j^2} - \frac{T(2b_j^2(\sigma_{UE}^2)^2 - \sigma_{UE}^2\sigma^2)}{(\sigma^2)^2} \end{bmatrix} \tag{4.16}$$

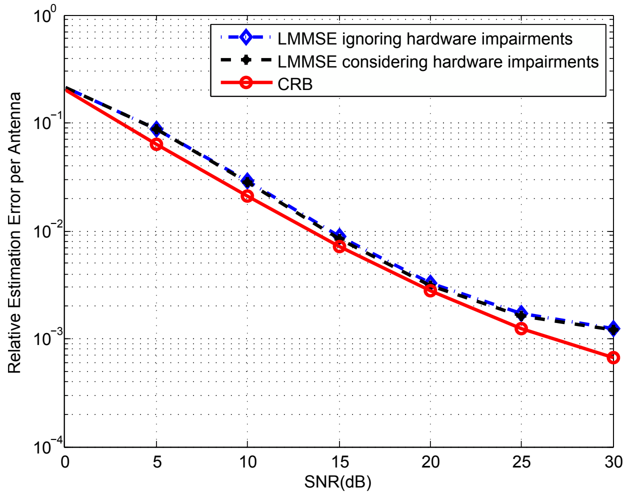
The $2N_r \times 2N_r$ information matrix for the signals at all the antennas becomes:

$$I(h) = \begin{bmatrix} I(a_1, b_1) & 0 & \dots & 0 \\ 0 & I(a_2, b_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I(a_{N_r}, b_{N_r}) \end{bmatrix} \tag{4.17}$$

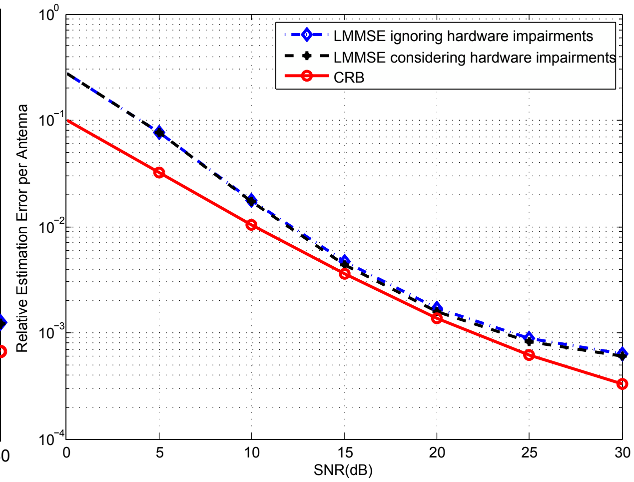
Finally, the Cramer-Rao bound is given by:

$$\hat{h} \geq I(h)^{-1} \tag{4.18}$$

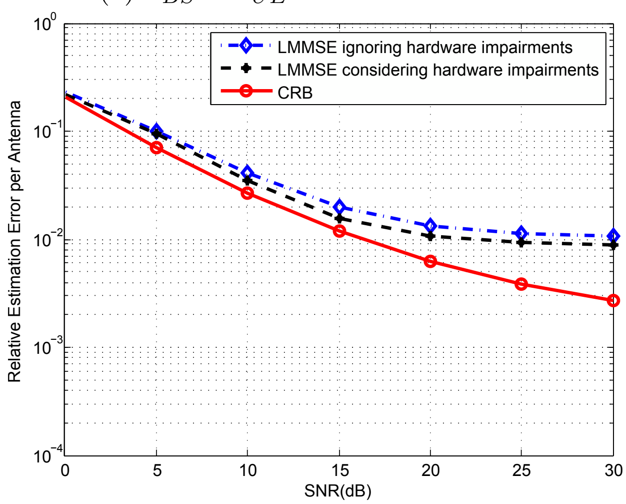
In Figure 4.1, the LMMSE estimator from [5], along with the conventional, impairment ignoring, LMMSE estimator are compared to the CRB for a system with 50 antennas at the base station and one single-antenna user as SNR varies from $0dB$ to $30dB$. The covariance matrix, R of the channel is a symmetric toeplitz matrix with value 1 at the diagonal and correlation factor equal to 0.5. The level of the impairments is chosen to be 0.05^2 and 0.15^2 as the lowest and the upper bound of the parameter ,respectively, [11]. The length of the training pilots are either $T = 5$ or $T = 10$. The simulation was performed over 1000 Monte-Carlo realizations of the channel, the distortion on both ends and the Gaussian noise. The results of the simulations show that both LMMSE estimators achieve a performance close to the Cramer-Rao bound.



