Abstract

Pilot contamination is known to be one of the main bottlenecks in multi-cell multi-input multi-output (MIMO) networks. In this study, we investigate the effect of angular spread (AS) on the ergodic rate of a multi-cell network under pilot contamination, considering correlated MIMO channels and moderate antenna array sizes. We derive an exact expression for the achievable rate of the eigen-beamforming (EBF) precoding for an arbitrary antenna array size. Computer simulations for various scenarios are shown to match exactly with our analytical rate derivation under pilot contamination. They show that even though larger AS values make the pilot contamination effect for the EBF precoder stronger, this effect usually appears to be non-dominant on the achievable rates for moderate array sizes. We also show that the achievable rates for both the EBF and the regularized zero-forcing (RZF) precoders generally follow a non-monotonic behavior as the AS increases, which may be used to improve the aggregate network throughput with effective user scheduling algorithms.

Index Terms

Correlated multi-input multi-output (MIMO), pilot contamination, eigen-beamforming (EBF), regularized zero-forcing (RZF), uniform linear array (ULA).

I. INTRODUCTION

Massive multi-input multi-output (MIMO) is a recent technology that can significantly improve the spectral/energy efficiency of future wireless networks [1]–[3], and hence can help to address the exponentially growing traffic demand from mobile users. When the user density gets larger...
in a massive MIMO network, which is especially the case in urban areas, the scarcity of pilot resources necessitates the reuse of same pilot resources by the user equipments (UEs) in different cells. The resulting effect, known as the pilot contamination, is a fundamental problem in multi-cell time division duplex (TDD) networks, impairing the orthogonality of the downlink transmissions from different base stations (BSs), and hence diminishing the aggregate capacity.

The pilot contamination effect has been studied extensively in the literature recently, e.g., see the survey [4] and the references therein. In particular, the pioneering papers [5], [6] define and investigate the pilot contamination problem over an uncorrelated MIMO channel. In [7], analytical rate expressions are derived for correlated MIMO channels under pilot contamination. However, this analysis is done for the asymptotic regime considering very large antenna array sizes. In [8], the pilot contamination is considered over a correlated MIMO channel, and effects of a varying angular spread (AS) are investigated. Considering an asymptotic analysis, the adverse impacts of the pilot contamination are discussed to be completely eliminated with large antenna arrays. This is achieved when the uplink (UL) beams have non-overlapping angular support, which can only be satisfied with small AS values. In a follow-up work [9], the power-domain separation of the desired and the interfering user channels is considered (also studied in [10]), without any explicit investigation related to the effects of the AS.

Even though pilot contamination effect is one of the main bottlenecks for multi-cell correlated MIMO channels including millimeter-wave (mmWave) transmissions [11], [12], the current literature lacks a rigorous analysis for the effect of the AS on the user rates under pilot contamination, especially when the array size is not large (practically more feasible at lower frequencies). In this study, we therefore investigate the impact of the AS on the achievable data rates for a TDD based transmission over correlated MIMO channels under pilot contamination, with a special emphasis on the moderate antenna array size regime, as a rigorous extension of [13]. Our specific contributions can be summarized as follows:

i. The exact non-asymptotic analytical expression for the achievable rate is derived considering eigen-beamforming (EBF) precoding and correlated MIMO channels. As opposed to the earlier work in the literature [7], [8], this analysis is valid for any antenna array size, which is verified to match perfectly with the simulation results under various settings.

ii. We show analytically that a larger AS leads to stronger pilot contamination for the EBF precoding by impairing the interference channel orthogonality. However, the high pilot contamination effect usually does not always dominate the ergodic rate for moderate array
sizes. As a result, achievable rates do not saturate with respect to AS in the moderate array size regime, unless the interference from multiple cells is in a worst-case condition. In particular, when the AS gets larger, we observe that the channel power fluctuation around its long-term mean reduces (similar to the so-called channel hardening effect [14]–[16]), and this in turn improves achievable rates of the EBF.

iii. We show that the achievable rates of the EBF and the regularized zero-forcing (RZF) precoders exhibit a non-monotonic behavior with respect to the AS. Since the AS that results in a rate minima/maxima depends on the relative positions of the UEs and their serving BSs, user-cell pairing algorithms can be developed to maximize the aggregate network throughput.

The rest of the paper is organized as follows. Section II introduces the system model for a multi-cell, TDD-based correlated MIMO network under pilot contamination. The UL and downlink (DL) transmission schemes are provided in Section III-A and Section III-B, respectively, while the achievable DL rates for the EBF precoding are derived for a given AS and an arbitrary array size in Section III-C. The individual power terms constituting the achievable rate expression are further investigated for the EBF precoding in Section IV, in order to develop insights on the overall behavior of the ergodic rate as a function of the AS. Extensive numerical results are provided in Section V, and Section VI provides some concluding remarks.

Notations: Bold and uppercase letters represent matrices whereas bold and lowercase letters represent vectors. \( \mathbf{A}(m,n) \) denotes the \( m \)th row and \( n \)th column element of matrix \( \mathbf{A} \). \( \| \cdot \| \), \( | \cdot | \), \( (\cdot)^T \), \( (\cdot)^H \), \( (\cdot)^* \), \( \text{tr} (\cdot) \), \( \otimes \), \( \text{Var}\{\cdot\} \) and \( \mathbb{E}\{\cdot\} \) represent the Euclidean norm, absolute-value norm, transpose, Hermitian transpose, complex conjugation, trace of a matrix, Kronecker product, statistical variance and expectation operators, respectively. \( \mathcal{CN}(\mathbf{m},\mathbf{C}) \) denotes the complex-valued multivariate Gaussian distribution with the mean vector \( \mathbf{m} \) and the covariance matrix \( \mathbf{C} \), and \( \mathcal{U}[a,b] \) denotes the continuous Uniform distribution over the interval \([a,b]\). \( \mathbf{I}_M \) is the \( M \times M \) identity matrix, and \( \delta(a,b) \) is the Kronecker delta function taking 1 if \( a = b \), and 0 otherwise. \( \overset{p}{\underset{\rightarrow}{\longrightarrow}} \) denotes the convergence in probability.

II. SYSTEM MODEL

We consider a multi-cell scenario with \( N_L \) cells where each cell includes a single BS equipped with a uniform linear antenna array (ULA) of size \( M \). In each cell, a total of \( K \) UEs with single antennas are being served by their respective BSs under perfect time-synchronization.
Fig. 1: The multi-cell network consisting of hexagonal cells with the side length $r_2$. All the UEs are dropped at a distance of $r_1$ from their serving BSs. The $i$th and $j$th UEs are located at $\theta$ and $0^\circ$, respectively, with respect to the horizontal axis. $\Delta \theta_i$ denotes the angle between the directions from $i$th BS to the $i$th and $j$th UEs.

Since we are dealing with pilot contamination with varying AS, we assume that there is one UE in each cell that employs the same pilot sequence with other UEs in other cells during the UL channel estimation. By this way, all the users in this multi-cell layout are contributing to the pilot contamination, and the scenario where multiple UEs employ non-orthogonal pilot sequences in each cell remains as a straightforward extension. Note that the perfect time-synchronization assumption is arguably the worst condition in terms of the pilot contamination as any synchronization problem will make the pilot sequences more orthogonal, and hence reduce the pilot contamination effect [8].

In our analysis, we assume a TDD protocol consisting of subsequent UL training and DL transmission phases, where the interaction between two adjacent cells are sketched for the DL transmission in Fig. 1. In the UL training phase, all UEs transmit the same pilot sequences to all the BSs in the network. Based on the received pilot sequence, each BS first estimates the channel to its own UE, and then computes the precoding vector based on this channel estimate. During
the DL transmission phase, each BS transmits the desired data to its own UE after applying the precoding vector computed based on the UL training. Note that, the DL beamforming in each cell does not align well with its desired UE, which is shown in Fig. 1 as the different directions for the DL propagation and UE channel vectors, when the interference channels are not statistically orthogonal to the desired UE channel. As a result, the UL channel estimation and precoder computation gets impaired by this non-orthogonal interference, which is directly related to the AS and reduces the overall data rates, as discussed in the subsequent sections.

