PILOT PRECODING AND COMBINING IN MULTIUSER MIMO NETWORKS

Nima N. Moghadam[⋆] *Hossein Shokri-Ghadikolaei*[⋆] *Gabor Fodor*[⋆,†] *Mats Bengtsson*[⋆] *Carlo Fischione*[⋆]

[⋆]Electrical Engineering School, KTH Royal Institute of Technology, Stockholm, Sweden
[†]Ericsson Research, Stockholm, Sweden

ABSTRACT

Although the benefits of precoding and combining of data streams are widely recognized, the potential of precoding the pilot signals at the user equipment (UE) side and combining them at the base station (BS) side has not received adequate attention. This paper considers a multiuser multiple input multiple output (MU-MIMO) cellular system in which the BS acquires channel state information (CSI) by means of uplink pilot signals and proposes pilot precoding and combining to improve the CSI quality. We first evaluate the channel estimation performance of a baseline scenario in which CSI is acquired with no pilot precoding. Next, we characterize the channel estimation error when the pilot signals are precoded by spatial filters that asymptotically maximize the channel estimation quality. Finally, we study the case when, in addition to pilot precoding at the UE side, the BS utilizes the second order statistics of the channels to further improve the channel estimation performance. The analytical and numerical results show that, specially in scenarios with large number of antennas at the BS and UEs, pilot precoding and combining has a great potential to improve the channel estimation quality in MU-MIMO systems.

***Index Terms*—** multiuser MIMO, channel estimation, minimum mean squared error, transceiver design.

1. INTRODUCTION

The existing multiple-input multiple-output (MIMO) systems employ an order of magnitude greater number of antenna ports at wireless access points than in the early releases of wireless standards. The 3rd Generation Partnership Project (3GPP), for example, is currently studying the details of technology enablers and performance benefits of deploying large scale antenna systems supporting up to 64 antenna ports at cellular base stations (BSs) [1]. Higher frequency bands, such as millimeter-wave (mmWave), are even more appealing for these large scale antenna systems, since the physical array size can be greatly reduced due to the decrease in wavelength [2].

In addition to the BSs, user equipments (UEs) compliant with the existing and emerging wireless standards are also employing a growing number of receive and transmit antennas. Today, UEs of the 3GPP Long Term Evolution systems, for example, can employ up to 4 antennas for transmit/receive diversity and for spatial multiplexing [3]. Clients of the IEEE 802.11ac can employ 8 antenna elements [4]. 5G systems, especially at the mmWave bands, will include high-end UEs supporting higher number of transmit/receive antennas [5]. Nevertheless, a majority of the studies in massive MIMO systems typically assume that BSs equipped with a large number of antennas serve a lower number of single-antenna UEs. These studies consider spatial multiplexing and beamforming of the user data streams to boost the achieved spectral efficiency and the per-stream signal-to-noise-and-interference ratio (SINR) and thereby the overall system capacity [6,7]. While transmit beamforming for the downlink transmission is key to achieve high capacity, considering only a single antenna UE prevents any beamforming in the uplink direction - either for data or pilot transmission.

Acquiring accurate channel state information at the transmitter for precoding in the downlink and channel state information at the receiver for demodulation in the uplink are among the main bottlenecks of massive MIMO systems. Depending on the deployment scenario and duplexing scheme, channel state information acquisition faces three main challenges: 1) scaling the number of downlink pilots (in frequency division duplexing (FDD) systems), 2) ensuring sufficiently high signal-to-noise ratio (SNR) for uplink pilot signals, and 3) mitigating the negative effects of pilot contamination [8–10]. References [11–14] highlight the importance of the channel estimation quality for the system performance without analyzing the potential of precoding the pilot signals. The works reported in [10, 15] show that the channel estimation quality is improved if the UEs can be separated in the angular domain, i.e., if their angle of arrivals (AoAs) do not overlap. However, the impact of multiple UE antennas on the spatial separability is not studied in these works.

In this paper, we consider a MU-MIMO network and focus on the problem of ensuring sufficiently high SNR for the *uplink pilot signals*. We suggest employing multiple transmit antennas at the UEs and use *pilot precoding and combining at the UEs and BS, respectively*. As we shall see, pilot precoding and combining not only improves the channel estimation quality, but it also has the potential to mitigate the effects of pilot contamination. Specifically, our analytical and numerical results suggest that precoded pilots reduce the variance of the channel estimation error by a factor that is proportional to the number of UE antennas and substantially outperforms a system with no pilot precoding. We show that pilot precoding and combining, besides improving the channel estimation quality, facilitates the implementation of massive MIMO in FDD systems and also substantially alleviates pilot contaminations. We believe that both our analytical and numerical results are important contributions to future wireless networks.