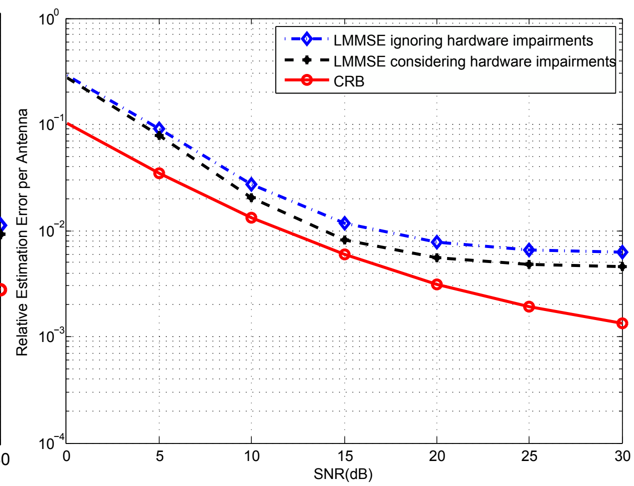
(a) $\kappa_{BS} = \kappa_{UE} = 0.05^2$ and $T = 5$



(b) $\kappa_{BS} = \kappa_{UE} = 0.05^2$ and $T = 10$



(c) $\kappa_{BS} = \kappa_{UE} = 0.15^2$ and $T = 5$



(d) $\kappa_{BS} = \kappa_{UE} = 0.15^2$ and $T = 10$

Figure 4.1: Estimation error per antenna over SNR. Conventional LMSSE estimator and LMMSE estimator considering hardware impairments compared to the CRB for a single-user system with 50 antennas at the base station.

4.2 Cramer-Rao bound for a multi-user Large MIMO system

In the multi-user case the system model is given by:

$$Y = H(X + I_{UE}) + I_{BS} + N \tag{4.19}$$

where $X \in \mathbb{C}^{N_t \times T}$ is the deterministic pilot signal; T is the length of the training sequence. The channel H is a deterministic $N_r \times N_t$ matrix. The additive Gaussian noise $N \in \mathbb{C}^{N_r \times T}$ has zero mean and covariance matrix $\sigma_N^2 I$. The distortion that occurs at the user-end, $I_{UE} \in \mathbb{C}^{N_t \times T}$, consists of N_t independent Gaussian random variables with zero mean and variance σ_{UE}^2 and $I_{BS} \in \mathbb{C}^{N_r \times T} \sim \mathcal{CN}(0, \sigma_{BS}^2)$ is the distortion that occurs at the base station. The received signal at the j_{th} antenna at the instant $t \in (0, T)$ is given by:

$$y_{jt} = \sum_{n=1}^{M} h_{jn}(x_{nt} + i_{UE_{nt}}) + i_{BS_{jt}} + n_{jt} \tag{4.20}$$

Thus, y_{jt} is a complex Gaussian variable; $\sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}) + \mathrm{i}(a_{jn}d_{ni} + b_{jn}c_{ni})$ and $\sigma_j^2 = \sum_{n=1}^{M} |h_{jn}|^2\sigma_{UE}^2 + \sigma_{BS}^2 + \sigma_N^2$ are its mean value and variance, respectively. In order to calculate the CRB it is preferable to use two real Gaussian variables y_{r_j} and y_{im_j} as in Section 3.1. Similarly to the single user case the received signals $y_{1i}, y_{2i} \dots y_{N_r i}$ are independent since the channel H is deterministic. The probability function of y is:

$$
\begin{aligned}
p(y_1, y_2, ..., y_N) &= p(y_1)p(y_2)\dots p(y_N) \\
p(y_1, y_2, ..., y_{N_r}) &= \prod_{j=1}^{N}\prod_{i=1}^{T} \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\sum_{i=1}^{T} \frac{(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))^2 + (y_{in_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))^2}{2\sigma_j^2}}
\end{aligned}
\tag{4.21}
$$

The natural logarithm of the probability function is:

$$
\log p(y_1, y_2, ..., y_{N_r}) = \log p(y_1) + \log p(y_2) + \cdots + \log p(y_{N_r}) \tag{4.22}
$$

