

# Decontaminating Pilots in Massive MIMO Systems

Haifan Yin<sup>†</sup>, David Gesbert<sup>†</sup>, Miltiades C. Filippou<sup>†</sup>, Yingzhuang Liu<sup>\*</sup>

<sup>†</sup>Mobile Communications Department, EURECOM, 06410 Biot, France

<sup>\*</sup>Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, 430074 Wuhan, China

Email: {yin, gesbert, filippou}@eurecom.fr, liuyz@mail.hust.edu.cn

**Abstract**—Pilot contamination is known to severely limit the performance of large-scale antenna (“massive MIMO”) systems due to degraded channel estimation. This paper proposes a two-fold approach to this problem. First we show analytically that pilot contamination can be made to vanish asymptotically in the number of antennas for a certain class of channel fading statistics. The key lies in setting a suitable condition on the second order statistics for desired and interference signals. Second we show how a coordinated user-to-pilot assignment method can be devised to help fulfill this condition in practical networks. Large gains are illustrated in our simulations for even small antenna array sizes.

**Index Terms**—massive MIMO, pilot contamination, channel estimation, scheduling, covariance information.

## I. INTRODUCTION

FULL reuse of the frequency across neighboring cells leads to severe interference, which in turn limits the quality of service offered to cellular users, especially those located at the cell edge. As service providers seek solutions to restore performance in low-SINR cell locations, several approaches aimed at mitigating inter-cell interference have emerged in the last few years. Among these, the solutions which exploit the additional degrees of freedom made available by the use of multiple antennas seem the most promising, particularly so at the base station side where such arrays are more affordable.

In an effort to solve this problem while limiting the requirements for user data sharing over the backhaul network, coordinated beamforming approaches have been proposed in which 1) multiple-antenna processing is exploited at each base station, and 2) the optimization of the beamforming vectors at all cooperating base stations is performed jointly. Coordinated beamforming does not require the exchange of user message information unlike in e.g. Network MIMO. Yet it still demands the exchange of channel state information across the transmitters on a fast time scale and low-latency basis, making almost as challenging to implement in practice as the above mentioned network MIMO schemes.

Fortunately a path towards solving some of the essential practical problems related to beamforming-based interference avoidance was suggested in [1]. In this work, it was pointed out that the need for exchanging CSIT between base stations could be alleviated by simply increasing the number  $M$  of antennas at each transmitter (so-called massive MIMO). This result is rooted in the law of large numbers, which predicts that, as the number of antennas increases, the vector channel for a desired terminal will tend to be more orthogonal to the vector channel

of a randomly selected interfering user. This makes it possible to reject interference at the base station side by simply aligning the beamforming vector with the desired channel (“Maximum Ratio Combining” or spatial matched filter). Hence in theory, a simple fully distributed per-cell beamforming scheme can offer performance scaling (with  $M$ ) similar to a more complex centralized optimization.

Unfortunately, the above conclusion relies on perfect channel estimation at the base station side. In reality channel information is acquired on the basis of finite length pilot sequences, and crucially, in the presence of inter-cell interference. In particular, it has been shown that pilot contamination effects [2] [3] [4] (i.e., the reuse of non-orthogonal pilot sequences across interfering cells) cause the interference rejection performance to quickly saturate with the number of antennas, thereby undermining the value of MIMO systems in cellular network scenarios.

In this paper, we address the problem of channel estimation in the presence of multi-cell interference generated from pilot contamination. We propose an estimation method which provides a substantial improvement in performance and relies on two key ideas. The first is the exploitation of dormant side-information lying in the second-order statistics of the user channels, both for desired and interfering users. We show that the channel estimation performance is a function of the degree to which dominant signal subspaces pertaining to the desired and interference channel covariance overlap with each other. In particular, we demonstrate a powerful result indicating that the exploitation of covariance information under certain subspace conditions on the covariance matrices can lead to a complete removal of pilot contamination effects in the large  $M$  limit. We then turn to a practical algorithm design where this concept is exploited. The key idea behind the new algorithm is the use of a covariance-aware pilot assignment strategy within the channel estimation phase itself. While diversity-based scheduling methods have been popularized for maximizing various throughput-fairness performance criteria [5] [6] [7] [8], the potential benefit of user-to-pilot assignment in the context of interference-prone channel estimation has received very little attention so far.