Notations: Capital bold letters denote matrices and lower bold letters denote vectors. The superscripts $\mathbf{X}^*$, $\mathbf{X}^{\mathrm{T}}$, $\mathbf{X}^{\mathrm{H}}$ stand for the conjugate, transpose, transpose conjugate of $\mathbf{X}$, respectively. $\mathbf{X} \circ \mathbf{Y}$, $\mathbf{X} \otimes \mathbf{Y}$, and $\mathbf{X} \odot \mathbf{Y}$ denote the Hadamard product, Kronecker product, and Khatri-Rao product of matrices $\mathbf{X}$ and $\mathbf{Y}$, respectively. $\mathbf{I}$ is the identity matrix with the appropriate size, and $\mathrm{vec}(\mathbf{X})$ and $\kappa(\mathbf{X})$ represent the vectorization and the condition number of matrix $\mathbf{X}$, respectively and $[x]^+ = \max(x, 0)$.

2. SYSTEM MODEL

Consider the uplink of a single cell network, where one BS with M antennas serves K multi-antenna UEs each equipped with N antennas.

978-1-5090-4117-6/17/$31.00 ©2017 IEEE

2.1. Channel Model

We assume a narrow-band block fading channel between each UE and the BS where the channel is relatively constant for T_c seconds and changes to a statistically independent value in the next block. Within one fading block, we assume a cluster channel model [16,17] with L paths between the BS and every UE. Let g_k^i be the complex gain of i-th path between the BS and UE k, which includes both path-loss and small scale fading. In particular, g_k^i's are independent and identically distributed and drawn from distribution $\mathcal{CN}(0,\sigma_k^2)$ where $1/\sigma_k^2$ is the path loss between the BS and UE k [11]. The uplink channel matrix between the BS and UE k is given by

$$\mathbf{H}_k=\sqrt{\delta M}\sum_{i=1}^{L} g_k^i\,\mathbf{b}\left(\theta_k^i\right)\mathbf{u}^{\mathrm{H}}\left(\phi_k^i\right)=\mathbf{B}_k\mathbf{G}_k\mathbf{U}_k^{\mathrm{H}}\in\mathbb{C}^{M\times N}, \quad (1)$$

where θ_k^i and ϕ_k^i are the AoA and angle of departure (AoD) of the i-th path and $\delta = N/L$. $\mathbf{b}\in\mathbb{C}^M$ and $\mathbf{u}\in\mathbb{C}^N$ are unit array response vectors of the BS's and UEs' antenna arrays, respectively, $\mathbf{B}_k=\left[\mathbf{b}(\theta_k^1),\ldots,\mathbf{b}(\theta_k^L)\right]$, $\mathbf{U}_k=\left[\mathbf{u}(\phi_k^1),\ldots,\mathbf{u}(\phi_k^L)\right]$, and $\mathbf{G}_k=\sqrt{\delta M}\,\mathrm{diag}\left(\mathbf{g}_k\right)$, where $\mathbf{g}_k=\left[g_k^1,\ldots,g_k^L\right]$. For the asymptotic performance analysis, we assume an antenna configuration at the BS that satisfies

$$\lim_{M\to\infty}\mathbf{b}(\theta)^{\mathrm{H}}\mathbf{b}(\phi)=\begin{cases}1 & \theta=\phi,\\ 0 & \text{otherwise.}\end{cases} \quad (2)$$

We assume that a similar condition holds for the antenna configuration at the UEs. These conditions automatically hold for uniform linear array antennas as well as randomly positioned antenna elements in the arrays, for example.

The vectorized channel between UE k and the BS can be presented as [18]

$$\mathrm{vec}\left(\mathbf{H}_k\right)=\left(\mathbf{U}_k^*\odot\mathbf{B}_k\right)\mathbf{g}_k\,. \quad (3)$$

We assume that the second order statistics of the channel (including $\{\theta_k^i\}$, $\{\phi_k^i\}$, and L) remain constant for many fading blocks.

Given $\{\theta_k^i\}$ and $\{\phi_k^i\}$, the channel is zero-mean circularly symmetric complex Gaussian and its covariance matrix is

$$\begin{aligned}\mathbf{R}_k&=\mathbb{E}\left[\mathrm{vec}\left(\mathbf{H}_k\right)\mathrm{vec}\left(\mathbf{H}_k\right)^{\mathrm{H}}\right],\\ &=\delta M\sigma_k^2\left(\mathbf{U}_k^*\odot\mathbf{B}_k\right)\left(\mathbf{U}_k^*\odot\mathbf{B}_k\right)^{\mathrm{H}}\,. \end{aligned}\quad (4)$$

The baseband signal received at the BS is

$$\mathbf{y}(t)=\sum_{k=1}^{K}\mathbf{H}_k\mathbf{x}_k(t)+\mathbf{z}(t)\,,$$

where $\mathbf{x}(t)\in\mathbb{C}^N$ and $\mathbf{z}(t)\sim\mathcal{CN}(\mathbf{0},\sigma_z^2\mathbf{I})$ represent the symbol vector transmitted by UE k and receiver noise at time t, respectively.