The derivatives with respect to a_{jm} and b_{jm} are:

$$
\begin{aligned}
\frac{d \log p(y_1, y_2, ..., y_{N_r})}{da_{jm}} = &- \frac{T}{2}\frac{2a_{jm}\sigma_{UE}^2}{\sigma_j^2} \\
&+ \sum_{i=1}^{T} \frac{c_{mi}(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))}{\sigma_j^2} \\
&+ \sum_{i=1}^{T} \frac{a_{jm}\sigma_{UE}^2(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))^2}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{a_{jm}\sigma_{UE}^2(y_{im_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))^2}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{d_{mi}(y_{im_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))}{\sigma_j^2}
\end{aligned}
\tag{4.23}
$$

$$\frac{d\log p(y_1, y_2, ..., y_{N_r})}{db_{jm}} = -\frac{T}{2}\frac{2b_{jm}\sigma_{UE}^2}{\sigma_j^2} - \sum_{i=1}^{T}\frac{d_{mi}(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))}{\sigma_j^2} + \sum_{i=1}^{T}\frac{b_{jm}\sigma_{UE}^2(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))^2}{(\sigma_j^2)^2} + \sum_{i=1}^{T}\frac{b_{jm}\sigma_{UE}^2(y_{im_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))^2}{(\sigma_j^2)^2} + \sum_{i=1}^{T}\frac{c_{mi}(y_{im_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))}{\sigma_j^2} \tag{4.24}$$

At this point, we should mention that the Information matrix of the signal at the j_{th} antenna for the multi-user case is a $2N_t \times 2N_t$ matrix and is given by:

$$I(a_j, b_j) = \begin{bmatrix} I_{a_{j1}a_{j1}} & I_{a_{j1}b_{j1}} & I_{a_{j1}a_{j2}} & I_{a_{j1}b_{j2}} & \dots & I_{a_{j1}a_{jM}} & I_{a_{j1}b_{jM}} \\ I_{b_{j1}a_{j1}} & I_{b_{j1}b_{j1}} & I_{b_{j1}a_{j2}} & I_{b_{j1}b_{j2}} & \dots & I_{b_{j1}a_{jM}} & I_{b_{j1}b_{jM}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ I_{b_{jM}a_{j1}} & I_{b_{jM}b_{j1}} & I_{b_{jM}a_{j2}} & I_{b_{jM}b_{j2}} & \dots & I_{b_{jM1}a_{jM}} & I_{b_{jM}b_{jM}} \end{bmatrix} \tag{4.25}$$

The second derivative of (4.23) with respect to a_{jm} is given by:

$$\begin{aligned}\frac{d\log p(y_1, y_2, ..., y_{N_r})}{da_{jm}da_{jm}} =& \frac{T(2a_{jm}^2(\sigma_{UE}^2)^2 - \sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \\ &+ \sum_{i=1}^{T} \frac{-c_{mi}^2\sigma_j^2 - 2a_{jm}c_{mi}\sigma_{UE}^2(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))}{(\sigma_j^2)^2} \\ &+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))^2}{(\sigma_j^2)^2} \\ &- \sum_{i=1}^{T} \frac{2a_{jm}c_{im}(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))}{(\sigma_j^2)^2} \\ &- \sum_{i=1}^{T} \frac{4a_{jm}^2(\sigma_{UE}^2)^2\sigma_j^2(y_{r_{ji}} - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))^2}{(\sigma_j^2)^4} \\ &+ \sum_{i=1}^{T} \frac{-d_{mi}^2\sigma_j^2 - 2a_{jm}d_{mi}\sigma_{UE}^2(y_{im_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))}{(\sigma_j^2)^2} \\ &+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2(y_{im_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))^2}{(\sigma_j^2)^2} \\ &- \sum_{i=1}^{T} \frac{2a_{jm}d_{im}(y_{im_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))}{(\sigma_j^2)^2} \\ &- \sum_{i=1}^{T} \frac{4a_{jm}^2(\sigma_{UE}^2)^2\sigma_j^2(y_{im_{ji}} - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))^2}{(\sigma_j^2)^4}\end{aligned} \tag{4.26}$$

Taking the expected value over y_j:

$$
\begin{aligned}
E[\frac{d\log p(y_1,y_2,...,y_{N_r})}{da_{jm}da_{jm}}] =& \frac{T(2a_{jm}^2(\sigma_{UE}^2)^2-\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \\
&+\sum_{i=1}^{T}\frac{-c_{mi}^2\sigma_j^2-2a_{jm}c_{mi}\sigma_{UE}^2(E[y_{r_{ji}}]-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))}{(\sigma_j^2)^2} \\
&+\sum_{i=1}^{T}\frac{\sigma_{UE}^2(E[y_{r_{ji}}]-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))^2}{(\sigma_j^2)^2} \\
&-\sum_{i=1}^{T}\frac{2a_{jm}c_{im}(E[y_{r_{ji}}]-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))}{(\sigma_j^2)^2} \\
&-\sum_{i=1}^{T}\frac{4a_{jm}^2(\sigma_{UE}^2)^2\sigma_j^2(E[y_{r_{ji}}]-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))^2}{(\sigma_j^2)^4} \\
&+\sum_{i=1}^{T}\frac{-d_{mi}^2\sigma_j^2-2a_{jm}d_{mi}\sigma_{UE}^2(E[y_{im_{ji}}]-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))}{(\sigma_j^2)^2} \\
&+\sum_{i=1}^{T}\frac{\sigma_{UE}^2(E[y_{im_{ji}}]-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))^2}{(\sigma_j^2)^2} \\
&-\sum_{i=1}^{T}\frac{2a_{jm}d_{im}(E[y_{im_{ji}}]-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))}{(\sigma_j^2)^2} \\
&-\sum_{i=1}^{T}\frac{4a_{jm}^2(\sigma_{UE}^2)^2\sigma_j^2(E[y_{im_{ji}}]-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))^2}{(\sigma_j^2)^4} \\
=& -\sum_{i=1}^{T}\frac{c_{mi}^2+d_{mi}^2}{\sigma_j^2}+\frac{T(2a_{jm}^2(\sigma_{UE}^2)^2-\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \qquad (4.27)
\end{aligned}
$$

