IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 24, 2016, accepted August 15, 2016, date of publication August 30, 2016, date of current version June 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2016.2604202

Two-Stage Time-Domain Pilot Contamination Elimination in Large-Scale Multiple-Antenna Aided and TDD Based OFDM Systems

XINYING GUO[1,2,3], JIANKANG ZHANG[1,2,3], (Member, IEEE), SHENG CHEN[3,4], (Fellow, IEEE), XIAOMIN MU[1], AND LAJOS HANZO[3,4], (Fellow, IEEE)

[1]School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China
[2]National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
[3]Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K.
[4]King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: Lajos Hanzo (lh@ecs.soton.ac.uk)

This work was supported in part by the National Natural Science Foundation of China under Grant 61271421, Grant 61301150, and Grant 61571401, in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University, China, under Grant 2016D02, in part by the European Research Council's Advanced Fellow Grant Beam-me-up, and in part by EPSRC's under Grant EP/Noo4558/1. The research data for this paper is available at http://doi.org/10.5258/SOTON/404960.

ABSTRACT Pilot contamination (PC) is a major impediment of large-scale multi-cell multiple-input multiple-output systems. Hence, we propose an optimal pilot design for time-domain channel estimation, which is capable of completely eliminating PC. More specifically, a sophisticated combination of downlink training and "scheduled" uplink training is designed with the aid of the optimal pilot set. Given the optimal pilot set, every user acquires its unique downlink time-domain channel state information (CSI) through downlink training. The estimated downlink CSIs are then embedded in the uplink training. As a result, PC can be completely eliminated, at the cost of a slight increase in training computational complexity. Our simulation results demonstrate the power of the proposed scheme. Most significantly, our scheme imposes a modest training overhead of $(L + 3)$, training-phase durations corresponding to the number of orthogonal frequency division multiplexing symbols, where L is the number of cells, which is substantially lower than that imposed by some of the existing PC elimination schemes. Therefore, it imposes a less stringent requirement on the channel's coherence time. Finally, our scheme does not need any information exchange between base stations.

INDEX TERMS Multi-cell systems, large-scale multiple-input multiple-output, orthogonal frequency division multiplexing, pilot contamination, time division duplexing, time-domain channel estimation.

I. INTRODUCTION

To meet the demand for ever-higher spectral efficiency and energy efficiency in cellular systems, large-scale or massive multiple-input multiple-output (MIMO) techniques are capable of offering dramatic data rate improvements and power savings [1]–[6]. In a massive MIMO system, each base station (BS) is equipped with a large number of antennas with the objective of serving a number of single-antenna mobile stations (MSs) simultaneously using the same time/frequency resource. Owing to the asymptotic orthogonality of the different users' channel vectors, low-complexity downlink (DL) transmit precoding and beamforming-aided uplink (UL) reception are capable of achieving both a high capacity and a high link reliability [1]. It is then worth emphasizing that the performance of a large-scale MIMO system critically relies on the accuracy of the channel state information (CSI) estimation. In frequency-division duplexing based systems, the DL CSI must be estimated by the MS receivers and signalled back to the BS in order to carry out transmit precoding based DL transmission. Therefore, the CSI signaling overhead scales linearly with the number of DL transmit antennas employed at the BS, which renders the use of large antenna arrays impractical. Hence, most massive MIMO systems rely on the time-division duplexing (TDD) protocol in order to exploit the reciprocity of the UL and DL channels. The overhead imposed by acquiring the CSI at a BS with the aid of UL training only scales linearly with the number of MSs, which is typically much smaller than the number of BS's antennas.

However, the number of available orthogonal pilot sequences that can be created is limited by the channel's coherence time. Therefore, the pilots have to be reused in multi-cell systems. As a result, the channel estimation (CE) at a BS is seriously contaminated by the MSs in other cells. This pilot contamination (PC) severely limits the achievable sum-rate in large-scale MIMO systems [1], [2], [7], [8].

Consequently, significant research efforts have been dedicated to this PC problem, and numerous solutions have been proposed for eliminating or reducing the PC effects, which can be divided into two categories. The first category of solutions focus on pilot design and pilot allocation. Based on maximizing the signal-to-interference ratio (SIR), a pilot design method was introduced in [9]. In [10], the authors proposed a cell sectorization-based pilot assignment scheme for mitigating PC. The authors of [11] formed a sectorized multi-cell massive MIMO system based on a sectorization method for increasing the achievable UL sum-rate. However, the optimal design of pilot reuse patterns for the sectors of each cell was not addressed in [11]. A soft pilot reuse and multi-cell block diagonalization precoding scheme was proposed in [12] for partitioning the users into cell-center and cell-edge groups. However, this class of solutions cannot completely eliminate the inter-cell interference caused by PC, since the re-use of the same group of pilots is still required.

The second category of solutions rely on sophisticated signal processing algorithms. A method using time-staggered pilots was proposed in [13], which eliminates PC by ensuring that the pilot transmissions of the different users exploiting the same pilot do not overlap in time. However, this scheme is costly and difficult to implement, because a central controller must manage the complex staggering of pilots in all cells by relying on tight coordination amongst the BSs. Secondly, the training duration is dramatically increased. Yin *et al.* [14] proposed a Bayesian estimator, which requires the second-order statistics of the channel coefficients at the BS. To further mitigate the inter-cell interference, a sophisticated pilot assignment scheme was also proposed in [14], which requires cooperation amongst the BSs. Although this smart non-linear signal processing algorithm is capable of significantly reducing the PC, its complexity is excessive, since it requires the channel covariance matrices of all the UL channels. Note that the acquisition of such a large amount of second-order statistics at the BSs is extremely time-consuming. Furthermore, sharing them among the BSs requires a huge amount of backhaul transmissions. Recently, a location-aware CE scheme was proposed by utilizing a fast Fourier transform (FFT) based post-processing after the conventional pilot aided CE [15], which is capable of distinguishing the users having different angles of arrival (AOA), even if they use the same pilot. To maximally benefit from this location-aware CE algorithm, a location-aware pilot assignment scheme was also proposed in [15] for ensuring that the non-overlapping AOAs of the users of different cells are allowed to adopt the same pilot. The advantage of this method is its appealing simplicity and efficiency. Moreover, it does not increase the training duration imposed, but it requires the knowledge of the users' AOAs.

More related to our current work, there exist two PC eliminating schemes based on sophisticated processing [16], [17]. The scheme of [16] consisted of an amalgam of DL and UL training phases. More specifically, a total of $(L + 3)$ training phases are required for an L-cell system. This scheme is capable of completely eliminating the PC at the expense of requiring a long channel coherence time. In particular, let U be the number of users per cell, which is assumed to be the length of the training sequences. Then, the channel's coherence interval (COHI) is required to be no shorter than the duration of $(L + 3)U$ orthogonal frequency division multiplexing (OFDM) symbols. In addition to a conventional simultaneous UL training phase, the scheme proposed in [17] consisted of consecutive pilot transmission phases in which each BS stays idle during one phase and repeatedly transmits pilot sequence in all the other phases. Thus, this scheme consists of a total of $(L+1)$ UL training phases and, therefore, it requires that the COHI is no shorter than the duration of $(L + 1)U$ OFDM symbols. These two schemes also increase the training computational complexity slightly.

Against the above background, an effective PC elimination scheme is proposed for multi-cell TDD based OFDM systems. Unlike [1], which assumes that the channel frequency response (CFR) is constant over certain frequency smoothness interval, we consider the generic case of varying CFR where frequency-domain (FD) pilot symbol (PS) have to occupy all the subcarriers. We also adopt a very different approach for analysing PC. Specifically, we approach the PC problem by considering the received signals of all the OFDM subcarriers together for each individual BS antenna, which constitutes an effective technique of performing CE. This is fundamentally different from all the existing schemes [13]–[17], which approach the PC problem by considering the received signals of all the target BS's antennas for each individual OFDM subcarrier. While this per-subcarrier processing is natural technique of performing detection or precoding by exploiting the orthogonality of the subcarriers, ironically, it is actually an ineffective approach for dealing with CE. Furthermore, we perform CE in time-domain (TD), and obtain the required FD channel transfer functions (FDCHTFs) with the aid of the FFT operation. Our contribution is twofold. Firstly, an optimal pilot set is designed, which can provide the desired orthogonality for differentiating all the channel impulse responses (CIRs) in a conventional simultaneous UL training phase under certain operating conditions. However, under more hostile environments, where the number of users per cell and/or the length of the CIRs is high, the optimal pilot set designed no longer has the desired orthogonality for all the users. This will result in PC in a conventional simultaneous UL training based scenario. Therefore, a strategy of two processing stages is proposed, which relies on a DL training phase, followed by appropriately scheduled UL training using the optimal pilot set. By exploiting the orthogonality of the optimal pilot set, each user acquires its unique DL TD CSI with the aid of DL training. The estimated DL TD CSIs are then embedded in the UL training. As a result, the PC can be completely eliminated. Most significantly, our scheme only imposes an overhead of $(L+3)$ OFDM symbols, compared to the $(L+1)U$ OFDM symbols required by the scheme of [17]. Therefore, it imposes a less stringent requirement on the COHI. In contrast to many existing PC reduction schemes, our scheme does not require any information exchange amongst the BSs.

Throughout our discussions, $\mathbb{C}$ denotes the complex number field and for $A \in \mathbb{C}$, we have $A = A_R + \mathrm{j}A_I$, where $\mathrm{j} = \sqrt{-1}$, while A_R and A_I are the real as well as imaginary parts of A, respectively. Boldface upper-case symbols denote matrices, e.g., $\mathbf{X}$, while underlined boldface upper-case symbols denote column vectors, e.g., $\underline{\mathbf{X}}$. The transpose, conjugate and Hermitian transpose operators are denoted by $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$ and $(\cdot)^{\mathrm{H}}$, respectively, while $\mathrm{diag}\{\underline{\mathbf{X}}\}$ denotes the diagonal matrix having diagonal entries equal to the elements of $\underline{\mathbf{X}}$. Furthermore, $\mathbf{X}_{[n,m]}$ denotes the entity of $\mathbf{X}$ at the n-th row and m-th column, and $\underline{\mathbf{X}}_{[n]}$ is the n-th element of $\underline{\mathbf{X}}$, while $\widehat{X}$ represents the estimate of X. The $(K \times K)$-element identity matrix is denoted by $\mathbf{I}_K$, and $\mathbf{0}_K$ denotes the $(K \times K)$-element matrix with all zero elements, while $\delta(t)$ represents the discrete Dirac delta function and $\mathbb{E}\{\cdot\}$ denotes the expectation operator. Furthermore, $\|\cdot\|_F$ represents the Frobenius norm, while $\lfloor\cdot\rfloor$ denotes the integer floor.