2.2. Pilot Transmission

To estimate the channel, UE k transmits training matrix $\mathbf{P}_k$ over $\tau\ll T_c$ channel uses with a total energy constraint

$$\mathrm{tr}\left(\mathbf{P}_k\mathbf{P}_k^{\mathrm{H}}\right)=\rho\,. \quad (5)$$

Note that the energy constraints in (5) stem from the regulatory constraints and the limitations on the battery consumption. Although the precoding is performed fully in the baseband, for the sake of mathematical convenience the per-antenna power constraint is not considered here. However, the general intuitions gained from this work are still valid taking the per-antenna power constraint into account.

UE k precodes its uplink pilot symbols using a spatial filter $\mathbf{V}_k$, and the BS then combines these symbols using $\mathbf{W}_k$. To design the spatial filters, in the next sections, we assume that the AoAs and AoDs are available at the BS and the UEs, respectively. In practice the AoAs can be estimated as it is shown in [19, 20]. Assuming that the channels are reciprocal, the UEs can also find their AoDs in the uplink by estimating their AoAs in the downlink. Collectively, the filtered received signal corresponding to UE j during the pilot transmission phase (τ channel uses) is

$$\mathbf{Y}_j=\mathbf{W}_j^{\mathrm{H}}\sum_{k=1}^{K}\mathbf{H}_k\mathbf{V}_k\mathbf{P}_k+\mathbf{W}_j^{\mathrm{H}}\mathbf{Z}\,, \quad (6)$$

where $\mathbf{Z}=\left[\mathbf{z}(1),\ldots,\mathbf{z}(\tau)\right]\in\mathbb{C}^{M\times\tau}$ is the receiver noise during pilot transmission. The dimensions of $\mathbf{Y}_j$, $\mathbf{W}_j$, $\mathbf{V}_j$, and $\mathbf{P}_j$ depend on whether or not pilot precoding and combing are used, as specified in the following section.

3. CHANNEL ESTIMATION

In this section, we investigate the channel estimation quality for three different pilot transmission scenarios: *non-precoded uncombined* pilots ($\mathtt{s}_1$), *precoded uncombined* pilots ($\mathtt{s}_2$), and *precoded combined* pilots ($\mathtt{s}_3$). In all the scenarios, the channels are estimated using minimum mean squared error (MMSE) estimators assuming the spatial filters are known by the BS. Similar to [21], from (6) the estimate of the channel $\mathbf{H}_j$ in scenario $\mathtt{x}\in\{\mathtt{s}_1,\mathtt{s}_2,\mathtt{s}_3\}$ can be computed as

$$\mathrm{vec}\left(\widehat{\mathbf{H}}_j^{(\mathtt{x})}\right)=\mathbf{R}_j\widetilde{\mathbf{P}}_j^{(\mathtt{x})}\left(\sum_{k=1}^{K}\left(\widetilde{\mathbf{P}}_{jk}^{(\mathtt{x})}\right)^{\mathrm{H}}\mathbf{R}_k\widetilde{\mathbf{P}}_{jk}^{(\mathtt{x})}+\sigma_z^2\mathbf{I}\right)^{-1}\mathrm{vec}\left(\mathbf{Y}_j^{(\mathtt{x})}\right), \quad (7)$$

where $\left(\widetilde{\mathbf{P}}_{jk}^{(\mathtt{x})}\right)^{\mathrm{H}}=\left(\mathbf{V}_k\mathbf{P}_k^{(\mathtt{x})}\right)^{\mathrm{T}}\otimes\mathbf{W}_j^{\mathrm{H}}$ and $\widetilde{\mathbf{P}}_j^{(\mathtt{x})}=\widetilde{\mathbf{P}}_{jj}^{(\mathtt{x})}$. Also, $\mathbf{Y}_j^{(\mathtt{x})}$ is the filtered received signal in scenario $\mathtt{x}$.

3.1. Non-precoded Uncombined Pilots ($\mathtt{s}_1$)

In this scenario, used as a benchmark, orthogonal training sequences are transmitted from UEs' antenna elements, and neither precoding at the UEs nor combining at the BS are used. Thus, the minimum number of training symbols needed to be transmitted from each antenna element (to guarantee orthogonality of the pilots of different antennas) is $\tau^{(\mathtt{s}_1)}=KN$. Considering (5), we have

$$\mathbf{P}_k^{(\mathtt{s}_1)}\left(\mathbf{P}_j^{(\mathtt{s}_1)}\right)^{\mathrm{H}}=\begin{cases}\frac{\rho}{N}\mathbf{I} & k=j\,,\\ \mathbf{0} & \text{otherwise}\,,\end{cases} \quad (8)$$

where $\mathbf{P}_k^{(\mathtt{s}_1)}\in\mathbb{C}^{N\times\tau^{(\mathtt{s}_1)}}$ is the training matrix of UE k in $\mathtt{s}_1$.