Similarly to (4.27) :

$$
E[\frac{d\log p(y_1,y_2,...,y_{N_r})}{db_{jm}db_{jm}}] = -\sum_{i=1}^{T}\frac{c_{mi}^2+d_{mi}^2}{\sigma_j^2}+\frac{T(2b_{jm}^2(\sigma_{UE}^2)^2-\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \qquad (4.28)
$$

The second derivative of (4.23) with respect to $a_{jm'}$ is given by:

$$\begin{aligned}
\frac{d\log p(y_1,y_2,...,y_{N_r})}{da_{jm}da_{jm'}} =& \frac{T(2a_{jm}a_{jm'}(\sigma_{UE}^2)^2-\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \\
&+\sum_{i=1}^{T}\frac{-c_{mi}c_{m'i}\sigma_j^2-2a_{jm'}c_{mi}\sigma_{UE}^2(y_{r_{ji}}-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))}{(\sigma_j^2)^2} \\
&+\sum_{i=1}^{T}\frac{\sigma_{UE}^2(y_{r_{ji}}-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))^2}{(\sigma_j^2)^2} \\
&-\sum_{i=1}^{T}\frac{2a_{jm}c_{im'}(y_{r_{ji}}-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))}{(\sigma_j^2)^2} \\
&-\sum_{i=1}^{T}\frac{4a_{jm}a_{jm'}(\sigma_{UE}^2)^2\sigma_j^2(y_{r_{ji}}-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))^2}{(\sigma_j^2)^4} \\
&+\sum_{i=1}^{T}\frac{-d_{mi}d_{m'i}\sigma_j^2-2a_{jm'}d_{mi}\sigma_{UE}^2(y_{im_{ji}}-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))}{(\sigma_j^2)^2} \\
&+\sum_{i=1}^{T}\frac{\sigma_{UE}^2(y_{im_{ji}}-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))^2}{(\sigma_j^2)^2} \\
&-\sum_{i=1}^{T}\frac{2a_{jm'}d_{im}(y_{im_{ji}}-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))}{(\sigma_j^2)^2} \\
&-\sum_{i=1}^{T}\frac{4a_{jm}a_{jm'}(\sigma_{UE}^2)^2\sigma_j^2(y_{im_{ji}}-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))^2}{(\sigma_j^2)^4} \qquad (4.29)
\end{aligned}$$

Taking the expected value over y_j:

$$
\begin{aligned}
E[\frac{d\log p(y_1, y_2, ..., y_{N_r})}{da_{jm}da_{jm'}}] =& \frac{T(2a_{jm}a_{jm'}(\sigma_{UE}^2)^2 - \sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{-c_{mi}[c_{m'i}\sigma_j^2 + 2a_{jm'}(E[y_{r_{ji}}] - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))]}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2(E[y_{r_{ji}}] - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))^2}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{2a_{jm}c_{im'}(E[y_{r_{ji}}] - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{4a_{jm}a_{jm'}(\sigma_{UE}^2)^2\sigma_j^2(E[y_{r_{ji}}] - \sum_{n=1}^{M}(a_{jn}c_{ni} - b_{jn}d_{ni}))^2}{(\sigma_j^2)^4} \\
&+ \sum_{i=1}^{T} \frac{-d_{mi}[d_{m'i}\sigma_j^2 + 2a_{jm'}(E[y_{im_{ji}}] - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))]}{(\sigma_j^2)^2} \\
&+ \sum_{i=1}^{T} \frac{\sigma_{UE}^2(E[y_{im_{ji}}] - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))^2}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{2a_{jm'}d_{im}(E[y_{im_{ji}}] - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))}{(\sigma_j^2)^2} \\
&- \sum_{i=1}^{T} \frac{4a_{jm}a_{jm'}(\sigma_{UE}^2)^2\sigma_j^2(E[y_{im_{ji}}] - \sum_{n=1}^{M}(a_{jn}d_{ni} + b_{jn}c_{ni}))^2}{(\sigma_j^2)^4} \\
=& -\sum_{i=1}^{T} \frac{c_{mi}c_{m'i} + d_{mi}d_{m'i}}{\sigma_j^2} + \frac{T(2a_{jm}a_{jm'}(\sigma_{UE}^2)^2 - \sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \qquad (4.30)
\end{aligned}
$$

