8-2016

Channel Estimation Overhead Reduction for Downlink FDD Massive MIMO Systems

Abderrahmane Mayouche
University of Arkansas, Fayetteville

Channel Estimation Overhead Reduction for Downlink FDD Massive MIMO Systems

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Electrical Engineering

by

Abderrahmane Mayouche
University of Sciences and Technology Houari Boumediene
Bachelor of science in Computer Science, 2011
University of Sciences and Technology Houari Boumediene
Master of Engineering in Telecommunications, 2013

August 2016
University of Arkansas, Fayetteville

This thesis is approved for recommendation to the Graduate Council

____
Dr. Jing Yang
Thesis director

____
Dr. Jingxian Wu
Committee Member

____
Dr. Roy McCann
Committee Member

Abstract

Massive multiple-input multiple-output (MIMO) is the concept of deploying a very large number of antennas at the base stations (BS) of cellular networks. Frequency-division duplexing (FDD) massive MIMO systems in the downlink (DL) suffer significantly from the channel estimation overhead. In this thesis, we propose a minimum mean square error (MMSE)-based channel estimation framework that exploits the spatial correlation between the antennas at the BS to reduce the latter overhead. We investigate how the number of antennas at the BS affects the channel estimation error through analytical and asymptotic analysis. In addition, we derive a lower bound on the spectral efficiency of the communication system. Close form expressions of the asymptotic MSE and the spectral efficiency lower bound are obtained. Furthermore, perfect match between theoretical and simulation results is observed, and results show the feasibility of our proposed scheme.

Contents

1 Introduction

1.1 Motivation

Massive multiple-input multiple-output (MIMO) is the concept of deploying a very large number of antennas at the base stations (BS) of cellular networks, and has been shown to dramatically mitigate the inter-user interference with simple low complexity precoders and receive combiners [1]. It is one of the highly considered candidates for the deployment of the fifth generation (5G) of cellular networks [2] [3]. It provides the next generation network with capabilities to help satisfy the demand of users by 2020 [4]. Massive MIMO systems (also called Large-Scale Antenna Systems, Large-Scale MIMO, ARGOS, Full-Dimension MIMO, or Hyper-MIMO) has shown the potential to improve the spectral efficiency by order of magnitude through relatively simple processing [5]. Although massive MIMO offers high spectral efficiency for next generation cellular systems, this monumental gain cannot be met without the knowledge of channel state information at the transmitter (CSIT) [6] [7].

Most current cellular systems are implemented in frequency-division duplexing (FDD). One apparent advantage of FDD is that it is generally regarded as more efficient in terms of delay sensitivity and traffic symmetry compared to the time-division duplexing (TDD) mode [8]. Meanwhile, there is much less research work on FDD massive MIMO than on TTD. One of the main reasons for this tendency to the TDD mode is the feasibility of obtaining CSIT in this mode. In TDD, channel state information at the BS can be obtained via the use of uplink pilots under the assumption of channel reciprocity [9]. However, FDD massive MIMO suffers greatly from the pilot and feedback overhead when conventional channel estimation techniques are used. As the number of antennas at the BS increases, the number of pilot symbols and the CSI feedback overhead become prohibitively large. Hence, obtaining CSI at the transmitter is essential to exploit the benefits of FDD massive MIMO [10].

We find in the literature some work to tackle the channel estimation overhead problem. The Joint Spatial Division and Multiplexing (JSDM) approach diminishes the channel estimation overhead [11]. It reduces the CSIT by exploiting the structure of channel correlation. Another recent method to reduce the latter overhead was proposed in [12]. The idea was to design downlink pilot sequences and uplink channel feedback codebooks in a way that it decreases the said overhead by exploiting the multiuser spatial channel correlation. In [13], a feedback overhead reduction technique was introduced based on the antenna grouping concept. Whereas in [14], a CSIT estimation method was proposed to reduce the pilot and feedback overhead by using the compressive sensing concept. It should be noted that this is not an exhaustive list of the methods used to reduce the channel estimation overhead in FDD massive MIMO systems.

The closest to our work, to the best of our knowledge, is presented in [15]. Adriana et al. proposed a strategy that exploits transmit antenna correlation, which is due to the short distance between the antennas in the massive MIMO system. Their idea is to acquire instantaneous CSI for a subset of antennas and use averaging over the received CSI to obtain the channel estimates at the remaining antennas.

1.2 Objectives

The focus of our work is to reduce the pilot and feedback overhead in downlink FDD massive MIMO systems. To this end, we use the same idea as in [15], namely, the BS sends pilots to the user from just a subset of the antennas, the user estimates the channel at the antennas that have sent pilots and feeds back the estimated coefficients to the BS, then the BS computes the channel estimates at the remaining antennas through exploiting the spatial correlation between the antennas at the BS. Our work differs from [15] in many ways. First, we do not use averaging over the estimated coefficients, instead we propose to use the minimum mean square error (MMSE) estimator to compute the channel at the antennas that did not send pilots. Second, we derived analytical as well as asymptotic channel estimation MSE of the system. We also derived a lower bound on the spectral efficiency of the system. Finally, our numerical results were obtained using system simulations and not Monte Carlo simulations. Our approach results in a 50% reduction of the channel estimation overhead.

It should be noted that our proposed channel estimation framework was inspired from the work in [16], where the context was in a temporally correlated single-input single-output (SISO) channel over a fast time-varying fading channel.

1.3 Outline

The thesis is presented as follows, chapter 1 outlines a literature review on overhead reductions schemes for FDD massive MIMO systems and our contributions, chapter 2 presents the theoretical analysis of our proposed scheme, where we first introduce the system model, then we present the derivations of the analytical and the asymptotic mean square error (MSE) of the channel estimation as well as the spectral efficiency of the system with imperfect channel estimation. Chapter 3 summarizes simulation results. Finally, in chapter 4 we terminate our thesis with conclusions and future work.