We consider that the UL channel between the $i$th UE and the $j$th BS is given as follows

$$h_{ij} = \frac{1}{\sqrt{N_P}} \sum_{p=1}^{N_P} \alpha_{ij,p} a(\phi_{ij,p}),$$

(1)

where $N_P$ is the number of unresolvable paths, $\alpha_{ij,p}$ is the complex path attenuation, $\phi_{ij,p}$ is the angle of arrival (AoA) of the $p$-th path, and $a(\phi_{ij,p})$ is the steering vector given as

$$a(\phi_{ij,p}) = \left[1, e^{-j2\pi \frac{D}{\lambda} \cos(\phi_{ij,p})}, \ldots, e^{-j2\pi \frac{D}{\lambda}(M-1) \cos(\phi_{ij,p})}\right]^T,$$

(2)

where $D$ is the element spacing in the ULA, and $\lambda$ is the wavelength.

The complex path attenuation $\alpha_{ij,p}$ and the AoA $\phi_{ij,p}$ are assumed to be uncorrelated over any of their indices, and with each other. In particular, $\alpha_{ij,p}$ is circularly symmetric complex Gaussian with $\alpha_{ij,p} \sim \mathcal{CN}(0, \beta_{ij})$, and the variance $\beta_{ij} = \frac{\zeta}{d_{ij}}$ captures the effect of the large-scale path loss, where $d_{ij}$ is the distance between $i$th UE and $j$th BS, $\gamma$ is the path loss exponent, and $\zeta$ is the normalization parameter to achieve a given signal-to-noise ratio (SNR). We consider uniform distribution for the AoA with $\phi_{ij,p} \sim \mathcal{U} \left[\bar{\phi}_{ij} - \Delta, \bar{\phi}_{ij} + \Delta\right]$, where $\bar{\phi}_{ij}$ is the line-of-sight (LoS) angle between $i$th UE and the $j$th BS, and $\Delta$ is the AS.

**III. UL/DL Transmission and Achievable DL Rates**

In this section, we first introduce the transmission schemes in the UL and the DL phases, and then consider the achievable rates in the DL data transmission after precoding.

**A. UL Training and Channel Estimation**

In the UL training phase, the UEs transmit the common pilot sequence of size $\tau$ denoted by $s = [s_1, s_2, \ldots, s_\tau]^T$, where each pilot symbol is chosen in an independent and identical (iid)
fashion from a discrete alphabet $A_{UL}$ consisting of unity norm entries. The $M \times \tau$ matrix of the received symbols at the $j$th BS is given as

$$Y_{UL}^j = \sum_{i=1}^{N_L} h_{ij} s^T + N_j ,$$

(3)

where $N_j$ is a $M \times \tau$ noise matrix consisting of circularly symmetric complex Gaussian entries with $\mathcal{CN}(0,\sigma^2)$. In an equivalent vector representation, (3) is given as

$$y_{UL}^j = S \sum_{i=1}^{N_L} h_{ij} + n_j ,$$

(4)

where $M\tau \times 1$ vectors $y_{UL}^j$ and $n_j$ are obtained by stacking all columns of $Y_{UL}^j$ and $N_j$, respectively, and $S = s \otimes I_M$ is the training matrix of size $M \tau \times \tau$ satisfying $S^H S = \tau I_M$. Following the convention of [5], the SNR is defined for this particular phase to be $1/\sigma^2$.

At the $j$th BS, the channel to the $i$th UE can be estimated using linear minimum mean square error (LMMSE) criterion as follows [8]

$$\hat{h}_{ij} = \hat{R}_{ij} S^H y_{UL}^j ,$$

(5)

where $\hat{R}_{ij}$ is the pilot-independent estimation filter given as

$$\hat{R}_{ij} = R_{ij} \left( \sigma^2 I_M + \tau \sum_{\ell=1}^{N_L} R_{\ell j} \right)^{-1} .$$

(6)

The covariance matrix $R_{ij} = \mathbb{E}\{h_{ij} h_{ij}^H\}$ in (6) is defined element-wise as follows

$$R_{ij}(m,n) = \beta_{ij} R_{ij}^\phi(m,n) = \beta_{ij} \int_0^{2\pi} \exp \left( -j 2\pi (m-n) \frac{D}{\lambda} \cos(\phi_{ij}) \right) p_\phi(\phi) d\phi ,$$

(7)

where $p_\phi(\phi)$ is the probability distribution function (pdf) of the AoA, and $R_{ij}^\phi = \mathbb{E}\{a(\phi_{ij}) a^H(\phi_{ij})\}$ is the angular covariance matrix of the steering vector. Employing (6) and (7), the resulting covariance matrix of the channel estimate in (5) is given as

$$\hat{R}_{jj} = \mathbb{E}\{\hat{h}_{jj} \hat{h}_{jj}^H\} = \tau \hat{R}_{jj} R_{jj} ,$$

(8)

where we present the detailed derivation steps for covariance matrices in Appendix A.
B. DL Data Transmission with Precoding

In the DL data transmission phase, each BS employs the channel estimate obtained in the UL training phase to compute the precoding vector for its own UE relying on the perfect reciprocity of the UL and the DL channels in the TDD protocol [5]. The received signal at the \( j \)th UE is therefore given as

\[
y_{j}^{\text{DL}} = \sqrt{\eta_{j}} h_{jj}^{\text{H}} w_{j} q_{j} + \sum_{i=1; i \neq j}^{N_{t}} \sqrt{\eta_{i}} h_{ij}^{\text{H}} w_{i} q_{i} + n_{j},
\]  

where \( w_{i} \) is the \( M \times 1 \) precoding vector of the \( i \)th BS for its own user, \( \eta_{i} = \left\{ \mathbb{E} \left\{ \text{tr} \left[ w_{i} w_{i}^{\text{H}} \right] \right\} \right\}^{-1} \) normalizes the average transmit power of the \( i \)th BS to achieve the same SNR in the UL training phase [7], \( q_{j} \) is the unit-energy data symbol transmitted from \( j \)th BS to its own UE and chosen from a discrete alphabet \( \mathcal{A}_{\text{DL}} \) in an iid fashion, and \( n_{j} \) is the circularly symmetric complex Gaussian noise with \( \mathcal{CN}(0, \sigma^{2}) \).

The beamforming strategy is assumed to be either the eigen-beamforming (EBF) or the regularized zero-forcing (RZF) [7], and is given at the \( j \)th BS as follows

\[
w_{j}^{\text{EBF}} = \hat{h}_{jj}, \quad \text{(EBF Precoder)} \tag{10}
\]

\[
w_{j}^{\text{RZF}} = \left( \hat{h}_{jj} \hat{h}_{jj}^{\text{H}} + \sigma^{2} I_{M} \right)^{-1} \hat{h}_{jj}. \quad \text{(RZF Precoder)} \tag{11}
\]

In the following, the impact of the AS is investigated through the achievable rates under both of these beamforming strategies, with a rigorous analytical rate derivation for the EBF precoder.

C. Achievable DL Rates with Precoding

We now consider achievable DL rates as a function of the AS over the underlying correlated channel with the EBF and the RZF precoding. In particular, we provide a non-asymptotic exact analytical expression for achievable rates with EBF precoding. In the TDD protocol under consideration, the BSs are assumed to have the CSI through the channel estimation performed in the UL training phase, and perform either the RZF or the EBF precoding prior to the DL transmission. On the other hand, the UEs are assumed to have no CSI at all, but know the long-term statistics of the effective channel. Under this assumption, the following rate at the \( j \)th
UE is achievable \[6\]

\[
R_j = \log_2 \left( 1 + \frac{\eta_j \mathbb{E}\{h_{jj}^H w_j\}^2}{1 + \eta_j \text{Var}\{h_{jj}^H w_j\} + \sum_{i=1; i \neq j}^{N_L} \eta_i \mathbb{E}\{|h_{ij}^H w_i|^2\}} \right),
\]  

(12)

where \(\eta_j \mathbb{E}\{h_{jj}^H w_j\}^2\) captures the desired signal power, \(\eta_j \text{Var}\{h_{jj}^H w_j\}\) is interpreted as the self-interference and arises from the lack of information on the instantaneous channel, and \(\sum_{i \neq j} \eta_i \mathbb{E}\{|h_{ji}^H w_i|^2\}\) is the intercell interference since it represents the interference from the other BS signals.

The rate approximation in (12) is arguably conservative, as discussed in [17], and can be interpreted as “self-interference limited” rate since the self-interference term dominates at high SNR regime for finite ULA sizes. However, since our focus in this study is to evaluate the impact of the AS on the correlated MIMO channels at fixed SNR, the rate approximation in (12) is used confidently. It is worth noting that, the achievable rates can also be evaluated by using the alternative expression suggested in [17, Eqn. (32)].

In the following theorem, considering that the EBF precoder in (10) is used in the DL transmission, we derive the analytical expressions of the first and the second order moments for the effective channel in order to be able to calculate the achievable rate in (12).