## II. SIGNAL AND CHANNEL MODELS

We consider a network of  $L$  time-synchronized cells, with full spectrum reuse. Note that assuming synchronization between uplink pilots provides a worst case scenario from a pilot

contamination point of view, since any lack of synchronization will tend to statistically decorrelate the pilots. Estimation of (block-fading) channels in the uplink is considered<sup>1</sup>, and all the base stations are equipped with  $M$  antennas. To simplify the notations, we assume the 1st cell is the target cell, unless otherwise notified. For ease of exposition we consider the case where a single user per cell transmits its pilot sequence to its serving base. The pilot sequence used in the  $l$ -th cell is denoted by:

$$\mathbf{s}_l = [s_{l1} \ s_{l2} \ \cdots \ s_{l\tau}]^T. \quad (1)$$

The powers of pilot sequences are assumed equal such that  $|s_{l1}|^2 + \cdots + |s_{l\tau}|^2 = \tau, l = 1, 2, \dots, L$ .

We denote the receive covariance matrix  $\mathbf{R}_l \in \mathbb{C}^{M \times M}$  as  $\mathbf{R}_l \triangleq \mathbb{E}\{\mathbf{h}_l \mathbf{h}_l^H\}$ , where  $\mathbf{h}_l$  is the channel vector between the  $l$ -th cell user and the target base station. Thus  $\mathbf{h}_1$  is the desired channel while  $\mathbf{h}_l, l > 1$  are interference channels. Channel vectors are assumed to be  $M \times 1$  complex Gaussian, undergoing correlation due to the finite multipath angle spread at the base station side as [9]:

$$\mathbf{h}_l = \mathbf{R}_l^{1/2} \mathbf{h}_{w_l}, \quad (2)$$

where  $\mathbf{h}_{w_l}$  is the spatially white  $M \times 1$  SIMO channel with  $\mathbf{h}_{w_l} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M), l = 1, 2, \dots, L$ , and  $\mathbf{I}_M$  is the  $M \times M$  identity matrix. In this paper, we make the assumption that covariance matrix  $\mathbf{R}_l \triangleq \mathbb{E}\{\mathbf{h}_l \mathbf{h}_l^H\}$  can be obtained separately from the desired and interference channels. During the pilot phase, the  $M \times \tau$  signal received at the target base station is

$$\mathbf{Y} = \sum_{l=1}^L \mathbf{h}_l \mathbf{s}_l^T + \mathbf{N}, \quad (3)$$

where  $\mathbf{N} \in \mathbb{C}^{M \times \tau}$  is the spatially and temporally white additive Gaussian noise with element-wise variance  $\sigma_n^2$ .

### III. COVARIANCE-BASED CHANNEL ESTIMATION

#### A. Pilot Contamination

Conventional channel estimation relies on correlating the received signal with the known pilot sequence (referred here as Least Squares (LS) estimate). Hence, using the model in (3), an LS estimator for the desired channel  $\mathbf{h}_1$  is

$$\hat{\mathbf{h}}_1^{\text{LS}} = \mathbf{Y} \mathbf{s}_1^* (\mathbf{s}_1^T \mathbf{s}_1^*)^{-1}. \quad (4)$$

The conventional estimator suffers from a lack of orthogonality between the desired and interfering pilots, an effect known as pilot contamination [2], [10], [11]. In particular, when the *same* pilot sequence is reused in all  $L$  cells, i.e.,  $\mathbf{s}_1 = \cdots = \mathbf{s}_L = \mathbf{s}$ , the estimator can be written as

$$\hat{\mathbf{h}}_1^{\text{LS}} = \mathbf{h}_1 + \sum_{l \neq 1}^L \mathbf{h}_l + \mathbf{N} \mathbf{s}^* / \tau. \quad (5)$$

As it appears in (5), the interfering channels leak directly into the desired channel estimate. The estimation performance is

<sup>1</sup>Similar ideas would be applicable for downlink channel estimation, provided the UE is equipped with multiple antennas as well, in which case the estimation would help resolve interferences originating from neighboring base stations.

then limited by the signal to interference ratio at the base station, which in turns limits the ability to design an effective interference-avoiding beamforming solution.

#### B. Bayesian Estimation

We hereby propose an improved channel estimator with the aim of reducing the pilot contamination effect, and taking advantage of the multiple antenna dimensions. We suggest to do so by exploiting side information lying in the second order statistics (covariance matrices) of the channel vectors. The role of covariance matrices is to capture structure information related to the distribution (mainly mean and spread) of the multi-path angles of arrival at the base station. Due to the typically elevated position of the base station, rays impinge on the antennas with a finite angle-of-arrival (AOA) spread and a user location-dependent mean angle. Note that covariance-aided channel estimation itself is not a novel idea, e.g. in [12]. In [13], the authors focused on optimal design of pilot sequences and they exploited the covariance matrices of desired channels and colored interference. The optimal training sequences were developed with adaptation to the statistics of disturbance. In our paper, however, the pilot design itself is shown not having an impact on interference reduction, since fully aligned pilots could be transmitted. Instead we focus on i) studying the limiting behavior of covariance-based estimates in the presence of interference and large-scale antenna arrays, and ii) how to *shape* covariance information for the full benefit of channel estimation quality.