2 Theoretical analysis

2.1 System Model

We consider a downlink FDD Massive MIMO system of N_t transmit antennas and a single receive antenna. Particularly, the deployment scenario of our system is according to the linear antenna array configuration [17], as depicted in Fig 1. Equivalently, our system can be called an FDD massive multiple-input single-output (MISO) system, which employs pilot-assisted channel estimation and experiences Rayleigh fading. The channel coefficients between the antennas at the BS are correlated, and therefore the fading channel is spatially correlated.

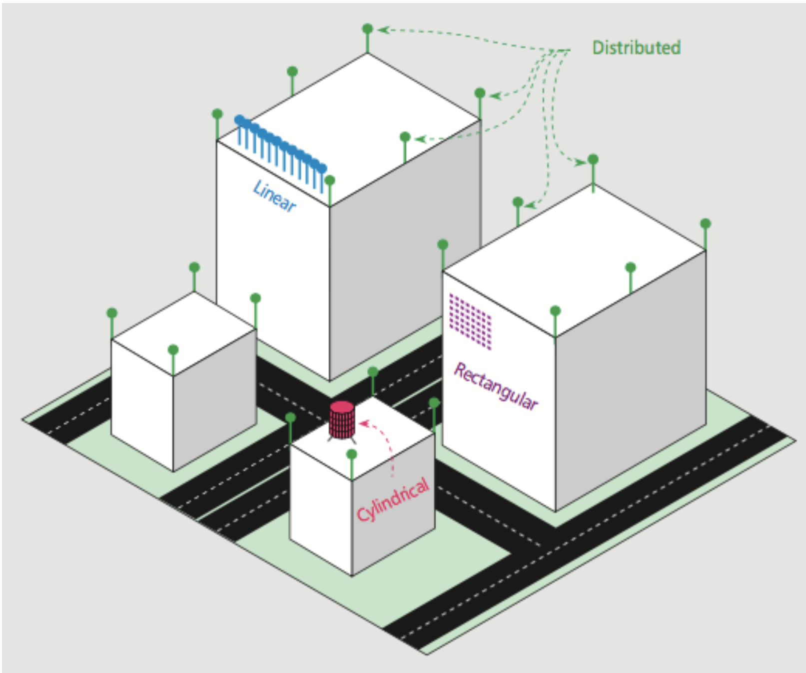

Figure 1: Some possible antenna configurations and deployment scenarios for a massive MIMO base station [17].

The BS sends T pilot signals using only N_p antennas out of N_t, see Fig. 2. In our work, we fix the value of N_p to $\frac{N_t}{2}$, which results in a *50%* reduction in the pilot and feedback overhead, and we consider $T = N_p$ for maximal overhead reduction. It should be noted that more than 50% reduction of the overhead results in unacceptable communication performance, hence choosing $N_p = \frac{N_t}{2}$.

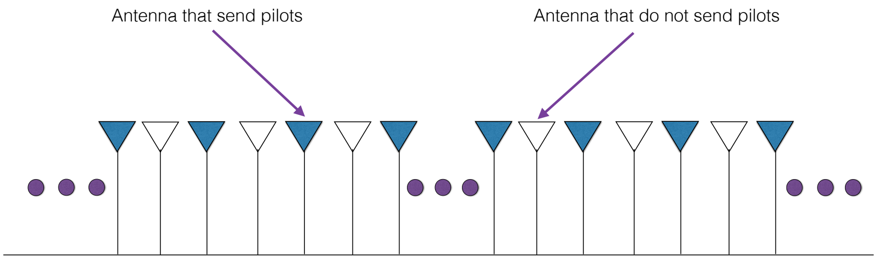

Figure 2: Antennas that send pilots in a linear massive MIMO system.

We also define the pilot percentage as $\delta = \frac{N_p}{N_t} = \frac{1}{2}$. The input-output relation can be written as in [18]:

$$\mathbf{y}_p = \mathbf{X}^H \mathbf{h}_p + \mathbf{z}_p \tag{1}$$

where $\mathbf{y}_p \in \mathcal{C}^{T\times 1}$ is the received signal, $\mathbf{X} = [\mathbf{x}[0] ... \mathbf{x}[T-1]] \in \mathcal{C}^{N_p \times T}$ is the transmitted pilot signals in a matrix form, each column $\mathbf{x}[i]$ represents one pilot signal sent from N_p transmit antennas, $[\cdot]^H$ denotes the matrix Hermitian operator, $\mathbf{h}_p \in \mathcal{C}^{N_p \times 1}$ are the channel fading coefficients, $\mathbf{z}_p \in \mathcal{C}^{T\times 1}$ is the additive white Gaussian noise vector at the receiver with covariance matrix $\mathbf{R}_z = \sigma_z^2 \mathbf{I}_N$.

Since we consider a spatially correlated block-fading channel, the channel vector $\mathbf{h}$ can be modeled as in [19]:

$$\mathbf{h} = \mathbf{R}^{\frac{1}{2}} \mathbf{h}_w \tag{2}$$

where $\mathbf{R} = \mathbb{E}[\mathbf{h}\mathbf{h}^H]$ is the spatial correlation matrix of the channel vector $\mathbf{h}$, and $\mathbf{h}_w \sim \mathcal{CN}(0, I_{N_p})$[1] is an independent and identically distributed (i.i.d.) complex Gaussian vector.

[1] $\mathcal{CN}$ denotes the circular symmetric complex Gaussian distribution

The correlation matrix $\mathbf{R}$ can be modeled in many ways depending on the spatial positioning of the antennas. The exponential model is commonly used to model spatial correlation matrices [20] [21] mainly because of its simplicity. Even with its simplicity, the exponential model accurately characterizes uniform linear array (ULA) antennas in different scenarios, as the experiments in [22] show.

The cross correlation between the antenna of index i and the antenna of index j channels according to the exponential model is given by

$$\rho(i-j) = r^{|i-j|} \tag{3}$$

where $0 \leq r \leq 1$. The correlation coefficient r depends mainly on the distance between the antennas at the BS and also on the user locations.

2.2 Channel estimation MSE

In this section, we investigate how the number of antennas at the base station affects the channel estimation error through analytical and asymptotic analysis. The channel estimation is performed in two steps: the channel coefficients at the antennas that send pilots are estimated, then the channel coefficients at the remaining antennas are obtained by the use of MMSE interpolation over the estimated coefficients.