II. SYSTEM MODEL AND OPTIMAL PILOT DESIGN

We consider a cellular network composed of L hexagonal cells, labelled by $l = 1, 2, \cdots, L$, where the BS of each cell is equipped with an array of Q antennas for supporting the U MSs, each employing a single antenna. All BSs and MSs are synchronized, and the system relies on a TDD protocol associated with radical unity frequency reuse (UFR). Furthermore, $Q \gg U$.

A. UPLINK TRAINING

All MSs of all cells commence by synchronously transmitting an OFDM PS to their serving BSs. The FD PS of user u in the l-th cell is given by $\underline{\mathbf{X}}_l^u = \left[X_l^u[1]\, X_l^u[2] \cdots X_l^u[N]\right]^T$, where N is the number of subcarriers and the power of each pilot $X_l^u[n]$ is unity. Let $H_{l,l',q}^u[n]$ be the UL FDCHTF linking the u-th user in cell l to the q-th antenna of the l'-th cell's BS, at the n-th subcarrier. Furthermore, let $Y_{l',q}[n]$ be the signal received by the q-th receive antenna of the l'-th BS at the n-th subcarrier, which can be expressed as

$$Y_{l',q}[n] = \sqrt{p_r}\sum_{u'=1}^{U} H_{l',l',q}^{u'}[n]X_{l'}^{u'}[n] + \sqrt{p_r}\sum_{l=1,l\neq l'}^{L}\sum_{u=1}^{U} H_{l,l',q}^{u}[n]X_l^u[n] + W_{l',q}[n] \quad (1)$$

for $1 \leq l' \leq L$ and $1 \leq q \leq Q$, where p_r is the average power of each user and $W_{l',q}[n]$ is the FD representation of the UL channel's additive white Gaussian noise (AWGN), denoted by $W_{l',q}[n] \sim \mathcal{CN}(0, \sigma_w^2)$ with σ_w^2 being the power of $W_{l',q}[n]$. The set of equations constituted by (1) for $1 \leq n \leq N$ can be written in the more compact form of

$$\underline{\mathbf{Y}}_{l',q} = \underbrace{\sqrt{p_r}\sum_{u'=1}^{U}\mathbf{X}_{l'}^{u'}\underline{\mathbf{H}}_{l',l',q}^{u'}}_{\text{Desired term}} + \underbrace{\sqrt{p_r}\sum_{l=1,l\neq l'}^{L}\sum_{u=1}^{U}\mathbf{X}_l^u\underline{\mathbf{H}}_{l,l',q}^u}_{\text{Inter-cell interference}} + \underline{\mathbf{W}}_{l',q}, \quad (2)$$

where $\mathbf{X}_l^u = \mathrm{diag}\{\underline{\mathbf{X}}_l^u\}$, while $\underline{\mathbf{Y}}_{l',q} \in \mathbb{C}^{N\times 1}$, $\underline{\mathbf{H}}_{l,l,q}^u \in \mathbb{C}^{N\times 1}$ and $\underline{\mathbf{W}}_{l',q} \in \mathbb{C}^{N\times 1}$ are the three column vectors hosting $Y_{l',q}[n]$, $H_{l,l,q}^u[n]$ and $W_{l',q}[n]$ for $1 \leq n \leq N$, respectively. It is worth emphasizing again that the signal vector (2) is collected over all the N OFDM subcarriers for an individual BS antenna. This is in contrast to all the existing approaches, which consider the signal vector over all the target BS's antennas for an individual subcarrier [1], [13]–[17]. Note that our approach to CE has a significant advantage. Specifically, let the COHI be measured in terms of the number of OFDM symbols. To implement the conventional simultaneous UL training, our approach only requires that the COHI is no shorter than a single OFDM-symbol duration, while all the existing schemes require that the COHI is no shorter than the duration of U OFDM symbols, because they all require each user to transmit a pilot sequence of length U.

Without loss of generality, assume that a uniformly spaced linear antenna array (ULA) is employed at the BS. Then the TD CIR vector $\underline{\mathbf{G}}_{l,l',q}^u \in \mathbb{C}^{K\times 1}$ is given by

$$\begin{aligned}\underline{\mathbf{G}}_{l,l',q}^u &= \left[G_{l,l',q}^u[1]\, G_{l,l',q}^u[2] \cdots G_{l,l',q}^u[K]\right]^{\mathrm{T}} \\ &= \Big[\alpha_{l,l',q,1}^u e^{-\mathrm{j}2\pi\frac{(q-1)D}{\lambda}\cos\left(\theta_{l,l',q,1}^u\right)} \ldots \\ &\qquad \alpha_{l,l',q,K}^u e^{-\mathrm{j}2\pi\frac{(q-1)D}{\lambda}\cos\left(\theta_{l,l',q,K}^u\right)}\Big]^{\mathrm{T}}, \end{aligned} \quad (3)$$

where K is the maximum duration of the CIR, while D and λ are the antenna spacing and the carrier's wavelength, respectively. In (3), $\theta_{l,l',q,k}^u$ is the AOA of the k-th CIR tap between the u-th MS in the l-th cell and the q-th antenna of the l'-th BS, while the complex-valued tap is given by

$$\alpha_{l,l',q,k}^u = \varrho_{l,l',q,k}^u e^{-\mathrm{j}\varphi_{l,l',q,k}^u}\sqrt{\beta_{l,l',q,k}^u}, \quad (4)$$

where the fast-fading channel gain $\varrho_{l,l',q,k}^u$ follows a Rayleigh distribution and the phase $\varphi_{l,l',q,k}^u$ is a random variable uniformly distributed in $[0,\ 2\pi)$, while the coefficient $\beta_{l,l',q,k}^u$ accounts for the effects of pathloss and shadow fading. Since $\beta_{l,l',q,k}^u$ only changes slowly as a function of distance [1], we may assume that $\beta_{l,l',q,k}^u = \beta_{l,l'}^u$ for $1 \leq k \leq K$ and $1 \leq q \leq Q$. All $\alpha_{l,l',q,k}^u$ are assumed to remain constant over the duration of the COHI. Then the FDCHTF vector $\underline{\mathbf{H}}_{l,l',q}^u$ can be expressed as

$$\underline{\mathbf{H}}_{l,l',q}^u = \mathbf{F}\underline{\mathbf{G}}_{l,l',q}^u, \quad (5)$$

where $\mathbf{F} \in \mathbb{C}^{N\times K}$ is the FFT matrix, whose elements are given by $\mathbf{F}_{[n,k]} = \frac{1}{\sqrt{K}}e^{-\mathrm{j}2\pi(n-1)(k-1)/N}$ for $1 \leq n \leq N$

and $1 \le k \le K$. Using (5), (2) can be rewritten as

$$\underline{\mathbf{Y}}_{l',q} = \sqrt{p_r}\sum_{u'=1}^{U}\mathbf{X}_{l'}^{u'}\mathbf{F}\underline{\mathbf{G}}_{l',l',q}^{u'} + \sqrt{p_r}\sum_{l=1,l\neq l'}^{L}\sum_{u=1}^{U}\mathbf{X}_{l}^{u}\mathbf{F}\underline{\mathbf{G}}_{l,l',q}^{u} + \underline{\mathbf{W}}_{l',q}. \quad (6)$$

B. TIME-DOMAIN CHANNEL ESTIMATION

Theorem 1: Let us now design an FD PS matrix set for all MSs in all cells according to [18] as

$$\mathbf{P} = \{\mathbf{P}[i], 1 \le i \le LU\} \quad (7)$$

with

$$\mathbf{P}[i] = \mathbf{X}_l^u, \quad (8)$$

where $i = (u-1)L + l$ is the index of the LU users for $1 \le u \le U$ and $1 \le l \le L$.

Specifically, the i-th element of the FD PS matrix set is generated from a reference $\mathbf{P}[1] = \mathbf{X}_1^1$ according to

$$\mathbf{P}[i] = \mathbf{\Phi}[i]\mathbf{P}[1], \ 1 \le i \le LU, \quad (9)$$

where the diagonal matrix

$$\mathbf{\Phi}[i] = \mathrm{diag}\{1, e^{\mathrm{j}2\pi\frac{(i-1)\zeta}{N}}, e^{\mathrm{j}2\pi\frac{2(i-1)\zeta}{N}}, \cdots, e^{\mathrm{j}2\pi\frac{(N-1)(i-1)\zeta}{N}}\}, \quad (10)$$

shifts the phase of the reference $\mathbf{P}[1]$ by a positive integer parameter ζ.

Provided that we have $\zeta = \lfloor\frac{N}{LU}\rfloor \ge K$, all the LU PS matrices $\mathbf{X}_l^u$ for $1 \le u \le U$ and $1 \le l \le L$ in this FD PS matrix set are mutually orthogonal.

