ANALYSIS OF THE PILOT CONTAMINATION EFFECT IN VERY LARGE MULTICELL MULTIUSER MIMO SYSTEMS FOR PHYSICAL CHANNEL MODELS

Hien Quoc Ngo * Thomas L. Marzetta† Erik G. Larsson*

* Department of Electrical Engineering (ISY), Linköping University, 581 83 Linköping, Sweden
† Bell Laboratories, Alcatel-Lucent, Murray Hill, NJ 07974, USA

ABSTRACT

We consider multicell multiuser MIMO systems with a very large number of antennas at the base station. We assume that the channel is estimated by using uplink training sequences, and we consider a physical channel model where the angular domain is separated into a finite number of directions. We analyze the so-called pilot contamination effect discovered in previous work, and show that this effect persists under the finite-dimensional channel model that we consider. We further derive closed-form bounds on the achievable rate of uplink data transmission with maximum-ratio combining, for a finite and an infinite number of base station antennas.

Index Terms— Pilot contamination, very large MIMO systems.

1. INTRODUCTION

Multiuser multiple-input multiple-output (MU-MIMO) systems can offer a spatial multiplexing gain without the requirement of multiple antennas at the users [1]. Most studies have assumed that the base station has some channel state information (CSI). The problem of not having an a priori CSI at the base station has been considered in [2,3], assuming that the channel estimation is done by using uplink pilots. This requires time-division duplex (TDD) operation. References [2,3] considered a single-cell setting. This is only reasonable when the pilot sequences available in each cell are orthogonal to those in other cells. However, in practical cellular networks, channel coherence times are not long enough to allow for orthogonality between the pilots in different cells. Therefore, non-orthogonal training sequences must be utilized and hence, the multicell setting should be considered.

In the multicell scenario with non-orthogonal pilots in different cells, channel estimates obtained in a given cell will be impaired by pilots transmitted by users in other cells. This effect, called “pilot contamination” has been analyzed in [4]. Recently, [5] considered the multicell MU-MIMO system with very large antenna arrays at the base station and showed that when the number of antennas increases without bound, uncorrelated noise and fast fading vanish and the pilot contamination effect dictates the ultimate limit on the system performance.

Most of the studies referred to above assume that the channels are independent [2–4] or the channel vectors for different users are asymptotically orthogonal [5]. However, in reality, the MIMO channel is generally correlated because the antennas are not sufficiently well separated or the propagation environment does not offer rich enough scattering. In this paper, we investigate the multicell MU-MIMO with large antenna arrays assuming a physical channel model. More precisely, we consider a finite-dimensional channel model in which the angular domain is partitioned into a large, but finite number of directions which is small relative to the number of base station antennas. The channels are estimated by using uplink training (assuming TDD operation, as in previous work). For such channels, the number of parameters to be estimated is fixed regardless of the number of antennas. We show that the pilot contamination effect persists under the finite-dimensional channel model. Furthermore, we derive a closed-form lower bound on the achievable rate of the uplink transmission, assuming maximum-ratio combining at the base station. This bound is valid for a large but finite number of antennas.

2. SYSTEM MODEL

Consider L cells, where each cell contains one base station equipped with M antennas and K single-antenna users. Assume that the L base stations share the same frequency band. We consider uplink transmission, where the lth base station receives signals from all users in all cells. See Fig. 1. Then, the \( M \times 1 \) received vector at the lth base station is given by

\[
y_l = \sqrt{p_u} \sum_{i=1}^{L} \mathbf{T}_{il} \mathbf{x}_i + \mathbf{n}_l
\]  

where \( \mathbf{T}_{il} \) represents the \( M \times K \) channel matrix between the lth base station and the K users in the ith cell, i.e., \( \mathbf{T}_{il} \) is the channel coefficient between the mth antenna of the lth base station and the kth user in the ith cell; \( \sqrt{p_u} \) is the K × 1 transmitted vector of K users in the ith cell (the average power used by each user is \( p_u \)); and \( \mathbf{n}_l \) contains \( M \times 1 \) additive white Gaussian noise (AWGN). We assume that the elements of \( \mathbf{n}_l \) are Gaussian distributed with zero mean and unit variance.

