# CS-Based CSIT Estimation for Downlink Pilot Decontamination in Multi-Cell FDD Massive MIMO

Zhen Gao

***Abstract*—Efficient channel state information at transmitter (CSIT) for frequency division duplex (FDD) massive MIMO can facilitate its backward compatibility with existing FDD cellular networks. To date, several CSIT estimation schemes have been proposed for FDD single-cell massive MIMO systems, but they fail to consider inter-cell-interference (ICI) and suffer from downlink pilot contamination in multi-cell scenario. To solve this problem, this letter proposes a compressive sensing (CS)-based CSIT estimation scheme to combat ICI in FDD multi-cell massive MIMO systems. Specifically, angle-domain massive MIMO channels exhibit the common sparsity over different subcarriers, and such sparsity is partially shared by adjacent users. By exploiting these sparsity properties, we design the pilot signal and the associated channel estimation algorithm under the framework of CS theory, where the channels associated with multiple adjacent BSs can be reliably estimated with low training overhead for downlink pilot decontamination. Simulation results verify the good downlink pilot decontamination performance of the proposed solution compared to its conventional counterparts in multi-cell FDD massive MIMO.**

***Index Terms*—Frequency division duplex (FDD), massive MIMO, channel estimation, compressive sensing, pilot contamination.**

## I. INTRODUCTION

Reliable channel state information at transmitter (CSIT) is essential to fully exploit potential advantages of massive MIMO. For time division duplex (TDD) massive MIMO, CSIT can be acquired in the uplink by leveraging the channel reciprocity, where the channels of dozens of users can be easily acquired at base station (BS) with hundreds of antennas [1]. However, CSIT for frequency division duplex (FDD) massive MIMO can be more challenging, since single-antenna users have to acquire and feedback the high-dimensional channels to the BS [2]–[8].

To date, there have been several CSIT estimation schemes proposed for FDD massive MIMO to facilitate its backward compatibility with current cellular networks dominated by FD-D [2]–[8]. Specifically, [2] proposed an open-loop and closed-loop downlink channel estimation scheme for FDD massive MIMO, but the long-term channel statistics known at the user is required and may be difficult in practice. [3]–[5] proposed the compressive sensing (CS)-based downlink channel estimation by assuming the delay-domain sparsity of massive MIMO channels, but such assumption may not hold in indoor scenarios due to rich scatterers at the user side. [6]–[8] proposed the CS-based CSIT estimation schemes by assuming the sparsity of angle-domain massive MIMO channels. However, [6], [7] are limited to narrow-band systems without considering practical broad-band systems, while [7], [8] only consider the signal-user CSIT estimation and fail to exploit the channel correlation of multiple adjacent users. Furthermore, existing schemes [2]–[8] only consider the single-cell scenario, and they may suffer from downlink pilot contamination due to inter-cell-interfere (ICI).

The author is with School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (E-mail: gaozhen010375@foxmail.com).

In this letter, we consider the practical multi-cell FDD massive MIMO systems. In such scenario, users in target cell will receive the downlink pilot from adjacent cells, which will contaminate the downlink channel estimation of the target cell and thus degrade the system performance. This phenomenon is termed as the *downlink pilot contamination of multi-cell FDD massive MIMO*, while conventional CSI acquisition schemes either only consider ICI in TDD massive MIMO or fail to consider ICI in FDD massive MIMO. To this end, we propose a CS-based CSIT estimation scheme to alleviate the pilot contamination in multi-cell FDD massive MIMO systems. Particularly, we observe that angle-domain massive MIMO channels exhibit the common sparsity over different subcarriers due to the limited number of scatterers seen from the BS and the very similar scatterers experienced by different subcarriers. Moreover, such sparsity is partially shared by adjacent users due to some common scatterers. By jointly exploiting these sparsity properties of massive MIMO channels in the angular domain, under the framework of CS theory, we design the pilot signal and CS-based channel estimator for multi-cell FDD massive MIMO. The proposed scheme can reliably acquire the channels associated with multiple adjacent BSs with low training overhead for downlink pilot decontamination. Simulation results confirm that the proposed solution outperforms existing schemes in multi-cell FDD massive MIMO systems.