Proof: Using the PS matrix set (8) in (6), we have

$$\underline{\mathbf{Y}}_{l',q} = \sqrt{p_r}\sum_{u'=1}^{U}\mathbf{P}[(u'-1)L+l']\mathbf{F}\underline{\mathbf{G}}_{l',l',q}^{u'} + \sqrt{p_r}\sum_{l=1,l\neq l'}^{L}\sum_{u=1}^{U}\mathbf{P}[(u-1)L+l]\mathbf{F}\underline{\mathbf{G}}_{l,l',q}^{u} + \underline{\mathbf{W}}_{l',q}. \quad (11)$$

Given $\zeta = \lfloor\frac{N}{LU}\rfloor \ge K$, we have to prove that

$$\left(\mathbf{P}[(u_1-1)L+l_1]\mathbf{F}\right)^H\left(\mathbf{P}[(u_2-1)L+l_2]\mathbf{F}\right) = \begin{cases}\mathbf{0}_K, & l_1 \neq l_2 \cup u_1 \neq u_2,\\ \frac{N}{K}\mathbf{I}_K, & l_1 = l_2 \cap u_1 = u_2.\end{cases} \quad (12)$$

Let us set

$$\mathbf{T} = (\mathbf{P}[i_1]\mathbf{F})^{\mathrm{H}}(\mathbf{P}[i_2]\mathbf{F}), \ 1 \le i_1, i_2 \le LU, \quad (13)$$

$$\mathbf{T}_{[k_1,k_2]} = \frac{1}{K}\sum_{n=1}^{N}\left(\mathbf{P}[i_1]_{[n,n]}\right)^*\mathbf{P}[i_2]_{[n,n]}e^{\frac{\mathrm{j}2\pi(n-1)(k_1-k_2)}{N}}, \quad 1 \le k_1, k_2 \le K. \quad (14)$$

Furthermore, for the integer $t = k_1 - k_2$, we have $|t| \le K-1$ and (14) can be rewritten as

$$\mathbf{T}_{[k_1,k_2]} = \frac{1}{K}\sum_{n=1}^{N}\left(\mathbf{P}[1]_{[n,n]}\right)^*\mathbf{P}[1]_{[n,n]}e^{\frac{-\mathrm{j}2\pi(n-1)(i_1-1)\zeta}{N}} \times e^{\frac{\mathrm{j}2\pi(n-1)(i_2-1)\zeta}{N}}e^{\frac{\mathrm{j}2\pi(n-1)t}{N}} = \frac{N}{K}\delta\big(t-(i_1-i_2)\zeta\big). \quad (15)$$

Under the condition of $i_1 > i_2$ and $t \equiv N+t$ if $t \le 0$, we have

$$\begin{cases}\zeta \le (i_1-i_2)\zeta \le (LU-1)\zeta,\\ 1 \le t \le K-1 \ \text{ or } \ N-K+1 \le t \le N.\end{cases} \quad (16)$$

To keep $\delta\big(t-(i_1-i_2)\zeta\big) = 0$, i.e., $t \neq (i_1-i_2)\zeta$, the inequalities in (16) should have no intersection. In other words, we must have $K-1 < (i_1-i_2)\zeta < N-K+1$, i.e.,

$$\begin{cases}\zeta \ge K,\\ LU\zeta - \zeta \le N-K.\end{cases} \quad (17)$$

If we restrict the range of ζ to

$$K \le \zeta \le \frac{N}{LU}, \quad (18)$$

which clearly meets the condition of (17), then $\mathbf{T} = \mathbf{0}_K$ can be achieved for all $i_1 > i_2$. For $i_1 < i_2$, we can arrive at the same conclusion. Hence, $\mathbf{T} = \mathbf{0}_K, \forall i_1 \neq i_2$, when $K \le \zeta \le \frac{N}{LU}$. We note that the ordered relationship of $i = (u-1)L + l$, $i_1 \neq i_2$ is equivalent to $l_1 \neq l_2 \cup u_1 \neq u_2$.

Let us now consider $i_1 = i_2$, which is equivalent to $l_1 = l_2 \cap u_1 = u_2$. If $t = k_1 - k_2 \neq 0$, we have $\mathbf{T}_{[k_1,k_2]} = 0$. If $(i_1 = i_2) \cap (k_1 = k_2)$, clearly $\mathbf{T}_{[k_1,k_2]} = \frac{N}{K}$. Thus $\mathbf{T} = \frac{N}{K}\mathbf{I}_K$ for $i_1 = i_2$, given that ζ meets the condition of (18).

It can be seen that the orthogonality of (12) actually holds for $K \le \zeta \le \frac{N}{LU}$. Then it also holds for $\zeta = \lfloor\frac{N}{LU}\rfloor \ge K$. ■

1) PILOT-CONTAMINATION-FREE SCENARIO

Given the network of L cells, the maximum number of users U supported per cell, the maximum duration K of the CIR and the number N of subcarrier resources available, we can always design the FD PS matrix set $\mathbf{P}$ of (7). According to Theorem 1, the identifiability of the unique TD CIR vectors $\underline{\mathbf{G}}_{l',l',q}^{u'}$ for all the MSs in all the cells is guaranteed under the condition of:

$$KLU \le N, \quad (19)$$

which implies a PC free scenario, where all the LU MSs of all the L cells can simultaneously transmit their PS matrices to their serving BSs for PC-free CE. The least squares (LS) CE of $\underline{\mathbf{G}}_{l',l',q}^{u'}$ is given by

$$\widehat{\underline{\mathbf{G}}}_{l',l',q}^{u'} = \frac{K}{N\sqrt{p_r}}\left(\mathbf{P}[(u'-1)L+l']\mathbf{F}\right)^{\mathrm{H}}\underline{\mathbf{Y}}_{l',q} = \underline{\mathbf{G}}_{l',l',q}^{u'} + \frac{K}{N\sqrt{p_r}}\left(\mathbf{P}[(u'-1)L+l']\mathbf{F}\right)^{\mathrm{H}}\underline{\mathbf{W}}_{l',q}. \quad (20)$$

The mean square error (MSE) of the estimator (20) is defined by

$$\overline{\Xi}_0 = \mathbb{E}\left\{\frac{1}{K}\|\underline{\Xi}_0\|_F^2\right\}, \tag{21}$$

where $\underline{\Xi}_0 = \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q} - \underline{\mathbf{G}}^{u'}_{l',l',q}$ is the CE error.

Theorem 2: The MSE of (20) is given by

$$\overline{\Xi}_0 = \frac{K\sigma_w^2}{Np_r}. \tag{22}$$

Proof: From (20) we have

$$\underline{\Xi}_0 = \frac{K}{N\sqrt{p_r}}\left(\mathbf{P}[(u'-1)L+l']\mathbf{F}\right)^{\mathrm{H}}\underline{\mathbf{W}}_{l',q}. \tag{23}$$

The k-th element of $\underline{\Xi}_0$ is

$$\underline{\Xi}_{0[k]} = \frac{\sqrt{K}}{N\sqrt{p_r}}\sum_{n=1}^{N} e^{\frac{\mathrm{j}2\pi(n-1)(k-1)}{N}}\left(\mathbf{P}[1]_{[n,n]}\right)^* \times e^{\frac{-\mathrm{j}2\pi\left((u'-1)L+l'-1\right)\zeta(n-1)}{N}} W_{l',q}[n]. \tag{24}$$

Since the noise obeys $W_{l',q}[n] \sim \mathcal{CN}(0,\sigma_w^2)$, $\underline{\Xi}_{0[k]}$ is Gaussian distributed with zero mean, while its power is given by

$$\mathbb{E}\left\{\left|\underline{\Xi}_{0[k]}\right|^2\right\} = \frac{K}{N^2p_r}N\sigma_w^2 = \frac{K\sigma_w^2}{Np_r}. \tag{25}$$

Clearly, the MSE of (20) attains the Cramer-Rao lower bound (CRLB) [19]. ■

2) PILOT-CONTAMINATION SCENARIO

If the condition (19) is not met, then the orthogonality (12) does not hold, i.e., the number of mutually orthogonal elements in $\mathbf{P}$ is less than LU, which results in PC during the UL training, when all MSs of all cells simultaneously transmit their PS matrices to their serving BSs. In this paper we propose an efficient TD two-stage scheme that is capable of eliminating PC with the aid of a combined DL and UL training sessions, which was developed from our previous work [16].

III. TWO-STAGE PILOT CONTAMINATION ELIMINATION SCHEME

The proposed PC elimination scheme consists of a carefully constructed amalgam of a DL training stage and an UL training stage. Furthermore, the DL training contains two phases, while the UL training contains $(L+1)$ training phases, where each phase occupies the duration of a single OFDM symbol. Thus, it is assumed that the channel is time-invariant for the duration of $(L+3)$ OFDM symbols. Before we detail the operations of the DL and UL training stages, we make the following observation. Usually, the number of cells is lower than the number of MSs. Hence, although we consider having insufficient subcarrier sources of $N < KLU$, the following condition is generally satisfied

$$KL < KU \le N < KLU. \tag{26}$$

A. THE DL TRAINING STAGE

From (26), we have $N > KL$. Thus we can design an FD PS matrix set $\mathbf{P}_{DL} = \{\mathbf{S}_l, 1 \le l \le L\}$, similarly to (9) and (10) by replacing LU with L, where $\mathbf{S}_l = \mathrm{diag}\{S_l[1], S_l[2], \cdots, S_l[N]\}$ and the power of each pilot $S_l[n]$ is unity. Since $\lfloor\frac{N}{L}\rfloor \ge K$ holds, $\mathbf{P}_{DL}$ contains the L orthogonal PS matrices associated with $\left(\mathbf{S}_{l'}\mathbf{F}\right)^H\left(\mathbf{S}_{l'}\mathbf{F}\right) = \frac{N}{K}\mathbf{I}_K$ and $\left(\mathbf{S}_{l'}\mathbf{F}\right)^H\left(\mathbf{S}_l\mathbf{F}\right) = \mathbf{0}_K$ for $1 \le l, l' \le L$ and $l' \ne l$. Hence, we can assign an orthogonal PS matrix to each cell.