2.2.1 MMSE Channel Estimation at Pilot Antennas

The BS sends a pilot signal of T symbols from its N_p antennas. The received signal $\mathbf{y}_p \in \mathcal{C}^{T\times 1}$ at the user is used to estimate the channel vector $\mathbf{h}_p$ by minimizing the average MSE $\sigma^2_{p,N_p} = \frac{1}{N_p}\mathbb{E}(\|\hat{\mathbf{h}}_p - \mathbf{h}_p\|^2)$ with $\|\cdot\|$ denoting the Euclidean norm, which can be written as

$$\hat{\mathbf{h}}_p = \mathbb{E}(\mathbf{h}_p|\mathbf{y}_p) = \mathbf{W}_p^H \mathbf{y}_p \tag{4}$$

where $\mathbb{E}$ is the mathematical expectation operator, $\hat{\mathbf{h}}_p \in \mathcal{C}^{N_p\times 1}$ is the estimate of $\mathbf{h}_p$, and $\mathbf{W}_p$ is the MMSE estimation matrix such that $\mathbf{W}_p^H = \Sigma_{\mathbf{h}_p\mathbf{y}_p}\Sigma_{\mathbf{y}_p\mathbf{y}_p}^{-1}$

We derive $\Sigma_{\mathbf{h}_p\mathbf{y}_p}$ and $\Sigma_{\mathbf{y}_p\mathbf{y}_p}$ as follows:

$$\begin{aligned}\Sigma_{\mathbf{h}_p\mathbf{y}_p} &= \mathbb{E}[\mathbf{h}_p\mathbf{y}_p^H] \\ &= \mathbb{E}[\mathbf{h}_p(\mathbf{h}_p^H\mathbf{X}+\mathbf{z}_p^H)] \\ &= \mathbb{E}[\mathbf{h}_p\mathbf{h}_p^H]\mathbf{X}+\mathbb{E}[\mathbf{h}_p]\mathbb{E}[\mathbf{z}_p^H] \\ &= \mathbf{R}_p\mathbf{X}\end{aligned} \tag{5}$$

Note that $\mathbb{E}[\mathbf{z}_p] = 0$ because the noise is zero mean.

$$\begin{aligned}\Sigma_{\mathbf{y}_p\mathbf{y}_p} &= \mathbb{E}[\mathbf{h}_p\mathbf{y}_p^H] \\ &= \mathbb{E}[(\mathbf{X}^H\mathbf{h}_p+\mathbf{z}_p)(\mathbf{X}^H\mathbf{h}_p+\mathbf{z}_p)^H] \\ &= \mathbb{E}[\mathbf{X}^H\mathbf{h}_p\mathbf{h}_p^H\mathbf{X}+\mathbf{X}^H\mathbf{h}_p\mathbf{z}_p^H+\mathbf{z}_p\mathbf{h}_p^H\mathbf{X}+\mathbf{z}_p\mathbf{z}_p^H] \\ &= \mathbf{X}^H\mathbb{E}[\mathbf{h}_p\mathbf{h}_p^H]\mathbf{X}+\mathbf{X}^H\mathbb{E}[\mathbf{h}_p]\mathbb{E}[\mathbf{z}_p^H]+\mathbb{E}[\mathbf{z}_p]\mathbb{E}[\mathbf{h}_p^H]\mathbf{X}+\mathbb{E}[\mathbf{z}_p\mathbf{z}_p^H] \\ &= \mathbf{X}^H\mathbf{R}_p\mathbf{X}+\sigma_z^2\mathbf{I}_{N_p}\end{aligned} \tag{6}$$

Hence, by combining (5) and (6), we obtain the following:

$$\mathbf{W}_p^H = \mathbf{R}_p\mathbf{X}\left(\mathbf{X}^H\mathbf{R}_p\mathbf{X}+\frac{1}{\gamma_0}\mathbf{I}_{N_p}\right)^{-1} \tag{7}$$

with $\mathbf{R}_p = \mathbb{E}[\mathbf{h}_p\mathbf{h}_p^H] \in \mathcal{C}^{N_p\times N_p}$ being the correlation matrix between antennas that send pilots, which is a symmetric Toeplitz matrix with its row defined as $[\rho(K\times 0),\rho(K\times 1),...,\rho(K\times(N_p-1))]$, such that $K=\frac{1}{\delta}$ where δ is the pilot percentage, ρ is defined in (3), γ_0 represents the SNR without fading, and $\mathbf{I}_{N_p}$ is a size-N_p identity matrix.

The error covariance matrix is defined as $\mathbf{R}_{ee}^{p} = \mathbb{E}[\mathbf{e}_p \mathbf{e}_p^H]$ with $\mathbf{e}_p = (\hat{\mathbf{h}}_p - \mathbf{h}_p)$ and $\hat{\mathbf{h}}_p$ being the estimate of $\mathbf{h}_p$, and can be computed as in [23] as:

$$\begin{aligned}
\mathbf{R}_{ee}^{p} &= \mathbb{E}[\mathbf{e}_p(\hat{\mathbf{h}}_p - \mathbf{h}_p)^H] \\
&= \mathbb{E}[\mathbf{e}_p \hat{\mathbf{h}}_p^H] - \mathbb{E}[\mathbf{e}_p \mathbf{h}_p^H] \\
&= -\mathbb{E}[(\hat{\mathbf{h}}_p - \mathbf{h}_p)\mathbf{h}_p^H] \\
&= -\mathbb{E}[\hat{\mathbf{h}}_p \mathbf{h}_p^H] + \mathbb{E}[\mathbf{h}_p \mathbf{h}_p^H] \\
&= \mathbf{R}_p - \mathbb{E}[\mathbf{W}_p^H \mathbf{y}_p \mathbf{h}_p^H] \\
&= \mathbf{R}_p - \mathbb{E}[\mathbf{W}_p^H (\mathbf{X}^H \mathbf{h}_p + \mathbf{z}_p)\mathbf{h}_p^H] \\
&= \mathbf{R}_p - \mathbf{W}_p^H \mathbf{X}^H \mathbb{E}[\mathbf{h}_p \mathbf{h}_p^H] - \mathbf{W}_p^H \mathbb{E}[\mathbf{z}_p]\mathbb{E}[\mathbf{h}_p^H] \\
&= \mathbf{R}_p - \mathbf{R}_p \mathbf{X}\left(\mathbf{X}^H \mathbf{R}_p \mathbf{X} + \frac{1}{\gamma_0}\mathbf{I}_{N_p}\right)^{-1} \mathbf{X}^H \mathbf{R}_p.
\end{aligned} \tag{8}$$

