Pilot Contamination Problem in Multi-Cell TDD Systems

Jubin Jose*, Alexei Ashikhmin†, Thomas L. Marzetta‡, and Sriram Vishwanath*

*Laboratory for Informatics, Networks and Communications (LINC)
Department of Electrical and Computer Engineering
The University of Texas at Austin, Austin, TX 78712
†Bell Laboratories, Alcatel-Lucent Inc., Murray Hill, NJ 07974

Abstract—This paper considers a multi-cell multiple antenna system with precoding at the base stations for downlink transmission. To enable precoding, channel state information (CSI) is obtained via uplink training. This paper mathematically characterizes the impact that uplink training has on the performance of multi-cell multiple antenna systems. When non-orthogonal training sequences are used for uplink training, it is shown that the precoding matrix used by the base station in one cell becomes corrupted by the channel between that base station and the users in other cells. This problem of pilot contamination is analyzed in this paper. A multi-cell MMSE-based precoding is proposed that, when combined with frequency/time/pilot reuse techniques, mitigate this problem.

I. INTRODUCTION

Wireless cellular networks with multiple antennas at the base stations are a central component of networks designed to enable through put intensive applications. In the single-cell setting known as multi-user multiple-input multiple output (MIMO), multiple antenna systems have been studied extensively (see [1]–[3] and references therein). It is well understood that channel state information (CSI) at the base station is critical for achieving high system performance. In TDD systems, an efficient method to obtain CSI at the base station is through uplink training by utilizing the reciprocity of the wireless medium [4]. In this paper, we consider uplink training and transmit precoding in a multi-cell scenario with $L$ cells, where each cell consists of a base station with $M$ antennas and $K$ users with single antenna each.

The impact of uplink training on the resulting channel estimate (and thus system performance) in the multi-cell scenario is significantly different from that in a single-cell scenario. In the multi-cell scenario, non-orthogonal training sequences (pilots) must be utilized, as orthogonal pilots would need length of at least $K \times L$ symbols. In practice, the short channel coherence times due to mobility do not allow for such long training sequences.

Pilot contamination is encountered only when analyzing a multi-cell MIMO system with training, and is lost when narrowing focus to a single-cell setting. Pilot contamination occurs when the channel estimate at the base station in one cell becomes polluted by users from other cells. The use of non-orthogonal training sequences causes this contamination. Our goal in this paper is to develop techniques (and thus achievable rates) to mitigate this contamination.

A technique (similar to the techniques used in [3]) can be applied to analyze any precoding method used at the base stations. For the setting with one user in every cell, we derive closed-form expressions for the achievable rates given by this technique. These closed form expressions allow us to determine the extent to which pilot contamination impacts system performance. In particular, we show that the achievable rates can saturate with the number of antennas at the base station $M$. This analysis allows one to determine the appropriate frequency/time/pilot reuse factor to maximize system throughput in the presence of pilot contamination.

In the multi-cell scenario, there has been significant research on utilizing coordination among base stations [5]–[8] when CSI is available. This existing body of work focuses on the gain that can be obtained through coordination of the base stations. Dirty paper coding based approaches and joint beamforming/precoding approaches are considered in [6]. Linear precoding methods for clustered networks with full intra-cluster coordination and limited inter-cluster coordination are proposed in [9]. These approaches generally require “good” channel estimates at the base stations. Due to non-orthogonal training sequences, the resulting channel estimate (of the channel between a base station and all users) can be shown to be rank deficient. We propose a multi-cell MMSE-based precoding that depends on the set of training sequences assigned to the users. Our approach does not need coordination between base stations required by the joint precoding techniques. When coordination is present, this approach can be applied at the inter-cluster level. In many training sequence allocations, numerical results show that our approach give significant gains over zero-forcing precoding on the users in every cell.

II. SYSTEM MODEL

We consider a cellular system with $L$ cells numbered $1, 2, \ldots, L$. Each cell comprises of one base station with $M$ antennas and $K(\leq M)$ single-antenna users. Let the average power (during transmission) at the base station be $p_f$ and
the average power at each user be \( p_r \). The propagation factor between the \( m^{th} \) base station antenna of the \( l^{th} \) cell and the \( k^{th} \) user of the \( j^{th} \) cell is \( \sqrt{\beta_{jk} h_{jkm}} \), where \( \{ \beta_{jk} \} \) are non-negative constants and assumed to be known to everybody, and \( \{ h_{jkm} \} \) are i.i.d. \( \mathcal{CN}(0,1) \) random variables. We assume channel reciprocity for the forward and reverse links, i.e., the propagation factor \( \sqrt{\beta_{jk} h_{jkm}} \) is same for both forward and reverse links, and block fading, i.e., \( \{ h_{jkm} \} \) remains constant for a duration of \( T \) symbols. The additive noises at all terminals are i.i.d. \( \mathcal{CN}(0,1) \) random variables. The system equations describing the signals received at the base station and the users are given in the next section.