1) FIRST PHASE

All the BSs broadcast their assigned OFDM PSs using a single antenna, say, the first antennas of the BSs. The signal vector received by MS u' in cell l', $\underline{\mathbf{Z}}^{u',(1)}_{l'} \in \mathbb{C}^{N\times 1}$, can be readily expressed by

$$\underline{\mathbf{Z}}^{u',(1)}_{l'} = \sqrt{p_f}\sum_{l=1}^{L}\mathbf{S}_l\mathbf{F}\underline{\mathbf{G}}^{u'}_{l',l,1} + \underline{\mathbf{V}}^{u',(1)}_{l'},\ 1 \le u' \le U, \tag{27}$$

where p_f denotes the average transmit power of each BS, and $\underline{\mathbf{G}}^{u'}_{l',l,1}$ represents the TD CIR vector between the first antenna of the l-th BS and the MS u' in cell l', while $\underline{\mathbf{V}}^{u',(1)}_{l'} = \left[V^{u',(1)}_{l'}[1]\ V^{u',(1)}_{l'}[2], \cdots V^{u',(1)}_{l'}[N]\right]^T$ is the FD representation of the channel's AWGN. The LS estimate of $\underline{\mathbf{G}}^{u'}_{l',l',1}$ is readily given by

$$\begin{aligned}\widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1} &= \frac{K\left(\mathbf{S}_{l'}\mathbf{F}\right)^H\underline{\mathbf{Z}}^{u',(1)}_{l'}}{N\sqrt{p_f}} \\ &= \underline{\mathbf{G}}^{u'}_{l',l',1} + \frac{K}{N\sqrt{p_f}}\left(\mathbf{S}_{l'}\mathbf{F}\right)^H\underline{\mathbf{V}}^{u',(1)}_{l'},\ 1 \le u' \le U.\end{aligned} \tag{28}$$

It can be seen that this LS estimator is free from PC.

2) SECOND PHASE

All the BSs active all the DL transmit antennas to broadcast the same FD PS matrices to the MSs and, therefore, the received OFDM signal $\underline{\mathbf{Z}}^{u',(2)}_{l'} \in \mathbb{C}^{N\times 1}$ of MS u' in cells l' is given by

$$\begin{aligned}\underline{\mathbf{Z}}^{u',(2)}_{l'} &= \sqrt{p_f}\sum_{l=1}^{L}\mathbf{S}_l\mathbf{F}\sum_{q=1}^{Q}\underline{\mathbf{G}}^{u'}_{l',l,q} + \underline{\mathbf{V}}^{u',(2)}_{l'} \\ &= \sqrt{p_f}\sum_{l=1}^{L}\mathbf{S}_l\mathbf{F}\underline{\mathbf{G}}^{u',\mathrm{sum}}_{l',l} + \underline{\mathbf{V}}^{u',(2)}_{l'},\ 1 \le u' \le U,\end{aligned} \tag{29}$$

where $\underline{\mathbf{G}}^{u'}_{l',l,q}$ denotes the TD CIR vector between the q-th antenna of the l-th BS and the MS u' in cell l', and $\underline{\mathbf{V}}^{u',(2)}_{l'} = \left[V^{u',(2)}_{l'}[1]\ V^{u',(2)}_{l'}[2]\cdots V^{u',(2)}_{l'}[N]\right]^T$ is the FD representation of the channel's AWGN, while $\underline{\mathbf{G}}^{u',\mathrm{sum}}_{l',l} = \sum_{q=1}^{Q}\underline{\mathbf{G}}^{u'}_{l',l,q}$ is the sum of the TD CIRs for all the links between the l-th BS and the MS u' in cell l'. Thus, MS u' in cell l' can estimate $\underline{\mathbf{G}}_{l',l'}^{u',\mathrm{sum}}$ using the LS estimate of

$$\widehat{\underline{\mathbf{G}}}_{l',l'}^{u',\mathrm{sum}} = \frac{K\left(\mathbf{S}_{l'}\mathbf{F}\right)^H \underline{\mathbf{Z}}_{l'}^{u',(2)}}{N\sqrt{p_f}} = \underline{\mathbf{G}}_{l',l'}^{u',\mathrm{sum}} + \frac{K}{N\sqrt{p_f}}\left(\mathbf{S}_{l'}\mathbf{F}\right)^H \underline{\mathbf{V}}_{l'}^{u',(2)},\ 1 \le u' \le U. \tag{30}$$

Again, this LS estimator is free from PC.

B. THE UL TRAINING STAGE

We assume the worst-case scenario of having an FD PS matrix set of only U orthogonal PS matrices $\mathbf{P}_{UL} = \left\{\mathbf{X}^u, 1 \le u \le U\right\}$, with $\mathbf{X}^u = \mathrm{diag}\left\{X^u[1], X^u[2], \cdots, X^u[N]\right\}$, and the power of each pilot $X^u[n]$ being unity. From (26), $N \ge KU$. Thus this FD PS matrix set can be designed similarly to (9) and (10) by replacing LU with U. Since $\lfloor \frac{N}{U} \rfloor \ge K$, $\left(\mathbf{X}^{u_1}\mathbf{F}\right)^H\left(\mathbf{X}^{u_1}\mathbf{F}\right) = \frac{N}{K}\mathbf{I}_K$ and $\left(\mathbf{X}^{u_1}\mathbf{F}\right)^H\left(\mathbf{X}^{u_2}\mathbf{F}\right) = \mathbf{0}_K$ for $1 \le u_1, u_2 \le U$ and $u_1 \ne u_2$. $\mathbf{P}_{UL}$ is reused in every cell. The UL training stage consists of the $(L+1)$ scheduled training phases.

1) INITIAL PHASE

During the initial phase of the UL training, which is indexed as 0, the MSs roaming in all the cells simultaneously transmit their pre-assigned FD PS matrices $\mathbf{X}_l^u = \mathbf{X}^u$ for $1 \le u \le U$ and $1 \le l \le L$ to their corresponding BSs. The signal vector received during phase 0 at the q-th receive antenna of the l'-th BS can be expressed according to (6) as

$$\underline{\mathbf{Y}}_{l',q}^{(0)} = \sqrt{p_r}\sum_{u'=1}^{U}\mathbf{X}^{u'}\mathbf{F}\underline{\mathbf{G}}_{l',l',q}^{u'} + \underbrace{\sqrt{p_r}\sum_{l=1,l\ne l'}^{L}\sum_{u=1}^{U}\mathbf{X}^{u}\mathbf{F}\underline{\mathbf{G}}_{l,l',q}^{u}}_{\text{Pilot contamination}} + \underline{\mathbf{W}}_{l',q}^{(0)}, \tag{31}$$

where $\underline{\mathbf{W}}_{l',q}^{(0)}$ is the corresponding UL channel AWGN vector. The LS estimate of $\underline{\mathbf{G}}_{l',l',q}^{u'}$ based on $\underline{\mathbf{Y}}_{l',q}^{(0)}$ is given by

$$\underline{\widetilde{\mathbf{Y}}}_{l',q}^{u',(0)} = \frac{K}{N\sqrt{p_r}}\left(\mathbf{X}^{u'}\mathbf{F}\right)^H\underline{\mathbf{Y}}_{l',q}^{(0)} = \underline{\mathbf{G}}_{l',l',q}^{u'} + \sum_{l=1,l\ne l'}^{L}\underline{\mathbf{G}}_{l,l',q}^{u'} + \underline{\widetilde{\mathbf{W}}}_{l',q}^{u',(0)}, \tag{32}$$

where $\underline{\widetilde{\mathbf{W}}}_{l',q}^{u',(0)} = \frac{K}{N\sqrt{p_r}}\left(\mathbf{X}^{u'}\mathbf{F}\right)^H\underline{\mathbf{W}}_{l',q}^{(0)}$ and the power of $\underline{\widetilde{\mathbf{W}}}_{l',q}^{u',(0)}$ is $\frac{\sigma_w^2}{p_r}$.

2) PHASE l'

During the l'-th phase of the UL training, where $1 \le l' \le L$, the MSs in cell l' transmit their own specifically power predistorted UL PS matrices $\bar{\mathbf{X}}_{l'}^{u'}$ given by

$$\bar{\mathbf{X}}_{l'}^{u'} = B_{l'}^{u'}\mathbf{X}^{u'},\ 1 \le u' \le U, \tag{33}$$

to the l'-th BS, where

$$B_{l'}^{u'} = \frac{\sum_{k=1}^{K}\widehat{\underline{\mathbf{G}}}_{l',l'}^{u',\mathrm{sum}}[k]}{\sum_{k=1}^{K}\widehat{\underline{\mathbf{G}}}_{l',l',1}^{u'}[k]}, \tag{34}$$

assuming that $\sum_{k=1}^{K}\widehat{\underline{\mathbf{G}}}_{l',l'}^{u',\mathrm{sum}}[k] \ne 0$ and $\sum_{k=1}^{K}\widehat{\underline{\mathbf{G}}}_{l',l',1}^{u'}[k] \ne 0$. As $\bar{\mathbf{X}}_{l'}^{u'}$ encapsulates the estimated DL TD CIR information $B_{l'}^{u'}$ obtained by the MS u' in cell l' during the DL training stage, it is distinct to this MS. At the same time, the MSs roaming in all the other cells simultaneously transmit their pre-assigned PS matrices $\mathbf{X}_l^{u'} = \mathbf{X}^{u'}$ to their BSs, where $1 \le u' \le U$, $1 \le l \le L$ and $l \ne l'$.