**Theorem 1:** Assuming that LMMSE channel estimation is used in the UL training, and that EBF precoding is used prior to DL transmission, the first order moment of the effective channel is given as

\[
\mathbb{E}\{h_{jj}^H w_j\} = \text{tr}\{\hat{R}_{jj}\}.
\]

Likewise, the second order moment is

\[
\mathbb{E}\{|h_{ij}^H w_i|^2\} = \tau^2 \sum_{m=1}^{M} \sum_{n=1}^{M} \sum_{m'=1}^{M} \sum_{n'=1}^{M} \tilde{R}_{ii}(m, n) \tilde{R}_{ii}^H(m', n') \times \\
\left[ E_{\phi}(m, n, m', n') + \sum_{k=1; k \neq j}^{N_L} R_{ji}(n', m) R_{ki}(n, m') \right] + \sigma^2 \text{tr}\{\tilde{S} \tilde{R}_{ii}^H \tilde{R}_{jj} \tilde{R}_{ii} \tilde{S}^H\},
\]

(14)
where $E_{\phi}(m, n, m', n')$ is defined as
\[
E_{\phi}(m, n, m', n') = \frac{\beta^2_{ji}}{N_p} \left[ 4 E_{ji}(n-m+n'-m') + (N_p-1) \left( E_{ji}(n-m)E_{ji}(n'-m') + E_{ji}(n'-m)E_{ji}(n-m') \right) \right],
\]
with $E_{ji}(m) = R_{ji}^\phi(n+m, n)$ for any $n \leq M-m$.

**Proof:** See Appendix B.

Note that once (13) and (14) are computed, the signal and the intercell interference terms in (12) are readily available by employing $\eta_j = \left[ \text{tr}\{\hat{R}_{jj}\} \right]^{-1}$, and the self-interference is given by
\[
\text{Var}\{h_j^H w_j\} = \mathbb{E}\left\{ |h_j^H w_j|^2 \right\} - \left( \mathbb{E}\{h_j^H w_j\} \right)^2.
\]

**IV. Individual Power Terms for EBF Precoder**

The achievable rate in (12) consists of the desired signal in its numerator, as well as the self-interference and the intercell interference power terms in its denominator. In this section, we investigate the behavior of these three power terms for the EBF precoding in more detail from statistical and geometrical perspectives. We note that any variation in the achievable rate appears due to a joint contribution from all these three terms, and taking into account any of them individually may be misleading in evaluating the overall rate results.

**A. Covariance Matrix Diagonalization**

We first consider the structure of the angular covariance matrices $R_{jj}^\phi$ and $R_{ij}^\phi$ in (7) with varying AS and antenna array size. Under the assumptions of Section II, the main diagonal entries of these covariance matrices are all 1, and the off-diagonal entries representing the angular correlation have non-zero norms smaller than 1. In Fig. 2, the Euclidean norms of the $(m, n)$th off-diagonal entries having the minimum and the maximum absolute separation of $|m-n|=1$ and $|m-n|=M-1$, respectively, are depicted along with the increasing AS for $M = \{10, 50\}$ with $\Delta \theta_j = 40^\circ$ in Fig. 1. We observe that each of these covariance matrices gets more diagonalized (the magnitude of the off-diagonal entries decreases) when the array size $M$ or the AS increases. The diagonalization rate increases with $M$ since both of the covariance matrices get diagonalized more quickly for larger $M$ values. Note that any BS in the multi-cell network will receive signals from a wide range of AoAs as the AS increases, which is similar to the uncorrelated rich-scattering environment where the possible AoAs span the whole $[0, 2\pi)$ angular domain.
Fig. 2: Norm of the entries of the angular covariance matrices $R_{jj}$ and $R_{ij}$ for $M = \{10, 50\}$.

The signal power variation with respect to the AS can be assessed through the diagonalization characteristic of the covariance matrices. Employing $\eta_j = \left[ \text{tr}\{\hat{R}_{jj}\} \right]^{-1}$ and the first order moment given in (13), the signal power can be expressed as $\eta_j |E\{h_{jj}^H w_j\}|^2 = \text{tr}\{\hat{R}_{jj}\}$, which can be further elaborated as follows

$$\text{tr}\{\tau \hat{R}_{jj} R_{jj}\} = \tau \sum_m \hat{R}_{jj}(m, m) R_{jj}(m, m) + \tau \sum_m \sum_n \hat{R}_{jj}(m, n) R_{jj}(n, m).$$

(17)

Since the estimation matrix $\hat{R}_{jj}$ in (6) becomes diagonal for larger AS values, similar to $R_{ij}$ and $R_{jj}$, and both the terms at the right hand side of (17) are real and positive, the second summation in (17) decreases when the AS increases. This leads the covariance matrix to become more diagonal. As a result, the signal power decreases with increasing AS through the diagonalization of the covariance matrices.

B. Fluctuation in Channel Power

The rate bound given in (12) is discussed to be achievable in [6] assuming that the UEs in the network do not know their instantaneous channels, but rather they only know the respective long-term means. This lack of information on the exact channel values is captured by the self-interference term $\eta_j \text{Var}\{h_{jj}^H w_j\}$ in (12). This term actually represents the power of the deviation between the instantaneous channel and the long-term mean, given equivalently as $\eta_j E\left\{|h_{jj}^H w_j - E\{h_{jj}^H w_j\}|^2\right\}$. Assuming EBF precoding in (10) with perfect CSI, this term
becomes equivalent to the variance of the channel power. We therefore note that as the fluctuation of the channel square-norm around the long-term mean decreases, which is similar to the phenomenon known as the channel hardening [14]–[16], the self-interference term should decrease accordingly. In the following, this fluctuation is shown to decrease monotonically for the EBF precoding when the AS increases, which then reduces the self-interference term.

Lemma 1: The channel considered in (1) hardens, such that $\|h_{ij}\|^2 / \mathbb{E}\{\|h_{ij}\|^2\} \xrightarrow{P} 1$ as $M \to \infty$, if we have $\mathcal{M}_{ij} \to 0$ as $M \to \infty$, where $\mathcal{M}_{ij}$ is the hardening measure given as

$$\mathcal{M}_{ij} = \frac{3}{N_P} + \frac{N_P - 1}{MN_P} \left(1 + \frac{2}{M} \sum_{m=1}^{M-1} |E_{ij}(m)|^2 \right), \quad (18)$$

with $E_{ij}(m) = R_{ij}^\phi(n+m,n)$ for any $n \leq M-m$.

Proof: See the derivation in Appendix C as an extension of [16].

We observe that the desired convergence $\mathcal{M}_{ij} \to 0$ is satisfied only if the number of paths $N_P$ is sufficiently large. Even when the number of paths $N_P$ or the antenna array size $M$ have moderate values in contrast to asymptotic approximations, the behavior of the self-interference power can still be assessed from (18). In Fig. 3, we depict $\mathcal{M}_{ij}$ along with AS for various $M$ under the assumption that $N_P = 50$ and $\phi_{ij} \sim \mathcal{U}[-\Delta, +\Delta]$ with $\Delta$ representing the AS. We observe that $\mathcal{M}_{ij}$ decreases monotonically for increasing AS for all cases, and gets even smaller values as the array size $M$ increases. Note that the smaller $\mathcal{M}_{ij}$ implies a better convergence of the channel square-norm to its long-term mean with high probability, and hence less fluctuation in the channel power around its long-term mean. Since the self-interference power is closely related to the fluctuation in the channel power, decaying behavior of $\mathcal{M}_{ij}$ with the increasing AS implies that the self-interference power reduces, as well.

C. Interference Channel Orthogonality

In the DL data transmission, the pilot contamination shows its adverse effect by impairing the orthogonality between the desired and the interfering user channels, which is basically captured by the intercell interference term $\sum_{i \neq j} \sqrt{\eta_i} h_{ji}^H w_i q_i$ in (9). From a geometrical perspective, the intercell interference involves the inner product between each interference channel $h_{ji}$ for $j \neq i$, and the precoder $w_i$, which is a function of the estimate of the desired user channel $h_{ii}$. One way to examine how the pilot contamination impairs the orthogonality, and hence amplify the intercell interference power, with varying AS is through a geometric interpreta-
Fig. 3: The channel hardening measure $M_{ij}$ for $M \in \{10, 50, 100, \infty\}$ and $N_P = 50$.

This can be done by analyzing the pdf of the random angle $\varphi_{ij}$ between $h_{ji}$ and $w_i$, where we leave the actual numerical evaluation to Section V. Note that, if the channels were perfectly known and spatially uncorrelated, the desired pdf would be given analytically as $f_{\varphi_{ij}}(\varphi) = 2 (N-1) \left( \sin \varphi \right)^{2N-3} \cos \varphi$ \[18\].