### III. Communication Scheme

The communication scheme consists of two phases: uplink training and data transmission.

#### A. Uplink Training

At the beginning of every coherence interval, all users (in all cells) synchronously transmit training sequences of length \( \tau \) symbols. Let \( \sqrt{\tau \psi_j l} \), \( \psi_j l \) = 1, be the training vector transmitted by the \( k^{th} \) user in the \( j^{th} \) cell. Consider the base station of the \( l^{th} \) cell. The vector received at the \( m^{th} \) antenna is

\[
y_m = \sum_{j=1}^{L} \sum_{k=1}^{K} \sqrt{p_r \beta_{jk} h_{jkm}} \psi_j l w_{jm} + w_{lm},
\]

where \( w_{jm} \) is the additive noise. Let \( \Lambda_l = [\psi_{j1}, \psi_{j2}, \ldots, \psi_{jM}] \), \( \Psi_j = [\psi_{j1}, \psi_{j2}, \ldots, \psi_{jK}] \), \( \mathbf{D}_{jl} = \text{diag}\{\beta_{j1}, \beta_{j2}, \ldots, \beta_{jK}\} \), and \( \mathbf{H}_{jl} = \begin{bmatrix} h_{j11} & \cdots & h_{j1M} \\ \vdots & \ddots & \vdots \\ h_{jK1} & \cdots & h_{jKM} \end{bmatrix} \).

From (1), the signal received at the \( j^{th} \) base station is

\[
Y_l = \sqrt{p_r \tau} \sum_{l=1}^{L} \Psi_j \mathbf{D}_{jl} \mathbf{H}_{jl} + \Lambda_l.
\]

The MMSE estimate of the channel \( \mathbf{H}_{jl} \) given \( Y_l \) in (2) is

\[
\hat{\mathbf{H}}_{jl} = \sqrt{p_r \tau} \mathbf{D}_{jl} \Psi_j^{-1} \left( I + p_r \tau \sum_{l=1}^{L} \Psi_j \mathbf{D}_{jl} \Psi_j^{-1} \right)^{-1} Y_l.
\]

Let \( \hat{\Lambda}_l = [\hat{\Lambda}_{l1}, \hat{\Lambda}_{l2}, \ldots, \hat{\Lambda}_{LM}] \).

#### B. Downlink Transmission

Consider the base station of the \( l^{th} \) cell. Let the information symbols to be transmitted to \( K \) users be \( q_j = [q_{j1} q_{j2} \ldots q_{jK}]^T \) and the \( M \times K \) linear precoding matrix be \( \mathbf{A}_l \). The signal vector transmitted by this base station is \( \mathbf{A}_l q_j \). We consider transmission symbols and precoding methods such that \( \mathbb{E}[q_j] = 0 \), \( \mathbb{E}[q_j q_j^H] = \mathbf{I} \), and \( \text{tr} \left( \mathbf{A}_l^H \mathbf{A}_l \right) = 1 \). These (sufficient) conditions imply that the average power constraint at the base station is satisfied.

Now, consider the users in the \( j^{th} \) cell. The noisy signal vector received by these users is

\[
x_j = \sum_{l=1}^{L} \sqrt{p_f} \mathbf{D}_{jl}^H \mathbf{H}_{jl} \mathbf{A}_l q_j + z_j,
\]

where \( z_j \) is the additive noise. From (4), we have

\[
x_{jk} = \sum_{l=1}^{L} \sqrt{p_f} \beta_{jk} h_{jkl1} h_{jkl2} \ldots h_{jklM} \mathbf{D}_{jl} q_{li} + z_{jk},
\]

where \( z_{jk} \) is the \( k^{th} \) element of the precoding matrix \( \mathbf{A}_l \) and \( z_{jk} \) is the \( k^{th} \) column of the precoding matrix \( \mathbf{A}_l \).