2.1. Physical Channel Model

Here we introduce the finite-dimensional channel model that is used throughout the paper. The angular domain is divided into a large but finite number of directions \( P \). \( P \) is fixed regardless of the number of base station antennas \( M < P \). Each direction, corresponding to the angle \( \phi_k, \phi_k \in [-\pi/2, \pi/2], k = 1, \ldots, P \), is associated with an \( M \times 1 \) array steering vector \( \mathbf{a}(\phi_k) \) which is given by

\[
\mathbf{a}(\phi_k) = \frac{1}{\sqrt{P}} \left[ e^{-jf_1(\phi_k)}, e^{-jf_2(\phi_k)}, \ldots, e^{-jf_M(\phi_k)} \right]^T
\]  

where \( f_i(\phi) \) is some function of \( \phi \). The channel vector from kth user in the ith cell to the lth base station is then a linear combination of the
steering vectors as follows: $\sum_{m=1}^{P} g_{ilkm} a(\phi_m)$, where $g_{ilkm}$ is the propagation coefficient from the $i$th user to the $l$th base station, associated with the physical direction $m$ (direction of arrival $\phi_m$). Let $G_{il} \triangleq [g_{ilk1} \cdots g_{ilkK}]$ be a $P \times K$ matrix with $g_{ilk} \triangleq [g_{ilk1} \cdots g_{ilkK}]^T$ that contains the path gains from the $i$th user to the $l$th user.

The propagation channel $G_{il}$ models independent fast fading, geometric attenuation, and log-normal shadow fading. Its elements $g_{ilkm}$ are given by

$$g_{ilkm} = h_{ilkm} \sqrt{\beta_{ilk}}, \quad m = 1, \ldots, P$$

where $h_{ilkm}$ is a fast fading coefficient assumed to be zero mean and have unit variance, and $\sqrt{\beta_{ilk}}$ models the path loss and shadowing which are assumed to be independent of the direction $m$ and to be constant and known a priori. This assumption is reasonable since the value of $\beta_{ilk}$ changes very slowly with time. Then, we have

$$G_{il} = H_{il} D_{il}^{1/2}$$

where $H_{il}$ is the $P \times K$ matrix of fast fading coefficients between the $K$ users in the $i$th cell and the $l$th base station, i.e., $[H_{il}]_{km} = h_{ilkm}$, and $D_{il}$ is a $K \times K$ diagonal matrix whose diagonal elements are given by $[D_{il}]_{kk} = \beta_{ilk}$. Therefore, (1) can be written as

$$y_l = \sqrt{p_l} A \sum_{i=1}^{L} g_{il} x_i + n_l = \sqrt{p_l} A \sum_{i=1}^{L} H_{il} D_{il}^{1/2} x_i + n_l$$

(6)

3. CHANNEL ESTIMATION

Channel estimation is performed by using training sequences received on the uplink. A part of the coherence interval is used for the uplink training. All users in all cells simultaneously transmit pilot sequences of length $\tau$ symbols. The assumption on synchronized transmission represents the worst case from the pilot contamination point of view, but it makes no fundamental difference to assume unsynchronized transmission [5]. We assume that the same set of pilot sequences is used in all $L$ cells. Therefore, the pilot sequences used in the $l$th cell can be represented by a $\tau \times K$ matrix $\sqrt{p_l} \Phi_l = \sqrt{p_l} \Phi$ ($\tau \geq K$), which satisfies $\Phi^H \Phi = I_K$, where $p_l = \tau p_M$. From (6), the received pilot matrix at the $l$th base station is

$$Y_{pl} = \sqrt{p_l} A \sum_{i=1}^{L} H_{il} D_{il}^{1/2} \Phi_l^H + N_l$$

(7)

where $N_l \sim \mathcal{N}_M, \{0, I_M, I_K\}$ is noise.