We assume the representative setting of Fig. 1 with $\theta = 200^\circ$, quadrature phase-shift keying (QPSK) symbols in the UL training phase with the sequence length $\tau = 1$, and the path-loss exponent $\zeta = 2$. In Fig. 4, we depict the pdf of the random angle $\varphi_{ij}$ for $M = \{10, 50\}$ and $\text{SNR} = \{0, 20\}$ dB. We observe that the desired and the interfering user channels are sufficiently orthogonal for small AS values since $\varphi_{ij}$ takes values close to $90^\circ$ with high probability, and the resulting orthogonality can even be stronger than the uncorrelated channel case. This geometrical interpretation agrees with \[8\] in the sense that the training beams of the UEs do not overlap in the UL transmission when the AS is sufficiently small, and the scenario becomes free from any pilot contamination effect. As a result, the random angle between the desired and the interfering user channels has the same pdf for the perfect CSI and the channel estimation scenarios when the beams are separated sufficiently, or equivalently the AS is small enough. When the AS starts
Fig. 4: The pdf of the angle $\varphi_{ij}$ between $h_{ji}$ and $w_i$ for $M = \{10, 50\}$ and SNR = \{0, 20\} dB.
to increase, the desired and the interfering user channels become correlated as the multipaths start to share spatially closer AoA domains. As a result, the desired orthogonality inherently gets impaired for larger AS values, even for the perfect CSI case. When the desired user channel is being estimated, this orthogonality gets impaired even more because of the pilot contamination effect. In addition, the orthogonality is impaired further when the SNR increases since larger SNR in each cell implies more interference power transferred to other cells. The effect of impairment in the channel orthogonality can be observed in Fig. 4 by the further movement of the pdf associated with the channel estimation away from $90^\circ$ as compared to the pdf for the perfect CSI. Finally, comparing Fig. 4(a) and Fig. 4(b), we observe for a given AS that the desired and the interfering users are getting more orthogonal as the antenna array size increases, which is one of the main goals of massive MIMO in the context of the intercell interference rejection [2].

V. Numerical Results

In this section, we present numerical results for the effect of varying AS on the achievable rates in a multi-cell network under pilot contamination and considering the EBF and RZF precoders. The theoretical derivations presented in Section III-C are employed for analytical evaluations, and the corresponding simulation data is generated through extensive Monte Carlo runs. Without any loss of generality, we assume QPSK modulated pilot symbols in the UL training of sequence length $\tau = 1$, the path-loss exponent $\zeta = 2$, SNR = 0 dB, and $N_P = 100$, together with the distances $r_1 = 40$ m and $r_2 = 50$ m shown in Fig. 1.

A. Two-Cell Scenario: Single UE Position

This section considers a two-cell scenario where the $i$th and $j$th cells in Fig. 1 are designated as the interfering and the desired cells, respectively, and the angular position of the $i$th UE is $\theta = 200^\circ$. In Fig. 5, the achievable data rates are presented for the EBF precoding for array sizes of $M = \{10, 20, 50, 100\}$. We observe that the analytical results follow the characteristic behavior of the simulation data in all cases of interest. We also observe that the EBF precoder does not exhibit a monotonic increase for $M = 10$, where the achievable rates make a minimum at $AS = 28^\circ$. As it will become clear in Section V-B, this undesired minimum data rate appears at a specific AS value. Indeed, the location of this minimum can be controlled through the deployment geometry, and, in particular, the angle of UE separation represented by $\Delta\theta$’s in Fig. 1. On the other hand, the real AS value of the propagation environment basically depends
on the carrier frequency the communication setting among the other features [11], [12]. Therefore this non-monotonic behavior of the achievable rates can be utilized to achieve higher aggregate throughput by discouraging the formation of user-cell pairs for which this unfavorable AS value of the minimum is close to the real AS value of the environment. Note that this non-monotonic behavior cannot be revealed through an asymptotic analysis because of the relatively small array size. Another important observation is that although user rates for large antenna array sizes exhibit a fast increase for small AS values, they tend to saturate eventually because of the pilot contamination, as discussed in Section IV-C. On the other hand, the small to moderate array sizes result in no such saturation, which points out the pilot contamination being a non-dominant effect over achievable rates in this array size regime.

In Fig. 6, the signal, the self-interference, and the intercell interference powers for the EBF precoder are depicted separately for the scenario of Fig. 5. We observe that the analytical results follow the simulation data successfully in all scenarios of interest. The signal power and the intercell interference are observed to exhibit relatively flat characteristics over a range of small

Fig. 5: Achievable rates for the EBF precoding in a 2-cell scenario with $M = \{10, 20, 50, 100\}$ and the interfering UE angular position at $\theta = 200^\circ$. 
Fig. 6: Signal, self-interference and intercell interference powers for the EBF precoding in a 2-cell scenario with $M = \{10, 20\}$ and the interfering UE angular position at $\theta = 200^\circ$.

AS values up to approximately $10^\circ$. In this region, the AS values are sufficiently small, and the UL training beams of the UEs are therefore well separated. This is the reason for zero intercell interference in this region, which implies no pilot contamination effect and agrees with the geometrical interpretation of Section IV-C. As the AS gets values larger than $10^\circ$, the intercell interference power starts to increase, which implies stronger pilot contamination effect, and saturates around $\text{AS} = 50^\circ$. Further, the DL transmission does not align with the desired signal direction any more, which appears as the decreasing trend in the signal power. Note that the self-interference power follows a decaying characteristic with the increasing AS as discussed in Section IV-B, and starts to decrease earlier and more rapidly for $M = 20$ as compared to $M = 10$, as shown in Fig. 6. This is actually the reason behind the formation of the non-monotonic rate behavior since the decrease in the signal power dominates initially for $M = 10$ which causes a minimum at $\text{AS} = 28^\circ$. On the other hand, since the self-interference power starts to decay quickly for a larger array size of $M = 20$, this effect dominates the relatively slower decrease in the signal power, and hence, we observe a monotonic increase in the data rates.
The achievable rates for the RZF precoder in the DL transmission is depicted in Fig. 7 for the scenario of Fig. 5 with the ULA array sizes of $M = \{10, 20, 50, 100\}$. We observe that the achievable rates of the RZF precoder is larger than that of the EBF precoder in Fig. 5 for the same array sizes at the expense of a larger computational complexity. We also observe a non-monotonic variation in achievable rates for all the array sizes of interest, where there is a minimum at $AS = 31^\circ$ for $M = \{10, 20\}$, and a maximum at $AS = \{15^\circ, 16^\circ\}$ for $M = \{50, 100\}$, respectively. Similar to the EBF precoder case, this non-monotonic behavior can be utilized to achieve a higher aggregate throughput by i) encouraging the formation of user-cell pairs for which the favorable AS value of the maximum is close to the real AS value of the environment, and similarly, ii) discouraging the user-cell pairs for which the unfavorable AS value of the minimum is close to the real AS value.

**B. Two-Cell Scenario: Multiple UE Positions**

In this section, we consider the effect of various angular positions of the $i$th interfering UE with $\theta = \{180^\circ, 200^\circ, 220^\circ\}$ in a 2-cell scenario. In Fig. 8, we present the achievable rates for the
Fig. 8: Achievable rates for the EBF and the RZF precoding in a 2-cell scenario with $M = \{10, 50\}$ and varying interfering UE position at $\theta = \{180^\circ, 200^\circ, 220^\circ\}$. 
EBF and the RZF precoding in a 2-cell scenario with $M = \{10, 50\}$ at a set of angular positions $\theta = \{180^\circ, 200^\circ, 220^\circ\}$ for the $i$th interfering UE. We observe that as the interfering UE gets closer to the desired UE, which is indicated by the increasing $\theta$ in Fig. 1, the achievable rate reduces for both the precoders. In addition, we observe either lower maxima or deeper minima located at smaller AS values, when $\theta$ increases. As discussed in Section IV, the intercell interference increases for larger $\theta$ due to the closer interfering UE, and this indicates more powerful pilot contamination. Furthermore, the signal power accordingly gets smaller when $\theta$ increases, since the DL transmission does not align well with the desired user any more, as shown in Fig. 9. Since the self-interference does not depend on $\theta$, increasing intercell interference and decreasing signal power results in the reduced user rates. This is also the reason behind the deeper minima for the EBF precoding. The intercell interference along with larger $\theta$ can be very strong at $M = 50$ such that the maxima of the RZF for $\theta = \{180^\circ, 200^\circ\}$ switch into the minimum for $\theta = 220^\circ$. In addition, the monotonically increasing rates of the EBF for $\theta = \{180^\circ, 200^\circ\}$ are destroyed with a formation of a minimum at $\text{AS} = 11^\circ$ when $\theta$ becomes $220^\circ$. 