Note that $\mathbb{E}[\mathbf{e}_p \hat{\mathbf{h}}_p^H] = 0$ because of the orthogonal principle. Hence, the average channel estimation MSE is calculated as

$$\sigma_{p,N_p}^2 = \frac{1}{N_p}\mathbb{E}(\|\mathbf{h}_p - \hat{\mathbf{h}}_p\|^2) = \frac{1}{N_p} tr(\mathbf{R}_{ee}^{p}) \tag{9}$$

where $tr(\cdot)$ represents the trace operation. The computation of the average MSE involves matrix inversion and trace operation. For this reason, it is not straightforward to analyze the impact of the number of antennas at the BS on the channel estimation MSE. Thus, we resort to asymptotic analysis by letting $N_p \to \infty$ and $N_t \to \infty$ while keeping a finite pilot percentage δ.

For simplicity of the derivation of the asymptotic MSE, we assume that the average energy of pilot signals is normalized to 1, $\mathbb{E}(\|\mathbf{x}\|^2) = 1$, and $\mathbf{X}^H \mathbf{X} = \mathbf{I}_T$, i.e., the pilot signals are orthogonal to each other. This assumption can be easily met using constant amplitude symbols, such as phase shift keying. It should be noted that data symbols are not subject to this assumption.

Consequently, the error covariance matrix reduces to

$$\mathbf{R}_{ee}^{p} = \mathbf{R}_{p} - \mathbf{R}_{p}\left(\mathbf{R}_{p} + \frac{1}{\gamma_0}\mathbf{I}_{N_p}\right)^{-1}\mathbf{R}_{p}. \tag{10}$$

Asymptotic results are presented as follows.

*Proposition*1 : When $N_p \to \infty$ and $N_t \to \infty$, while keeping a finite pilot percentage δ, the asymptotic channel estimation MSE, $\sigma_p^2 = \lim_{N_p\to\infty}\sigma_{p,N_p}^2$, of the estimated channel coefficients at the pilot antennas corresponds to results in [24, Eq. (11)] and can be represented as

$$\sigma_p^2 = \left[(1+\gamma_0)^2 + \frac{4\gamma_o r^{\frac{2}{\delta}}}{(1-r^{\frac{2}{\delta}})}\right]^{-\frac{1}{2}}. \tag{11}$$

Proof : By taking into consideration the pilot percentage δ, the reduced error covariance matrix becomes:

$$\mathbf{R}_{ee}^{p} = \mathbf{R}_{p} - \mathbf{R}_{p}\left(\mathbf{R}_{p} + \frac{\delta}{\gamma_0}\mathbf{I}_{N_p}\right)^{-1}\mathbf{R}_{p}. \tag{12}$$

The eigenvalue representation of $\mathbf{R}_{ee}^{p}$ is:

$$\begin{aligned}
\mathbf{R}_{ee}^{p} &= \frac{1}{N_p}\sum_{n=1}^{N_p}\left[\lambda_n - \left(\lambda_n + \frac{\delta}{\gamma_0}\right)^{-1}\lambda_n^2\right] \\
&= \frac{1}{N_p}\sum_{n=1}^{N_p}\left[\lambda_n - \left(\frac{\gamma_0\lambda_n + \delta}{\gamma_0}\right)^{-1}\lambda_n^2\right] \\
&= \frac{1}{N_p}\sum_{n=1}^{N_p}\left[\lambda_n - \frac{\gamma_0\lambda_n^2}{\gamma_0\lambda_n + \delta}\right] \\
&= \frac{1}{N_p}\sum_{n=1}^{N_p}\left[\frac{\delta\lambda_n}{\gamma_0\lambda_n + \delta}\right] \\
&= \frac{1}{N_p}\sum_{n=1}^{N_p}\left[\frac{\gamma_0\lambda_n + \delta}{\delta\lambda_n}\right]^{-1} \\
&= \frac{1}{N_p}\sum_{n=1}^{N_p}\left(\frac{1}{\lambda_n} + \frac{\gamma_0}{\delta}\right)^{-1}
\end{aligned} \tag{13}$$

Using Szego's Theorem [25], when $N_p \to \infty$, the average channel estimation MSE, σ^2_{p,N_p}, can be rewritten as

$$\sigma_p^2 = \lim_{N_p \to \infty} \sigma^2_{p,N_p} = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left[\frac{1}{\Lambda_p(\Omega)} + \frac{\gamma_0}{\delta} \right]^{-1} d\Omega \tag{14}$$

where $\Lambda_p(\Omega) = \sum_{-\infty}^{+\infty} \rho^{|n|} e^{-jn\Omega}$ is the discrete-time Fourier transform (DTFT) of sequence of the correlation coefficients $\{\rho^{|n|}\}_n$, and can be computed as follows:

$$\begin{aligned} \Lambda_p(\Omega) &= \sum_{-\infty}^{+\infty} \rho^{|n|} e^{-jn\Omega} \\ &= \sum_{-\infty}^{-1} \left(\rho e^{-j\Omega}\right)^{-n} + \sum_{0}^{+\infty} \left(\rho e^{-j\Omega}\right)^{n} \\ &= \sum_{1}^{+\infty} \left(\rho e^{-j\Omega}\right)^{-n} + \sum_{0}^{+\infty} \left(\rho e^{-j\Omega}\right)^{n} \\ &= \frac{\rho e^{-j\Omega}}{1 - \rho e^{-j\Omega}} + \frac{1}{1 - \rho e^{-j\Omega}} \\ &= \frac{1 - \rho^2}{1 - 2\rho \cos\Omega + \rho^2} \end{aligned} \tag{15}$$

Consequently, σ_p^2 becomes:

$$\begin{aligned} \sigma_p^2 &= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left[\frac{1 - 2\rho\cos\Omega + \rho^2}{1 - \rho^2} + \frac{\gamma_0}{\delta} \right]^{-1} d\Omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left(\frac{\delta(1-\rho^2)}{\delta(1+\rho^2) + \gamma_0(1-\rho^2) - 2\delta\rho\cos\Omega} \right) d\Omega. \end{aligned} \tag{16}$$

We solve the integral in (16) with [26, Eq. (2.553.3)], we obtain result in (11).