Previously reported results dealing with large scale antenna systems have revealed that even simple receivers such as spatial matched filters are enough to eliminate interference when  $M$  grows large [1]. We therefore present a channel estimation method for the desired channel only. A more general case which estimates the interference channels as well can be found in [14]. For ease of exposition, the worst case situation with a unique pilot sequence reused in all  $L$  cells is considered:

$$\mathbf{s} = [s_1 \ s_2 \ \cdots \ s_\tau]^T. \quad (6)$$

We define a training matrix  $\bar{\mathbf{S}} \triangleq \mathbf{s} \otimes \mathbf{I}_M$ . Note that  $\bar{\mathbf{S}}^H \bar{\mathbf{S}} = \tau \mathbf{I}_M$ . Then the vectorized received training signal at the target base station can be expressed as

$$\mathbf{y} = \bar{\mathbf{S}} \sum_{l=1}^L \mathbf{h}_l + \mathbf{n}, \quad (7)$$

where  $\mathbf{y} \triangleq \text{vec}(\mathbf{Y})$ ,  $\mathbf{n} \triangleq \text{vec}(\mathbf{N})$ . Exploiting the covariance information of the channels, a Bayesian estimator, which is equivalent to an MMSE estimator [15], is given by

$$\hat{\mathbf{h}}_1 = \mathbf{R}_1 \bar{\mathbf{S}}^H \left( \bar{\mathbf{S}} \left( \sum_{l=1}^L \mathbf{R}_l \right) \bar{\mathbf{S}}^H + \sigma_n^2 \mathbf{I}_{\tau M} \right)^{-1} \mathbf{y}. \quad (8)$$

Note that the expression (8) is similar to the traditional Bayesian estimate as shown in [15] [13]. The difference is that here identical pilot sequences are sent by users, and covariance information is assumed to be known.

A more convenient form can be written as:

$$\hat{\mathbf{h}}_1 = \mathbf{R}_1 \left( \sigma_n^2 \mathbf{I}_M + \tau \sum_{l=1}^L \mathbf{R}_l \right)^{-1} \bar{\mathbf{S}}^H \mathbf{y}. \quad (9)$$

In the section below, we examine the degradation caused by the pilot contamination on the estimation performance. In particular, we point out the role played by the use of covariance matrices in dramatically improving the pilot contamination effects under certain conditions on the rank structure.

### C. Performance Analysis

We are interested in the mean square error (MSE) of the proposed estimator, which can be defined as:

$$\mathcal{M}_1 \triangleq \mathbb{E} \left\{ \left\| \hat{\mathbf{h}}_1 - \mathbf{h}_1 \right\|_F^2 \right\}. \quad (10)$$

The MSE of the Bayesian estimator (9) with fully aligned pilots is given by

$$\mathcal{M}_1 = \text{tr} \left\{ \mathbf{R}_1 - \mathbf{R}_1^2 \left( \frac{\sigma_n^2}{\tau} \mathbf{I}_M + \sum_{l=1}^L \mathbf{R}_l \right)^{-1} \right\}. \quad (11)$$

Of course, it is clear from (11) that the MSE is not dependent on the specific design of the pilot sequence, but on the total power of it.

We can readily get the channel estimate of (9) obtained in an interference free scenario, by setting interference terms to zero:

$$\hat{\mathbf{h}}_1^{\text{no int}} = \mathbf{R}_1 (\sigma_n^2 \mathbf{I}_M + \tau \mathbf{R}_1)^{-1} \bar{\mathbf{S}}^H (\bar{\mathbf{S}} \mathbf{h}_1 + \mathbf{n}), \quad (12)$$

where superscript *no int* refers to the "no interference case", and the corresponding MSE:

$$\mathcal{M}_1^{\text{no int}} = \text{tr} \left\{ \mathbf{R}_1 \left( \mathbf{I}_M + \frac{\tau}{\sigma_n^2} \mathbf{R}_1 \right)^{-1} \right\}. \quad (13)$$

### D. Large Scale Analysis

We seek to analyze the performance for the above estimators in the regime of large antenna number  $M$ . For tractability, our analysis is based on the assumption of a uniform linear array with supercritical antenna spacing (i.e., less than or equal to half wavelength).