3.1. Minimum Mean-Square Error (MMSE) Estimation

We assume that the base station uses MMSE estimation. The received pilot matrix $Y_{pl}$ can be represented by $Y_{pl} = F_l^H Y_{pl} F_l^H$, where $F_l^H$ is the orthogonal complement of $F_l$. Since $Y_{pl} F_l^H$ only includes the noise part, $Y_{pl} F_l^H$ is a sufficient statistic for the estimation of $H_{il}$. Let $\tilde{Y}_{pl} = Y_{pl} F_l^H$. We have

$$\tilde{Y}_{pl} = \sqrt{p_l} A \sum_{i=1}^{L} H_{il} D_{il}^{1/2} + W_l$$

(8)

where $W_l \sim \mathcal{N}_M, \{0, I_M, I_K\}$. Since $H_{il}$ has independent columns, we can estimate each column of $H_{il}$ independently. Let $\tilde{y}_{pl,ln}$ be the $n$th column of $\tilde{Y}_{pl}$. Then

$$\tilde{y}_{pl,ln} = \sqrt{p_l} A h_{iln}\beta_{iln}^{1/2} + \sqrt{p_l} A \sum_{i \neq l} h_{iln}\beta_{iln}^{1/2} + w_{ln}$$

(9)

where $h_{iln}$ and $w_{ln}$ are the $n$th columns of $H_{il}$ and $W_l$, respectively. Denote by $z_{ln} = \sqrt{p_l} A h_{iln}^{1/2} + R_{iln}$. Then the MMSE estimate of $h_{iln}$ is given by

$$h_{iln} = \beta_{iln}^{1/2} \sqrt{p_l} A^H \left( p_l A^H A + I_p \right)^{-1} z_{iln}$$

(10)

where $R_{iln} = E \left\{ z_{ln} z_{ln}^H \right\}$. By using the matrix inversion lemma, we obtain

$$h_{iln} = \beta_{iln}^{1/2} \sqrt{p_l} A^H \left( p_l A^H A + I_p \right)^{-1} A^H \tilde{y}_{pl,ln}$$

(11)

The $k$th diagonal element of $p_l A^H A \sum_{i=1}^{L} \beta_{iln}$ in (11) equals $\sum_{i=1}^{L} \beta_{iln}$. Since the uplink is typically interference-limited [4], $\frac{M}{p_l} \sum_{i=1}^{L} \beta_{iln} \gg 1$. Therefore, $h_{iln}$ can be approximated as

$$h_{iln} \approx \beta_{iln}^{1/2} \sqrt{p_l} A^H \left( p_l A^H A + I_p \right)^{-1} A^H \tilde{y}_{pl,ln}$$

(12)

Thus, the MMSE estimate of $H_{il}$ is

$$\hat{H}_{il} = p \beta_{iln}^{1/2} A^H \left( p_l A^H A + I_p \right)^{-1} A^H \tilde{y}_{pl,ln}$$

(13)

where $D_{il} = \sum_{i=1}^{L} D_{il}$. Then, the estimate of the physical channel matrix between the $l$th base station and the $K$ users in the $l$th cell is given by

$$\hat{H}_{il} = A H_{il} D_{il}^{1/2} = p \beta_{iln}^{1/2} A^H \tilde{y}_{pl,ln} D_{il}^{1/2}$$

(14)
where \( \Pi_A \triangleq A (A^\dagger A)^{-1} A^\dagger \) is the orthogonal projection onto \( A \).
We can see that since post-multiplication of \( Y_{p,i} \) with \( \Phi^* \) means just multiplication with the pseudoinverse (\( \Phi^* \Phi = I_L \)), \( Y_{p,i} \) is the conventional least-squares channel estimate. The channel estimator that we derived thus performs conventional channel estimation and then projects the estimate onto the physical (beam-space) model for the array.

4. UPLINK DATA TRANSMISSION

We consider the uplink transmission represented by (1). The base station uses its channel estimate obtained by uplink training as in (14) to perform maximum-ratio combining.