Similarly to (4.30) :

$$
E[\frac{d\log p(y_1, y_2, ..., y_{N_r})}{db_{jm}db_{jm'}}] = \sum_{i=1}^{T} \frac{c_{mi}c_{m'i} + d_{mi}d_{m'i}}{\sigma_j^2} + \frac{T(2b_{jm}b_{jm'}(\sigma_{UE}^2)^2 - \sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2} \qquad (4.31)
$$

The second derivative of (4.23) with respect to b_{jm} is given by:

$$\begin{aligned}\frac{d\log p(y_1,y_2,...,y_{N_r})}{da_{jm}db_{jm}} =& \frac{T(2a_{jm}b_{jm}(\sigma_{UE}^2)^2-\sigma_{UE}^2\sigma_j^2)}{(\sigma_j^2)^2}\\ &+\sum_{i=1}^{T}\frac{c_{mi}d_{mi}\sigma_j^2-2a_{jm}c_{mi}\sigma_{UE}^2(y_{r_{ji}}-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))}{(\sigma_j^2)^2}\\ &+\sum_{i=1}^{T}\frac{2a_{jm}d_{im}(y_{r_{ji}}-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))}{(\sigma_j^2)^2}\\ &+\sum_{i=1}^{T}\frac{4a_{jm}b_{jm}(\sigma_{UE}^2)^2\sigma_j^2(y_{r_{ji}}-\sum_{n=1}^{M}(a_{jn}c_{ni}-b_{jn}d_{ni}))^2}{(\sigma_j^2)^4}\\ &-\sum_{i=1}^{T}\frac{c_{mi}d_{mi}\sigma_j^2-2a_{jm}d_{mi}\sigma_{UE}^2(y_{im_{ji}}-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))}{(\sigma_j^2)^2}\\ &-\sum_{i=1}^{T}\frac{2a_{jm}c_{im}(y_{im_{ji}}-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))}{(\sigma_j^2)^2}\\ &-\sum_{i=1}^{T}\frac{4a_{jm}b_{jm}(\sigma_{UE}^2)^2\sigma_j^2(y_{im_{ji}}-\sum_{n=1}^{M}(a_{jn}d_{ni}+b_{jn}c_{ni}))^2}{(\sigma_j^2)^4}\end{aligned} \tag{4.32}$$

The expectation of which is given by:

$$E[\frac{d\log p(y_1,y_2,...,y_{N_r})}{da_{jm}db_{jm}}]=\frac{2a_{jm}b_{jm}(\sigma_{UE}^2)^2T}{(\sigma_j^2)^2} \tag{4.33}$$

As mentioned previously, the second derivatives of $\log p(y_1,y_2,\cdots,y_{N_r})$ are continuous functions. As a result, according to the Schwartz's theorem [26], its partial derivations are commutative. Consequently, we get:

$$E[\frac{d\log p(y_1,y_2,...,y_{N_r})}{db_{jm}da_{jm}}]=E[\frac{d\log p(y_1,y_2,...,y_{N_r})}{da_{jm}db_{jm}}] \tag{4.34}$$

The $2N_t \times 2N_t$ information matrix of the signal at the j_{th} antenna from (4.25):

$$I(a_j, b_j) = \begin{bmatrix} I_{a_{j1}a_{j1}} & I_{a_{j1}b_{j1}} & I_{a_{j1}a_{j2}} & I_{a_{j1}b_{j2}} & \dots & I_{a_{j1}a_{jM}} & I_{a_{j1}b_{jM}} \\ I_{b_{j1}a_{j1}} & I_{b_{j1}b_{j1}} & I_{b_{j1}a_{j2}} & I_{b_{j1}b_{j2}} & \dots & I_{b_{j1}a_{jM}} & I_{b_{j1}b_{jM}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ I_{a_{jM}a_{j1}} & I_{a_{jM}b_{j1}} & I_{a_{jM}a_{j2}} & I_{a_{jM}b_{j2}} & \dots & I_{a_{jM}a_{jM}} & I_{a_{jM}b_{jM}} \\ I_{b_{jM}a_{j1}} & I_{b_{jM}b_{j1}} & I_{b_{jM}a_{j2}} & I_{b_{jM}b_{j2}} & \dots & I_{b_{jM1}a_{jM}} & I_{b_{jM}b_{jM}} \end{bmatrix}$$