Hence we have the following multipath model

$$\mathbf{h}_i = \frac{1}{\sqrt{P}} \sum_{p=1}^P \mathbf{a}(\theta_{ip}) \alpha_{ip}, \quad (14)$$

where  $P$  is the arbitrary number of i.i.d. paths,  $\alpha_{ip} \sim \mathcal{CN}(0, \delta_i^2)$  is independent over channel index  $i$  and path index  $p$ , where  $\delta_i$  is the  $i$ -th channel's average attenuation.  $\mathbf{a}(\theta)$  is the steering vector, as shown in [16]:

$$\mathbf{a}(\theta) \triangleq \begin{bmatrix} 1 \\ e^{-j2\pi \frac{D}{\lambda} \cos(\theta)} \\ \vdots \\ e^{-j2\pi \frac{(M-1)D}{\lambda} \cos(\theta)} \end{bmatrix}, \quad (15)$$

where  $D$  is the antenna spacing at the base station and  $\lambda$  is the signal wavelength, such that  $D \leq \lambda/2$ .  $\theta_{ip} \in [0, \pi]$  is a random AOA. Note that we can limit angles to  $[0, \pi]$  because any  $\theta \in [-\pi, 0]$  can be replaced by  $-\theta$  giving the same steering vector.

Below, we momentarily assume that the selected users exhibit multipath AOAs that do not overlap with the AOAs of the desired user, i.e., the AOA spread and user locations are such that multipath for the desired user are confined to a region of space where interfering paths are very unlikely to exist. Although the asymptotic analysis below makes use of this condition, it will be shown in Section IV how such a structure can be shaped implicitly by the coordinated pilot assignment. Finally, simulations reveal in Section V the robustness with respect to an overlap between AOA in desired and interference channel (for instance in the case of Gaussian AOA distribution).

Our main result is as follows:

**Theorem 1.** Assume the multipath angle of arrival  $\theta$  yielding channel  $\mathbf{h}_j, j = 1, \dots, L$ , in (14), is distributed according to an arbitrary density  $p_j(\theta)$  with bounded support, i.e.,  $p_j(\theta) = 0$  for  $\theta \notin [\theta_j^{\min}, \theta_j^{\max}]$  for some fixed  $\theta_j^{\min} \leq \theta_j^{\max} \in [0, \pi]$ . If the  $L - 1$  intervals  $[\theta_i^{\min}, \theta_i^{\max}]$ ,  $i = 2, \dots, L$  are strictly non-overlapping with the desired channel's AOA interval<sup>2</sup>  $[\theta_1^{\min}, \theta_1^{\max}]$ , we have

$$\lim_{M \rightarrow \infty} \hat{\mathbf{h}}_1 = \hat{\mathbf{h}}_1^{\text{no int}}. \quad (16)$$

From the channel model (14), we get

$$\mathbf{R}_i = \frac{\delta_i^2}{P} \sum_{p=1}^P \mathbb{E} \{ \mathbf{a}(\theta_{ip}) \mathbf{a}(\theta_{ip})^H \} = \delta_i^2 \mathbb{E} \{ \mathbf{a}(\theta_i) \mathbf{a}(\theta_i)^H \},$$

where  $\theta_i$  has the PDF  $p_i(\theta)$  for all  $i = 1, \dots, L$ . The proof of Theorem 1 relies on three intermediate lemmas which explicit the eigenstructure of the covariance matrices. Due to lack of space, the detailed proofs of these lemmas are omitted and can be found in [14].

**Lemma 1.** Define  $\boldsymbol{\alpha}(x) \triangleq [1 \ e^{-j\pi x} \ \dots \ e^{-j\pi(M-1)x}]^T$  and  $\mathcal{A} \triangleq \text{span}\{\boldsymbol{\alpha}(x), x \in [-1, 1]\}$ . Given  $b_1, b_2 \in [-1, 1]$  and  $b_1 < b_2$ , define  $\mathcal{B} \triangleq \text{span}\{\boldsymbol{\alpha}(x), x \in [b_1, b_2]\}$ , then

- $\dim\{\mathcal{A}\} = M$
- $\dim\{\mathcal{B}\} \sim (b_2 - b_1)M/2$  when  $M$  grows large.

**Lemma 2.** When  $M$  grows large,

$$\text{rank}(\mathbf{R}_i) \leq d_i M,$$

where  $d_i$  is defined as

$$d_i \triangleq (\cos(\theta_i^{\min}) - \cos(\theta_i^{\max})) \frac{D}{\lambda}.$$

<sup>2</sup>This condition is just one example of a practical scenario leading to non-overlapping signal subspaces between the desired and the interference covariances, however, more general multipath scenarios could be used.

