Research Article

Smart Switching Strategy-Based Supervision Rule to Mitigate the Problem of Pilot Contamination in Massive MIMO Systems

Abdelfettah Belhabib,1 Mohamed Boulouird,1,2 and Moha M’Rabet Hassani1

1Instrumentation, Signals and Physical Systems (I2SP) Group, Faculty of Sciences Semlalia, Cadi Ayyad University, Marrakesh, Morocco
2National School of Applied Sciences of Marrakesh (ENSA-M), Cadi Ayyad University, Marrakesh, Morocco

Correspondence should be addressed to Abdelfettah Belhabib; abdelfettah.belhabib@edu.uca.ac.ma

Received 3 March 2020; Revised 27 June 2020; Accepted 15 July 2020; Published 17 August 2020

Academic Editor: Pavlos I. Lazaridis

Despite the large benefits that can be fulfilled through the exploitation of Massive Multi-input Multioutput (M-MIMO), this technology still constrained by a well-known constraint, called as pilot contamination problem (PCP), which is the main consequence of, simultaneously, reusing the same set of orthogonal pilot sequences (OPSs) for the users of several cells. Due to the scarcity of the OPS resources, the reuse of the same set of the OPSs for the users of different cells is unavoidable. Hence, this work proposes a novel decontaminating strategy, which is aimed at guaranteeing a trade-off between the use of the OPSs and the mitigation of the PCP. Specifically, to make the use of the available OPSs better, we propose the consolidation of two powerful decontaminating strategies. Under a derived supervision rule (SR), these strategies are the soft pilot reuse-based multicell block diagonalization precoding (SPR-MBDP) and the weighted graph coloring-based pilot assignment (WGC-PA). The SR enables the switching mechanism between the two strategies, which leads to address the PCP with a fewer number of the OPSs compared to the SPR-MBDP, therefore boosting the per-cell achievable rate. Simulation results prove the effectiveness of our proposed strategy.

1. Introduction

M-MIMO technology was introduced in 2010 by [1], and it is aimed at serving, simultaneously, tens of the users (K user per cell). This is by employing a massive number (i.e., hundreds) of antennas (i.e., M antennas) at the Base Stations (BSs). This relevant technology has large benefits [2]. Thus, as the number of antennas deployed in the BSs M goes to the infinity, the noisy components and the intracell interferences are made small [1], and the transmit power is reduced through a beamforming technique, which contributes to boosting the energy efficiency [3]. The M-MIMO technology has been matured and adapted to the fifth-generation (5G); this is with a finite number of antennas $M = 64$, and it has been commercialized in the sub-6 GHz bands [4]. However, this is not the end of the M-MIMO, because its large benefits are not yet fully achieved, and still constrained by the PCP [4] (Figure 1).

The reuse of the same set of the OPSs, across different cells, leads to the problem of intercell interferences. To avoid this problem, each user should be assigned a specific pilot sequence; in other words, the users of all cells must be assigned with a set of OPSs. However, this is not possible due to the scarcity of the available set of OPSs; consequently, the reuse of the same set of OPSs across different cells is unavoidable; hence, the PCP will emerge between the users that employ the same OPSs. Consequently, the channel state information (CSI) will be, inaccurately, estimated and the whole process of communication between the users and the BSs will be dramatically damaged [5].

Since the appearance of M-MIMO, a large body of research has been carried out to deal with the PCP. A time-shifted-based strategy has been proposed in [6, 7], which proposes to insert a delay between the phases of communication of the adjacent cells; meanwhile, some of them perform in the data transmission phase; the others are obliged to perform in
the pilot transmission phase. However, the estimated CSI at the BSs becomes contaminated by the transmitted data in adjacent cells instead of being contaminated by the pilot sequences. Based on the singular value decomposition (SVD), the null space of interference was exploited to schedule the data signals toward their desired target (i.e., users) [8, 9]. Indeed, extra pilot resources are required to accurately identify the null space of interferences. The space signatures were suggested to spatially identify the desired signals based on their angle of arrivals (AoAs); therefore, the same OPSs are allocated to the users having nonoverlapped AoAs [10, 11]. Nevertheless, a nonoverlapping AoA is assumed which does not necessarily hold in practice.

Optimization-based decontamination strategy was suggested to boost the quality-of-services (QoS) for the overall users [12–14]. Specifically, the proposal of [12] is aimed at finding out the optimal pilot assignment scheme, that is, through maximizing the sum rate within cells, while [13] focuses on assigning OPSs to the users of adjacent cells, which are suffering from a severe influence of the PCP; in [14], the users within cells are divided, based on their signal-to-interference-plus-noise ratio (SINR), and those having a low SINR are optimally assigned with OPSs.