Substituting pilot symbols (8) and spatial filters of $\mathtt{s}_1$ (i.e., identity matrices) into (7), we have[1]

$$\mathrm{vec}\left(\widehat{\mathbf{H}}_j^{(\mathtt{s}_1)}\right)=\mathbf{R}_j\widetilde{\mathbf{P}}_j^{(\mathtt{s}_1)}\left(\left(\widetilde{\mathbf{P}}_j^{(\mathtt{s}_1)}\right)^{\mathrm{H}}\mathbf{R}_j\widetilde{\mathbf{P}}_j^{(\mathtt{s}_1)}+\sigma_z^2\mathbf{I}\right)^{-1}\mathrm{vec}\left(\mathbf{Y}^{(\mathtt{s}_1)}\right), \quad (9)$$

where $\left(\widetilde{\mathbf{P}}_j^{(\mathtt{s}_1)}\right)^{\mathrm{H}}=\left(\mathbf{P}_j^{(\mathtt{s}_1)}\right)^{\mathrm{T}}\otimes\mathbf{I}$. Note that $\widetilde{\mathbf{P}}_j^{(\mathtt{s}_1)}$'s inherit the orthogonality property of $\mathbf{P}_j^{(\mathtt{s}_1)}$'s and therefore the signals received from UE $k\neq j$ can be canceled out in the process of channel estimation for UE j.

Define the channel estimation error matrix as $\widetilde{\mathbf{H}}_k^{(\mathtt{s}_1)}=\mathbf{H}_k^{(\mathtt{s}_1)}-\widehat{\mathbf{H}}_k^{(\mathtt{s}_1)}$ and its covariance as $\widetilde{\mathbf{R}}_k^{(\mathtt{s}_1)}=\mathbb{E}\left[\mathrm{vec}\left(\widetilde{\mathbf{H}}_k^{(\mathtt{s}_1)}\right)\mathrm{vec}\left(\widetilde{\mathbf{H}}_k^{(\mathtt{s}_1)}\right)^{\mathrm{H}}\right]$. Then the following proposition characterizes the accuracy of channel estimation in this scenario:

Proposition 1. *Consider the system model of $\mathtt{s}_1$. The channel estimation error is bounded as*

$$\frac{1}{1+\zeta_k M}\left[1-\frac{\epsilon^{(\mathtt{s}_1)}-1}{\zeta_k M}\right]^{+}\leq\frac{\mathrm{tr}\left(\widetilde{\mathbf{R}}_k^{(\mathtt{s}_1)}\right)}{\mathrm{tr}\left(\mathbf{R}_k\right)}\leq\frac{1}{1+\zeta_k M}\,, \quad (10)$$

[1] Due to space limitation, proofs of the propositions and the corollaries are not presented here. Please find the proofs in the full version of the paper in [22].

where $\zeta_k = \frac{\rho\sigma_k^2}{L\sigma_z^2}$ *and*

$$\epsilon^{(\mathtt{s}_1)} = \frac{\left(1+\kappa\left(\mathbf{I}+\zeta_k M \mathbf{R}_{\mathbf{U}_k}^{\mathrm{T}} \circ \mathbf{R}_{\mathbf{B}_k}\right)\right)^2}{4\kappa\left(\mathbf{I}+\zeta_k M \mathbf{R}_{\mathbf{U}_k}^{\mathrm{T}} \circ \mathbf{R}_{\mathbf{B}_k}\right)} \geq 1\,, \qquad (11)$$

with $\mathbf{R}_{\mathbf{U}_k} = \mathbf{U}_k^H \mathbf{U}_k$, *and* $\mathbf{R}_{\mathbf{B}_k} = \mathbf{B}_k^H \mathbf{B}_k$.

Corollary 1. *As* $N \to \infty$ *(so* $\mathbf{R}_{\mathbf{U}_k} \to \mathbf{I}$*) or* $M \to \infty$ *(so* $\mathbf{R}_{\mathbf{B}_k} \to \mathbf{I}$*), both upper and lower bounds in* (10) *become tighter, and*

$$\mathrm{tr}\big(\widetilde{\mathbf{R}}_k^{(\mathtt{s}_1)}\big)/\mathrm{tr}\left(\mathbf{R}_k\right) \to (1+\zeta_k M)^{-1}\ . \qquad (12)$$

3.2. Precoded Uncombined Pilots ($\mathtt{s}_2$)

In the second pilot transmission scenario, the pilots are precoded using spatial filters at the UEs but there is no combiner at the BS. Assuming that $\mathbf{U}_k$ is available at UE k, the spatial filters can help focusing the training energy to the strongest paths between the UEs and the BS, thus boosting the SNR in the training phase. Noting that there are $L \leq N$ paths between each UE and the BS, $\tau^{(\mathtt{s}_2)} = KL$ training symbols suffices for transmitting orthogonal training sequences through all the paths. In other words, unlike scenario $\mathtt{s}_1$ where orthogonal pilots are assigned to the UE antennas, scenario$\mathtt{s}_2$ assigns orthogonal pilots to the paths. The orthogonality of training sequences and their energy constraints implies that

$$\mathbf{P}_k^{(\mathtt{s}_2)}\left(\mathbf{P}_j^{(\mathtt{s}_2)}\right)^{\mathrm{H}} = \begin{cases} \frac{\rho}{L}\mathbf{I} & k=j\,, \\ \mathbf{0} & \text{otherwise}\,, \end{cases} \qquad (13)$$