**Lemma 3.** *The null space  $\text{null}(\mathbf{R}_i)$  includes a certain set of unit-norm vectors:*

$$\text{null}(\mathbf{R}_i) \supset \text{span} \left\{ \frac{\mathbf{a}(\Phi)}{\sqrt{M}}, \forall \Phi \notin [\theta_i^{\min}, \theta_i^{\max}] \right\}, \text{ as } M \rightarrow \infty.$$

Lemma 1 characterizes the number of dimensions a linear space has, which is spanned by  $\alpha(x)$ , in which  $x$  plays the role of spatial frequency. Lemma 2 indicates that for large  $M$ , there exists a null space  $\text{null}(\mathbf{R}_i)$  of dimension  $(1 - d_i)M$ . Interestingly, related eigenstructure properties of the covariance matrices were independently derived in [17] [18] for the purpose of reducing the overhead of downlink channel estimation and CSI feedback in massive MIMO for FDD systems. Lemma 3 indicates that multipath components with AOA outside the AOA region for a given user will tend to fall in the null space of its covariance matrix in the large number of antenna case.

We now return to the proof of theorem 1. When  $M$  is large,  $\mathbf{R}_i$  can be decomposed into

$$\mathbf{R}_i = \mathbf{U}_i \boldsymbol{\Sigma}_i \mathbf{U}_i^H, \quad (17)$$

where  $\mathbf{U}_i$  is the signal eigenvector matrix of size  $M \times m_i$ , in which  $m_i \leq d_i M$ .  $\boldsymbol{\Sigma}_i$  is an eigenvalue matrix of size  $m_i \times m_i$ . Due to lemma 3 and the fact that densities  $p_i(\theta)$  and  $p_1(\theta)$  have non-overlapping supports, we have

$$\mathbf{U}_i^H \mathbf{U}_1 = 0, \forall i \neq 1, \text{ as } M \rightarrow \infty. \quad (18)$$

Combining the channel estimate (9) and the channel model (7), we obtain

$$\hat{\mathbf{h}}_1 = \mathbf{R}_1 \left( \sigma_n^2 \mathbf{I}_M + \tau \sum_{l=1}^L \mathbf{R}_l \right)^{-1} \bar{\mathbf{S}}^H \left( \bar{\mathbf{S}} \sum_{i=1}^L \mathbf{h}_i + \mathbf{n} \right).$$

According to (18), matrices  $\mathbf{R}_1$  and  $\sum_{l=2}^L \mathbf{R}_l$  span orthogonal subspaces in the large  $M$  limit. Therefore we place ourselves in the asymptotic regime for  $M$ , in this case  $\tau \sum_{l=2}^L \mathbf{R}_l$  can be eigen-decomposed according to:

$$\tau \sum_{l=2}^L \mathbf{R}_l = \mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^H, \quad (19)$$

where  $\mathbf{W}$  is the eigenvector matrix such that  $\mathbf{W}^H \mathbf{W} = \mathbf{I}$  and  $\text{span}\{\mathbf{W}\}$  is included in the orthogonal complement of  $\text{span}\{\mathbf{U}_1\}$ . Now denote  $\mathbf{V}$  the unitary matrix corresponding to the orthogonal complement of both  $\text{span}\{\mathbf{W}\}$  and  $\text{span}\{\mathbf{U}_1\}$  so that the  $M \times M$  identity matrix can now be decomposed into:

$$\mathbf{I}_M = \mathbf{U}_1 \mathbf{U}_1^H + \mathbf{W} \mathbf{W}^H + \mathbf{V} \mathbf{V}^H. \quad (20)$$

Thus, for large  $M$ :

$$\begin{aligned} \hat{\mathbf{h}}_1 \sim & \mathbf{U}_1 \boldsymbol{\Sigma}_1 \mathbf{U}_1^H \left( \sigma_n^2 \mathbf{U}_1 \mathbf{U}_1^H + \sigma_n^2 \mathbf{V} \mathbf{V}^H + \sigma_n^2 \mathbf{W} \mathbf{W}^H \right. \\ & \left. + \tau \mathbf{U}_1 \boldsymbol{\Sigma}_1 \mathbf{U}_1^H + \mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^H \right)^{-1} \left( \tau \sum_{i=1}^L \mathbf{h}_i + \bar{\mathbf{S}}^H \mathbf{n} \right). \end{aligned}$$

Due to asymptotic orthogonality between  $\mathbf{U}_1$ ,  $\mathbf{W}$  and  $\mathbf{V}$ ,

$$\begin{aligned} \hat{\mathbf{h}}_1 & \sim \mathbf{U}_1 \boldsymbol{\Sigma}_1 (\sigma_n^2 \mathbf{I}_{m_1} + \tau \boldsymbol{\Sigma}_1)^{-1} \mathbf{U}_1^H \left( \tau \sum_{i=1}^L \mathbf{h}_i + \bar{\mathbf{S}}^H \mathbf{n} \right) \\ & \sim \mathbf{U}_1 \boldsymbol{\Sigma}_1 (\sigma_n^2 \mathbf{I}_{m_1} + \tau \boldsymbol{\Sigma}_1)^{-1} \tau \left( \mathbf{U}_1^H \mathbf{h}_1 + \sum_{i=2}^L \mathbf{U}_1^H \mathbf{h}_i + \frac{\bar{\mathbf{S}}^H \mathbf{n}_1}{\tau} \right). \end{aligned}$$