4.1. Maximum Ratio Combining

The \( l \)th base station processes its received signal by multiplying it by the conjugate-transpose of the channel estimate. From (6), (8) and (14), we have

\[
\mathbf{r}_l = \sum_{k \neq l} \sum_{n=1}^{L} \mathbf{g}_{ln}^* \mathbf{A} \mathbf{g}_{lk}^* + \mathbf{W}_l^\dagger \mathbf{A}_l^\dagger \mathbf{G}_{jl} x_j + \mathbf{n}_l.
\]

As \( M \to \infty \), the products of uncorrelated quantities can be ignored [5]. Then (15) becomes

\[
\frac{1}{\sqrt{p_l M}} \mathbf{r}_l \to \mathbf{D}_l \mathbf{D}_l^{-1} \left( \sum_{l=1}^{L} \mathbf{G}_{jl} \right) \frac{\mathbf{A}_l^\dagger \mathbf{A}}{M} \left( \sum_{j=1}^{L} \mathbf{G}_{jl} x_j \right) \quad (16)
\]

We can see that for an unlimited number of antennas, the effect of uncorrelated noise disappears. In particular, the pilot contamination effect, which is due to the interference from users in other cells, persists under the finite-dimensional channel model.

We now consider the special case of a uniform array. Then the response vector is given by

\[
\mathbf{a}(\phi_k) = \frac{1}{\sqrt{p}} \left[ 1, e^{-j 2 \pi \frac{d}{\lambda} \sin \phi_k}, ..., e^{-j 2 \pi \frac{d}{\lambda} (M-1) \sin \phi_k} \right]^T
\]

where \( d \) is the antenna spacing, and \( \lambda \) is the wavelength. For \( k \neq l \),

\[
\frac{1}{\sqrt{p}} \mathbf{a}_l^\dagger (\phi_k) \mathbf{a} (\phi_l) = \frac{1}{\sqrt{p}} \sum_{m=0}^{M-1} e^{-j 2 \pi \frac{d}{\lambda} (\sin \phi_k - \sin \phi_l) m} \to \frac{1}{\sqrt{p}} \quad (17)
\]

For \( k = l \), \( \frac{1}{\sqrt{p}} \mathbf{a}_l^\dagger (\phi_k) \mathbf{a} (\phi_l) = \frac{1}{\sqrt{p}} \). Therefore,

\[
\frac{1}{\sqrt{p}} \mathbf{a}_l^\dagger \mathbf{A}_l^\dagger \mathbf{A} \to \frac{1}{\sqrt{p}} \mathbf{I}_P. \quad (18)
\]

Substitution of (19) into (16) yields

\[
\frac{1}{\sqrt{p_l M}} \mathbf{r}_l \to \mathbf{D}_l \mathbf{D}_l^{-1} \left( \sum_{l=1}^{L} \mathbf{G}_{jl} \right) \left( \sum_{j=1}^{L} \mathbf{G}_{jl} x_j \right). \quad (20)
\]

Since the elements of \( \mathbf{G}_{jl} \) are independent, the above result reveals that the performance of the system under the finite-dimensional channel model with \( P \) angular bins and with an unlimited number of base station antennas is the same as the performance under an uncorrelated channel model with \( P \) antennas.

4.2. Analysis of the Pilot Contamination Effect

To gain further insight into the pilot contamination effect, we derive lower bounds on the achievable rate for a finite and an infinite number of base station antennas (the analysis requires that \( M > P \)). To obtain these lower bounds we use the techniques of [3, 4, 6]. We assume a uniform array at the base station, and the elements of \( \mathbf{H}_d \) are i.i.d Gaussian random variables. Let \( r_{ln} \) be the \( n \)th element of the received \( K \times 1 \) vector \( \mathbf{r}_l \). Denote \( \tilde{r}_{ln} \triangleq \frac{\sqrt{\pi} \sum_{l=1}^{L} \beta_{ln} \beta_{il}}{\beta_{il}} r_{ln} \). Then from (15), we have

\[
\tilde{r}_{ln} = \left( \sum_{l=1}^{L} \mathbf{g}_{il} + \mathbf{w}_{ln} \right) \mathbf{A} \left( \sum_{j=1}^{L} \mathbf{G}_{jl} x_j + \mathbf{n}_l \right) \quad (21)
\]