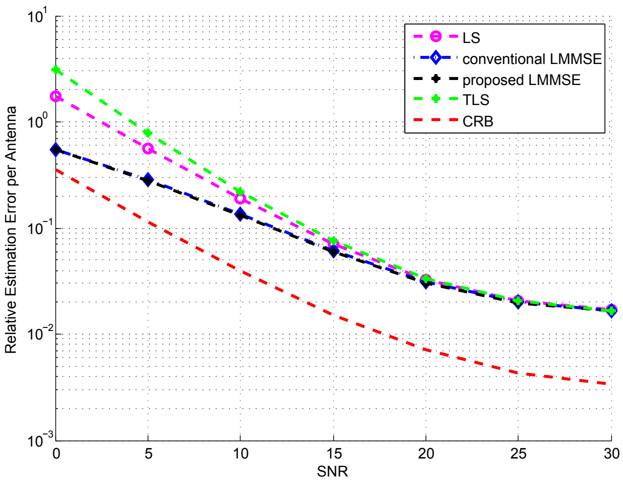
Where $I_{a_{jm}a_{jm}} = -E[\frac{d\log p(y_1,y_2,\ldots,y_N)}{da_{jm}\,da_{jm}}]$, $I_{a_{jm}a_{jm'}} = -E[\frac{d\log p(y_1,y_2,\ldots,y_N)}{da_{jm}\,da_{jm'}}]$, $I_{a_{jm}b_{jm}} = -E[\frac{d\log p(y_1,y_2,\ldots,y_N)}{da_{jm}\,db_{jm}}]$, $I_{b_{jm}b_{jm}} = -E[\frac{d\log p(y_1,y_2,\ldots,y_N)}{db_{jm}\,db_{jm}}]$. The $2N_tN_r \times 2N_tN_r$ information matrix for the signals at all the antennas is as (4.17):

$$I(H) = \begin{bmatrix} I(a_1, b_1) & 0 & \dots & 0 \\ 0 & I(a_2, b_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I(a_{N_r}, b_{N_r}) \end{bmatrix}$$

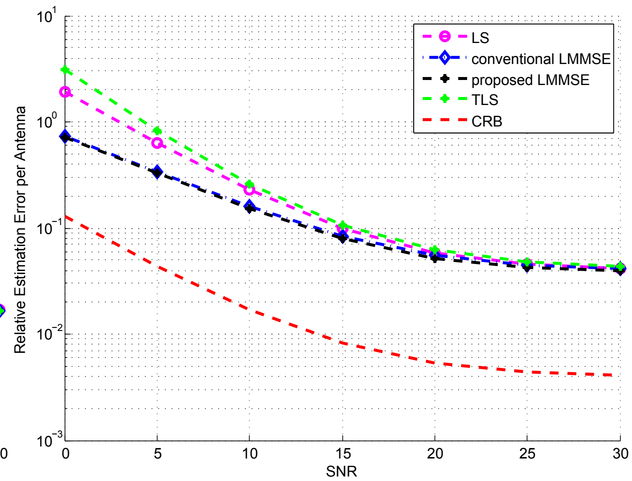
Finally, the Cramer-Rao bound is given by:

$$var(\hat{H}) \geq I(H)^{-1} \tag{4.35}$$

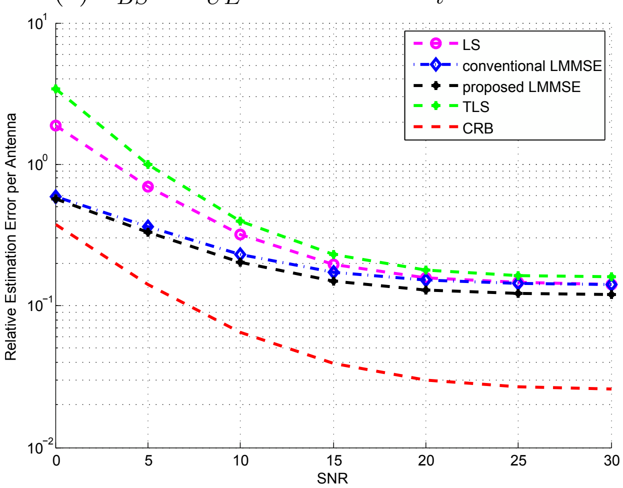
Figure 4.2 presents the LMMSE estimator proposed in Chapter 3 with the LS, the TLS and the conventional, impairment ignoring, LMMSE estimator relatively to the Cramer-Rao bound for a multi-user system with 50 antennas at the base station and 5 or 15 single antenna users as SNR varies from $0dB$ to $30dB$. The level of the impairments is chosen to be 0.05^2 and 0.15^2 as the lower bound and the upper bound of the parameter, respectively, [11]. The simulation was performed over 1000 Monte-Carlo realizations of the channel, the distortion on both ends and the Gaussian noise. Contrary to the single-user scenario, when having multiple users, the estimators do not achieve a performance close to the Cramer-Rao bound as shown in the simulations below.