However, since  $\mathbf{h}_i \in \text{span}\{\mathbf{a}(\theta), \forall \theta \in [\theta_i^{\min}, \theta_i^{\max}]\}$ , we have from lemma 3 that  $\frac{\|\mathbf{U}_1^H \mathbf{h}_i\|}{\|\mathbf{U}_1^H \mathbf{h}_1\|} \rightarrow 0$ , for  $i \neq 1$  when  $M \rightarrow \infty$ . Therefore

$$\lim_{M \rightarrow \infty} \hat{\mathbf{h}}_1 = \tau \mathbf{U}_1 \boldsymbol{\Sigma}_1 (\sigma_n^2 \mathbf{I}_{m_1} + \tau \boldsymbol{\Sigma}_1)^{-1} \left( \mathbf{U}_1^H \mathbf{h}_1 + \frac{\bar{\mathbf{S}}^H \mathbf{n}}{\tau} \right),$$

which is identical to  $\hat{\mathbf{h}}_1^{\text{no int}}$  if we apply the EVD decomposition (17) for  $\mathbf{R}_1$  in (12). This proves theorem 1.  $\square$

We also believe that, although antenna calibration is needed as a technical assumption in the theorem, orthogonality of covariance's signal subspaces will occur in non-tightly calibrated settings provided the AOA regions do not overlap.

#### IV. COORDINATED PILOT ASSIGNMENT

In the ideal case where the desired and the interference covariances span distinct subspaces, we have demonstrated that the pilot contamination effect tends to vanish in the large antenna array case. In this section we make use of this property by designing a suitable coordination protocol for assigning pilot sequences to users in the  $L$  cells. The role of the coordination is to *shape* covariance matrices in an effort to try and satisfy the non-overlapping AOA constraint of Theorem 1. We assume that in all  $L$  cells, the considered pilot sequence will be assigned to one (out of  $K$ ) user in each of the  $L$  cells. Let  $\mathcal{G} \triangleq \{1, \dots, K\}$ , then  $\mathcal{K}_l \in \mathcal{G}$  denotes the index of the user in the  $l$ -th cell who is assigned the pilot sequence  $\mathbf{s}$ . The set of selected users is denoted by  $\mathcal{U}$  in what follows. For a user set  $\mathcal{U}$ , we define a network utility function

$$F(\mathcal{U}) \triangleq \sum_{j=1}^{|\mathcal{U}|} \frac{\mathcal{M}_j(\mathcal{U})}{\text{tr}\{\mathbf{R}_{j_j}(\mathcal{U})\}}, \quad (21)$$

where  $|\mathcal{U}|$  is the cardinal number of the set  $\mathcal{U}$ .  $\mathcal{M}_j(\mathcal{U})$  is the estimation MSE for the desired channel at the  $j$ -th base station, with a notation readily extended from  $\mathcal{M}_1$  in (11), where this time cell  $j$  is the target cell when computing  $\mathcal{M}_j$ .  $\mathbf{R}_{j_j}(\mathcal{U})$  is the covariance matrix of the desired channel at the  $j$ -th cell.

The principle of the coordinated pilot assignment is exploiting covariance information at all cells in order to minimize the sum MSE metric. Hence,  $L$  users are assigned to an identical pilot sequence when the corresponding  $L^2$  covariance matrices exhibit the most orthogonal signal subspaces. To reduce the complexity, a classical greedy approach is proposed:

- 1)  $\mathcal{U} = \emptyset$
- 2) For  $l = 1, \dots, L$  do:
  - $\mathcal{K}_l = \arg \min_{k \in \mathcal{G}} F(\mathcal{U} \cup \{k\})$
  - $\mathcal{U} \leftarrow \mathcal{U} \cup \{\mathcal{K}_l\}$
- End

The coordination can be interpreted as follows: To minimize the estimation error, a base station tends to assign a given pilot to the user whose spatial feature has most differences with the interfering users assigned this same pilot.

## V. NUMERICAL RESULTS

In order to preserve fairness between users and avoid having high-SNR users being systematically assigned the pilot  $s$ , we consider a symmetric multicell network where the users are all distributed on the cell edge and have the same distance with their base stations.<sup>3</sup> We adopt the model of a cluster of synchronized and hexagonally shaped cells. We keep the basic simulation parameters in Table I, unless otherwise stated.