As a result, the UL signal received by the q-th antenna of the l'-th BS at phase l' is readily expressed as

$$\underline{\mathbf{Y}}_{l',q}^{(l')} = \sqrt{p_r}\sum_{u'=1}^{U}B_{l'}^{u'}\mathbf{X}^{u'}\mathbf{F}\underline{\mathbf{G}}_{l',l',q}^{u'} + \underbrace{\sqrt{p_r}\sum_{l=1,l\ne l'}^{L}\sum_{u=1}^{U}\mathbf{X}^{u}\mathbf{F}\underline{\mathbf{G}}_{l,l',q}^{u}}_{\text{Pilot contamination}} + \underline{\mathbf{W}}_{l',q}^{(l')}, \tag{35}$$

where $\underline{\mathbf{W}}_{l',q}^{(l')}$ is the corresponding UL AWGN vector. The LS estimate of $\underline{\mathbf{G}}_{l',l',q}^{u'}$ based on $\underline{\mathbf{Y}}_{l',q}^{(l')}$ is given by

$$\underline{\widetilde{\mathbf{Y}}}_{l',q}^{u',(l')} = \frac{K}{N\sqrt{p_r}}\left(\mathbf{X}^{u'}\mathbf{F}\right)^H\underline{\mathbf{Y}}_{l',q}^{(l')} = \frac{\sum_{k=1}^{K}\widehat{\underline{\mathbf{G}}}_{l',l'}^{u',\mathrm{sum}}[k]}{\sum_{k=1}^{K}\widehat{\underline{\mathbf{G}}}_{l',l',1}^{u'}[k]}\mathbf{I}_{N\times N}\underline{\mathbf{G}}_{l',l',q}^{u'} + \sum_{l=1,l\ne l'}^{L}\underline{\mathbf{G}}_{l,l',q}^{u'} + \underline{\widetilde{\mathbf{W}}}_{l',q}^{u',(l')}, \tag{36}$$

where $\underline{\widetilde{\mathbf{W}}}_{l',q}^{u',(l')} = \frac{K}{N\sqrt{p_r}}\left(\mathbf{X}^{u'}\mathbf{F}\right)^H\underline{\mathbf{W}}_{l',q}^{(l')}$, and the power of $\underline{\widetilde{\mathbf{W}}}_{l',q}^{u',(l')}$ is $\frac{\sigma_w^2}{p_r}$.

3) PILOT CONTAMINATION ELIMINATION

Let $\underline{\check{\mathbf{Y}}}_{l',q}^{u'} = \underline{\widetilde{\mathbf{Y}}}_{l',q}^{u',(l')} - \underline{\widetilde{\mathbf{Y}}}_{l',q}^{u',(0)}$ and $\underline{\check{\mathbf{W}}}_{l',q}^{u'} = \underline{\widetilde{\mathbf{W}}}_{l',q}^{u',(l')} - \underline{\widetilde{\mathbf{W}}}_{l',q}^{u',(0)}$, where the power of $\underline{\check{\mathbf{W}}}_{l',q}^{u'}$ is $\frac{2\sigma_w^2}{p_r}$. From (32) and (36), we readily arrive at

$$\underline{\check{\mathbf{Y}}}_{l',q}^{u'} = \left(\frac{\sum_{k=1}^{K}\widehat{\underline{\mathbf{G}}}_{l',l'}^{u',\mathrm{sum}}[k]}{\sum_{k=1}^{K}\widehat{\underline{\mathbf{G}}}_{l',l',1}^{u'}[k]} - 1\right)\underline{\mathbf{G}}_{l',l',q}^{u'} + \underline{\check{\mathbf{W}}}_{l',q}^{u'},\quad 1 \le l' \le L,\ 1 \le q \le Q. \tag{37}$$

Observe in (37) that the PC is completely eliminated.

$$\begin{bmatrix} \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[1] & \check{\underline{\mathbf{Y}}}^{u'}_{l',2}[1] & \cdots & \check{\underline{\mathbf{Y}}}^{u'}_{l',Q}[1] \\ \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[2] & \check{\underline{\mathbf{Y}}}^{u'}_{l',2}[2] & \cdots & \check{\underline{\mathbf{Y}}}^{u'}_{l',Q}[2] \\ \vdots & \vdots & \vdots & \vdots \\ \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[K] & \check{\underline{\mathbf{Y}}}^{u'}_{l',2}[K] & \cdots & \check{\underline{\mathbf{Y}}}^{u'}_{l',Q}[K] \end{bmatrix} = \left(\frac{\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u',\mathrm{sum}}_{l',l'}[k]}{\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k]} - 1 \right) \begin{bmatrix} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[1] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[1] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[1] \\ \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[2] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[2] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[2] \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[K] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[K] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[K] \end{bmatrix}$$

$$= \frac{\sum\limits_{q=2}^{Q}\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}[k]}{\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k]} \begin{bmatrix} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[1] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[1] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[1] \\ \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[2] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[2] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[2] \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[K] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[K] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[K] \end{bmatrix}. \tag{38}$$

$$\begin{bmatrix} \sum\limits_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[k] & \sum\limits_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',2}[k] & \cdots & \sum\limits_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',Q}[k] \end{bmatrix}$$

$$= \frac{\sum\limits_{q=2}^{Q}\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}[k]}{\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[K]} \begin{bmatrix} \sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k] & \sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[k] & \cdots & \sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[k] \end{bmatrix}$$

$$= \begin{bmatrix} \sum\limits_{q=2}^{Q}\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}[k] & \frac{\sum\limits_{q=2}^{Q}\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}[k]}{\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k]} \sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[k] & \cdots & \frac{\sum\limits_{q=2}^{Q}\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}[k]}{\sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k]} \sum\limits_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[k] \end{bmatrix}. \tag{39}$$

In order to extract the estimates of the MS-specific TD CIR vectors $\underline{\mathbf{G}}^{u'}_{l',l',q}$ for $1 \leq u' \leq U$ and $1 \leq q \leq Q$, we expand (37). Specifically, we do not distinguish the DL and UL TD CIR estimates, since they may be assumed to be identical, owing to the channel's reciprocity. We then proceed by substituting $\underline{\mathbf{G}}^{u'}_{l',l',q}$ in (37) by their estimates $\widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}$ in order to express it in the element-based form (38), as shown at the top of this page. By summing the elements in each column for both the left-hand and the right-hand sides of (38), we arrive at (39), as shown at the top of this page, given at the bottom of the page. It can readily be shown that

$$\sum_{q=2}^{Q}\sum_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}[k] = \sum_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[k]. \tag{40}$$

By summing the last $(Q-1)$ elements for both the left-hand and the right-hand sides of (39), we arrive at

$$\begin{aligned} \sum_{q=2}^{Q}\sum_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',q}[k] &= \frac{\left(\sum_{q=2}^{Q}\sum_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}[k]\right)^2}{\sum_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k]} \\ &= \frac{\left(\sum_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[k]\right)^2}{\sum_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k]} \end{aligned} \tag{41}$$

or

$$\sum_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k] = \frac{\left(\sum_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[k]\right)^2}{\sum_{q=2}^{Q}\sum_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',q}[k]}. \tag{42}$$

With the aid of (40) and (42), we have

$$\frac{\sum_{q=2}^{Q}\sum_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',q}[k]}{\sum_{k=1}^{K} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[k]} = \frac{\sum_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[k]}{\frac{\left(\sum_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',1}k]\right)^2}{\sum_{q=2}^{Q}\sum_{k=1}^{K} \check{\underline{\mathbf{Y}}}^{u'}_{l',q}[k]}} = \widehat{Y}_{TSE}, \tag{43}$$

and (38) can be rewritten as

$$\begin{bmatrix} \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[1] & \check{\underline{\mathbf{Y}}}^{u'}_{l',2}[1] & \cdots & \check{\underline{\mathbf{Y}}}^{u'}_{l',Q}[1] \\ \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[2] & \check{\underline{\mathbf{Y}}}^{u'}_{l',2}[2] & \cdots & \check{\underline{\mathbf{Y}}}^{u'}_{l',Q}[2] \\ \vdots & \vdots & \vdots & \vdots \\ \check{\underline{\mathbf{Y}}}^{u'}_{l',1}[K] & \check{\underline{\mathbf{Y}}}^{u'}_{l',2}[K] & \cdots & \check{\underline{\mathbf{Y}}}^{u'}_{l',Q}[K] \end{bmatrix} = \widehat{Y}_{TSE}$$

$$\times \begin{bmatrix} \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[1] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[1] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[1] \\ \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[2] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[2] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[2] \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',1}[K] & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',2}[K] & \cdots & \widehat{\underline{\mathbf{G}}}^{u'}_{l',l',Q}[K] \end{bmatrix}. \tag{44}$$

From (44) we obtain the estimates of $\underline{\mathbf{G}}^{u'}_{l',l',q}[k]$ as

$$\underline{\widehat{\mathbf{G}}}^{u'}_{l',l',q}[k]=\frac{\check{\mathbf{Y}}^{u'}_{l',q}[k]}{\widehat{Y}_{TSE}},\ 1\le k\le K, 1\le q\le Q, 1\le u'\le U. \tag{45}$$

Then the estimates of $\underline{\mathbf{H}}^{u'}_{l',l',q}$ are readily given by

$$\underline{\widehat{\mathbf{H}}}^{u'}_{l',l',q}=\mathbf{F}\underline{\widehat{\mathbf{G}}}^{u'}_{l',l',q},\ 1\le q\le Q, 1\le u'\le U. \tag{46}$$

REMARK

It can be seen that the proposed two-stage scheme completely eliminates the PC. However, the noise effects in the DL training stage are propagated back to the UL training stage. Furthermore, the signal processing operations in the UL training stage also amplify the noise. Therefore the MSE of this estimator becomes considerably higher than the CRLB.

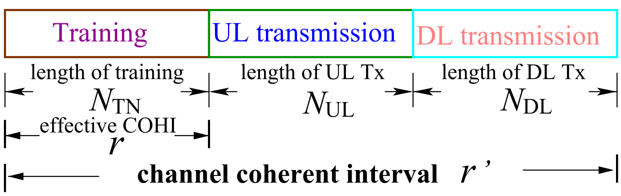


FIGURE 1. TDD protocol frame structure and its relationship with the channel coherence time.

C. COMPARISON WITH EXISTING SCHEMES

Fig. 1 illustrates the TDD protocol's frame structure and its relationship to the channel's coherence time. Let us denote the channel's coherence time as $t_{coherent}$ and the OFDM symbol duration as T_{symbol}. We can define the overall COHI as the ratio of $t_{coherent}$ over T_{symbol}

$$r' = \left\lfloor \frac{t_{coherent}}{T_{symbol}} \right\rfloor, \tag{47}$$

which specifies the maximum number of OFDM symbols during which the CIRs remain near-constant. Since the system has to carry out both the training as well as the UL and DL data transmissions within r', we define the effective COHI for performing CE as

$$r = r' - N_{UL} - N_{DL}, \tag{48}$$

where N_{UL} and N_{DL} represent the numbers of OFDM symbols transmitted during the UL and DL data transmissions, respectively. In other words, the duration of training must satisfy $N_{TN} \le r$.