2.2.2 MMSE Channel Interpolation

Once the channel coefficients are estimated at the user, the latter feeds back the estimated coefficients, $\hat{\mathbf{h}}_p$, to the BS. These estimated coefficients can be interpolated to obtain the channel estimates at the remaining antennas as follows

$$\hat{\mathbf{h}}_n = \mathbb{E}(\mathbf{h}_n|\hat{\mathbf{h}}_p) = \Sigma_{\mathbf{h}_n\hat{\mathbf{h}}_p}\Sigma_{\hat{\mathbf{h}}_p\hat{\mathbf{h}}_p}\hat{\mathbf{h}}_p \tag{17}$$

We derive $\Sigma_{\mathbf{h}_n\hat{\mathbf{h}}_p}$ and $\Sigma_{\hat{\mathbf{h}}_p\hat{\mathbf{h}}_p}$ as follows:

$$\begin{aligned}\Sigma_{\mathbf{h}_n\hat{\mathbf{h}}_p} &= \mathbb{E}[\mathbf{h}_n\hat{\mathbf{h}}_p^H] \\ &= \mathbb{E}[\mathbf{h}_n(\mathbf{h}_p^H\mathbf{X}+\mathbf{z}_p^H)\mathbf{W}_p] \\ &= \mathbb{E}[\mathbf{h}_n\mathbf{h}_p^H]\mathbf{X}\mathbf{W}_p + \mathbb{E}[\mathbf{h}_n]\mathbb{E}[\mathbf{z}_p^H]\mathbf{W}_p \\ &= \mathbf{R}_{np}\mathbf{X}\mathbf{W}_p \end{aligned} \tag{18}$$

where $\hat{\mathbf{h}}_n \in \mathcal{C}^{N_p\times 1}$ is the estimate of $\mathbf{h}_n$, $\mathbf{R}_{np} = \mathbb{E}(\mathbf{h}_n\mathbf{h}_p^H) \in \mathcal{R}^{N_p\times N_p}$ is the cross-correlation matrix between channel fading vectors $\mathbf{h}_n$ and $\mathbf{h}_p$, which is a Toeplitz matrix with its first row being $[\rho(1), \rho(K-1), ..., \rho((N_p-1)\times K-1))]$ and the first column being $[\rho(1), \rho(K+1), ..., \rho((N_p-1)\times K+1))]^T$ such that K = 1 where δ is the pilot percentage, with $[\cdot]^T$ denoting the transpose operator.

$$\begin{aligned}\Sigma_{\hat{\mathbf{h}}_p\hat{\mathbf{h}}_p} &= \mathbb{E}[\hat{\mathbf{h}}_p\hat{\mathbf{h}}_p^H] \\ &= \mathbb{E}[\mathbf{W}_p^H(\mathbf{X}^H\mathbf{h}_p+\mathbf{z}_p)(\mathbf{X}^H\mathbf{h}_p+\mathbf{z}_p)^H\mathbf{W}_p] \\ &= \mathbf{W}_p^H\mathbb{E}[\mathbf{X}^H\mathbf{h}_p\mathbf{h}_p^H\mathbf{X}+\mathbf{X}^H\mathbf{h}_p\mathbf{z}_p^H+\mathbf{z}_p\mathbf{h}_p^H\mathbf{X}+\mathbf{z}_p\mathbf{z}_p^H]\mathbf{W}_p \\ &= \mathbf{W}_p^H\left[\mathbf{X}^H\mathbb{E}[\mathbf{h}_p\mathbf{h}_p^H]\mathbf{X}+\mathbf{X}^H\mathbb{E}[\mathbf{h}_p]\mathbb{E}[\mathbf{z}_p^H]+\mathbb{E}[\mathbf{z}_p]\mathbb{E}[\mathbf{h}_p^H]\mathbf{X}+\mathbb{E}[\mathbf{z}_p\mathbf{z}_p^H]\right]\mathbf{W}_p \\ &= \mathbf{W}_p^H\left[\mathbf{X}^H\mathbf{R}_p\mathbf{X}+\sigma_z^2\mathbf{I}_{N_p}\right]\mathbf{W}_p \end{aligned} \tag{19}$$

Hence, by combining (18) and (20), we obtain the following:

$$
\begin{aligned}
\hat{\mathbf{h}}_n &= \Sigma_{\mathbf{h}_n\hat{\mathbf{h}}_p}\Sigma_{\hat{\mathbf{h}}_p\hat{\mathbf{h}}_p}\hat{\mathbf{h}}_p \\
&= \mathbf{R}_{np}\mathbf{X}\mathbf{W}_p\big(\mathbf{W}_p^H\big[\mathbf{X}^H\mathbf{R}_p\mathbf{X}+\sigma_z^2\mathbf{I}_{N_p}\big]\mathbf{W}_p\big)^{-1}\mathbf{W}_p^H\mathbf{y}_p \\
&= \mathbf{R}_{np}\mathbf{X}\left(\mathbf{X}^H\mathbf{R}_p\mathbf{X}+\frac{1}{\gamma_0}\mathbf{I}_{N_p}\right)^{-1}\mathbf{y}_p
\end{aligned} \tag{20}
$$

The corresponding error correlation matrix of the channel estimation at non-pilot antennas, $\mathbf{R}_{ee}^n \triangleq \mathbb{E}\big[(\hat{\mathbf{h}}_n-\mathbf{h}_n)(\hat{\mathbf{h}}_n-\mathbf{h}_n)^H\big]$, can be computed in the same way as $\mathbf{R}_{ee}^p$ as

$$
\mathbf{R}_{ee}^n = \mathbf{R}_{nn}-\mathbf{R}_{np}\mathbf{X}\left(\mathbf{X}^H\mathbf{R}_p\mathbf{X}+\frac{1}{\gamma_0}\mathbf{I}_{N_p}\right)^{-1}\mathbf{X}^H\mathbf{R}_{pn} \tag{21}
$$

where $\mathbf{R}_{nn}=\mathbb{E}(\mathbf{h}_n\mathbf{h}_n^H)=\mathbf{R}_p$ is the auto-correlation matrix of the fading vector $\mathbf{h}_n$, and $\mathbf{R}_{pn}=\mathbf{R}_{np}^H$.