TABLE I  
BASIC SIMULATION PARAMETERS

Cell radius	1 km
Cell edge SNR	20 dB
Number of users per-cell	10
Distance from a user to its BS	800 m
Path loss exponent	3
Carrier frequency	2 GHz
Antenna spacing	$\lambda/2$
Number of paths	50
Pilot length	10

Two types of AOA distributions are considered here, a bounded one (uniform) and a non-bounded one (Gaussian). Two performance metrics are used to evaluate the proposed channel estimation scheme: the channel estimation error normalized by the F-norm of the channel coefficients, and the per-cell rate obtained assuming standard maximal ratio combining (MRC) beamformer based on the channel estimate.

In the following part, "LS" stands for conventional LS channel estimation. "CB" denotes the Covariance-aided Bayesian estimation (without coordination), and "CPA" is the proposed Coordinated Pilot Assignment-based Bayesian estimation.

We first validate theorem 1 in Fig. 1 with a 2-cell network, where the two users' positions are fixed. AOA of desired channels are uniformly distributed with a mean of 90 degrees, and an angle spread  $\theta_{\Delta} = 20$  degrees for both users, yielding no overlap between desired and interfering multipaths. The pilot contamination is quickly eliminated with increasing the number of antennas.

In Fig. 2, the estimation MSE versus the BS antenna number is illustrated where the AOA have uniform distribution. The performance of CPA estimator improves quickly with  $M$  from 2 to 10. In the 2-cell network, the proposed pilot assignment policy has the ability of avoiding the overlap between AOA for the desired and interference channels.

We then examine the impact of RMS angle spread of uniformly distributed AOA on the estimation error. We can see in Fig. 3 that the estimation error is a monotonically increasing function of the RMS angle spread. In contrast, an angle spread tending toward zero will cause the channel

<sup>3</sup>In practice, users with greater average SNR levels (but equal across cells) can be assigned together on a separate pilot pool.

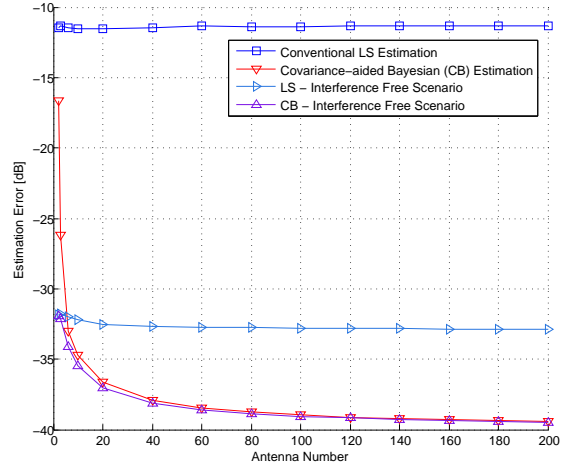


Fig. 1. Estimation MSE vs. BS antenna number, 2-cell network, fixed positions of two users, uniformly distributed AOA with  $\theta_{\Delta} = 20$  degrees, non-overlapping multipath.

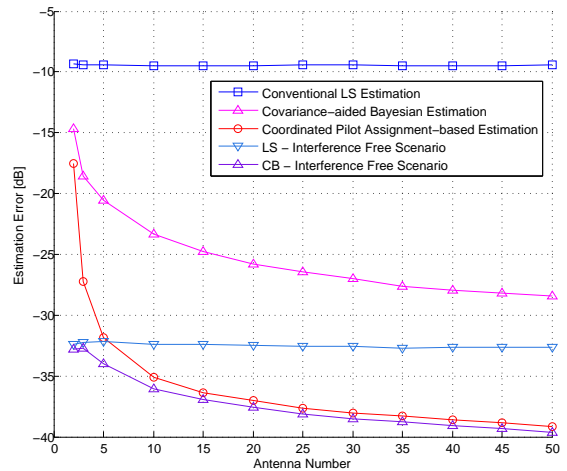


Fig. 2. Estimation MSE vs. antenna number, uniformly distributed AOA with  $\theta_{\Delta} = 10$  degrees, 2-cell network.

direction to collapse into a deterministic quantity, yielding large gains for covariance-based channel estimation.

Fig. 4 depicts the downlink per-cell rate achieved by the MRC beamforming strategy and suggests large gains when the Bayesian estimation is used in conjunction with the proposed coordinated pilot assignment strategy and intermediate gains when it is used alone. Interestingly it is shown that the rate performance almost saturates with  $M$  in the classical LS case (due to pilot contamination) while it increases quickly with  $M$  for the proposed estimators, indicating the full benefits of massive MIMO systems are exploited.