Notation: the boldface lower and upper-case symbols denote column vectors and matrices, respectively. The Moore-Penrose inversion, transpose, and conjugate transpose operators are given by $(\cdot)^\dagger$, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^*$, respectively. $|\Gamma|_c$ is the cardinality of the set $\Gamma$. $\mathrm{E}\{\cdot\}$ is the expectation operator. $(\mathbf{a})_\Gamma$ denotes the entries of $\mathbf{a}$ whose indices are defined by $\Gamma$, while $(\mathbf{A})_{:,k}$ denotes the $k$th column of the matrix $\mathbf{A}$. $[\mathbf{a}]_i$ denotes the $i$th entry of the vector $\mathbf{a}$, and $[\mathbf{A}]_{i,j}$ denotes the $i$th-row and $j$th-column element of the matrix $\mathbf{A}$. Finally, $\Omega^c$ is the complementary set of $\Omega$.

## II. SYSTEM MODEL

We consider a multi-cell FDD massive MIMO system composed of $L$ hexagonal cells, and each cell consists of a central $M$-antenna BS and $N$ single-antenna users with $N \ll M$ [1]. For the $k$th user in the $\tilde{l}$th cell, the received downlink signal of the $p$th subcarrier can be expressed as

$$y_{k,\tilde{l},p}=\mathbf{x}_{\tilde{l},p}^{\mathrm{T}}\mathbf{h}_{k,\tilde{l},p}+\sum\nolimits_{l=0,l\neq\tilde{l}}^{L-1}\mathbf{x}_{l,p}^{\mathrm{T}}\mathbf{h}_{k,l,p}+v_{k,\tilde{l},p}, 1\leq p\leq P, \quad (1)$$

where $\mathbf{h}_{k,l,p}\in\mathbb{C}^{M\times 1}$ denotes the downlink channel of the $p$th subcarrier between the $k$th user and the $l$th BS, $\mathbf{x}_{l,p}\in\mathbb{C}^{M\times 1}$ is the transmitted signal from the $l$th BS, $v_{k,\tilde{l},p}$ is additive white Gaussian noise (AWGN), and $P$ is the size of one OFDM symbol. From (1), it can be observed that the reliable estimation of $\mathbf{h}_{k,\tilde{l},p}$ is challenging due to two following reasons. First, the estimation of $M$-dimensional $\mathbf{h}_{k,\tilde{l},p}$ will lead to the prohibitively

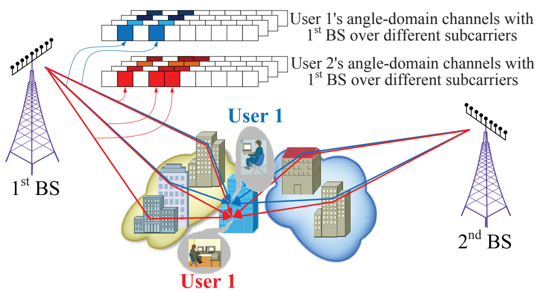


Fig. 1. Illustration of the angle-domain sparsity of massive MIMO channels.

high training overhead. Second, the $k$th user of the $\tilde{l}$th cell suffers from ICI, i.e., $\sum_{l=0,l\neq\tilde{l}}^{L-1}\mathbf{x}_{l,p}^{\mathrm{T}}\mathbf{h}_{k,l,p}$.

For massive MIMO systems as shown in Fig. 1, the BS is usually elevated high with few scatterers around, while users are located at low elevation with relatively rich local scatterers, which leads the multipath components of the channels associated with one user to concentrate on the limited angle seen from the BS side [6]–[8]. Based on this phenomenon, [6]–[8] assume that the angle-domain massive MIMO channel vectors $\tilde{\mathbf{h}}_{k,l,p}=\mathbf{F}^{*}\mathbf{h}_{k,l,p}$ exhibit the sparsity, where $\mathbf{F}\in\mathbb{C}^{M\times M}$ is the unitary matrix representing the transformation matrix of the angular domain at the BS side. Such sparsity indicates that only a small part of elements of the angle-domain channel vector $\tilde{\mathbf{h}}_{k,l,p}$ contain almost all the multipath components between the $l$th BS and the $k$th user, i.e., $|\Omega_{k,l,p}|_c \ll M$, where

$$\Omega_{k,l,p}=\mathrm{supp}\{\tilde{\mathbf{h}}_{k,l,p}\}=\left\{m:\left\|[\tilde{\mathbf{h}}_{k,l,p}]_m\right\|_2>p_{\mathrm{th}},1\leq m\leq M\right\}, \quad (2)$$