The average channel estimation MSE of the spatial interpolation is therefore given by

$$
\sigma_{n,N_p}^2 = \frac{1}{N_p}\mathbb{E}(\|\mathbf{h}_n-\hat{\mathbf{h}}_n\|^2) = \frac{1}{N_p}tr(\mathbf{R}_{ee}^n). \tag{22}
$$

However, when assuming orthogonal pilot signals, the error correlation matrix of the channel estimation at the non-pilot antennas reduces to

$$\mathbf{R}_{ee}^{n} = \mathbf{R}_{nn} - \mathbf{R}_{np}\left(\mathbf{R}_{p} + \frac{1}{\gamma_0}\mathbf{I}_{N_p}\right)^{-1}\mathbf{R}_{pn} \tag{23}$$

Hence, the asymptotic channel estimation MSE at non-pilot antennas results are presented in the following proposition.

*Proposition*2 : When $N_p \to \infty$ and $N_t \to \infty$, while keeping a finite pilot percentage δ, the asymptotic channel estimation MSE , $\sigma_n^2 = \lim_{N_p \to \infty} \sigma_{n,N_p}^2$, of the estimated channel coefficients at the non-pilot antennas can be expressed as

$$\sigma_n^2 = \left(\frac{1}{\gamma_0} + \frac{1 - r^{\frac{1}{\delta}}}{1 + r^{\frac{1}{\delta}}}\right)^{\frac{1}{2}} \left(\frac{1}{\gamma_0} + \frac{1 + r^{\frac{1}{\delta}}}{1 - r^{\frac{1}{\delta}}}\right)^{-\frac{1}{2}}. \tag{24}$$

Results were derived in exact same way as in σ_p^2, equation (11). Note that the above equation has the same form as in [24, Eq. (17)], where the asymptotic MSE of interpolation was derived in the context of distortion-tolerant wireless sensor networks.

2.3 Spectral efficiency with imperfect channel estimation

The estimated channel coefficient $\hat{h}_i$ is complex Gaussian random variable with zero mean and a variance of $1-\sigma_e^2$, namely $\hat{h}_i \sim \mathcal{CN}(0, 1-\sigma_e^2)$ where σ_e^2 represents the channel estimation MSE. $|\hat{h}_i|^2$ is therefore exponentially distributed with parameter $(1-\sigma_e^2)^{-1}$. Hence, the average capacity lower bound of the MISO system with per-antenna constraint can be expressed as

$$\bar{C}_{low} = \int_0^{+\infty} \log_2\left(1 + \frac{1}{\sigma_e^2 + \frac{1}{\gamma_0}}\alpha\right) P_\alpha(\alpha) d\alpha \tag{25}$$

with $\alpha = \sum_{i=1}^{N_t} |\hat{h}_i|^2$, which is the sum of correlated exponential random variables, and $P_\alpha(\alpha)$ is the probability density function of α [27] and is defined as

$$P_\alpha(\alpha) = \sum_{i=1}^{N_t} \prod_{\substack{j=1 \\ j\neq i}}^{N_t} \frac{(\lambda_i)^{N_t-1}}{\lambda_j - \lambda_i} \exp\left(\frac{\alpha}{\lambda_i}\right) \tag{26}$$

where λ_i and λ_j are the eigenvalues of the covariance matrix $R_{\hat{h}}$ that is defined as $\mathbb{E}(\hat{\mathbf{h}}\hat{\mathbf{h}}^H)$. Note that the eigenvalues have to be different.

The spectral efficiency lower bound, $\eta_{low} = (1-\delta)\bar{C}_{low}$, is expressed as

$$\eta_{low} = \frac{(1-\delta)}{\ln 2} \sum_{i=1}^{N_t} \exp(\beta) Ei(-\beta) \prod_{\substack{j=1 \\ j\neq i}}^{N_t} \frac{(\lambda_i)^{N_t-1}}{\lambda_j - \lambda_i} \tag{27}$$

with $\beta = \frac{1}{\lambda_i}(\sigma_e^2 + \frac{1}{\gamma_0})$, and $Ei(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$ is the exponential integral.

3 Simulation results

In this section, we will verify and validate the feasibility of the proposed channel estimation scheme through analytical and simulation results.

3.1 Impact of number of antennas on the channel estimation MSE

In Fig. 3, we plot the analytical, asymptotic, and simulated channel estimation MSE at both antennas that send pilots and the ones that do not as a function of the number of antennas N_t. The parameters used in the simulation are $r = 0.9$, and $\gamma_0 = 10\ dB$. The analytical results are calculated according to (9) and (22), and the asymptotic results at the pilot antennas and the non-pilot antennas are defined in (11) and (24), respectively. We observe that the channel estimation MSE decreases as the total number of antennas, N_t, increases and it converges to the asymptotic MSE, with the non-pilot MSE values greater than the pilot MSE values. Increasing the number of antennas increases leads to more coefficients to estimate the channel, which results in better channel estimation, and hence the decrease of the channel estimation MSE.