## VI. DISCUSSIONS

In this paper, we assumed the individual covariance matrices can be estimated separately. This could be done in practice by exploiting resource blocks where the desired user and interference users are known to be assigned pilot sequences at different times. In future networks, one may imagine a

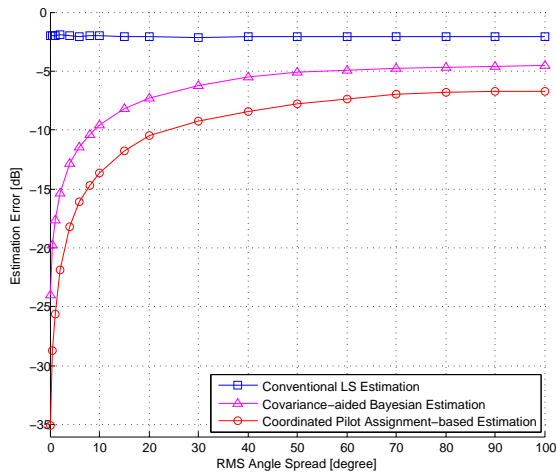


Fig. 3. Estimation MSE vs. RMS angle spread, uniformly distributed AOs,  $M = 10$ , 7-cell network.

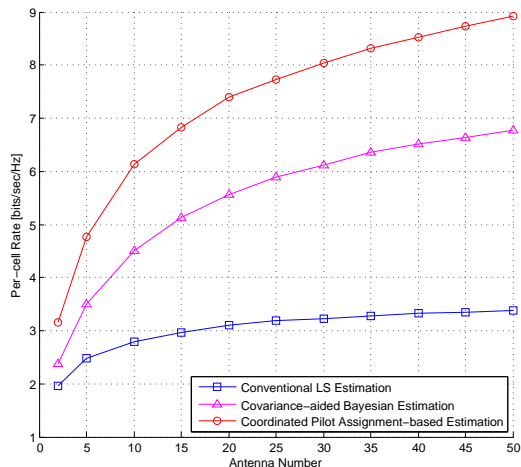


Fig. 4. Per-cell rate vs. antenna number, Gaussian distributed AOs with standard deviation  $\sigma = 10$  degrees, 7-cell network.

specific training design for learning second-order statistics. Since covariance information varies much slower than fast fading, such training may not consume a substantial amount of resources.

## VII. CONCLUSIONS

This paper proposed a covariance-aided channel estimation framework in the context of interference-limited multi-cell multiple antenna systems. We proposed a Bayesian estimator and demonstrated analytically the efficiency of such an approach for large-scale antenna systems, leading to a complete removal of pilot contamination effects in the case covariance matrices satisfy a certain non-overlapping condition on their dominant subspaces. We suggested a coordinated pilot assignment strategy that helps shape covariance matrices toward satisfying the needed condition and showed that channel estimation performance is close to interference-free scenarios.

## ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under grant No. 60972015 and 61231007, by the European research projects SAPHYRE and HARP under the FP7 ICT Objective 1.1 - The Network of the Future, and by the French national ANR-VERSO funded project LICORNE. Discussions with Dirk Slock and Laura Cottatellucci are gratefully acknowledged.

## REFERENCES

- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [2] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination problem in multi-cell TDD systems," in *Proc. IEEE International Symposium on Information Theory (ISIT09)*, Seoul, Korea, Jun. 2009, pp. 2184–2188.
- [3] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *Proc. 2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2011, pp. 545–550.
- [4] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell tdd systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [5] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE International Conference on Communications*, vol. 1, Seattle, WA, USA, Jun. 1995, pp. 331–335.
- [6] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.
- [7] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.
- [8] D. Gesbert, M. Kountouris, R. Heath, C. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sep. 2007.
- [9] A. J. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge University Press, 2003.
- [10] H. Q. Ngo, T. L. Marzetta, and E. G. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP11)*, Prague, Czech Republic, May 2011, pp. 3464–3467.
- [11] B. Gopalakrishnan and N. Jindal, "An analysis of pilot contamination on multi-user MIMO cellular systems with many antennas," in *2011 IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2011, pp. 381–385.
- [12] A. Scherb and K. Kammeyer, "Bayesian channel estimation for doubly correlated MIMO systems," in *Proc. IEEE Workshop Smart Antennas*, 2007.
- [13] E. Bjornson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1807–1820, Mar. 2010.
- [14] 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.
- [15] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.
- [16] J. A. Tsai, R. M. Buehrer, and B. D. Woerner, "The impact of AOA energy distribution on the spatial fading correlation of linear antenna array," in *Proc. IEEE Vehicular Technology Conference, (VTC 02)*, vol. 2, May 2002, pp. 933–937.
- [17] J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information," in *46th Annual Conference on Information Sciences and Systems, (CISS 2012)*, Princeton University, NJ, USA, Mar. 2012, pp. 1–6.
- [18] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing," *CoRR*, 2012. [Online]. Available: <http://arxiv.org/abs/1209.1402>