#### C. Achievable Rate

We provide a lower bound on the rate that is achievable using the above mentioned communication scheme by assuming worst-case Gaussian noise. This method is widely used (for e.g. [2]) as a lower bounding technique. Consider the \( k^{th} \) user in the \( j^{th} \) cell. From (5), the signal received by this user can be written as

\[
x_{jk} = \sum_{l=1}^{M} \sum_{i=1}^{K} g_{jk}^{li} q_{li} + z_{jk} = \mathbb{E}[g_{jk} q_{jk}] + \sum_{(i,j) \neq (j,k)} g_{jk}^{li} q_{li} + z_{jk},
\]

where \( g_{jk}^{li} = \sqrt{p_f \beta_{jk} h_{jkl1} h_{jkl2} \ldots h_{jklM}} \). We denote the effective noise (last three terms) in (6) by \( z_{jk}' \). Since \( q_{jk} \) and \( z_{jk}' \) are uncorrelated and \( \mathbb{E}[|z_{jk}'|^2] = \text{var} \{ |g_{jk}^{li}|^2 \} + \sum_{(i,j) \neq (j,k)} \mathbb{E}[|g_{jk}^{li}|^2] + 1 \), it is clear that the rate

\[
R_{jk} = C \left( 1 + \frac{\mathbb{E}[|g_{jk} q_{jk}|^2]}{\sum_{(i,j) \neq (j,k)} \mathbb{E}[|g_{jk}^{li}|^2]^2} \right)
\]

is achievable by this user. Here, \( C(\theta) = \log_2(1 + \theta) \).

### IV. Pilot Contamination Analysis

We analyze the pilot contamination problem in the following setting: one user per cell (\( K = 1 \)), same training sequence used by all users (\( \psi_{j1} = \psi_{j2} = \ldots \psi_{jK} \)) and zero-forcing (ZF) precoding. We consider this setting as it captures the primary effect of pilot contamination which is the correlation between the precoding matrix and the channel to users in other cells. We provide simple and insightful analytical results in this setting. To simplify notation, we drop the subscripts associated with the users in every cell, and we use \( h_{jl}, \mathbf{h}_{jl}, \mathbf{a}_l \), for \( \mathbf{H}_{jl}, \hat{\mathbf{H}}_{jl}, \mathbf{A}_l \) respectively. For ZF precoding, the precoding vector used the \( l^{th} \) cell is given by \( \mathbf{a}_l = \frac{h_{jl}}{\sqrt{h_{jl}^H h_{jl}}} \). The user in the \( j^{th} \) cell receives signal from its base station and from other base stations. From (4), this received signal is

\[
x_j = \sqrt{p_f} \beta_{jl} \mathbf{h}_{jl} q_j + \sum_{l \neq j} \sqrt{p_f} \beta_{jl} \mathbf{h}_{jl} \mathbf{a}_l q_l + z_j.
\]

We compute first and second order moments of the effective channel gain and the inter-cell interference, and use these to obtain an expression for the achievable rate given by (7).
In the setting considered here, the MMSE estimate of $\hat{h}_{jl}$ based on $Y_j$ given by (3) can be simplified using matrix inversion lemma as follows:

$$\hat{h}_{jl} = \frac{\sqrt{p_r^T \beta_{jl} \psi^T \left( I + \psi \left( p_r^T \sum_{i=1}^{L} \beta_{il} \right) \psi \right)^{-1}} Y_j}{\| \psi \|^2 Y_j}, \forall j.$$ \hspace{2cm} (10)

Using (10), we obtain

$$\hat{h}_{jl} q_j = \hat{h}_{jl} \frac{\hat{h}_{jl}}{\| \hat{h}_{jl} \|^2} = \hat{h}_{jl} \frac{\hat{h}_{jl}^\dagger}{\| \hat{h}_{jl} \|^2} + \hat{h}_{jl} \frac{\hat{h}_{jl}^\dagger}{\| \hat{h}_{jl} \|^2}.$$ \hspace{2cm} (11)

where $\hat{h}_{jl} = \bar{h}_{jl} - \hat{h}_{jl}$. From the properties of MMSE estimation, we know that $\bar{h}_{jl}$ is independent of $\hat{h}_{jl}$, $\bar{h}_{jl}$ is $CN \left( \sum_{i=1}^{L} \beta_{il}, I \right)$, and $\hat{h}_{jl}$ is $CN \left( \sum_{i=1}^{L} \beta_{jl}, \frac{1}{1+p_r^T \sum_{i=1}^{L} \beta_{il}} I \right)$. From (11), we get