Fig. 9: Signal, self-interference and intercell interference powers for the EBF precoding in a 2-cell scenario with $M = 10$ and varying interfering UE positions at $\theta = \{180^\circ, 200^\circ, 220^\circ\}$. 
C. Multi-Cell Scenario

Finally, we consider the impact of AS in a multi-cell setting with $N_L = \{2, 3, 5\}$ at $\theta = 200^\circ$. To this end, a multi-cell setting is generated by considering five cells as shown in Fig. 1, where the interfering UEs are located at $200^\circ, 160^\circ, 360^\circ, 60^\circ$ with respect to the horizontal axis for the $i$th, $k$th, $\ell$th, and $m$th cells, respectively. This layout provides almost the worst condition in terms of the interference power, and hence the pilot contamination. The effect of this multi-cell setting on the achievable rates of the EBF and the RZF precoders is presented in Fig. 10 for $M = \{10, 50\}$ and $\theta = 200^\circ$. We observe that as we consider more cells, the resulting interference makes achievable rates reduce together with much lower maxima or deeper minima. We even observe the formation of an additional maximum for the EBF at $\text{AS} = 46^\circ$ and minimum for the RZF at $\text{AS} = 5^\circ$ when $N_L = 5$. We also observe that adding more cells strengthens the intercell interference very rapidly and impairs the desired signal power proportionally, as shown in Fig. 11. This impairing effects eventually reduce and even saturate achievable rates, as shown in Fig. 10.

VI. CONCLUDING REMARKS

We investigated the impact of AS on the achievable rates in a multi-cell environment, considering correlated MIMO channels and pilot contamination effects. An exact analytical expression for achievable rate is derived for the EBF precoding with arbitrary antenna array sizes. Subsequently, non-monotonic behavior (with respect to the AS) of the achievable rates is revealed for both the EBF and the RZF precoders. The AS values for the possible minima and maxima depend on the relative positions of the UEs and their serving BSs, and hence, such a knowledge can be utilized to maximize the aggregate network throughput via effective user-cell pairing strategies.

When the AS gets larger, we showed through rigorous analyses that the covariance matrices tend to have a more diagonalized structure, the channel power fluctuation diminishes in a similar way as in the channel hardening, and the orthogonality of the interference channel gets impaired which gives rise to the pilot contamination effect. Our numerical results show that increase in the pilot contamination (due to increasing AS) does not introduce a dominant effect over the achievable rates. On the other hand, larger AS values reduce the channel power fluctuation (similar to the channel hardening effect), which in turn improves the achievable rates.

We believe that the insights developed in this manuscript on the behavior of the ergodic rate with respect to the AS will be especially important for millimeter wave (mmWave)-MIMO systems, due to the limited AS at mmWave spectrum. Note that mmWave systems can normally
Fig. 10: Achievable rates for the EBF and the RZF precoding in a multi-cell scenario with $N_L = \{2, 3, 5\}$, $M = \{10, 50\}$, and the interfering UE position at $\theta = 200^\circ$. 
Fig. 11: Signal, self-interference and intercell interference powers for the EBF precoding in a multi-cell scenario with $N_L = \{2, 3, 5\}$, $M = 10$, and the interfering UE position at $\theta = 200^\circ$.

accommodate large number of antennas due to reduced inter-element spacing. On the other hand, applications where there is limited space for placing antennas (e.g., Internet-of-things (IoT) and virtual/augmented reality devices with small form factors), operating in the 10-100 transmit antenna configuration may be essential, where our results in this paper will be highly applicable.

APPENDIX A

COVARIANCE MATRIX DERIVATION

The covariance matrix of the channel vector $h_{ij}$ in (1) is given as

$$
R_{ij} = \frac{1}{N_p} \sum_{p=1}^{N_p} \sum_{p'=1}^{N_p} E \{ a(\phi_{ij,p}) a^H(\phi_{ij,p'}) \} E \{ \alpha_{ij,p} \alpha^*_{ij,p'} \},
$$

(19)

$$
= \frac{\beta_{ij}}{N_p} \sum_{p=1}^{N_p} E \{ a(\phi_{ij,p}) a^H(\phi_{ij,p}) \},
$$

(20)

$$
= \beta_{ij} E \{ a(\phi_{ij}) a^H(\phi_{ij}) \},
$$

(21)
where (20) employs $\mathbb{E}\{\alpha_{ij,p} \alpha_{ij,p'}^*\} = \beta_{ij} \delta(p, p')$, and (21) follows from the fact that the distribution of AoA $\phi_{ij,p}$ is identical for any choice of the path index $p$. Defining the angular covariance matrix of the steering vector as $\mathbf{R}_\phi^{ij} = \mathbb{E}\{\mathbf{a}(\phi_{ij}) \mathbf{a}^H(\phi_{ij})\}$, and employing (2), the element-wise angular correlation is given as follows

$$
\mathbf{R}_\phi^{ij}(m, n) = \mathbb{E}\left\{\exp\left(-j2\pi(m - n)\frac{D}{\lambda} \cos(\phi_{ij})\right)\right\},
$$

$$
= \int_0^{2\pi} \exp\left(-j2\pi(m - n)\frac{D}{\lambda} \cos(\phi_{ij})\right) p_\phi(\phi) \, d\phi,
$$

(22)

where $p_\phi(\phi)$ is the probability distribution function (pdf) of the AoA distribution. In particular, assuming the uniform distribution for AoA with $\mathcal{U}[\phi_{ij} - \Delta, \phi_{ij} + \Delta]$, (22) can be given as

$$
\mathbf{R}_\phi^{ij}(m, n) = \frac{1}{2\Delta} \int_{\phi_{ij} - \Delta}^{\phi_{ij} + \Delta} \exp\left(-j2\pi(m - n)\frac{D}{\lambda} \cos(\phi_{ij})\right) d\phi.
$$

(23)

Employing (22), each entry of the covariance matrix in (21) is given by (7).

We now derive the covariance matrix of the channel estimate $\hat{\mathbf{h}}_{ij}$, denoted by $\hat{\mathbf{R}}_{ij} = \mathbb{E}\{\hat{\mathbf{h}}_{ij} \hat{\mathbf{h}}_{ij}^H\}$. Employing the definition of $\hat{\mathbf{h}}_{ij}$ in (5), and the UL signal model in (4), $\hat{\mathbf{R}}_{ij}$ is given as

$$
\hat{\mathbf{R}}_{ij} = \hat{\mathbf{R}}_{ij} \mathbf{S}^H \mathbb{E}\left\{\mathbf{y}_j^{UL} (\mathbf{y}_j^{UL})^H\right\} \mathbf{S} \hat{\mathbf{R}}_{ij}^H,
$$

$$
= \hat{\mathbf{R}}_{ij} \mathbf{S}^H \mathbb{E}\left\{\left(\mathbf{S} \sum_{k=1}^{N_L} \mathbf{h}_{kj} + \mathbf{n}_j\right) \left(\mathbf{S} \sum_{k=1}^{N_L} \mathbf{h}_{kj} + \mathbf{n}_j\right)^H\right\} \mathbf{S} \hat{\mathbf{R}}_{ij}^H,
$$

(24)

$$
= \tau^2 \sum_{k=1}^{N_L} \sum_{\ell=1}^{N_L} \hat{\mathbf{R}}_{ij} \mathbb{E}\{\mathbf{h}_{kj} \mathbf{h}_{j\ell}^H\} \hat{\mathbf{R}}_{ij}^H + \tau \sigma^2 \hat{\mathbf{R}}_{ij} \hat{\mathbf{R}}_{ij}^H
$$

$$
+ \tau \sum_{k=1}^{N_L} \hat{\mathbf{R}}_{ij} \mathbb{E}\{\mathbf{n}_j \mathbf{n}_j^H\} \mathbf{S} \hat{\mathbf{R}}_{ij}^H + \tau \sum_{k=1}^{N_L} \hat{\mathbf{R}}_{ij} \mathbf{S}^H \mathbb{E}\{\mathbf{n}_j \mathbf{n}_j^H\} \hat{\mathbf{R}}_{ij}^H,
$$

(25)