and $p_{\mathrm{th}}$ is a threshold according to AWGN [4]. Moreover, since channels of different subcarriers experience the very similar scatterers, they share the same sparsity pattern [3], [8], i.e.,

$$\Omega_{k,l,1}=\Omega_{k,l,2}=\cdots=\Omega_{k,l,P}=\Omega_{k,l}. \quad (3)$$

Additionally, for a group of $K$ users physically close to each other as illustrated in Fig. 1, their angle-domain channels share the partially common sparsity [6], which can be expressed as

$$\bigcap_{k=1}^{K}\Omega_{k,l}=\Omega_c\neq\phi. \quad (4)$$

It should be pointed out that $N$ users served by the BS using the same time-frequency resource usually come from different user groups for the improved performance [1].

## III. Proposed CS-Based CSIT Estimation Scheme

The proposed scheme includes the CS-based design of downlink multi-cell pilot and channel estimation algorithm, and both of them are significant for downlink pilot decontamination. By leveraging the angle-domain sparsity of massive MIMO channels, the proposed scheme can jointly acquire the channels of multiple adjacent BSs with low training overhead, which can mitigate the downlink pilot contamination.

### *A. Pilot Training for CSIT Estimation in Multi-Cell Scenario*

In the proposed scheme, each BS transmits the off-line designed downlink pilot for CSIT estimation, and the received downlink pilot signal at users can be fed back to their respective BSs via the uplink feedback channels. Here the uplink feedback channels are assumed to be AWGN channels after the uplink channel estimation and equalization [6]. For the $k$th user of the central target cell ($l=0$) in the $t$th time slot, the received pilot signal fed back to the BS can be expressed as

$$\begin{aligned} r_{k,p}^{t} &= \sum_{l=0}^{L-1}(\mathbf{s}_{l,p}^{t})^{\mathrm{T}}\mathbf{h}_{k,l,p}+w_{k,p}^{t}=\sum_{l=0}^{L-1}(\mathbf{s}_{l,p}^{t})^{\mathrm{T}}\mathbf{h}_{k,l,p}\delta\left(\rho_{k,l}>\rho_{\mathrm{th}}\right) \\ &\quad +\sum_{l=0}^{L-1}(\mathbf{s}_{l,p}^{t})^{\mathrm{T}}\mathbf{h}_{k,l,p}\delta\left(\rho_{k,l}\leq\rho_{\mathrm{th}}\right)+w_{k,p}^{t} \\ &= \sum_{l=0}^{L-1}(\mathbf{s}_{l,p}^{t})^{\mathrm{T}}\mathbf{h}_{k,l,p}\delta\left(\rho_{k,l}>\rho_{\mathrm{th}}\right)+\tilde{w}_{k,p}^{t}, \end{aligned} \quad (5)$$

where $\delta(\cdot)$ is Dirac delta function, $\mathbf{s}_{l,p}^{t}$ is the downlink pilot of the $l$th cell in the $t$th time slot, $\rho_{\mathrm{th}}$ is a predefined signal-to-noise-ratio (SNR) threshold, $\rho_{k,l}$ is the $k$th user's SNR associated with the $l$th BS, $w_{k,p}^{t}$ is the effective noise including the downlink channel and uplink feedback channel [6], and $\tilde{w}_{k,p}^{t}=\sum_{l=0}^{L-1}(\mathbf{s}_{l,p}^{t})^{\mathrm{T}}\mathbf{h}_{k,l,p}\delta\left(\rho_{k,l}\leq\rho_{\mathrm{th}}\right)+w_{k,p}^{t}$.

Due to the angle-domain sparsity of massive MIMO channel vectors as discussed in Section II, (5) can be rewritten as

$$\begin{aligned} r_{k,p}^{t} &= \sum_{l\in\Pi_k}(\mathbf{s}_{l,p}^{t})^{\mathrm{T}}\mathbf{F}\tilde{\mathbf{h}}_{k,l,p}+\tilde{w}_{k,p}^{t}=\sum_{l\in\Pi_k}\boldsymbol{\phi}_{l,p}^{t}\tilde{\mathbf{h}}_{k,l,p}+\tilde{w}_{k,p}^{t} \\ &= \boldsymbol{\theta}_{k,p}^{t}\bar{\tilde{\mathbf{h}}}_{k,p}+\tilde{w}_{k,p}^{t}, \end{aligned} \quad (6)$$