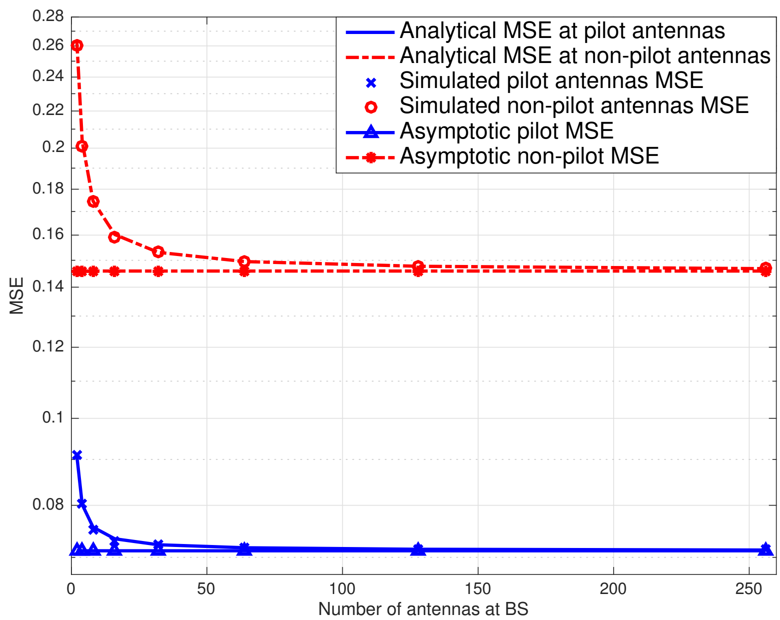

Figure 3: The MSE of the channel estimation as a function of N_t.

3.2 Impact of the SNR on the channel estimation MSE

Fig. 4 depicts the channel estimation MSE as a function of the SNR. The values used for the simulation are $r = 0.9$ and $N_t = 256$. The analytical results are calculated according to (9) and (22), and the asymptotic results at the pilot antennas and the non-pilot antennas are defined in (11) and (24), respectively. We observe that the channel estimation MSE at pilot-antennas decreases linearly with the SNR, because the latter MSE converges to zero as the SNR increases, as (10) indicates, i.e. when SNR tends ∞, the term inside the parentheses reduces to $\mathbf{R}_p$, which leads to 0 in the channel estimation MSE overall equation. However, the non-pilot MSE decreases with the SNR and remains constant after a certain value of γ_0. This can be explained by examining the channel estimation MSE expression in (23), the term with the γ_0 will vanish with high SNR values, but a constant term will remain mainly due to the cross correlation between the channel coefficients.

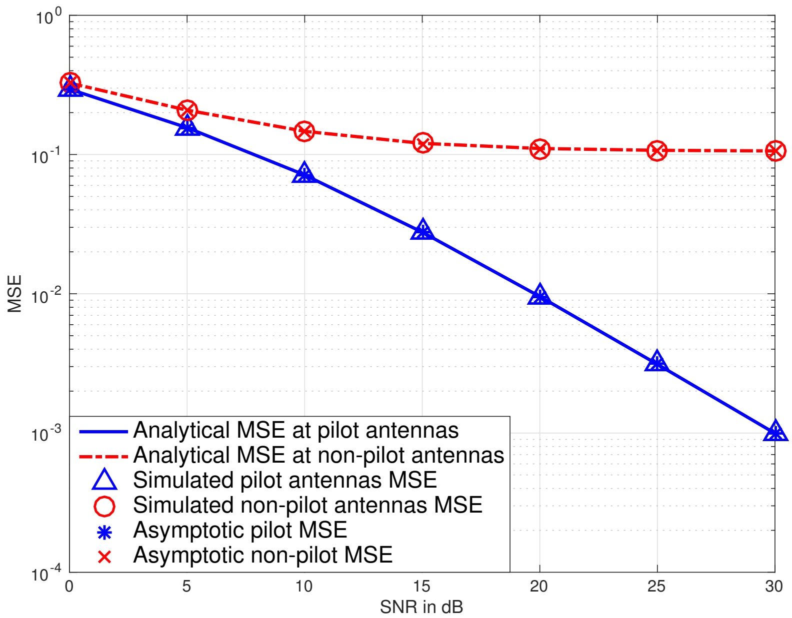

Figure 4: The channel estimation MSE as a function of the SNR.

3.3 Effect of the channel estimation MSE and SNR on the Spectral Efficiency

Fig. 5 plots the spectral efficiency lower bound η_{low}, defined in (27), as a function of SNR. One curve is computed using the pilot channel estimation MSE and the other one using the non-pilot channel estimation MSE, defined in equations (11) and (24), respectively. The values used for the simulation are $r = 0.8$ and $N_t = 128$. η_{low} with the pilot MSE increases linearly with the SNR because η_{low} is inversely proportional to the channel estimation MSE. However, η_{low} with the non-pilot MSE increases with the SNR and remains constant after a certain value of γ_0, behaving in accordance with the non-pilot channel estimation MSE. Hence, transmitting data through the antennas that send pilots leads to higher spectral efficiency when compared to data transmission using the non-pilot antennas.

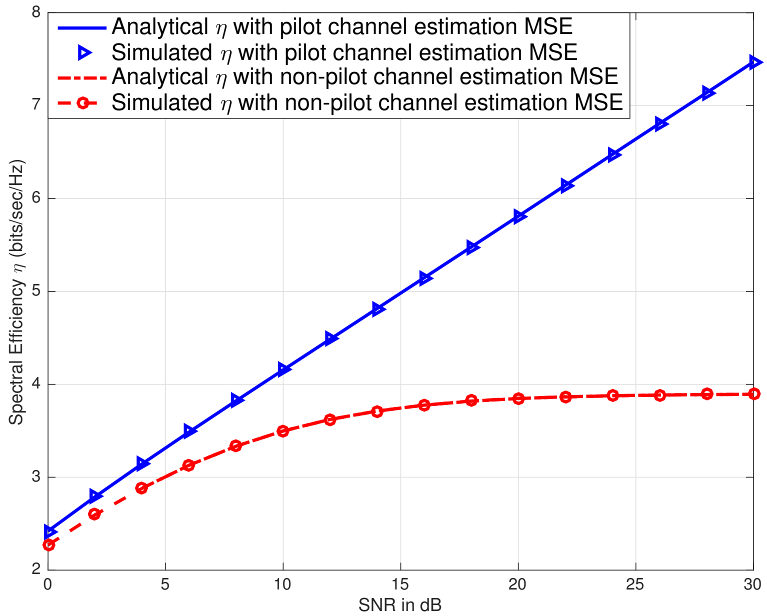

Figure 5: he spectral efficiency η as a function of the SNR.