Equation (21) can be rewritten as

\[
\tilde{r}_{ln} = \mathbf{a}_l^T \mathbf{x}_l + \sum_{j \neq l} \mathbf{a}_j^T \mathbf{x}_j + z_{ln}. \quad (22)
\]

where \( \mathbf{a}_l^T \triangleq \frac{\sqrt{\pi} \sum_{l=1}^{L} \beta_{il} \beta_{il}^*}{\beta_{il}} \mathbf{g}_{il} \mathbf{A}_{il} \), and \( z_{ln} \triangleq \left( \frac{\sqrt{\pi} \sum_{l=1}^{L} \beta_{il} \beta_{il}^*}{\beta_{il}} \mathbf{w}_{ln} \right) \mathbf{A} \mathbf{A}_l^\dagger \mathbf{A}_l. \) We have

\[
\mathbb{E} \left\{ \mathbf{a}_l^T \mathbf{x}_l \right\} = \mathbb{E} \left\{ \mathbf{a}_l^T \right\} \mathbf{A}_l^\dagger \mathbf{a}_l^T \mathbf{x}_l \quad (23)
\]

where \( \mathbf{e}_n \) is the \( n \)th column of the \( K \times K \) identity matrix. By adding and subtracting \( \mathbb{E} \left\{ \mathbf{a}_l^T \right\} \) from \( \mathbf{a}_l^T \) in (22), and using (23), we obtain

\[
\tilde{r}_{ln} = \mathbb{E} \left\{ \mathbf{a}_l^T \right\} \mathbf{x}_l + \left( \mathbf{a}_l^T - \mathbb{E} \left\{ \mathbf{a}_l^T \right\} \right) \mathbf{x}_l + \sum_{j \neq l} \mathbf{a}_j^T \mathbf{x}_j + z_{ln}
\]

\[
= \sqrt{p_l p_u} \mathbb{E} \left\{ \mathbf{g}_{il} \mathbf{A}_{il} \mathbf{g}_{il}^* \right\} \mathbf{x}_l + \left( \mathbf{a}_l^T - \mathbb{E} \left\{ \mathbf{a}_l^T \right\} \right) \mathbf{x}_l + \sum_{j \neq l} \mathbf{a}_j^T \mathbf{x}_j + z_{ln}. \quad (24)
\]

Let \( C (x) \triangleq \log_2 (1 + x) \). The \( n \)th user in the \( l \)th cell can achieve an uplink rate of at least

\[
R_{ln} \geq C \left( \sqrt{p_l p_u} \mathbb{E} \left\{ \mathbf{g}_{il} \mathbf{A}_{il} \mathbf{g}_{il}^* \right\} \right)^2 / \left( \sum_{j=1}^{L} \left\| \mathbf{a}_l^T - \mathbb{E} \left\{ \mathbf{a}_l^T \right\} \right\|^2 + \mathbb{E} \left\{ \left| z_{ln} \right|^2 \right\} \right) \quad (25)
\]

Proposition 1 The achievable uplink rate of the \( n \)th user in the \( l \)th cell in (25) is lower bounded by (26), shown at the top of the next page, where \( \beta_{il} \triangleq \sum_{k=1}^{K} \beta_{lk}, \beta_{il}^2 \triangleq \sum_{k=1}^{K} \beta_{lk}^2, \) and

\[
f (A) = \sum_{m=1}^{\varphi (A) \tau (A)^{ \tau (A)}} \chi_{m,n}(A) \mathcal{L}_{n+1} (\mathcal{L}_{m}) \quad (27)
\]

where \( \mathcal{A} \triangleq \text{diag} \left( \lambda_1, \lambda_2, ..., \lambda_P \right) \), with \( \lambda_k \), \( k = 1, ..., P, \) is the \( k \)th eigenvalue of \( \mathcal{A}^\dagger \mathcal{A} \), \( \varphi (A) \) is the number of distinct eigenvalues, \( \lambda_1, \ldots, \lambda_{\varphi (A)} \), are the distinct eigenvalues in decreasing order, \( \tau (A) \) is the multiplicity of \( \lambda_{\varphi (A)} \), and \( \chi_{m,n}(A) \) is the \( (m,n) \)th characteristic coefficient of \( A \) which is defined in [7].