where

$$\begin{cases} \Pi_k=\{l:\rho_{k,l}>\rho_{\mathrm{th}},0\leq l\leq L-1\}, \\ \boldsymbol{\phi}_{l,p}^{t}=(\mathbf{s}_{l,p}^{t})^{\mathrm{T}}\mathbf{F}\in\mathbb{C}^{1\times M}, \\ \boldsymbol{\theta}_{k,p}^{t}=[\boldsymbol{\phi}_{\Pi_k(1),p}^{t},\boldsymbol{\phi}_{\Pi_k(2),p}^{t},\cdots,\boldsymbol{\phi}_{\Pi_k(|\Pi_k|_c),p}^{t}]\in\mathbb{C}^{1\times M|\Pi_k|_c}, \\ \bar{\tilde{\mathbf{h}}}_{k,p}=[\tilde{\mathbf{h}}_{k,\Pi_k(1),p}^{\mathrm{T}},\tilde{\mathbf{h}}_{k,\Pi_k(2),p}^{\mathrm{T}},\cdots,\tilde{\mathbf{h}}_{k,\Pi_k(|\Pi_k|_c),p}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{C}^{M|\Pi_k|_c\times 1}, \end{cases} \quad (7)$$

$\Pi_k(i)$ denotes the $i$th element of the set $\Pi_k$, which can be acquired by comparing the received SNRs associated with different BSs and $\rho_{\mathrm{th}}$ at the $k$th user, and then fed them back to BSs for the following CSIT estimation.

Moreover, due to the temporal channel correlation, the channel $\mathbf{h}_{k,l,p}$ is considered to be unchanged in $G$ successive OFDM symbols within the channel coherence time [4]. By jointly collecting the feedback pilots in $G$ successive OFDM symbols, we can obtain the aggregate feedback signal

$$\mathbf{r}_{k,p}^{[G]}=\boldsymbol{\Theta}_{k,p}^{[G]}\bar{\tilde{\mathbf{h}}}_{k,p}+\tilde{\mathbf{w}}_{k,p}^{[G]}, \quad (8)$$

where we have

$$\begin{cases} \mathbf{r}_{k,p}^{[G]}=[(r_{k,p}^{1})^{\mathrm{T}},(r_{k,p}^{2})^{\mathrm{T}},\cdots,(r_{k,p}^{G})^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{C}^{G\times 1}, \\ \boldsymbol{\Theta}_{k,p}^{[G]}=[(\boldsymbol{\theta}_{k,p}^{1})^{\mathrm{T}},(\boldsymbol{\theta}_{k,p}^{2})^{\mathrm{T}},\cdots,(\boldsymbol{\theta}_{k,p}^{G})^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{C}^{G\times M|\Pi_k|_c}, \\ \tilde{\mathbf{w}}_{k,p}^{[G]}=[\tilde{w}_{k,p}^{1},\tilde{w}_{k,p}^{2},\cdots,\tilde{w}_{k,p}^{G}]^{\mathrm{T}}\in\mathbb{C}^{G\times 1}. \end{cases} \quad (9)$$

### *B. CS-Based CSIT Estimation Algorithm*

To reliably acquire the channel vector $\bar{\tilde{\mathbf{h}}}_{k,p}$ from (8), the training overhead $G$ required by conventional algorithms, e.g., the least squares (LS) algorithm, is usually proportional to $M|\Pi_k|_c$, the dimension of $\bar{\tilde{\mathbf{h}}}_{k,p}$. Usually, $G\geq M|\Pi_k|_c$ is required, which leads $G$ to be much larger than the channel coherence time, and otherwise results in the poor channel estimation performance [6].

Fortunately, the angle-domain sparsity of massive MIMO channel $\tilde{\mathbf{h}}_{k,l,p}$ implies that the aggregate angle-domain channel $\bar{\tilde{\mathbf{h}}}_{k,p}$ also has the sparsity according to (7), which motivates us to leverage the CS theory to estimate high-dimensional $\bar{\tilde{\mathbf{h}}}_{k,p}$ from low-dimensional $\mathbf{r}_{k,p}^{[G]}$ in (8). Moreover, the common sparsity shared by $\{\bar{\bar{\mathbf{h}}}_{k,p}\}_{p=1}^{P}$ for the $k$th user and the partially common sparsity shared by $\{\bar{\bar{\mathbf{h}}}_{k,p}\}_{k=1}^{K}$ for $K$ users in the same group can be leveraged for the further improved performance. Specifically, we consider the partially common support shared by the $K$ users physically close to each other, i.e.,