where $\mathbf{P}_k^{(\mathtt{s}_2)} \in \mathbb{C}^{L\times\tau^{(\mathtt{s}_2)}}$ is the training matrix transmitted by UE k in $\mathtt{s}_2$. Note although we are investigating the uplink channel estimation in this paper, pilot transmission scenario $\mathtt{s}_2$ entails the same complexity for downlink channel estimation. In contrast, the complexity of downlink channel estimation in scenario $\mathtt{s}_1$ is substantially higher than that of uplink channel estimation if $M \gg KN$. As a result, in massive MIMO systems, $\mathtt{s}_1$ *is only suitable for the time division duplexing (TDD) scheme (where the channel reciprocity principle holds), whereas* $\mathtt{s}_2$ *can be used in both TDD and FDD schemes.*

In this paper, we choose the precoding matrix $\mathbf{V}_k = \mathbf{U}_k$ for each UE k, which simplifies mathematical analysis and, at the same time, is asymptotically optimal in terms of maximizing the SNR [23]. No combining filter is considered at the BS in this scenario, so $\mathbf{W}_k = \mathbf{I}$. Substituting the training matrix and spatial filters of scenario $\mathtt{s}_2$ into (7), a new MMSE estimate for the channel is computed as

$$\mathrm{vec}\Big(\widehat{\mathbf{H}}_j^{(\mathtt{s}_2)}\Big) = \mathbf{R}_j \widetilde{\mathbf{P}}_j^{(\mathtt{s}_2)}\Big(\Big(\widetilde{\mathbf{P}}_j^{(\mathtt{s}_2)}\Big)^{\mathrm{H}}\mathbf{R}_j\widetilde{\mathbf{P}}_j^{(\mathtt{s}_2)} + \sigma_z^2\mathbf{I}\Big)^{-1}\mathrm{vec}\Big(\mathbf{Y}^{(\mathtt{s}_2)}\Big)\,, \qquad (14)$$

where $\big(\widetilde{\mathbf{P}}_k^{(\mathtt{s}_2)}\big)^{\mathrm{H}} = \big(\mathbf{U}_k\mathbf{P}_k^{(\mathtt{s}_2)}\big)^{\mathrm{T}} \otimes \mathbf{I}$.

Similar to $\mathtt{s}_1$, let us define the channel estimation error matrix and its covariance as $\widetilde{\mathbf{H}}_k^{(\mathtt{s}_2)} = \mathbf{H}_k - \widehat{\mathbf{H}}_k^{(\mathtt{s}_2)}$ and $\widetilde{\mathbf{R}}_k^{(\mathtt{s}_2)} = \mathbb{E}\big[\mathrm{vec}\big(\widetilde{\mathbf{H}}_k^{(\mathtt{s}_2)}\big)\mathrm{vec}\big(\widetilde{\mathbf{H}}_k^{(\mathtt{s}_2)}\big)^{\mathrm{H}}\big]$, respectively. The normalized estimation error in $\mathtt{s}_2$ can be bounded as:

Proposition 2. *Consider system model of* $\mathtt{s}_2$. *The channel estimation error is bounded as*

$$\frac{\lambda_{\max}^{-1}}{1+\delta\zeta_k M}\left[1-\frac{\epsilon^{(\mathtt{s}_2)}-1}{\delta\zeta_k M}\right]^{+} \leq \frac{\mathrm{tr}\big(\widetilde{\mathbf{R}}_k^{(\mathtt{s}_2)}\big)}{\mathrm{tr}\left(\mathbf{R}_k\right)} \leq \frac{\lambda_{\min}^{-1}}{1+\delta\zeta_k M}\,, \qquad (15)$$

where $\delta = N/L$,

$$\epsilon^{(\mathtt{s}_2)} = \frac{\left(1+\kappa\left(\mathbf{I}+\delta\zeta_k M(\mathbf{R}_{\mathbf{U}_k}^2)^{\mathrm{T}} \circ \mathbf{R}_{\mathbf{B}_k}\right)\right)^2}{4\kappa\left(\mathbf{I}+\delta\zeta_k M(\mathbf{R}_{\mathbf{U}_k}^2)^{\mathrm{T}} \circ \mathbf{R}_{\mathbf{B}_k}\right)} \geq 1\,, \qquad (16)$$

and $\lambda_{\min}$ *and* $\lambda_{\max}$ *represent the minimum and maximum eigenvalues of* $\mathbf{R}_{\mathbf{U}_k}$, *respectively.*

Corollary 2. *As* $N \to \infty$ *(so* $\mathbf{R}_{\mathbf{U}_k} \to \mathbf{I}$*), both upper and lower bounds in* (15) *become tighter, and*

$$\mathrm{tr}\big(\widetilde{\mathbf{R}}_k^{(\mathtt{s}_2)}\big)/\mathrm{tr}\left(\mathbf{R}_k\right) \to (1+\delta\zeta_k M)^{-1}\ . \qquad (17)$$

3.3. Precoded and Combined Pilots ($\mathtt{s}_3$)

In the third scenario, in addition to pilot precoding at the UEs, the received signals at the BS are also combined using the available information about the AoAs. Exploiting the spatial filters at the BS, given a large number of BS antennas, can lead to a sufficiently good spatial separation of the UEs. Therefore, a combiner at the BS may enable us to use non-orthogonal pilots for different UEs, if their orthogonality can be maintained in the spatial domain. In this scenario, non-orthogonal sequences with $\tau^{(\mathtt{s}_3)} < KL$ symbols are transmitted from each antenna element. Note that, similar to $\mathtt{s}_2$, $\mathtt{s}_3$ enables realization of massive MIMO using both TDD and FDD schemes.