4 Conclusions and future work

An MMSE-based channel estimation scheme for downlink FDD massive MIMO systems was introduced in this thesis to reduce the pilot and feedback overhead by *50%*, through exploiting the spatial correlation between the antennas at the base station. Analytical and asymptotic channel estimation MSE along with the spectral efficiency lower bound of the system were derived. A perfect match was observed between analytical and simulated results. Moreover, simulation results show that the channel estimation MSE for the interpolated channel coefficients is less than 0.2, which indicates an acceptable communication performance.

As for future work, experimentation of our proposed scheme might be performed to assess the practicality of the scheme in deploying future cellular communication systems. Furthermore, a complexity analysis for our proposed channel estimation framework can me carried out in comparison to other channel estimation schemes.

References

[1] T. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3590–3600, Nov. 2010.

[2] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, “What will 5G be?,” *IEEE J. Sel. Areas in Commun.*, vol. 32, pp. 1065–1082, June 2014.

[3] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Commun. Mag.*, vol. 52, pp. 74–80, Feb. 2014.

[4] T. E. Bogale and L. B. Le, “Massive MIMO and millimeter wave for 5G wireless HetNet: potentials and challenges,” vol. arXiv preprint arXiv:1510.06359, 2015.

[5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: Benefits and challenges,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, pp. 742–758, Oct 2014.

[6] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and ChenHe, “An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel,” *IEEE Trans. Signal Process.*, vol. 59, pp. 4331–4340, Sept. 2011.

[7] T. Bogale and L. Vandendorpe, “Weighted sum rate optimization for downlink multiuser mimo coordinated base station systems: Centralized and distributed algorithms,” *IEEE Trans. Signal Process.*, vol. 60, pp. 1876–1889, April 2012.

[8] P. Chan, E. Lo, R. Wang, E. Au, V. Lau, R. Cheng, W. H. Mow, R. Murch, and K. Letaief, “The evolution path of 4G networks: FDD or TDD?,” *IEEE Commun. Mag.*, vol. 44, pp. 42–50, Dec. 2006.

[9] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?,” *IEEE J. Sel. Areas Commun.*, vol. 31, pp. 160–171, Feb. 2013.

[10] T. L. Marzetta, G. Caire, M. Debbah, I. Chih-Lin, and S. K. Mohammed, “Special issue on massive MIMO,” *IEEE J. Commun., Networks*, vol. 15, pp. 333–337, Aug. 2013.

[11] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, “Joint spatial division and multiplexing - the large-scale array regime,” *IEEE Trans. Inf. Theory*, vol. 59, pp. 6441–6463, Oct 2013.

[12] Z. Jiang, A. Molisch, G. Caire, and Z. Niu, “On the achievable rates of FDD massive MIMO systems with spatial channel correlation,” in *in Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 276–280, 2014.

[13] B. Lee, J. Choi, J.-Y. Seol, D. Love, and B. Shim, “Antenna grouping based feedback compression for FDD-based massive MIMO systems,” *IEEE Trans. Commun.*, vol. 63, pp. 3261–3274, Sept 2015.

[14] X. Rao and V. Lau, “Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems,” *IEEE Trans. Signal Process.*, vol. 62, pp. 3261–3271, June 2014.

[15] A. Garcia-Rodriguez and C. Masouros, “Exploiting the tolerance of massive MIMO to incomplete CSI for low-complexity transmission,” *in Proc. IEEE Global Commun. Conf. (GLOBECOM)*, pp. 1–6, 2015.

[16] N. Sun and J. Wu, “Maximizing spectral efficiency for high mobility systems with imperfect channel state information,” *IEEE Trans. Wireless Commun.*, vol. 13, pp. 1462–1470, Mar. 2014.

[17] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, pp. 186–195, Feb. 2014.

[18] J. Choi, D. Love, and P. Bidigare, “Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, pp. 802–814, Oct. 2014.

[19] J. Choi and D. Love, “Bounds on eigenvalues of a spatial correlation matrix,” *IEEE Commun. Lett.*, vol. 18, pp. 1391–1394, Aug. 2014.

[20] B. Clerckx, G. Kim, and S. Kim, “MU-MIMO with channel statistics-based codebooks in spatially correlated channels,” *in Proc. IEEE Global Commun. Conf. (GLOBECOM)*, pp. 1–5, 2008.

[21] J. Choi, V. Raghavan, and D. Love, “Limited feedback design for the spatially correlated multi-antenna broadcast channel,” *in Proc. IEEE Global Commun. Conf. (GLOBECOM)*, pp. 3481–3486, 2013.

[22] B. Maharaj, J. Wallace, L. Linde, and M. Jensen, “Frequency scaling of spatial correlation from co-located 2.4 and 5.2 GHz wideband indoor MIMO channel measurements,” *IEEE Electronics Lett.*, vol. 41, pp. 336–337, March 2005.

[23] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory*. Prentice-Hall, 1993.

[24] J. Wu and N. Sun, “Optimum sensor density in distortion-tolerant wireless sensor networks,” *IEEE Trans. Wireless Commun.*, vol. 11, pp. 2056–2064, June 2012.

[25] H. Gazzah, P. Regalia, and J.-P. Delmas, “Asymptotic eigenvalue distribution of block toeplitz matrices and application to blind SIMO channel identification,” *IEEE Trans. Inf. Theory*, vol. 47, pp. 1243–1251, Mar 2001.

[26] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, 6nd edition ed., 2000.

[27] I. S. Ansari, F. Yilmaz, M. S. Alouini, and O. Kucur, "On the sum of gamma random variates with application to the performance of maximal ratio combining over Nakagami-m fading channels," *in Proc. IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, pp. 394–398, 2012.