(26)

where we employ the relations $\mathbb{E}\{\mathbf{n}_j \mathbf{n}_j^H\} = \sigma^2 \mathbf{I}_M$ and $\mathbf{S}^H \mathbf{S} = \tau \mathbf{I}_M$. Since the noise and the channel vectors are uncorrelated and zero-mean, the expectations in (26) cancel, and we have

$$
\hat{\mathbf{R}}_{ij} = \tau^2 \sum_{k=1}^{N_L} \hat{\mathbf{R}}_{ij} \mathbb{E}\{\mathbf{h}_{kj} \mathbf{h}_{j\ell}^H\} \hat{\mathbf{R}}_{ij}^H + \tau^2 \sum_{k=1}^{N_L} \sum_{\ell=1}^{N_L} \hat{\mathbf{R}}_{ij} \mathbb{E}\{\mathbf{h}_{kj} \mathbf{h}_{j\ell}^H\} \hat{\mathbf{R}}_{ij}^H + \tau \sigma^2 \hat{\mathbf{R}}_{ij} \hat{\mathbf{R}}_{ij}^H,
$$

(27)

where the second term in (27) vanishes since $\mathbb{E}\{\mathbf{h}_{kj} \mathbf{h}_{j\ell}^H\} = 0$ for $k \neq \ell$. Then $\hat{\mathbf{R}}_{ij}$ becomes
\[ \hat{R}_{ij} = \sum_{k=1}^{N_t} \tau^2 \tilde{R}_{kj} R_{kj}^H + \tau \sigma^2 \hat{R}_{ij} R_{ij}^H, \]

\[ = \tau \tilde{R}_{kj} \left( \tau \sum_{k=1}^{N_t} R_{kj} + \sigma^2 I_M \right) R_{ij}^H = \tau R_{ij} R_{ij}^H, \tag{28} \]

where we employ the definition of \( \tilde{R}_{ij} \) in (6). Using the Hermitian symmetry, we obtain (8).

**APPENDIX B**

**FIRST AND SECOND ORDER MOMENT DERIVATION**

In this section, we derive the first and second order moments \( \mathbb{E} \{ h_j^H w_j \} \) and \( \mathbb{E} \{ |h_j^H w_i|^2 \} \), respectively, for the EBF precoding given in (10). Before the analysis, we define the following property which is used throughout this section while evaluating the mean of the quadratic and the double-quadratic forms involving random vectors.

**Lemma 2:** Assume that \( \{ u_i \}_{i=1}^4 \) be a set of zero-mean random vectors of arbitrary sizes where each of them may be individually correlated with the arbitrary covariance matrices \( \{ C_i \}_{i=1}^4 \). For the given coefficient matrices \( A \) and \( B \) of the appropriate sizes and with arbitrary entries, the quadratic form \( u_1^H A u_2 \) and the double-quadratic form \( u_1^H A u_3 u_4^H B u_4 \) are zero-mean if at least one of these random vectors are uncorrelated with the others.

**Proof:** Assuming that \( u_1 \) is uncorrelated with the others, without any loss of generality, regardless of whether \( \{ u_i \}_{i=2}^4 \) are correlated with each other or not, we have

\[ \mathbb{E} \{ u_1^H A u_2 \} = \sum_m \sum_n A(m, n) \mathbb{E} \{ u_{1,m}^* \} \mathbb{E} \{ u_{2,n} \} = 0, \]

\[ \mathbb{E} \{ u_1^H A u_3 u_4^H B u_4 \} = \sum_m \sum_n \sum_k \sum_\ell A(m, n) B(k, \ell) \mathbb{E} \{ u_{1,m}^* \} \mathbb{E} \{ u_{2,n} u_{3,k} u_{4,\ell} \} = 0, \]

where \( u_{i,m} \) denotes the \( m \)th entry of \( u_i \). □

**A. First Order Moment**

Employing the UL signal model in (4) and the channel estimate in (5), the first order moment of the desired signal is given as follows

\[ \mathbb{E} \{ h_j^H w_j \} = \mathbb{E} \left\{ h_j^H \tilde{R}_{jj} S^H \left( S \sum_{i=1}^{N_t} h_{ij} + n_j \right) \right\}, \]

\[ = \tau \sum_{i=1}^{N_t} \mathbb{E} \{ h_{ij}^H \tilde{R}_{jj} h_{ij} \} + \mathbb{E} \{ h_{jj}^H \tilde{R}_{jj} S^H n_j \}, \tag{29} \]
where the last line employs $S^H S = \tau I_M$. Since $h_{j,j}$ and $n_j$ are zero-mean and uncorrelated, the second expectation in (29) vanishes as per Lemma 2. The desired expectation becomes

$$\mathbb{E}\{h_{j,j}^H w_j\} = \tau \mathbb{E}\{h_{j,j}^H \tilde{R}_{j,j} h_{j,j}\} + \tau \sum_{i=1 \atop i \neq j}^{N_L} \mathbb{E}\{h_{j,j}^H \tilde{R}_{j,j} h_{j,i}\}, \quad (30)$$

$$= \tau \mathbb{E}\{h_{j,j}^H \tilde{R}_{j,j} h_{j,j}\}, \quad (31)$$

where the second expectation in (30) is similarly zero as per Lemma 2 since the channels of $i$th and $j$th UEs to the $j$th BS, denoted by $h_{i,j}$ and $h_{j,j}$, respectively, are uncorrelated from each other, and zero-mean by definition. Finally, representing (31) by using the trace operator, employing the Hermitian symmetry of the covariance matrix, and incorporating the covariance matrix of the channel estimate in (8) yield the desired expression given in (13).

B. Second Order Moment

Employing (4) and (5), as in Appendix B-A, the second order moment of the desired signal is given as follows

$$\mathbb{E}\{|h_{j,i}^H w_i|^2\} = \mathbb{E}\{h_{j,i}^H \tilde{R}_{j,i} h_{j,i}\}, \quad (32)$$

$$= \mathbb{E}\{h_{j,i}^H \tilde{R}_{j,i} S^H \left(S \sum_{k=1}^{N_L} h_{k,i} + n_i\right) \left(S \sum_{\ell=1}^{N_L} h_{\ell,i} + n_i\right)^H \tilde{R}_{j,i} h_{j,i}\}, \quad (33)$$

$$= \tau^2 \sum_{k=1}^{N_L} \sum_{\ell=1}^{N_L} \mathbb{E}\{h_{j,i}^H \tilde{R}_{j,i} S^H h_{k,i} h_{\ell,i}^H \tilde{R}_{j,i} h_{j,i}\} + \mathbb{E}\{h_{j,i}^H \tilde{R}_{j,i} S^H n_i n_i^H \tilde{R}_{j,i} h_{j,i}\} \quad (34)$$

$$+ \tau \sum_{k=1}^{N_L} \mathbb{E}\{h_{j,i}^H \tilde{R}_{j,i} S^H n_i h_{k,i}^H \tilde{R}_{j,i} h_{j,i}\} + \tau \sum_{k=1}^{N_L} \mathbb{E}\{h_{j,i}^H \tilde{R}_{j,i} n_i h_{k,i}^H \tilde{R}_{j,i} h_{j,i}\}, \quad (35)$$

where the expectations in (35) vanishes in accordance with Lemma 2 since $n_i$ is uncorrelated with $h_{j,i}$ and $h_{k,i}$, and the second order moment becomes

$$\mathbb{E}\{|h_{j,i}^H w_i|^2\} = \tau^2 \sum_{k=1}^{N_L} \sum_{\ell=1}^{N_L} \mathbb{E}\{h_{j,i}^H \tilde{R}_{j,i} S^H h_{k,i} h_{\ell,i}^H \tilde{R}_{j,i} h_{j,i}\} + \mathbb{E}\{h_{j,i}^H \tilde{R}_{j,i} S^H n_i n_i^H \tilde{R}_{j,i} h_{j,i}\}. \quad (36)$$