We assume that the AoAs are known at the BS, and the combining matrix for UE k is $\mathbf{W}_k = \mathbf{B}_k$. To avoid further coordination between the BS and UEs, we assume that all the UEs use the same pilots, therefore there is a contamination of the pilots at the BS side. This is very similar to a multi-cell network where pilot reuse in neighboring cells causes the pilot contamination problem. Notice that scenario $\mathtt{s}_3$ addresses the pilot contamination problem, though we have a single cell network setting. In the following, we are interested to analyze how precoding and combing of the pilot signals alleviate the pilot contamination problem.

Substituting the filters $\mathbf{V}_k$ and $\mathbf{W}_k$ into (6), the received signal in $\mathtt{s}_3$ is

$$\begin{aligned}\mathrm{vec}\Big(\mathbf{Y}_j^{(\mathtt{s}_3)}\Big) = \Big(\widetilde{\mathbf{P}}_j^{(\mathtt{s}_3)}\Big)^{\mathrm{H}}\mathrm{vec}\left(\mathbf{H}_j\right) + \Big(\mathbf{I}\otimes\mathbf{B}_j^{\mathrm{H}}\Big)\mathrm{vec}\left(\mathbf{Z}\right) \\ + \underbrace{\sum_{k=1,k\neq j}^{K}\Big(\widetilde{\mathbf{P}}_{jk}^{(\mathtt{s}_3)}\Big)^{\mathrm{H}}\mathrm{vec}\left(\mathbf{H}_k\right)}_{\text{pilot contamination}}\,, \qquad (18)\end{aligned}$$

where $\big(\widetilde{\mathbf{P}}_{jk}^{(\mathtt{s}_3)}\big)^{\mathrm{H}} = \big(\mathbf{U}_j\mathbf{P}^{(\mathtt{s}_3)}\big)^{\mathrm{T}} \otimes \mathbf{B}_k^{\mathrm{H}}$ with $\mathbf{P}^{(\mathtt{s}_3)} = \mathbf{P}_1^{(\mathtt{s}_3)} = \ldots = \mathbf{P}_K^{(\mathtt{s}_3)} \in \mathbb{C}^{L\times\tau^{(\mathtt{s}_3)}}$ being the training matrix transmitted by UE j in $\mathtt{s}_3$.

We define the covariance matrix of the pilot contamination term and the covariance matrix of received signal without pilot contamination respectively by

$$\begin{aligned}\mathbf{Q}_j^{(\mathrm{PC})} &= \sum_{k=1,k\neq j}^{K}\Big(\widetilde{\mathbf{P}}_{jk}^{(\mathtt{s}_3)}\Big)^{\mathrm{H}}\mathbf{R}_k\widetilde{\mathbf{P}}_{jk}^{(\mathtt{s}_3)}\,, \\ \mathbf{Q}_j &= \Big(\widetilde{\mathbf{P}}_j^{(\mathtt{s}_3)}\Big)^{\mathrm{H}}\mathbf{R}_j\widetilde{\mathbf{P}}_j^{(\mathtt{s}_3)} + \sigma_z^2\left(\mathbf{I}\otimes\mathbf{R}_{\mathbf{B}_j}\right).\end{aligned}$$

Now, it is straightforward to show that the MMSE estimate of the channel in scenario $\mathtt{s}_3$ is

$$\mathrm{vec}\left(\widehat{\mathbf{H}}_j^{(\mathtt{s}_3)}\right) = \mathbf{R}_j\widetilde{\mathbf{P}}_j^{(\mathtt{s}_3)}\left(\mathbf{Q}_j + \mathbf{Q}_j^{(\mathrm{PC})}\right)^{-1}\mathrm{vec}\left(\mathbf{Y}_j^{(\mathtt{s}_3)}\right)\ . \qquad (19)$$

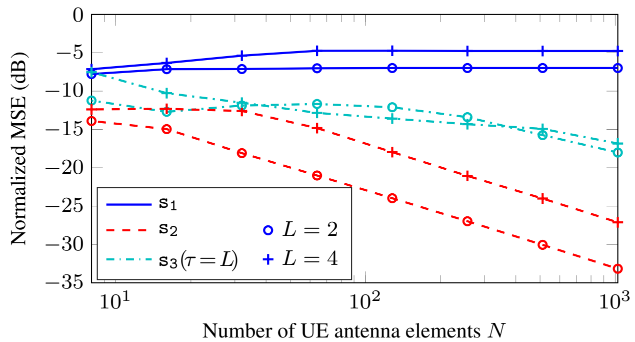
Corollary 3. *As* $M \to \infty$, $\mathbf{Q}_j^{(PC)} \to \mathbf{0}$, *namely the pilot contamination from other UEs goes to zero.*

Corollary 3 implies that the combiner substantially reduces the pilot contamination term, and asymptotically makes it zero.