$$\mathbf{R}_p^{[G]} = \mathbf{\Theta}_p^{[G]}\bar{\bar{\mathbf{H}}}_p + \tilde{\mathbf{W}}_p^{[G]}, 1 \leq p \leq P, \tag{10}$$

where we have

$$\begin{cases} \mathbf{R}_p^{[G]} = \left[\mathbf{r}_{1,p}^{[G]}, \mathbf{r}_{2,p}^{[G]}, \cdots, \mathbf{r}_{K,p}^{[G]}\right] \in \mathbb{C}^{G\times K}, \\ \Pi_1 = \Pi_2 = \cdots = \Pi_K = \Pi, \\ \mathbf{\Theta}_{1,p}^{[G]} = \mathbf{\Theta}_{2,p}^{[G]} = \cdots = \mathbf{\Theta}_{K,p}^{[G]} = \mathbf{\Theta}_p^{[G]} \in \mathbb{C}^{G\times M|\Pi|_c}, \\ \bar{\bar{\mathbf{H}}}_p = \left[\bar{\bar{\mathbf{h}}}_{1,p}, \bar{\bar{\mathbf{h}}}_{2,p}, \cdots, \bar{\bar{\mathbf{h}}}_{K,p}\right] \in \mathbb{C}^{M|\Pi|_c\times K}, \\ \tilde{\mathbf{W}}_p^{[G]} = \left[\tilde{\mathbf{w}}_{1,p}^{[G]}, \tilde{\mathbf{w}}_{2,p}^{[G]}, \cdots, \tilde{\mathbf{w}}_{K,p}^{[G]}\right]^{\mathrm{T}} \in \mathbb{C}^{G\times K}. \end{cases} \tag{11}$$

Note that since $K$ users in the same group are physically close to each other and their received signals from the same BS experience very similar large-scale fading, we can approximately obtain $\rho_{l,1} = \rho_{l,2} = \cdots = \rho_{l,K}$, and thus the second and third equations in (11) hold.

Given the measurements (10) and the sparse constraints (3) and (4), the CSI matrix $\{\bar{\bar{\mathbf{H}}}_p\}_{p=1}^{P}$ can be acquired by solving the following optimization problem

$$\begin{aligned} &\min_{\bar{\bar{\mathbf{H}}}_p, 1\leq p\leq P} \textstyle\sum_{p=1}^{P} \left\|\bar{\bar{\mathbf{H}}}_p\right\|_{0,2} = \min_{\bar{\bar{\mathbf{H}}}_p, 1\leq p\leq P} \textstyle\sum_{p=1}^{P} \left(\sum_{k=1}^{K}\left\|\bar{\bar{\mathbf{h}}}_{k,p}\right\|_0^2\right)^{1/2} \\ &\quad \text{s.t. } \mathbf{R}_p^{[G]} = \mathbf{\Theta}_p^{[G]}\bar{\bar{\mathbf{H}}}_p, \Omega_{k,l,p} = \Omega_{k,l}, \forall p, \bigcap_{k=1}^{K}\Omega_{k,l} \neq \phi. \end{aligned} \tag{12}$$

To solve the optimization problem (12), developed from the classical CS algorithm orthogonal matching pursuit (OMP), as shown in **Algorithm 1**, we propose a joint multi-user multi-carrier orthogonal matching pursuit (J-MUMC-OMP) algorithm. Specifically, lines 1-3 initialize the variables; lines 6 identifies the most possible angle-domain element by leveraging the sparsity constraints (3) and (4); lines 7-8 estimate the elements according to updated support set; lines 9-10 imply that if all $K$ users' $\rho$th angle-domain elements are dominated by AWGN, the iteration stops since the channel sparsity level is over-estimated; while in lines 11-12, for users whose $\rho$th angle-domain elements are dominated by AWGN, we delete the index $\rho$ and re-estimate the associated elements; lines 14-15 update the residue; line 16 indicates that if the residue of the current iteration is larger than that of the last iteration, stopping the iteration can help the algorithm to acquire the good mean square error (MSE) performance.