$$\mathbb{E}[\hat{h}_{jl} q_j] = \mathbb{E}[\| \hat{h}_{jl} \|^2] = \sqrt{\frac{p_r^T \beta_{jl}}{1+p_r^T \sum_{i=1}^{L} \beta_{il}}} \mathbb{E}[\theta], \quad \text{where} \quad \theta = \sum_{m=1}^{M} |u_m|^2.$$ \hspace{2cm} (12)

where $\theta = \sum_{m=1}^{M} |u_m|^2$ and $\{u_m\}$ is i.i.d. $CN(0, 1)$. From (11), we also have $\mathbb{E}[\| \hat{h}_{jl} \|^2]$ and variance $\text{var} \left( \| \sqrt{p_f^T \beta_{jl}} \hat{h}_{jl} q_j \| \right)$.

$$\mathbb{E}[\| \hat{h}_{jl} \|^2] = \left[ \frac{p_r^T \beta_{jl}}{1+p_r^T \sum_{i=1}^{L} \beta_{il}} \right] \mathbb{E}[\theta],$$

$$\text{var} \left( \| \sqrt{p_f^T \beta_{jl}} \hat{h}_{jl} q_j \| \right) = \left( \frac{p_r^T \beta_{jl}}{1+p_r^T \sum_{i=1}^{L} \beta_{il}} \right) \mathbb{E}[\theta].$$

Lemma 1: The effective channel gain in (8) has expectation

$$\mathbb{E}[\sqrt{p_f^T \beta_{jl}} \hat{h}_{jl} q_j] = \left( \frac{p_r^T \beta_{jl}}{1+p_r^T \sum_{i=1}^{L} \beta_{il}} \right)^{1/2} \mathbb{E}[\theta].$$

and variance $\text{var} \left( \| \sqrt{p_f^T \beta_{jl}} \hat{h}_{jl} q_j \| \right)$.

Lemma 2: For both signal and interference terms in (8), the first and second order moments are as follows: $\mathbb{E}[\sqrt{p_f^T \beta_{jl}} \hat{h}_{jl} q_j] = 0$, $\mathbb{E}[\| \sqrt{p_f^T \beta_{jl}} \hat{h}_{jl} q_j \|^2] = p_f^T \beta_{jl}$, $\mathbb{E}[\theta^2] = \frac{1}{1+p_r^T \sum_{i=1}^{L} \beta_{il}}$. Here, $\Gamma(\cdot)$ is the Gamma function. For large $M$, the achievable rate can be simplified as

$$\lim_{M \to \infty} R_j = C \left( \frac{\beta_{jl}^2}{\frac{1}{1+p_r^T \sum_{i=1}^{L} \beta_{il}} + \frac{1}{1+p_r^T \sum_{i=1}^{L} \beta_{il}}} \right).$$ \hspace{2cm} (14)

For large $M$, the value of $\text{var} \left( \theta \right)$ is insignificant compared to $M$. The results of the above theorem show that the performance does saturate with $M$. Typically, the reverse link is interference-limited, i.e., $p_r^T \sum_{i=1}^{L} \beta_{il} \gg 1, \forall j$. The term $\sum_{i=1}^{L} \beta_{ij}$ is the expected sum of squares of the propagation coefficients between the base station in the $j^{th}$ cell and all users. Therefore, $\sum_{i=1}^{L} \beta_{ij}$ is generally constant with respect to $j$. Using these approximations in (14), we get

$$R_j \approx C \left( \frac{\beta_{jl}^2}{\frac{1}{1+p_r^T \sum_{i=1}^{L} \beta_{il}}} \right).$$

This clearly show that the effect pilot contamination can be very significant if cross gains (between cells) are of the same order of direct gains (within the same cell).

V. MULTI-CELL MMSE PRECODING

The analysis in the previous section suggests that frequency/time/pilot reuse techniques are required to mitigate the pilot contamination problem. In this section, we develop a precoding method that can be combined with these techniques. In particular, we formulate an optimization problem which leads to the frequency reuse 1 (FR-1) multi-cell MMSE-based precoding. The generalization to other reuse schemes is straightforward.