The conventional simultaneous training scheme only requires a training duration of $N_{TN} = U$, but it suffers from serious PC. By contrast, our proposed scheme is capable of completely eliminating the PC and it imposes a training duration of $N_{TN} = L+3$ OFDM symbols. By comparison, the PC-free scheme of [16] requires a training duration of $N_{TN} = (L+3)U$, while the PC-free scheme of [17] imposes a training duration of $N_{TN} = (L+1)U$ OFDM symbols. It can be seen that a significant advantage of our proposed PC-free scheme over the existing PC-free schemes of [16] and [17] is that our scheme imposes a much less stringent requirement on the channel coherence's time.

The proposed scheme has a similar computational complexity as the scheme of [16], while it imposes a marginally higher complexity than the scheme of [17]. Compared with the conventional simultaneous training scheme, our proposed scheme as well as the schemes of [16] and [17] are capable of eliminating PC at the cost of imposing higher computational complexity.

TABLE 1. Default parameters used in the simulated multiple-antenna aided and TDD based OFDM system.

Number of cells L	7
Radius of each cell	1000 m
Number of MSs per-cell U	8
Number of antennas at each BS Q	100
Antenna spacing D	$\frac{\lambda}{2}$
Average transmit power per subcarrier at each MS p_r	0 dB
Average transmit power per subcarrier at each BS p_f	10 dB
Number of subcarriers N	1024
Length of CIR K	54
Path loss exponent	3
Mean of CIR paths' AOAs θ	90°
Standard deviation of CIR paths' AOAs σ_{AOA}	90°

IV. SIMULATION RESULTS

The default values of the various parameters for our simulated multi-cell TDD system are listed in Table 1. Unless otherwise specified, these default parameter values were used throughout. The UFR regime was assumed and the pre-assigned $\mathbf{P}_{DL}$ and $\mathbf{P}_{UL}$ were employed for the DL CE and UL CE of our proposed TD two-stage scheme, respectively. The signal-to-noise ratio (SNR) of the system was defined as E_s/N_0, where E_s denoted the energy per symbol and $N_0 = \sigma_w^2$ denoted the power of the channel's AWGN. All the CIR paths' AOAs $\theta^u_{l',l,q,k}$ were independently identically distributed (i.i.d.) Gaussian random variables with a mean of $\bar{\theta} = 90°$ and the standard deviation of $\sigma_{AOA} = 90°$.

A. ESTIMATION PERFORMANCE

We first compare the effective COHI required for different channel estimators. It can be seen from Fig. 2 that our simultaneous UL training scheme having sufficient subcarrier resources of $N = 3072 > KLU$ requires a minimum effective COHI of $r = 1$, while the conventional simultaneous training scheme can be invoked, provided that the effective COHI is no shorter than $r = U = 8$. Our proposed two-stage channel estimator requires the effective COHI $r = L+3 = 10$, which is much better than the effective COHI of $r = (L+1)U = 64$ required by the scheme of [17] and the effective COHI of $r = (L+3)U = 80$ imposed by the scheme of [16].

We then verify the efficiency of our proposed design by examining the normalized MSE (NMSE) of the CE, which is

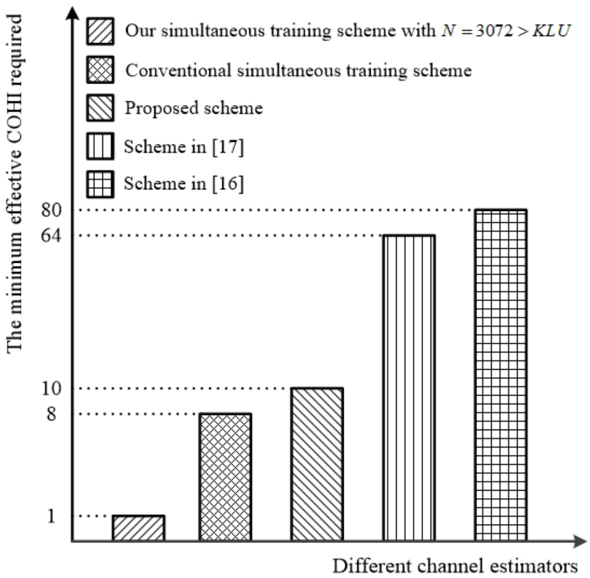


FIGURE 2. The minimum effective COHI required by different channel estimators.

defined as

$$\mathrm{NMSE} = \frac{\sum\limits_{l=1}^{L}\sum\limits_{u=1}^{U}\sum\limits_{q=1}^{Q}\sum\limits_{n=1}^{N}\left|\widehat{H}_{l,l,q}^{u}[n]-H_{l,l,q}^{u}[n]\right|^2}{\sum\limits_{l=1}^{L}\sum\limits_{u=1}^{U}\sum\limits_{q=1}^{Q}\sum\limits_{n=1}^{N}\left|H_{l,l,q}^{u}[n]\right|^2}, \quad (49)$$

where $H_{l,l,q}^{u}[n]$ and $\widehat{H}_{l,l,q}^{u}[n]$ represent the true channel and its estimation, respectively. All the NMSE simulation results were obtained by averaging over 100 channel realizations.

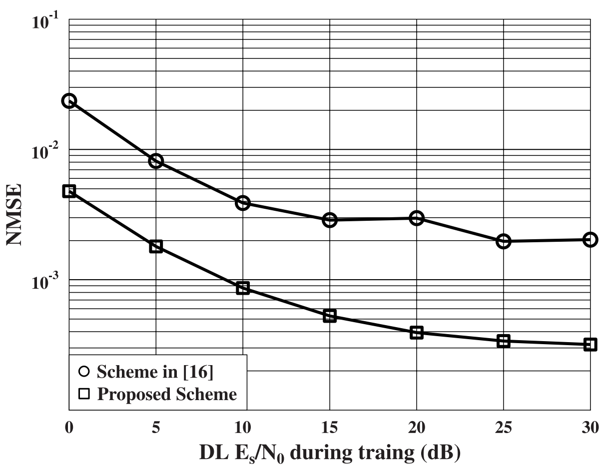


FIGURE 3. NMSE performance of two channel estimators as the functions of the DL training SNR, given the UL training SNR of $E_s/N_0 = 30$ dB.

Since our proposed scheme and the scheme of [16] consist of both the DL training and UL training stages, we commence by investigating the NMSEs of the two estimators as the functions of the DL training SNR, given the UL SNR of $E_s/N_0 = 30$ dB. The simulation results are depicted in Fig. 3, where it can be seen that the estimation accuracies of both schemes are determined by the DL training SNR. Our advocated scheme performs better than the scheme of [16], because in the DL training stage, the power of the noise in our scheme is N-times lower than that in the scheme of [16].

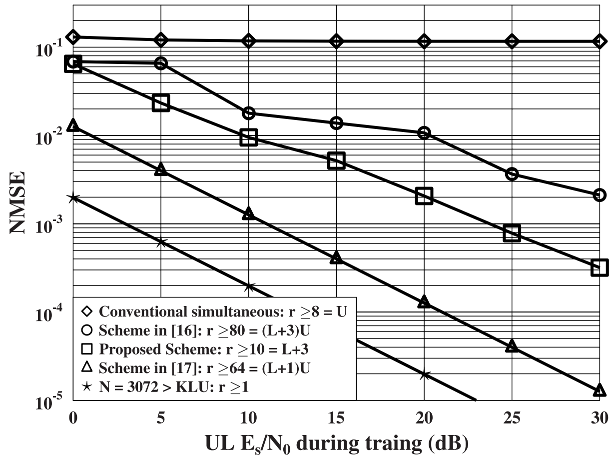


FIGURE 4. NMSE performance of different channel estimators as the functions of the UL training SNR. For the proposed scheme and the scheme of [16], the DL training SNR is set to $E_s/N_0 = 30$ dB.

We next compare the NMSE of our proposed scheme as a function of the UL training SNR to those of the conventional simultaneous UL training scheme as well as to the schemes of [16] and [17]. Since both the proposed scheme of this treatise and the scheme of [16] require a DL training stage, we set the DL training SNR to $E_s/N_0 = 30$ dB for the both schemes. The results obtained are shown in Fig. 4, where the NMSE performance of CE achieved by our simultaneous UL training scheme for a sufficiently high number of subcarriers given by $N = 3072 > KLU$ is also provided for comparison. As expected, the NMSE performance of the conventional simultaneous UL training scheme is the worst, exhibiting the highest NMSE floor of 0.1, because it seriously suffers from PC. The rest four schemes are all capable of eliminating PC, whereby two-stage scheme proposed here is better than the scheme of [16], which is due to the reduced noise power in its DL training stage. However, the NMSE performance of our two-stage scheme is worse than that of the scheme of [17] because for our two-stage scheme the CE error in the DL training stage is fed back to the UL training stage, which inevitably decreases the UL CE accuracy. However, it is worth emphasizing again that our scheme is capable of operating under the network condition of an effective COHI of $r \geq L+3 = 10$, while the scheme of [17] can only operate for an effective COHI of $r \geq (L+1)U = 64$. Also observe from Fig. 4 that the NMSE performance of our simultaneous UL training scheme associated with $N = 3072 > KLU$ is the best. As presented in Section II-B, given a sufficient number of subcarriers of $N \geq KLU$, our purposefully designed pilot set is always capable of achieving a PC-free estimation in the simultaneous UL training, which only requires an effective COHI of $r \geq 1$. Moreover, the MSE of this simultaneous UL training scheme attains the CRLB, which is inversely proportional to N.