The proposed J-MUMC-OMP algorithm has several distinctive features as follows. First, the proposed J-MUMC-OMP algorithm can jointly estimate the sparse signals $\{\bar{\bar{\mathbf{h}}}_{k,p}\}_{p=1,k=1}^{P,K}$, $\forall k$ by exploiting their common sparsity over different subcarriers. Second, the partially common sparsity of $K$ users' sparse channels $\{\bar{\bar{\mathbf{h}}}_{k,p}\}_{k=1}^{K}$, $\forall p$ is also considered for the further improved performance. Third, we provide the stopping criteria to adaptively acquire the sparsity level of channels. By contrast, the classical orthogonal matching pursuit (OMP) algorithm requires the sparsity level without considering these sparsity properties, while the joint-OMP algorithm proposed in [6] fails to leverage the common sparsity over different subcarriers.

**Algorithm 1** Proposed J-MUMC-OMP Algorithm.

**Input:** Noisy measurement matrix $\mathbf{R}_p^{[G]}$, sensing matrix $\mathbf{\Theta}_p^{[G]}$, $\forall p$, and the termination threshold $\gamma_{\mathrm{th}}$.
**Output:** The estimation of channel matrix $\hat{\bar{\bar{\mathbf{H}}}}_p$, $\forall p$.
1: $i = 0$; {Initialize the iteration index $i$}
2: $\{\Omega_k^i\}_{k=1}^{K} = \phi$; {Initialize the support sets of $K$ users' aggregate channel vectors}
3: $\mathbf{Z}_p^i = \mathbf{R}_p^{[G]}$; {Initialize the residue}
4: **repeat**
5: $\quad i = i + 1$;
6: $\quad \rho = \arg\max_{\tilde{\rho}} \left\{\sum_{p=1}^{P}\sum_{k=1}^{K}\left\|\left[\left(\mathbf{\Theta}_p^{[G]}\right)^*\left[\mathbf{Z}_p^{i-1}\right]_{:,k}\right]_{\tilde{\rho}}\right\|_2^2\right\}$;
7: $\quad \Omega_k^i = \Omega_k^{i-1} \cup \rho, \forall k$;
8: $\quad (\mathbf{g}_{k,p})_{\Omega_k^i} = (\mathbf{\Theta}_p^{[G]})_{\Omega_k^i}^{\dagger}\left[\mathbf{R}_p^{[G]}\right]_{:,k}$, $(\mathbf{g}_{k,p})_{(\Omega_k^i)^c} = \mathbf{0}$, $\forall k, p$;
9: $\quad$ **if** $\sum_{p=1}^{P}\left\|[\mathbf{g}_{k,p}]_{\rho}\right\|_2^2/P < \gamma_{\mathrm{th}}$, $\forall k$ **then**
10: $\quad\quad$ Quit iteration;
11: $\quad$ **else if** there exists $k$ meeting $\sum_{p=1}^{P}\left\|[\mathbf{g}_{k,p}]_{\rho}\right\|_2^2/P < \gamma_{\mathrm{th}}$ **then**
12: $\quad\quad \Omega_k^i = \Omega_k^{i-1}$, $(\mathbf{g}_{k,p})_{\Omega_k^i} = (\mathbf{\Theta}_p^{[G]})_{\Omega_k^i}^{\dagger}\left[\mathbf{R}_p^{[G]}\right]_{:,k}$, $(\mathbf{g}_{k,p})_{(\Omega_k^i)^c} = \mathbf{0}$, $\forall p$; for $k$ satisfy the above condition;
13: $\quad$ **end if**
14: $\quad \mathbf{G}_p^i = [\mathbf{g}_{1,p}, \mathbf{g}_{2,p}, \cdots, \mathbf{g}_{K,p}]$, $\forall p$;
15: $\quad \mathbf{Z}_p^i = \mathbf{R}_p^{[G]} - \mathbf{\Theta}_p^{[G]}\mathbf{G}_p^i$, $\forall p$;
16: **until** $\sum_{p=1}^{P}\left\|\mathbf{Z}_p^i\right\|_F \geq \sum_{p=1}^{P}\left\|\mathbf{Z}_p^{i-1}\right\|_F$;
17: $\hat{\bar{\bar{\mathbf{H}}}}_p = \mathbf{G}_p^{i-1}$, $\forall p$;