In the following, we will elaborate the two expectations, $E_1$ and $E_2$, in (36), separately. We
start with $E_1$ as follows

$$
E_1 = \tau^2 \sum_{k=1}^{N_L} \mathbb{E} \left\{ h_{ji}^H \tilde{R}_{ji} h_{ki}^H \tilde{R}_{ki}^H h_{ji} \right\} + \tau^2 \sum_{k=1}^{N_L} \sum_{\ell=1}^{N_L} \sum_{\ell \neq k} \mathbb{E} \left\{ h_{ji}^H \tilde{R}_{ji} h_{ki}^H \tilde{R}_{ki}^H h_{ji} \right\},
$$  \hspace{1cm} (37)

where $h_{ki}$ and $h_{\ell i}$ in the second expectation are obviously uncorrelated as $k \neq \ell$. Note that, $h_{ki}$ is uncorrelated with $h_{\ell i}$ and $h_{ji}$ when $j = \ell$, and $h_{\ell i}$ is uncorrelated with $h_{ki}$ and $h_{ji}$ when $j = k$, and finally all $h_{ki}$, $h_{\ell i}$ and $h_{ji}$ are uncorrelated when $j \neq \{k, \ell\}$. As a result, in any case, we have at least one zero-mean vector uncorrelated with the others, and the second expectation in (37) is therefore zero in accordance with Lemma 2. As a result, $E_1$ in (37) becomes

$$
E_1 = \tau^2 \sum_{k=1}^{N_L} \mathbb{E} \left\{ h_{ji}^H \tilde{R}_{ji} h_{ki}^H \tilde{R}_{ki}^H h_{ji} \right\},
$$

and the first expectation $E_{11}$ can be expressed in weighted sum of scalars as follows

$$
E_{11} = \sum_{m=1}^{M} \sum_{n=1}^{M} \sum_{m'=1}^{M} \sum_{n'=1}^{M} \tilde{R}_{i}(m,n) R_{ji}^H(m',n') \mathbb{E} \left\{ h_{ji,m}^* h_{ji,n} h_{ji,m'}^* h_{ji,n'} \right\},
$$  \hspace{1cm} (39)

where $h_{ji,m}$ is the $m$th element of the channel vector $h_{ji}$, and is given by employing (1) and (2) as follows

$$
h_{ji,m} = \frac{1}{\sqrt{N_P}} \sum_{p=1}^{N_P} \alpha_{ji,p} \exp \left\{ -j2\pi \frac{D}{\lambda} (m-1) \cos (\phi_{ji,p}) \right\}.
$$  \hspace{1cm} (40)

By (40), the expectation at the right-hand side of (39) can be further elaborated as follows

$$
E_\phi(m,n,m',n') = \frac{1}{N_P^2} \sum_{p_1=1}^{N_P} \sum_{p_2=1}^{N_P} \sum_{p_3=1}^{N_P} \sum_{p_4=1}^{N_P} \mathbb{E}_\alpha \mathbb{E} \left\{ \exp \left( -j2\pi \frac{D}{\lambda} \sum_{\nu=1}^{4} (-1)^\nu (u_\nu-1) \cos (\phi_{ji,p_\nu}) \right) \right\},
$$  \hspace{1cm} (41)

with $\{u_\nu\}_{\nu=1}^{4} = \{m, n, m', n'\}$ and $\mathbb{E}_\alpha = \mathbb{E} \left\{ \alpha_{ji,p_1}^* \alpha_{ji,p_2} \alpha_{ji,p_3}^* \alpha_{ji,p_4} \right\}$. Note that, $\mathbb{E}_\alpha$ is nonzero only when 1) $p_1 = p_2 = p_3 = p_4$, 2) $p_1 = p_2, p_3 = p_4$, or 3) $p_1 = p_4, p_2 = p_3$, and zero otherwise, since $\alpha_{ji,p}$ is zero-mean and uncorrelated over the path index $p$. In the following, we analyze these three conditions to have a closed-form expression for (41).

Remark 1: Note that, the other possibilities for the path indices $\{p_\nu\}_{\nu=1}^{4}$ for which $\mathbb{E}_\alpha$ is zero,
The identity $E_\alpha$ can easily be shown to be zero as $\alpha$ in cases i) and ii), the expectation $E_\alpha$ is zero-mean. For the cases i) and ii), the expectation $E_\alpha$ involves the term $\mathbb{E}\{\alpha_{ji,p}\}^\kappa$ with $\kappa \geq 1$ which is zero since $\alpha_{ji,p}$ is zero-mean, and hence yields $E_\alpha = 0$. The case iii) yields $E_\alpha = \mathbb{E}\{\alpha_{ji,p}^2\}^2$ which can easily be shown to be zero as $\alpha_{ji,p}$ has uncorrelated real and imaginary parts which are zero-mean.

**Case 1:** Assuming $p_1 = p_2 = p_3 = p_4$, the desired expectation $E_{\phi}(m, n, m', n')$ in (41) becomes

$$E_{\phi}(m, n, m', n') = \frac{1}{N_p^2} \sum_{p=1}^{N_p} \mathbb{E}\{\alpha_{ji,p}^4\} \mathbb{E}\left\{\exp \left(-j2\pi \frac{D}{\lambda} (n-m+n'-m') \cos (\phi_{ji,p})\right)\right\},$$

$$\quad (a) \quad \frac{1}{N_p} \mathbb{E}\{\alpha_{ji}\} \mathbb{E}\left\{\exp \left(-j2\pi \frac{D}{\lambda} (n-m+n'-m') \cos (\phi_{ji})\right)\right\},$$

$$\quad (b) \quad \frac{2\beta^2_{ji}}{N_p} E_{ji}(n-m+n'-m'),$$

where (a) follows from the uncorrelatedness of $\alpha_{ji}$ and $\phi_{ji}$ over the path index $p$, and (b) employs the identity $\mathbb{E}\{\alpha_{ji}\} = 2\beta^2_{ji}$ and the definition $E_{ji}(m) = \mathbb{E}\left\{\exp \left(-j2\pi \frac{D}{\lambda} (m) \cos (\phi_{ji})\right)\right\}$ which is equal to $R^2_{ji}(n+m,n)$ for any $n \leq M-m$ in (23). Note that $\phi_{ij}$ do not have identical distributions with the same parameters over the various subscripts representing the UE and the BS of interest, and we therefore keep the indices in $E_{ji}$.

**Case 2:** Assuming $p_1 = p_2$, $p_3 = p_4$, the desired expectation $E_{\phi}(m, n, m', n')$ becomes

$$E_{\phi}(m, n, m', n') = \frac{\beta^2_{ji}}{N_p^2} \sum_{p_1=1}^{N_p} \sum_{p_3=1}^{N_p} \mathbb{E}\left\{\exp \left(-j2\pi \frac{D}{\lambda} \left[(n-m) \cos (\phi_{ji,p_1}) + (n'-m') \cos (\phi_{ji,p_3})\right]\right)\right\},$$

$$\quad = \frac{\beta^2_{ji}}{N_p^2} \sum_{p_1=1}^{N_p} \mathbb{E}\left\{\exp \left(-j2\pi \frac{D}{\lambda} (n-m+n'-m') \cos (\phi_{ji,p_1})\right)\right\}$$

$$\quad + \frac{\beta^2_{ji}}{N_p^2} \sum_{p_1=1, p_3 \neq p_1}^{N_p} \sum_{p_3=1}^{N_p} \mathbb{E}\left\{\exp \left(-j2\pi \frac{D}{\lambda} \left[(n-m) \cos (\phi_{ji,p_1}) + (n'-m') \cos (\phi_{ji,p_3})\right]\right)\right\},$$

$$\quad = \frac{\beta^2_{ji}}{N_p} E_{ji}(n-m+n'-m') + \frac{\beta^2_{ji}}{N_p} (N_p - 1) E_{ji}(n-m) E_{ji}(n'-m'),$$

where the last line employs the fact that $\phi_{ji,p_1}$ and $\phi_{ji,p_3}$ are uncorrelated for $p_1 \neq p_3$.

**Case 3:** Assuming $p_1 = p_4$, $p_2 = p_3$, the desired expectation $E_{\phi}(m, n, m', n')$ is obtained by
following the derivation steps of Case 2 which yields

\[ E_\phi(m, n, m', n') = \frac{\beta_j^2}{N_p} E_{ji}(n-m+n'-m') + \frac{\beta_j^2}{N_p} (N_p-1)E_{ji}(n'-m)E_{ji}(n-m'). \]  

(44)

Incorporating (42), (43), and (44) yields the desired expression of \( E_\phi(m, n, m', n') \) in (15), and \( E_{11} \) can be computed by employing (15) in (39).