Consider the $j^{th}$ cell. We use the following notation: $F_{jl} = \sqrt{p_f^T D_j^T \beta_{jl}}\tilde{H}_{jl}$, $\tilde{F}_{jl} = \sqrt{p_f^T D_j^T \beta_{jl}}\tilde{H}_{jl}$ and $\tilde{F}_{jl} = F_{jl} - \tilde{F}_{jl}$ for all $j$ and $l$. The signal received by the users in this cell given by (4) is a function of all the precoding matrices (used at all the base stations). Therefore, the MMSE-based precoding methods for single-cell setting considered in [3] does not extend (directly) to this setting. Let us consider the signal and interference terms corresponding to the base station in the $j^{th}$ cell. Based on these...
Here, we consider only the case of two cells. The general case rigorous proof of this fact). Assumptions can only increase the sum capacity (we omit a

We further assume that the effective channel matrices and other similar techniques. Since our main concern is the multi-cell MMSE precoding matrix ($A^{\text{opt}}$).

Next, we obtain a closed-form expression for $A^{\text{opt}}$. The following lemma is required for obtaining the optimal solution to the problem (15).

**Lemma 3:** Let $A_{jl} = \left(I + p_r \gamma \sum_{i \neq j} \Psi_i D_i \Psi_i^\dagger\right)^{-1}$. For all $j$ and $l$, $E[f_{jl} | j] = \delta_{jl} I_M$, where

$$
\delta_{jl} = p_f \text{tr} \left\{ D_{jl} \left( I_K + p_r \gamma D_j^2 \Psi_j^\dagger A_j \Psi_j D_j^2 \right)^{-1} \right\}.
$$

(16)

**Theorem 2:** The optimal solution to the problem (15) is

$$
A^{\text{opt}}_l = \frac{1}{\alpha^{\text{opt}}} \left( \hat{F}^\dagger_l \hat{F}_l + \gamma^2 \sum_{j \neq l} \hat{F}_{jl} \hat{F}_j + \eta I_M \right)^{-1} \hat{F}^\dagger_l,
$$

where $\eta = \delta_{ll} + \gamma^2 \sum_{j \neq l} \delta_{jl} + K$, $\delta_{jl}$ is given by (16) and $\alpha^{\text{opt}}$ is such that $\|A^{\text{opt}}_l\|^2 = 1$.

The precoding described above is suited for maximizing the minimum of the rates achieved by all the users. When the performance metric of interest is sum rate, this precoding can be straightforwardly combined with power control, scheduling, and other similar techniques. Since our main concern is the inter-cell interference, and to avoid too complicated systems, we do not use that possibility in this paper. In Section VII, all comparisons are performed without power control.

**VI. Upper Bound**

In order to derive an upper bound, we assume that all channel matrices $\hat{F}_{jl}$ are available to all the base stations. We further assume that the effective channel matrices $F_{jl} A_j$ are known to all the users. It is easy to see that these two assumptions can only increase the sum capacity (we omit a rigorous proof of this fact).

Unlike the cooperative MIMO scenario, we assume that all base stations transmit data only to users located in their cells. Here, we consider only the case of two cells. The general case of $L$ cells can be treated similarly. If precoding matrices $A_1$ and $A_2$ are used, then, under the above assumptions, the total sum rate is

$$
R(F_{11}, F_{12}, F_{21}, F_{22}, A_1, A_2) = \sum_{j=1}^{K} \left( \frac{\frac{1}{K} \gamma \sum_{r \neq j} |F_{11r} A_j|^2}{\sigma^2 T_1} + \frac{\frac{1}{K} \gamma \sum_{r \neq j} |F_{21r} A_j|^2}{\sigma^2 T_2} \right) + \sum_{j=1}^{K} \left( \frac{\frac{1}{K} \gamma \sum_{r \neq j} |F_{22r} A_j|^2}{\sigma^2 T_2} + \frac{\frac{1}{K} \gamma \sum_{r \neq j} |F_{22r} A_j|^2}{\sigma^2 T_2} \right),
$$

where $T_i = \text{tr}(A_j^\dagger A_j)$, $i = 1, 2$, $F_{jl}^\dagger$ is the $j$th row of $F_{jl}$ and $\bar{a}_{il}$ is the $i$th row of $A_j$. We would like to maximize the function

$$
\mathbb{E}[f_{jl} | j] \left[ R(F_{11} + \hat{F}_{11}, \hat{F}_{12} + \hat{F}_{12}, \hat{F}_{21} + \hat{F}_{11}, \hat{F}_{22} + \hat{F}_{22}, A_1, A_2) \right].
$$