Corollary 1 For an unlimited number of base station antennas, the lower bound on the achievable uplink rate of the \( n \)th user in the \( l \)th cell becomes

\[
R_{ln} \to C \left( \frac{\beta_{il}^2}{\sum_{j \neq l} \beta_{jl}^2 + \frac{1}{\beta_{il}^2} \sum_{j=1}^{L} \beta_{jl}^2} \right) \quad (28)
\]
\[
R_{\text{ln}} = C \left( \sum_{j=1}^{L} \left( \frac{f(A^j)}{M^2} \beta^2_{jln} \right) + \frac{p_0 p_u \beta^2_{jln}}{M^2 \lambda} \beta_{jl} \beta_{ln} \sum_{k \neq l} \left( \frac{f(A^k)}{M^2} \beta^2_{kl} \right) - p_0 p_u \beta^2_{jln} + \frac{p_n \beta^2_{jln}}{M^2} \right) \right)
\]  

(26)

Fig. 2. Lower bound on the sum-rate of uplink transmission as a function of the number of base station antennas \( M \).

Remark 1 If \( P \) also goes to infinity, the lower bound on the achievable rate becomes \( R_{\text{ln}}^\infty = C \left( \frac{\beta^2_{jln}}{\sum_{j \neq l} \beta^2_{jln}} \right) \). This equals the exact value for the rate obtained in [5]. This is due to the fact that when \( P \) is large, things that were random before become deterministic and hence, the lower bound approaches the exact value.

5. NUMERICAL RESULTS

We consider a system with 4 cells, and \( K = 10 \) users per cell. The training sequence length is \( \tau = K \), and the number of physical directions is \( P = 20 \). We consider a uniform array at the base station with \( \frac{\sigma}{\tau} = 0.3 \) and \( \phi_k = -\pi/2 + (k - 1) \pi/P \). We further assume that all direct gains are equal to 1 and all cross gains are equal to \( a \), i.e., \( \beta_{jlk} = 1 \), and \( \beta_{jlk} = a \), \( \forall j \neq l, k = 1, \ldots, K \). The lower bound on the sum-rate of \( l \)th cell is defined as \( R_l = \sum_{n=1}^{K} R_{ln} \).

Figure 2 shows the lower bound on the sum-rate of the uplink transmission in the \( l \)th cell versus the number of antennas \( M \), at \( a = 0.1 \) and for different average transmit powers per user \( p_u = -10, -5, 0, \) and 10 dB. We see that at low SNR (since the noise power is unity, the SNR is equal to \( p_u \)), using a large number of base station antennas significantly improves the achievable rate. Furthermore, we can see that as \( M \rightarrow \infty \), the sum-rates approach \( R_{ln}^\infty \). This is the asymptotic value of sum-rate with an unlimited number of base station antennas and it is independent of the SNR (cf. (28)).

Figure 3 depicts the lower bound on the sum-rate of uplink transmission versus the cross gain at an SNR of \( p_u = 0 \) dB for different \( M \) = 20, 50 and 100. We can see that the effect of pilot contamination can be very significant if the value of the cross gain is close to the value of the direct gain, regardless of \( M \).

6. CONCLUDING REMARKS

This paper has analyzed the pilot contamination effect in multicell MU-MIMO systems with very large antenna arrays for a physical channel model with a finite number of scattering centers visible from the base station. We showed that the pilot contamination effect discovered in [5] persists under the finite-dimensional channel model. We also derived a closed-form lower bound on the achievable uplink rate for finite and infinite number of base station antennas.

7. REFERENCES