A weighted graph coloring-based pilot assignment (WGC-PA) decontaminating strategy is proposed in the work of [15], which is aimed at reducing the strength of the PCP upon the users that employ the same OPSs; this is through the employment of the graph coloring theory. The WGC-PA has been coupled with the soft pilot reuse (SPR) strategy in the work of [16, 17], which was, only, focused on the uplink phase. Specifically, [16] proposed to enhance the performance of edge users based on the SPR, while the slight contamination among the center users is addressed based on the WGC-PA. In [17], users within cells are split into three groups, and the problem of the PCP upon the two groups is separately addressed based on the WGC-PA and the SPR strategies, while the PCP upon the group of the center users is neglected.

Regarding the fact that the center users are close to the BSs, they benefit from a high QoS compared to the edge users, which are facing a severe degree of the PCP that comes from the edge users of the adjacent cells. To guarantee the fairness between the users of each cell, it is necessary to enhance the QoS of the edge users. Accordingly, a soft pilot reuse-based multicell block diagonalization precoding (SPR-MBDP) was proposed in [9], which suggests selecting the edge users of the overall cells; therefore, each one of the selected users is assigned with a unique OPSs, while the center users of the overall cells are obliged to reuse the same set of the OPSs (i.e., full pilot reuse); therefore, by exploiting the singular value decomposition (SVD) tool, data symbols are scheduled toward the target users. That is through a null subspace of the PCP. However, the problem of pilot overhead can damage the fulfilled performance of the SPR-MBDP [9], due to its requirement for extra OPSs.

In the existing decontaminating strategies, less attention has been paid to the problem of pilot overhead, which is—without any doubt—a crucial side that must be respected in any proposed decontaminating strategy; stand from that remark, this paper provides a new decontaminating strategy to boost the performance of the M-MIMO. Specifically, to ensure a trade-off between the use of the OPSs and the mitigation of the PCP, we propose the consolidation of the SPR-MBDP and the WGC-PA. This is under a derived supervision rule (SR) that economizes and controls the use of the precious OPSs while trying to overcome the PCP. The proposed strategy is referred to as a smart switching strategy-based supervision rule (S3SR). In a nutshell, the prime contribution of the present paper can be summarized as follows:

(i) A novel S3SR strategy is proposed to deal with the PCP in M-MIMO

(ii) To ensure a trade-off between the use of OPSs and the mitigation of the PCP, the proposed S3SR is automatically updated based on the SR, which is a new approach that can be exploited to deal with the PCP

(iii) Significant improvements are achieved and proved through simulation results

To the best of our knowledge, no prior work has analyzed the S3SR strategy based on the consolidation of the SPR-MBDP and the WGC-PA strategies in the multicell M-MIMO system.

This paper is organized as follows: Section 2 presents the system model adopted in this work. Section 3 sheds light on the communication schemes and reviews of the influence of PCP upon M-MIMO systems; therefore, Section 4 presents the proposed S3SR strategy. Section 5 provides the simulation results which prove the effectiveness of the proposed S3SR strategy compared to WGC-PA, SPR-MBDP, and the conventional strategy [1]. Finally, Section 6 summarizes the prime contribution of this work.

To facilitate the numerical analysis, the following terminologies are considered: the cardinality of a set, the transpose conjugate (i.e., Hermitian), the transpose, the conjugate, and the pseudoinverse are, respectively, denoted as card \{\cdot\}, (\cdot)^H, \cdot^\dagger, and \cdot^{-1}.
\((\cdot)^T\), \((\cdot)^*\), and \((\cdot)^\dagger\). The \(K \times K\) identity matrix is denoted as \(I_K\), while \((j, k)\) refers to the \(k^{th}\) user of the \(j^{th}\) cell. The Frobenius norm of a matrix is denoted as \(\|\cdot\|_F\).