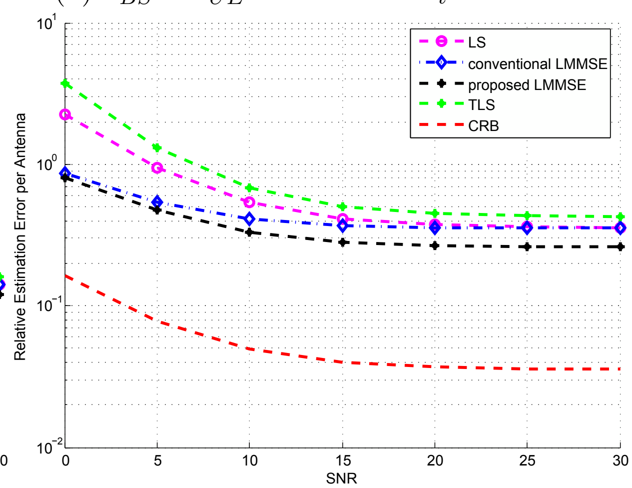
(a) $\kappa_{BS} = \kappa_{UE} = 0.05^2$ and $N_t = 5$ users



(b) $\kappa_{BS} = \kappa_{UE} = 0.05^2$ and $N_t = 15$ users



(c) $\kappa_{BS} = \kappa_{UE} = 0.15^2$ and $N_t = 5$ users



(d) $\kappa_{BS} = \kappa_{UE} = 0.15^2$ and $N_t = 15$ users

Figure 4.2: Estimation error per antenna over SNR. LS, TLS, conventional and proposed LMMSE estimators compared to Cramer-Rao bound for a multi-user system with 50 antennas at the base station.

Chapter 5

Summary and Future Research

5.1 Summary

In this thesis, we considered the problem of channel estimation under hardware impairments for a single and for a multi-user Large MIMO system. As we explained in Chapter 2, Large MIMO systems usually use inexpensive and low quality hardware components which may be imperfect. Instead of modelling the distortion from each impaired component separately, we modelled the total distortion at the base station and the users as Gaussian noise following the work of [31].

We used the LMMSE estimator from [5] for a single-user system, which takes into consideration the impairments on both ends, as a basis to develop the proposed LMMSE estimator for a multi-user system. In order to evaluate the performance of the proposed estimator we compared it with the conventional LMMSE, the LS and the TLS estimators. Furthermore, we calculated the Cramer-Rao bound of the proposed system model.

Under this framework, we showed that the proposed LMMSE channel estimator achieves better performance in terms of relative estimation error per antenna over Signal-to-Noise ratio (SNR) compared to the other estimators, especially as the level of impairments grows. Moreover, the LMMSE channel estimator considering hardware impairments was shown to be close to the Cramer-Rao bound for the single-user case.

5.2 Open Problems and Future Research

Although the proposed estimator achieves better performance compared to estimators used for conventional MIMO there is still room for improvement, as the Cramer-Rao bound shows. The research group in [5] remarks that there may be non-linear estimators which achieve better performance. Finding a non-linear estimator for Large MIMO and evaluating its trade-off between performance and complexity is a problem for future research.

Moreover, a more general problem is channel estimation for Large MIMO under hardware impairments for FDD. As mentioned in Chapter 2, there are research efforts in [15] exploring the subject of FDD for Large MIMO but perfect hardware is assumed. Approaching the problem considering imperfect hardware can be a challenging subject for future work.

Bibliography

[1] A. Assalini, S. Pupolin, and E. Dall'Anese, "Linear MMSE MIMO channel estimation with imperfect channel covariance information," in *Proc. of IEEE International Conference on Communications.*, 2009.

[2] A. Hakkarainen, M. Valkama, J. Werner, and K.R. Dandekar, "Widely-linear beamforming and RF impairment suppression in massive antenna arrays," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 383–397, 2013.