The second expectation \( E_{12} \) in (38) can be expressed as a weighted sum of scalars as follows

\[ E_{12} = \sum_{m=1}^{M} \sum_{n=1}^{M} \sum_{m'=1}^{M} \sum_{n'=1}^{M} \hat{R}_{ii}(m, n) \hat{R}_{ii}^H(m', n') E \left\{ h_{ji,n}^* h_{ji,m} h_{ki,n} h_{ki,m'}^* \right\}, \]

(45)

where the last line follows from the fact that \( k \neq j \) as imposed by the summation in (38). Employing (39) and (45), we obtain \( E_1 \) given in (38) as follows

\[ E_1 = \tau^2 \sum_{m=1}^{M} \sum_{n=1}^{M} \sum_{m'=1}^{M} \sum_{n'=1}^{M} \hat{R}_{ii}(m, n) \hat{R}_{ii}^H(m', n') \left[ E_\phi(m, n, m', n') + \sum_{k=1; k \neq j}^{N_t} R_{ji}(n', m) \hat{R}_{ki}(n, m') \right], \]

(46)

which can be computed by means of \( E_\phi(m, n, m', n') \) given in (15). Finally, we consider the expectation \( E_2 \) in (36). Defining \( v = S \hat{R}_{ji}^H h_{ji} \) and \( C_n = E \left\{ n_n^H \right\} \), \( E_2 \) is given as

\[ E_2 = E \left\{ v^H n_n^H v \right\} = \sum_{m=1}^{M} \sum_{n=1}^{M} C_n(m, n) E \left\{ v_n^* v_n \right\} = \sigma^2 \sum_{m=1}^{M} E \left\{ |v_m|^2 \right\}, \]

(47)

where we employ \( C_n = \sigma^2 I_M \), and substituting \( v = S \hat{R}_{ii}^H h_{ji} \) back in (47) yields

\[ E_2 = \sigma^2 \text{tr} \left\{ S \hat{R}_{ji}^H \hat{R}_{ji} S^H \right\}. \]

(48)

As a result, substituting (46) and (48) in (36) yields the desired second order moment in (14).

**APPENDIX C**

**MEASURE OF CHANNEL HARDENING**

The channel hardening measure \( \mathcal{M}_{ij} \) is defined in [16] as

\[ \mathcal{M}_{ij} = \frac{\text{Var} \left\{ \|h_{ij}\|^2 \right\}}{E \left\{ \|h_{ij}\|^2 \right\}} = \frac{E \left\{ \|h_{ij}\|^4 \right\}}{(E \left\{ \|h_{ij}\|^2 \right\})^2} - 1, \]

(49)
with the second order moment \( \mathbb{E} \{ \| h_{ij} \|^2 \} = \text{tr} \{ R_{ij} \} = \beta_{ij} M \), and the fourth order moment is
\[
\mathbb{E} \{ \| h_{ij} \|^4 \} = \frac{1}{N_p^2} \sum_{p_1} \sum_{p_2} \sum_{p_3} \sum_{p_4} \mathbb{E}_a \mathbb{E} \{ a^H(\phi_{ij,p_1}) a(\phi_{ij,p_2}) a^H(\phi_{ij,p_3}) a(\phi_{ij,p_4}) \},
\]
with \( \mathbb{E}_a = \mathbb{E} \{ \alpha_{ij,p_1}^* \alpha_{ij,p_2} \alpha_{ij,p_3}^* \alpha_{ij,p_4} \} \). Similar to the discussion for (41), \( \mathbb{E}_a \) is nonzero only when
1) \( p_1 = p_2 = p_3 = p_4 \), 2) \( p_1 = p_2, p_3 = p_4 \), or 3) \( p_1 = p_4, p_2 = p_3 \), and we therefore have
\[
\mathbb{E} \{ \| h_{ij} \|^4 \} = \frac{(2 + N_p)^2 \beta_{ij}^2 M^2}{N_p^2} + \frac{\beta_{ij}^2}{N_p^2} \sum_{p_1} \sum_{p_2} \mathbb{E} \{ a^H(\phi_{ij,p_1}) a(\phi_{ij,p_2}) a^H(\phi_{ij,p_3}) a(\phi_{ij,p_4}) \},
\]
where we employ \( \mathbb{E} \{ |\alpha_{ij}|^4 \} = 2 \beta_{ij}^2 \) and \( a^H(\phi_{ij,p}) a(\phi_{ij,p}) = M \). Substituting the second and the fourth order moments in (49), the channel hardening measure becomes
\[
\mathcal{M}_{ij} = \frac{2}{N_p} + \frac{1}{N_p^2 M^2} \sum_{p_1} \sum_{p_2} \sum_{p_3} \sum_{p_4} \mathbb{E} \{ a^H(\phi_{ij,p_1}) a(\phi_{ij,p_2}) a^H(\phi_{ij,p_3}) a(\phi_{ij,p_4}) \},
\]
\[
= \frac{2}{N_p} + \frac{1}{N_p^2 M^2} \sum_{p_1=1}^{N_p} \sum_{p_2=1}^{N_p} \sum_{n=1}^{M} \sum_{m=1}^{M} \mathbb{E} \left\{ \exp \left( -j 2 \pi \frac{D}{\lambda} \left[ (n-m) \cos(\phi_{ij,p_2}) - \cos(\phi_{ij,p_1}) \right] \right) \right\}.
\]
Taking the terms for the equality of \( p_1 = p_2 \) and \( m = n \) out of the summation, and employing the angular covariance matrix \( R_{ij}^\phi \) given in (22), we end up with
\[
\mathcal{M}_{ij} = \frac{3}{N_p} + \frac{N_p - 1}{N_p M} + \frac{1}{N_p^2 M^2} \sum_{p_1=1}^{N_p} \sum_{p_2=p_1}^{N_p} \sum_{n=1}^{M} \sum_{m=1}^{M} R_{ij}^\phi(n,m) R_{ij}^\phi(m,n),
\]
\[
= \frac{3}{N_p} + \frac{N_p - 1}{N_p M} + \frac{1}{N_p M^2} \sum_{n=1}^{M} \sum_{m=1}^{M} R_{ij}^\phi(n,m) R_{ij}^\phi(m,n), \tag{50}
\]
where the last line employs the uncorrelatedness of \( R_{ij}^\phi \) over the path index. Finally, employing \( E_{ij}(m) = R_{ij}^\phi(n+m,n) \) for any \( n \leq M - m \), and the conjugate symmetry of \( R_{ij}^\phi \) in (50), we obtain the channel hardening measure in (18).

REFERENCES

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” IEEE Commun. Mag., vol. 52, no. 2, pp. 186–195, Feb. 2014.
[2] T. L. Marzetta, “Massive MIMO: An introduction,” Bell Labs Tech. J., vol. 20, pp. 11–22, Mar. 2015.
[3] E. Björnson, E. G. Larsson, and T. L. Marzetta, “Massive MIMO: Ten myths and one critical question,” IEEE Commun. Mag., vol. 54, no. 2, pp. 114–123, Feb. 2016.
[4] O. Elijah, C. Y. Leow, T. A. 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, 2016.
[5] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Tran. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[6] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Trans. on Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[7] J. Hoydis, S. t. 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, no. 2, pp. 160–171, Feb. 2013.

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

[9] H. Yin, L. Cottatellucci, D. Gesbert, R. R. Müller, and G. He, “Robust pilot decontamination based on joint angle and power domain discrimination,” *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2990–3003, Jun. 2016.

[10] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, “Blind pilot decontamination,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[11] M. K. Samimi and T. S. Rappaport, “3-D millimeter-wave statistical channel model for 5G wireless system design,” *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 7, pp. 2207–2225, Jul. 2016.

[12] Technical Specification Group Radio Access Network, “Study on channel model for frequency spectrum above 6 GHz,” 3rd Generation Partnership Project (3GPP), Tech. Rep. 3GPP TR38.900 v14.1.0, 2016.

[13] J. Iscar, I. Güvenç, Ş. Dikmeş, and A. S. Ibrahim, “Optimal angular spread of the multipath clusters in mmWave systems under pilot contamination,” *Proc. IEEE Veh. Technol. Conf. (VTC2017-Fall)*, Sep. 2017, accepted for publication.

[14] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, “Multiple-antenna channel hardening and its implications for rate feedback and scheduling,” *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[15] D. Bai, P. Mitran, S. S. Ghassemzadeh, R. R. Miller, and V. Tarokh, “Rate of channel hardening of antenna selection diversity schemes and its implication on scheduling,” *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4353–4365, Oct. 2009.

[16] H. Q. Ngo and E. G. Larsson, “No downlink pilots are needed in TDD massive MIMO,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, May 2017.

[17] Z. Li, N. Rupasinghe, O. Y. Bursalioglu, C. Wang, H. Papadopoulos, and G. Caire, “Directional Training and Fast Sector-based Processing Schemes for mmWave Channels,” *ArXiv e-prints*, May 2017. [Online]. Available: https://arxiv.org/abs/1611.00453

[18] S. Loyka and F. Gagnon, “Performance analysis of the V-BLAST algorithm: An analytical approach,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1326–1337, Jul. 2004.