### C. CS-Based Downlink Pilot Design in Multi-Cell Scenario

The design of measurement matrices $\mathbf{\Theta}_p^{[G]}$ for different $p$'s in (12) are important to ensure the reliable channel estimation in CS theory. Owing to $\mathbf{\Theta}_p^{[G]} = [(\boldsymbol{\theta}_p^1)^{\mathrm{T}}, (\boldsymbol{\theta}_p^2)^{\mathrm{T}}, \cdots, (\boldsymbol{\theta}_p^G)^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{\theta}_p^t = [\boldsymbol{\phi}_{\Pi(1),p}^t, \boldsymbol{\phi}_{\Pi(2),p}^t, \cdots, \boldsymbol{\phi}_{\Pi(|\Pi|_c),p}^t]$, and $\boldsymbol{\phi}_{l,p}^t = (\mathbf{s}_{l,p}^t)^{\mathrm{T}}\mathbf{F}$, we observe that $\mathbf{\Theta}_p^{[G]}$, $\forall p$ are only determined by the pilot signals $\{\mathbf{s}_{l,p}^t\}_{l=0,p=1,t=1}^{L-1,P,G}$.

According to [9], a measurement matrix whose elements follow an independent identically distributed (i.i.d.) Gaussian distribution can achieve the good performance for sparse signal recovery. Furthermore, diverifying measurement matrices $\mathbf{\Theta}_p^{[G]}$, $\forall p$ can further improve the recovery performance of sparse signals when multiple sparse signals with the (partially) common sparsity are jointly recovered [9]. Specifically, we consider each element of pilot signals can be off-line designed as

$$\left[\mathbf{s}_{l,p}^t\right]_m = e^{\mathrm{j}\theta_{m,l,p,t}}, 1 \leq m \leq M, 1 \leq t \leq G, 1 \leq p \leq P, 0 \leq l \leq L-1, \tag{13}$$

where $\theta_{m,l,p,t}$ follows the i.i.d. uniform distribution in $[0,\ 2\pi)$. It is straightforward to prove that the designed pilot signals (13) can ensure elements of $\mathbf{\Theta}_p^{[G]}$, $\forall p$, to obey the i.i.d. complex Gaussian distribution with zero mean and unit variance. Hence, the proposed pilot signal design is optimal for the reliable compression and recovery of sparse angle-domain channels under the framework of CS theory.

### D. Multi-Cell Joint Precoding

In Section III-A, B, and C, we can use the low training overhead to estimate CSIT, which can be leveraged to perform multi-cell joint precoding to combat ICI. Specifically, we consider: 1) each BS uses zero forcing (ZF) precoding to serve multiple users; 2) multiple users served by the BS

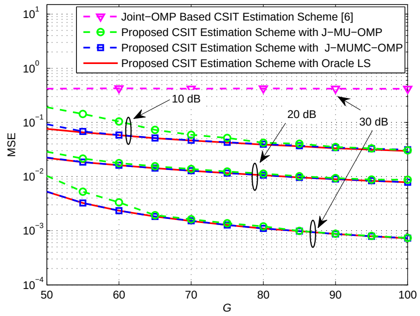


Fig. 2. Comparison of channel estimation MSE performance of different CSIT estimation solutions versus $G$ at different $\rho_{\rm edge}$'s.

using the same time-frequency resource should come from different user groups to reduce the correlation of different users' channel vectors and enhance the system capacity; 3) each user is jointly served by multiple adjacent BSs according to the channel quality. The multi-cell joint precoding can be integrated with the emerging cloud radio access network (C-RAN), where BS can be considered as the remote radio header (RRH) and a baseband unit (BBU) can be used to perform multi-cell joint precoding with centralized processing.

## IV. Simulation Results

In this section, we investigate the performance of the proposed CS-based CSIT estimation scheme for downlink pilot decontamination in multi-cell FDD massive MIMO. In simulations, we consider $L = 7$ hexagonal cells, each BS has $M = 128$ antennas to simultaneously serve $N$ users, the carrier frequency is $f_c = 2$ GHz, the system bandwidth is $f_s = 10$ MHz, the maximum delay spread is $\tau_{\max} = 5$ $\mu$s for typical urban scenario [8], $P = f_s\tau_{\max} = 50$, the cell radius is 1 km, and the distance-based path loss between the $l$th BS and the $k$th user is $\beta_{\rm PL} = 1/(d^\alpha)$, where $d$ is the geographical distance between the $l$th BS and the $k$th user, and the path loss exponent $\alpha$ is 3.8 dB/km. Moreover, cell-edge users are considered in simulations since they suffer from the most severe downlink pilot contamination. Specifically, we consider $K$ adjacent users as a group, and they are randomly distributed at the cell-edge of the central target BS with the geographical distance of 1 km. For the proposed J-MUMC-OMP algorithm, $\rho_{\rm th}$ is set as 3, 5, 10, 10, and 10, $\gamma_{\rm th}$ is set as 0.006, 0.004, 0.002, 0.0004 and 0.0003 for $\rho_{\rm edge} = 10$ dB, 15 dB, 20 dB, 25 dB, and 30 dB, respectively, where $\rho_{\rm edge}$ is the cell-edge SNR associated with the central target BS. The joint-OMP based CSIT estimation scheme [6] only considering the single-cell scenario is provided for comparison. Besides, we also provide the so-called J-MU-OMP algorithm, which is a special case of the proposed J-MUMC-OMP algorithm when only the partially common sparsity among different users is considered.