(17)

To maximize (17), we generalize the approach in [11]. Since $\hat{F}_{jl}$ are the MMSE estimates, the entries of $\hat{F}_{jl}$ are zero-mean complex Gaussian random variables with known variances. Let $\hat{F}_{jl}^\dagger$, $j, l = \{1, 2\}; i = \{1, \ldots, n\}$, be random matrices generated according to the corresponding distributions. Define $F_{jl}^\dagger = \hat{F}_{jl}^\dagger + \hat{F}_{jl}^\dagger$. Then, (17) is equal to

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} R(F_{11}^i, F_{12}^i, F_{21}^i, F_{22}^i, A_1, A_2).
$$

(18)

By taking partial derivatives of (18) with respect to entries of $A_1$ and $A_2$ and zeroing them out, we obtain the following result.

**Theorem 3:** The maximum in (18) is achieved only if

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left( F_{11}^i \Delta_1^i - (F_{12}^i)^\dagger \gamma_1 (F_{12}^i)^\dagger + F_{11}^i D_1^i \right) + \left( \gamma^2 \text{tr}(D_1^i) - \text{tr}(\Delta_1^i I_M) \right) A_1 + F_{21}^i \Delta_2^i - (F_{21}^i)^\dagger \gamma_2 (F_{21}^i)^\dagger + F_{22}^i D_2^i A_2 = 0,
$$

where $\Delta_1^i, D_1^i, \gamma_1, D_2^i, \gamma_2$ are $K \times K$ diagonal matrices defined via matrices $A_1, A_2, F_{jl}^\dagger, j, l = 1, 2$ (we omit their precise definitions because of limited space).

By choosing sufficiently large $n$, say $n = 1000$, we can well approximate (17). For a finite $n$ the equation from Theorem 3 can be solved numerically (we omit details) with respect to $A_1$ and $A_2$. Using these $A_1$ and $A_2$ in (18), we obtain an approximate upper bound on the average sum rate. The proofs (omitted due to lack of space) of lemmas and theorems are given in [10].

**VII. Numerical Results**

We consider the following two examples. In the first example, we consider a system with $p_f = 10$ dB, $p_r = -10$ dB, $L = 2$ cells, $K$ users in every cell and training length of $\tau = K$. The same set of orthogonal training sequences are used in both the cells. We keep all the direct gains to be 1 and all cross gains to be $a$, i.e., for all $k$, $\beta_{ijk} = 1$ if $j = l$ and $\beta_{ijk} = a$ if $j \neq l$.

In the second example, we consider a multi-cell system with $L = 4$ cells, $M = 8$ antennas at all base stations, $K = 2$
values of are reused in well “separated” cells. We keep the propagation in this example, we model a scenario where training sequences are used in the $1^{st}$ and $2^{nd}$ cells. The training sequences used in the $1^{st}$ ($2^{nd}$) cell are used in the $3^{rd}$ ($4^{th}$) cell. In this example, we model a scenario where training sequences are reused in well “separated” cells. We keep the propagation factors as follows: for all $k$, $\beta_{jlk} = 1$ if $j = l$, $\beta_{jlk} = a$ if $(j, l) \in \{(1, 2), (2, 1), (3, 4), (4, 3)\}$, and $\beta_{jlk} = b$ for all other values of $j$ and $l$.

We define the sum rate per cell as $R_{\text{sum}} = \frac{1}{L} \sum_{j,k} R_{jk}$ and the minimum rate achieved by all users by $R = \min_{j,k} R_{jk}$. We plot sum rates per cell for different precoding schemes for first and second examples in Fig. 1 and Fig. 2 respectively. GPS denotes precoding suggested in [3], which uses only pilots from users located in the same cell as the base station. ZF precoding has the same meaning. In all the examples, the Multi-Cell MMSE precoding is used with FR-1.

In Fig. 3, we plot minimum rates for the second example for ZF and multi-cell MMSE precoding. We observe improved performance of multi-cell MMSE precoding for wide range of values for $a$ and $b$. The main improvement comes from the fact that the base stations form precoding matrices that do not create significant interference to other cells. In Fig. 4, we plot minimum rates for the second example for GPS and multi-cell MMSE precodings as functions of $M$.

**REFERENCES**