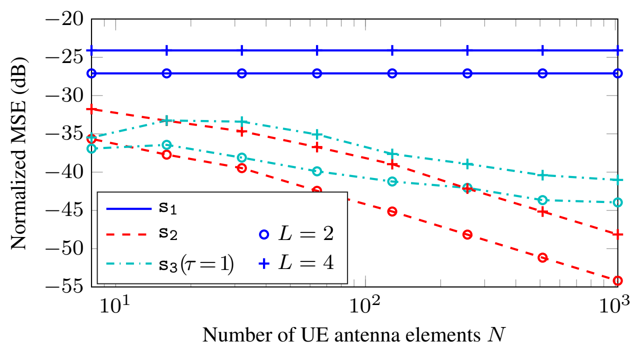
3.4. Design Insights

The analysis of the three channel estimation approaches suggests the following conclusions, which will be further numerically validated in Section 4:

- In $\mathtt{s}_2$, each UE focuses its pilot transmission energy to its strongest paths toward the BS, and thereby improves the channel estimation quality compared to $\mathtt{s}_1$. Alternatively, to achieve the same performance as $\mathtt{s}_1$, we substantially reduce the transmitted energy in and compensate for its effects by pilot precoding.



(a) Small antenna array at the BS, $M = 8$.



(b) Large antenna array at the BS, $M = 1024$

Fig. 1: Impact of number of antennas at the BS and UE side on the channel estimation performance. Three pilot transmission scenarios are compared: non-precoded uncombined pilots ($\mathtt{s}_1$), precoded uncombined pilots ($\mathtt{s}_2$), precoded combined pilots ($\mathtt{s}_3$).

- Although sending downlink pilot signals seems infeasible in $\mathtt{s}_1$ when $KN \ll M$, uplink and downlink pilot transmissions have identical complexity in $\mathtt{s}_2$ and in $\mathtt{s}_3$. Therefore, pilot precoding and combining facilitates realization of massive MIMO in both TDD and FDD settings.
- In $\mathtt{s}_3$, pilot combining increases spatial separability of different UEs such that the same pilots can be reused for all UEs (of all cells) once the number of receiver antennas grows large. Pilot precoding also reduces the potential interference among pilots of different antennas of a UE such that the same pilot can be reused for all antennas of one UE once the number of transmit antennas grows large.
- In the asymptotic regime of large number of transmit and receive antennas, only one pilot with a minimal energy may suffice to estimate the channel of all antennas of all UEs in the entire network.

4. NUMERICAL RESULTS

In this section, we illustrate numerically the impact of pilot precoding and combining using the second order statistics of the channel on the channel estimation performance. We consider a simple scenario where $K = 2$ UEs are located at equal distances from the BS with normalized path losses. Uniform linear arrays with half a wavelength spacing between the antenna elements are exploited at the UEs and BS. The AoDs for two users are equally spaced between $[-\pi/6, \pi/6]$. The AoAs for the signals received from UE 1 and 2 are equally spaced between $[-\pi/6, \pi/6]$ and $[0, \pi/3]$, respectively. We also assume that $\sigma_z^2 = 1$. Note that, the bounds found in Proposition 1 and 2 are rather tight and therefore we did not plot them here for the sake of clarity of the figures.

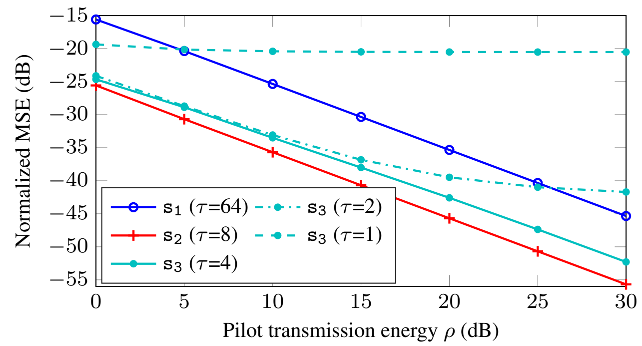


Fig. 2: The impact of ρ and τ on the channel estimation performance in the three pilot transmission scenarios.

Fig. 1 illustrates the impact of number of antennas at both UE and BS side on the normalized MSE for the three scenarios when $L = 2, 4$. We observe that by increasing N, while the estimation error in $\mathtt{s}_1$ does not change significantly, it decreases dramatically in $\mathtt{s}_2$ due to pilot precoding. Another observation from this figure is that despite the non-orthogonal pilot transmission in $\mathtt{s}_3$, still $\mathtt{s}_3$ outperforms $\mathtt{s}_1$ even at low N, where pilot precoding gain is small, owing to spatial filtering at the BS.