[3] A.B. Gershman and M. Biguesh, "MIMO channel estimation: Optimal training and tradeoffs between estimation techniques," in *Proc. of IEEE International Conference on Communications,* 2004.

[4] A.B. Gershman and M.Biguesh, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," in *Proc. of IEEE Transactions on Signal Processing,* 2006.

[5] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Transactions on Information Theory,* vol. 60, no. 11, pp. 7112-7139, 2014.

[6] E. Björnson, E.G. Larsson, and T.L. Marzetta. "Massive MIMO: Ten myths and one critical question," *To be published on IEEE Communications Magazine*, 2016.

[7] A. Chockalingam, "Low-complexity algorithms for Large MIMO detection," in *Proc. of 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, 2010.

[8] T. Datta, N. Srinidhi, A. Chockalingam, and B.S. Rajan, "Random-restart reactive Tabu search algorithm for detection in large MIMO systems," *IEEE Communications Letters,* vol. 14, no. 12, pp. 1107-1109, 2010.

[9] E. Larsson, O. Edfors, F. Tufvesson, and T.N. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine,* vol. 52, no. 2, pp. 186-195, 2014.

[10] H. Gene and C. F. Van Loan, "An analysis of the total least squares problem." in *Proc. of Society for Industrial and Applied Mathematics*, 1980.

[11] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, Wiley, 2011.

[12] H.Q. Ngo, T.L. Marzetta and E.G. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 2011.

[13] A. Khlifi and R. Bouallegue, "Performance analysis of LS and LMMSE channel estimation techniques for LTE downlink systems," *International Journal of Wireless and Mobile Networks* vol. 3, no. 5, 2011.

[14] A. Lauri and V. Mikko, "On circularity of receiver front-end signals under RF impairments," in *Proc. of Wireless Conference on Sustainable Wireless Technologies*, 2011.

[15] B. Lee, J. Choi, D.J. Love, and Ji-Yun Seol, "Antenna grouping based feedback compression for FDD-based massive MIMO systems," in *Proc. of IEEE International Conference on Communications*, 2015.

[16] P. Li and R.D. Murch, "Multiple output selection-LAS algorithm in large MIMO systems," *IEEE Communications Letters,* vol. 14, no. 5, pp. 399-401, 2008.

[17] S.L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Communications Letters*, vol. 5, no. 9, 2002.

[18] M. Sedaghat, A. Mohammad, R. Mueller, and G. Fischer, "A novel single-RF transmitter for massive MIMO," in *Proc. of 18th International ITG Workshop on Smart Antennas (WSA),* 2014.

[19] S. Mohammed and E. Larsson, "Per-antenna constant envelope precoding for Large multi-user MIMO systems," *IEEE Trans. on Communications*, vol. 61, no. 3, 2013.

[20] N.G. Ngo, E.G. Larsson, and T.L. Marzetta, "Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots," in *Proc. of Allerton Conference on Communication, Control, and Computing*, 2013.

[21] A.J. Paulraj, D.A. Gore, H. Bolcskei, and R.U. Nabar, "An overview of MIMO communications - a key to gigabyte wireless," in *Proc. of the IEEE Conference on Communications*, 2004.

[22] Y. Pei, T.H. Pahm, and Y.C. Liang, "How many RF chains are optimal for large-scale MIMO systems when circuit power is considered?" in *Proc. of Global Communications Conference (GLOBECOM)*, 2012.

[23] H.V. Poor, *An Introduction to signal Detection and Estimation*, Springer, 1994.

[24] F. Rusek, D. Persson, B. Lau, E.G. Larsson, T.L. and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40-60, 2013.

[25] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 68-73, 2004.

[26] J. Schwartz and N. Dunford, *Linear Operators : General Theory*, Wiley, 1988.

[27] N. Srinidhi, S. K. Mohammed, A. Chockalingam, and B.S. Rajan, "Near-ML signal detection in large-dimension linear vector channels using reactive tabu search," *IEEE Transactions on Communications,* vol. 59, no. 11, 2009.

[28] C. Studer, M. Weng, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *Proc. of ITG/IEEE Workshop on Smart Antennas*, 2010.

[29] K. Vardhan, S.K. Mohammed, A. Chockalingam, and B.S. Rajan, "A low-complexity detector for large MIMO systems and multi-carrier CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 3, pp. 473-485, 2008.

[30] H. Yang, M. Hill, and T.L. Marzetta, "Total energy efficiency of cellular large scale antenna system multiple access mobile networks," in *Proc. of IEEE Online Conference on Green Communications*, 2013.

[31] C. Shepard, H. Yu, N. Anand, L.E. Li, T.L. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proc. of International Conference on Mobile Computing and Networking*, 2012.