B. ACHIEVABLE SUM-RATE PERFORMANCE WITHOUT CONSIDERING THE TRAINING DURATION

The achievable UL sum-rates of different channel estimators are subsequently investigated. Specifically, during the UL transmission, the system's SNR is set to the same UL training SNR value, and each BS performs maximum-ratio combining (MRC) aided detection by multiplying its received signal with the conjugate-transpose of the channel estimate obtained during the training. The per-cell UL sum-rate is defined by

$$C_{UL} = \frac{1}{L} \sum_{l=1}^{L} \sum_{u=1}^{U} \log_2 \left(1 + \mathrm{SINR}_{l,u}\right), \tag{50}$$

where $\mathrm{SINR}_{l,u}$ is the desired signal-to-interference-plus-noise ratio of user u in cell l, obtained by the MRC based on the estimated channels. The per-cell UL sum-rate performances are compared in Fig. 5. Observe from Fig. 5 that the achievable UL sum-rates of the schemes proposed in [16] and [17] and the TD two-stage scheme associated in this treatise as well as our simultaneous UL training scheme associated with a sufficient high number of subcarriers of $N = 3072$ are very similar and they are all close to the perfect-CSI bound. This is because all these schemes are capable of eliminating PC. By contrast, an obvious difference of about 14 bits/sec/Hz is seen between the UL sum-rate of the conventional simultaneous training scheme and the other schemes. This simply confirms that the conventional simultaneous training scheme suffers from serious PC.

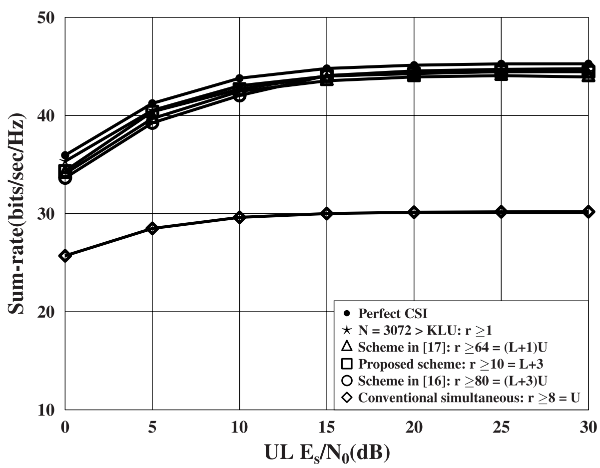


FIGURE 5. **Achievable per-cell UL sum-rate performance as the functions of the UL transmission SNR by different estimators with the UL training SNR equal to the UL transmission SNR. Additionally, for our proposed scheme and the scheme of [16], the DL training SNR is set to $E_s/N_0 = 30$ dB.**

Let us now study the achievable DL sum-rate performance. Specifically, given the channel estimate obtained at the UL training SNR of $E_s/N_0 = 30$ dB as well as additionally the DL training SNR of $E_s/N_0 = 30$ dB for our two-stage scheme and the scheme of [16], each BS carries out its DL transmission by invoking zero-forcing (ZF) transmit precoding based on the UL CE. The per-cell DL sum-rate, C_{DL}, is defined similarly to C_{UL} of (50). Fig. 6 shows the achievable per-cell DL sum-rate versus the DL system's SNR. It can be seen that the DL sum-rate performance is similar to the UL sum-rate performance. We also find that the influence of the DL system's SNR on the DL sum-rate remains modest for all the schemes. Again, a difference of about 14 bits/sec/Hz can be seen in the DL sum-rate between the conventional simultaneous training scheme and the other schemes. The results of Fig. 6 also confirm that all the PC-free schemes attain similar DL sum-rate performances that are very close to the perfect-CSI bound.

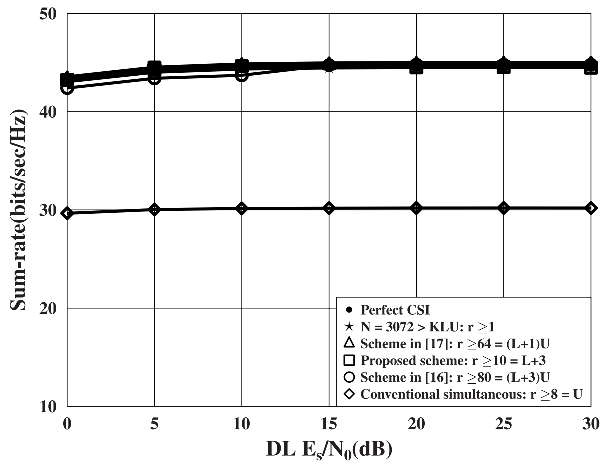


FIGURE 6. **Achievable per-cell DL sum-rate performance as the functions of the DL transmission SNR by different estimators where the UL training SNR is fixed to $E_s/N_0 = 30$ dB. Additionally, for our proposed scheme and the scheme of [16], the DL training SNR is set to $E_s/N_0 = 30$ dB.**

Moreover, we investigate the effect of the number of antennas Q. In particular, we fix both the UL training SNR and the DL training SNR to 30 dB for our proposed scheme of this treatise and the scheme of [16]. Then we vary the number of antennas Q and estimate the corresponding channel matrix. Each BS carries out ZF transmit precoding based DL transmission using the channel estimate obtained for the DL system's SNR of $E_s/N_0 = 30$ dB. Fig. 7 portrays the per-cell DL sum-rate performance versus the number of antennas Q. As expected, increasing the number of antennas enhances the achievable sum-rate. Fig. 7 confirms again that all the PC-free schemes attain very similar performances that are close to the perfect CSI bound, but there exists a performance gap of about 12 bits/sec/Hz between these schemes and the conventional simultaneous UL training scheme.

C. EFFECTIVE SUM-RATE PERFORMANCE WITH TRAINING OVERHEAD ADJUSTMENT

The achievable UL sum-rate C_{UL} and DL sum-rate C_{DL} shown in Section IV-B do not take into account the training overhead imposed and therefore they only represent the ideal performance achievable in the limit case of $r \to \infty$. Referring to Fig. 1, the training overhead reduces the system's effective throughput and therefore reduces the achievable

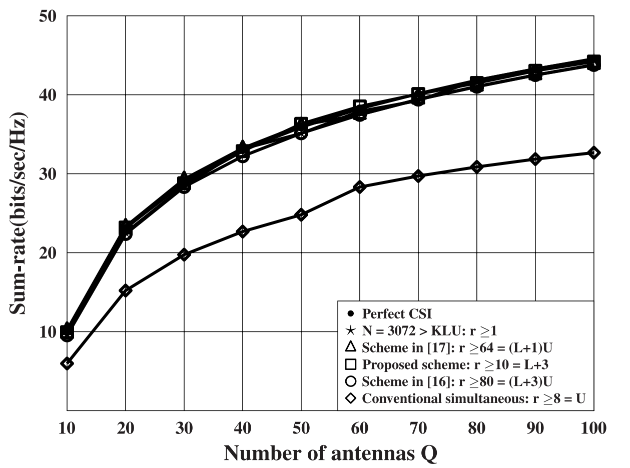


FIGURE 7. Achievable per-cell DL sum-rate performance as the functions of the number of antennas Q by different estimators given the DL transmission SNR of 30 dB. The UL training SNR is fixed to $E_s/N_0 = 30$ dB and additionally, for our proposed scheme and the scheme of [16], the DL training SNR is set to $E_s/N_0 = 30$ dB.

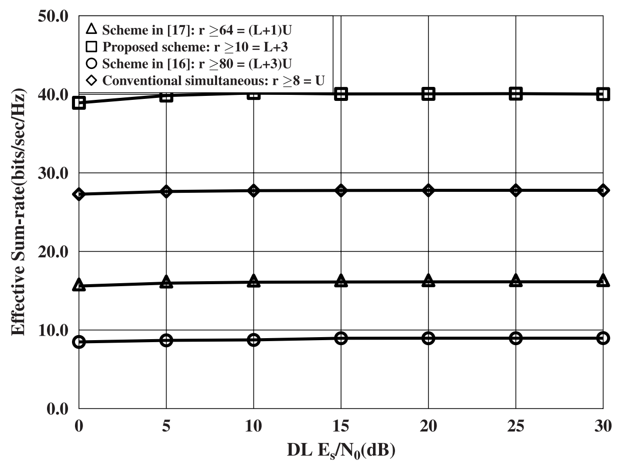


FIGURE 9. Effective per-cell DL sum-rate performance as the functions of the DL transmission SNR by different estimators, assuming that the COHI is $r = 100$. The UL training SNR is fixed to $E_s/N_0 = 30$ dB. Additionally, for our proposed scheme and the scheme of [16], the DL training SNR is set to $E_s/N_0 = 30$ dB.

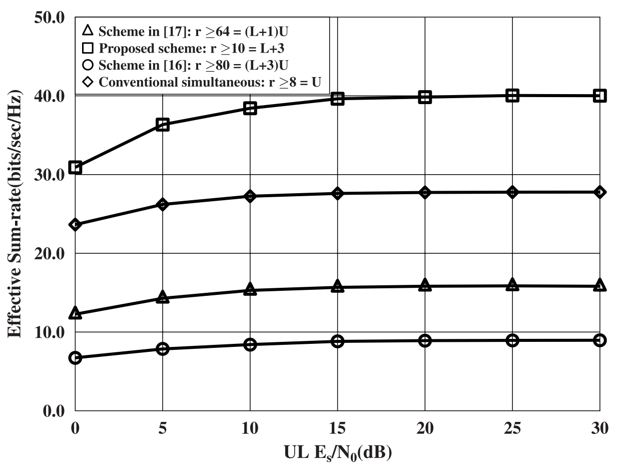


FIGURE 8. Effective per-cell UL sum-rate performance as the functions of the UL transmission SNR by different estimators with the UL training SNR equal to the UL transmission SNR, assuming that the COHI is $r = 100$. Additionally, for our proposed scheme and the scheme of [16], the DL training SNR is set to $E_s/N_0 = 30$ dB.