Comparing Fig. 1(a) and Fig. 1(b), we can conclude that as M grows large due to higher spatial selectivity at the BS, the performance of $\mathtt{s}_3$ improves relative to the other scenarios. Large N is also beneficial to $\mathtt{s}_3$. On one hand, increasing N leads to energy gain by focusing the pilot energy into the paths between the BS and the UEs. On the other hand, the interference between the training sequences transmitted from different antennas of a UE, can be mitigated by increasing N. Fig. 1(a) shows the case where the training sequences transmitted from different antenna elements of the same UE are orthogonal while in Fig. 1(b) all the UEs' antenna elements are transmitting the same pilot symbol.

In Fig. 2 the impact of pilot energy ρ and number of pilot symbols τ on the normalized MSE is investigated where $L = 4$, $N = 32$ and $M = 128$. The first observation from this figure is that, increasing the transmission power improves the channel estimation quality in all the scenarios. Another observation is that applying the pilot precoding at the UEs leads to 10 dB performance improvement which is close to the predicted precoding gain in the asympotic case, i.e. δ. This figure also shows that $\mathtt{s}_3$ achieves a performance close to $\mathtt{s}_2$ by transmitting half as many pilot symbols as in scenario $\mathtt{s}_2$.

5. CONCLUSIONS

This paper analyzed the benefits of pilot precoding and combining in MU-MIMO networks with multiple antenna UEs. We showed that, compared to the baseline non-precoded uncombined pilot transmissions, precoding the pilot signals at the transmitter and combining them at the receiver substantially improve the channel estimation quality, enable realization of massive MIMO in both TDD and FDD settings, have the potential of mitigating the effects of pilot contamination, and reduce the number of orthogonal pilots we need for the channel estimation phase.

6. REFERENCES

[1] 3GPP, "Study on elevation beamforming / full dimension (FD) multiple input multiple output (MIMO) for LTE, (release 13)," Technical Report TR 36.897, Technical Specification Group Radio Access Networks, Jun. 2015.

[2] S. Han, C.-l. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

[3] 3GPP, "User equipment (ue) radio transmission and reception (release 14)," Jun. 2016.

[4] O. Bejarano, E. Knightly, and M. Park, "IEEE 802.11 ac: from channelization to multi-user MIMO," *IEEE Commun. Mag.*, vol. 51, no. 10, pp. 84–90, Oct. 2013.

[5] A. Osseiran, J. Monserrat, and P. Marsch, Eds., *5G Mobile and Wireless Communications Technology*, Cambridge University Press, 2016, ISBN 978-1-107-13009-8.

[6] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO - opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, pp. 40–60, Jan. 2013.

[7] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[8] O. Elijah, C. Leow, T. Rahman, S. Nunoo, and S. Z-Iliya, "A comprehensive survey of pilot contamination in massive MIMO - 5G system," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 905–923, Nov. 2015.

[9] E. Björnson, E. G. Larsson, and T. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, pp. 114–123, Feb. 2016.

[10] B. Liu, S. Cheng, and X. Yuan, "Pilot contamination elimination precoding in multi-cell massive MIMO systems," in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2015 IEEE 26th Annual International Symposium on*, Aug. 2015, pp. 320–325.

[11] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[12] S. Rangan, T.S. Rappaport, and E. Erkip, "Millimeter wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[13] K. Guo, Y. Guo, G. Fodor, and G. Ascheid, "Uplink power control with MMSE receiver in multi-cell MU-massive-MIMO systems," in *Proc. IEEE International Conference on Communications (ICC)*, 2014, pp. 5184–5190.

[14] H. Shokri-Ghadikolaei, F. Boccardi, C. Fischione, G. Fodor, and M. Zorzi, "Spectrum sharing in mmWave cellular networks via cell association, coordination, and beamforming," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 11, pp. 2902–2917, Nov. 2016.

[15] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[16] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. Heath Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[17] F. Athley, A. Derneryd, J. Fridén, L. Manholm, and A. Stjernman, "MIMO performance of realistic UE antennas in LTE scenarios at 750 MHz," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 1337–1340, Nov. 2011.

[18] J. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. 25, no. 9, pp. 772–781, Sept. 1978.

[19] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[20] A. Kangas and T. Wigren, "Angle of arrival localization in LTE using MIMO pre-coder index feedback," *IEEE Commun. Lett*, vol. 17, no. 8, pp. 1584–1587, Aug. 2013.

[21] E. Björnson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with rician disturbance," *IEEE Trans. Signal Processing*, vol. 58, no. 3, pp. 1807–1820, Mar. 2010.

[22] N. N. Moghadam, H. Shokri-Ghadikolaei, G. Fodor, M. Bengtsson, and C. Fischione, "Pilot precoding and combining in multiuser MIMO networks," arXiv preprint arXiv:1611.01950v2.

[23] O. E. Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "The capacity optimality of beam steering in large millimeter wave MIMO systems," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPACWC)*, Jun. 2012, pp. 100–104.