Fig. 2 compares the channel estimation MSE performance of different CIST estimation schemes, where $K = 10$, $|\Omega_{k,l}|_c = 6$, $\{\Omega_{k,l}\}_{k=1}^{K/2} = \Omega_l^1$, $\{\Omega_{k,l}\}_{k=K/2+1}^{K} = \Omega_l^2$, and $|\Omega_l^1 \cap \Omega_l^2|_c = 4$, $\forall l$ were considered. The oracle LS estimator with the known $\{\Omega_{k,l}\}_{k=1}^{K}$ for $l \in \Pi$ was adopted as the performance bound. From Fig. 2, it can be observed that the joint-OMP based CSIT scheme [6] suffers from downlink pilot contamination and works poorly. In contrast, J-MUMC-OMP and J-MU-OMP algorithms can effectively solve this issue thanks to the CS-based pilot design and CSIT estimation algorithm in multi-cell scenario. Especially, compared with J-MU-OMP algorithm, the proposed J-MUMC-OMP is capable of approaching the oracle LS performance bound when $G \geq 55$, since the common sparsity of angle-domain massive MIMO channels over different subcarriers is also leveraged for further improved performance.

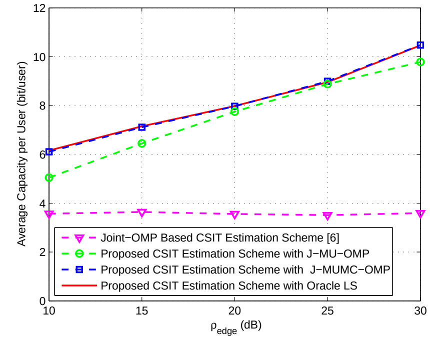


Fig. 3. Comparison of downlink average throughput per user with multi-cell joint ZF precoding when $G = 55$.

Fig. 3 compares the downlink average throughput per user (bit/user) with different CSIT estimation schemes, where ZF precoding is used with $N = 24$, and each user is jointly served by three best BSs according to their channel quality. It can be observed that the proposed J-MUMC-OMP based CSIT estimation scheme outperforms its counterparts, and its average throughput per user is capable of approaching that of the performance bound achieved by the oracle LS estimator.

## V. Conclusions

In this paper, we have proposed the CS-based CSIT estimation scheme for downlink pilot decontamination in multi-cell FDD massive MIMO systems, while existing schemes only consider the single-cell scenario and suffer from ICI. We have exploited the common sparsity of angle-domain massive MIMO channels over different subcarriers and the partially common sparsity shared by adjacent users. By exploiting these sparsity features, we design the pilot signal and channel estimation algorithm under the CS framework. The proposed scheme can reliably estimate multiple adjacent BSs' channels for downlink pilot decontamination. Simulation results confirm that the proposed solution outperforms existing schemes in multi-cell FDD massive MIMO with low training overhead.

## References


[1] E. Bjornson, 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.

[2] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 802–814, Oct. 2014.

[3] J. W. Choi, B. Shim, and Seok-Ho Chang, "Downlink pilot reduction for massive MIMO systems via compressed sensing," *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 1889-1892, Nov. 2015.

[4] W. Shen, L. Dai, Y. Shi, B. Shim, and Z. Wang, "Joint channel training and feedback for FDD massive MIMO systems," to appear in *IEEE Trans. Veh. Techn.*

[5] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 104-118, Jan. 2015.

[6] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, Jun. 2014.

[7] V. K. N. Lau, S. Cai, and A. Liu, "Closed-loop compressive CSIT estimation in FDD massive MIMO systems with 1 bit feedback,"*IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2146–2155, Apr. 2016.

[8] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169-6183, Dec. 2015.

[9] M. Duarte, S. Sarvotham, D. Baron, M. Wakin, and R. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, 2005, pp. 1537-1541.