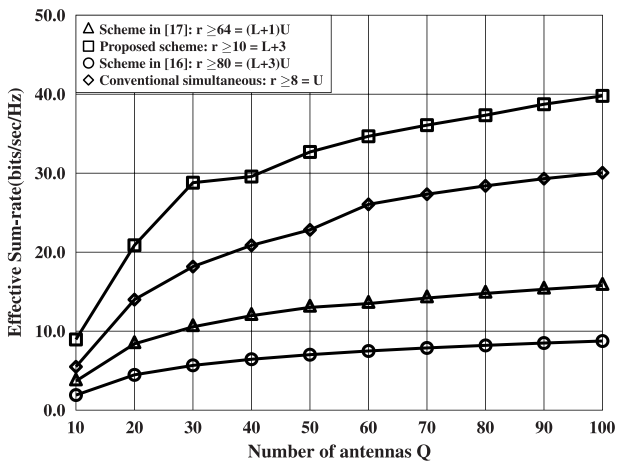


FIGURE 10. Effective per-cell DL sum-rate performance as the functions of the number of antennas Q by different estimators given the DL transmission SNR of 30 dB. The UL training SNR is fixed to $E_s/N_0 = 30$ dB and additionally, for our proposed scheme and the scheme of [16], the DL training SNR is set to $E_s/N_0 = 30$ dB.

sum-rate. Given C_{UL} and C_{DL}, the effective UL sum-rate C_{UL}^{ef} and effective DL sum-rate C_{DL}^{ef} are obtained respectively as

$$C_{UL}^{ef} = \frac{N_{UL}}{\frac{1}{2}N_{TN} + N_{UL}} C_{UL}, \tag{51}$$

$$C_{DL}^{ef} = \frac{N_{DL}}{\frac{1}{2}N_{TN} + N_{DL}} C_{DL}, \tag{52}$$

by taking into account the training overhead adjustment of $\frac{1}{2}N_{TN}$, where the factor $\frac{1}{2}$ is owing to the fact that the channel estimate obtained by training is used in both the UL and DL transmissions.

We consider a very slow fading system associated with the COHI $r = 100$ so that the scheme of [16] can be implemented in conjunction with $N_{UL} = N_{DL} = 10$ OFDM symbols. Since our proposed two-stage scheme requires $N_{TN} = 10$, it can implement the UL and DL transmissions with the aid of $N_{UL} = N_{DL} = 45$. On the other hand, the conventional simultaneous UL training scheme only requires $N_{TN} = 8$, and it can support UL and DL transmissions with the aid of $N_{UL} = N_{DL} = 46$. Under the identical experimental conditions of Figs. 5 to 7, the effective or training-overhead-adjusted sum-rate performance are depicted in Figs. 8 to 10, respectively. Observe from Figs. 8 and 9 that our proposed two-stage training scheme achieves the highest effective sum-rate performance, which is more than 10 bits/sec/Hz higher than that of the scheme of [17]. This clearly demonstrates the superiority of our scheme. Also observe that the performance gap between the scheme of [16] and the conventional simultaneous UL training scheme is halved, from about 14 bits/sec/Hz in Figs. 5 and 6 to about 7 bits/sec/Hz in Figs. 8 and 9, since the conventional simultaneous UL training scheme benefits from a much smaller training overhead.

V. CONCLUSIONS

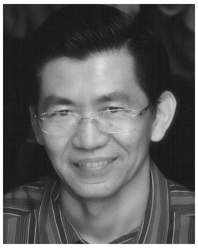
A novel pilot contamination elimination scheme has been proposed for multi-cell TDD and OFDM based massive MIMO systems. Our first contribution has been to prove that under the condition of having a sufficiently high bandwidth for accommodating sufficient subcarriers, an optimal orthogonal pilot set can readily be designed, which can completely eliminate the pilot contamination in a conventional simultaneous UL training session that only imposes a minimum training duration of one OFDM symbol. Our second and main contribution has been to tackle the more hostile operational environment associated with insufficient subcarrier resources. Specifically, we have proposed a two-stage time-domain scheme that combines a downlink training stage with an uplink training stage. Our proposed scheme is capable of completely eliminating PC and it imposes a training overhead of $(L + 3)$, with L being number of cells, which is significantly lower than that of the existing PC-free schemes. Consequently, our scheme imposes a much less stringent requirement on the channel's coherence time. Moreover, the effective sum-rate performance achievable by our scheme is considerably higher than that of the existing schemes, which has been demonstrated by the extensive simulation results.

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] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
[4] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
[5] L. Zhao, K. Li, K. Zheng, and M. O. Ahmad, "An analysis of the tradeoff between the energy and spectrum efficiencies in an uplink massive MIMO-OFDM system," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 62, no. 3, pp. 291–295, Mar. 2015.
[6] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.
[7] 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.
[8] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
[9] A. Hu, T. Lv, H. Gao, Y. Lu, and E. Liu, "Pilot design for large-scale multi-cell multiuser MIMO systems," in *Proc. ICC*, Budapest, Hungary, Jun. 2013, pp. 5381–5385.
[10] Z. Zhao, Z. Chen, and Y. Liu, "Cell sectorization-based pilot assignment scheme in massive MIMO systems," in *Proc. WTS*, New York, NY, USA, Apr. 2015, pp. 1–5.
[11] L. Q. He, L. Xiao, X. Xu, and S. Zhou, "Uplink sum-throughput evaluation of sectorized multi-cell massive MIMO system," in *Proc. ICCW*, London, U.K., Jun. 2015, pp. 1143–1148.
[12] Z. Zhu, *et al.*, "Soft pilot reuse and multicell block diagonalization precoding for massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3285–3298, May 2016.
[13] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192–201, Feb. 2013.
[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] Z. Wang, C. Qian, L. Dai, J. Chen, C. Sun, and S. Chen, "Location-based channel estimation and pilot assignment for massive MIMO systems," in *Proc. ICC Workshop*, London, U.K., Jun. 2015, pp. 1264–1268.
[16] J. Zhang, B. Zhang, S. Chen, X. Mu, M. El-Hajjar, and L. Hanzo, "Pilot contamination elimination for large-scale multiple-antenna aided OFDM systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 759–772, Oct. 2014.
[17] T. X. Vu, T. A. Vu, and T. Q. S. Quek, "Successive pilot contamination elimination in multiantenna multicell networks," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 617–620, Dec. 2014.
[18] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.
[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 2. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

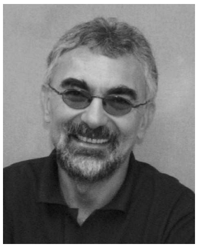
XINYING GUO received the B.Sc. degree in electronic and information engineering from the North University of China, Taiyuan, China, in 2011. She is currently pursuing the Ph.D. degree with the School of Information Engineering, Zhengzhou University. Since 2014, she has been a Visiting Ph.D. Student with Prof. L. Hanzo and Prof. S. Chen with the School of Electronics and Computer Science, University of Southampton, U.K. Her research interests are in the areas of wireless communications and signal processing, including channel estimation, multiuser detection, beamforming/precoding, and pilot contamination elimination/ reduction techniques.

JIANKANG ZHANG (S'08–M'12) received the B.Sc. degree in mathematics and applied mathematics from the Beijing University of Posts and Telecommunications in 2006, and the Ph.D. degree in communication and information systems from Zhengzhou University in 2012. He has been a Lecturer with the School of Information Engineering, Zhengzhou University. From 2009 to 2011, he was a Visiting Researcher in electronics and computer science, University of Southampton, U.K., in 2003. His research interests are in the areas of wireless communications and signal processing, including channel estimation, multiuser detection, beamforming/precoding, and optimization algorithms.



SHENG CHEN (M'90–SM'97–F'08) received the B.Eng. degree in control engineering from the East China Petroleum Institute, Dongying, China, in 1982, the Ph.D. degree in control engineering from City University, London, in 1986, and the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2005. From 1986 to 1999, he held a research and academic appointments with the University of Sheffield, the University of Edinburgh, and the University of Portsmouth, U.K. Since 1999, he has been with the School of Electronics and Computer Science, University of Southampton, where he is currently a Professor of Intelligent Systems and Signal Processing. His research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, evolutionary computation methods, and optimization. He has authored over 550 research papers. He is a fellow of IET, and a Distinguished Adjunct Professor with King Abdulaziz University, Jeddah, Saudi Arabia. He is an ISI Highly Cited Researcher in engineering in 2004. He is a fellow of the United Kingdom Royal Academy of Engineering.

XIAOMIN MU received the B.E. degree from the Beijing Institute of Technology, Beijing, China, in 1982. She is currently a Full Professor with the School of Information Engineering, Zhengzhou University, China. She has authored many papers in signal processing and co-authored two books. Her research interests include signal processing in communication systems, wireless communications, and cognitive radio.



LAJOS HANZO (F'04) received the D.Sc. degree in electronics, the Ph.D. degree, and the Doctor Honoris Causa degree from the Technical University of Budapest in 1976, 1983, and 2009, respectively. During his 38-year career in telecommunications, he has held various research and academic positions in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, U.K., as the Chair in telecommunications. He has successfully supervised 100 Ph.D. students, co-authored 20 John Wiley/IEEE Press books in mobile radio communications totaling in excess of 10,000 pages, authored over 1500 research entries at the IEEE Xplore, acted as the TPC Chair and the General Chair of the IEEE conferences, presented keynote lectures, and received a number of distinctions. He is directing a 100 strong Academic Research Team, working on a range of research projects in wireless multimedia communications sponsored by the industry, the Engineering and Physical Sciences Research Council, U.K., the European Research Council's Advanced Fellow Grant, and the Royal Society's Wolfson Research Merit Award. He is an Enthusiastic Supporter of industrial and an Academic Liaison and offers a range of industrial courses. He is also a fellow of the Royal Academy of Engineering, the Institution of Engineering and Technology, and the European Association for Signal Processing. He is the Governor of the IEEE VTS. From 2008 to 2012, he was the Editor-in-Chief of the IEEE Press and a Chaired Professor with Tsinghua University, Beijing. He has over 24,000 citations.