2. System Model

This section is aimed at presenting the system model, which is adopted in this paper. Hence, a multicell M-MIMO system is considered which is composed of \(L\) hexagonal cells. Each cell is equipped with a centered BS of \(M\) antennas, and \(K < M\) single-antenna users [1, 9, 15]. The channel vector that links the \(k^{th}\) user of the \(l^{th}\) cell (i.e., \((l, k)\)) to the BS of the \(j^{th}\) cell can be expressed as follows:

\[
g_{(l,k),j} = h_{(l,k),j}^* \beta_{(l,k),j}^\dagger
\]

where \(h_{(l,k),j} \in \mathbb{C}^{M \times 1}\) denotes the small scale fading (SSF) vector. The SSF vectors of the \(K\) users are statistically independent and follow a complex Gaussian distribution of zero-mean vector and a covariance matrix \(I_M\); therefore, \(h_{(l,k),j} \sim \mathcal{CN}(0, I_M)\). The second coefficient \(\beta_{(l,k),j}\) refers to the large-scale fading (LSF) coefficient, which accounts for the path loss and the shadow fading. The LSF in (1) can be expressed [2, 4, 9, 10, 15, 18] as follows:

\[
\beta_{(l,k),j} = \frac{z_{(l,k),j}}{r_{(l,k),j}^{\alpha}},
\]

where \(r_{(l,k),j}\) denotes the distance between \((l, k)\) and the BS of the \(j^{th}\) cell, \(z_{(l,k),j}\) refers to the shadow-fading factor, and it has a log-normal distribution (i.e., \(10\log_{10}(z_{(l,k),j})\)) that follows a zero-mean Gaussian distribution, which is characterized by a standard deviation \(\sigma_{shadow}\), while \(R\) denotes the cell radius and \(\alpha\) represents the path loss exponent. The channel matrix between the \(K\) users of the \(l^{th}\) cell to the BS of the \(j^{th}\) cell can be expressed as

\[
G_{(l,j)} = \begin{bmatrix} g_{(l,1),j} & g_{(l,2),j} & \cdots & g_{(l,K),j} \\ h_{(l,1),j} & h_{(l,2),j} & \cdots & h_{(l,K),j} \end{bmatrix} D_{(j)}^{1/2},
\]

where \(D_{(j)} = \text{diag}(\beta_{(l,1),j}, \beta_{(l,2),j}, \cdots, \beta_{(l,K),j})\) is a diagonal matrix, which includes the LSF coefficients of the \(K\) users of the \(l^{th}\) cell to the BS of the \(j^{th}\) cell.

2.1. The Pilot Sequences and Their Utility. In order to shed light on the concept of the OPSs, which is one of the fundamentals of the wireless communication systems, let’s assume that a user \(k\) of cell \(j\) is assigned with a pilot sequences \(\phi_k = [\phi_k^1, \phi_k^2, \cdots, \phi_k^\tau] \in \mathbb{C}^{1 \times \tau}\) of length \(\tau\), whereas a user \(k’\) of the same cell is assigned with the pilot sequences \(\phi_{k’}\) of length \(\tau’\). To estimate the channel vector of the two users at the BS of cell \(j\), the two users uplink their corresponding pilot sequences; therefore, the BS of cell \(j\) exploits the received pilot signals and the orthogonality condition to estimate the channel of each user. This is through correlating the received pilot signals with the pilot sequences. The orthogonality condition is expressed as follows:

\[
\phi_k^\dagger \phi_{k’}^H = \delta_{kk’},
\]

where \(\delta_{kk’}\) is the delta function, which is expressed as

\[
\delta_{kk’} = \begin{cases} 1, & \text{if } k = k’; \\ 0, & \text{otherwise}. \end{cases}
\]

Hence, to avoid the PCP, each user—of the overall cells—should be assigned with an OPS, which would require \(KL\) OPSs to serve the \(KL\) users of the considered \(L\) cells. Due to the problem of the pilot overhead, using that number of pilots is not allowed.

2.2. The Problem of Pilot Overhead. Under the block-fading channel model, based time-division-duplex (TDD) protocol, the communication between the users and their supported BSs can be split into coherence time intervals (i.e., frames) (Figure 2), where the channel responses remain constant. Therefore, at the beginning of each frame, the users uplink their pilot sequences of length \(\tau\). To avoid the intracell interference, the users that belong to the same cell are allocated with OPSs, which requires \(K = \tau\) orthogonal pilot sequences.

To reduce the PCP, extra orthogonal pilot sequences can be used at the expense of shrinking the bandwidth left for data transmission, this problem is what we call the problem of pilot overhead, and it is expressed by the ratio \(\mu = \tau / T\). In other words, increasing the number of OPSs—used for channel training—leads to reduce the subframe of the data transmission (Figure 2).

2.3. Asymptotic Orthogonality. Let \(g_{(l,k),j}^H g_{(l,k’),j}^H\) be two \(M \times 1\) vectors, which are mutually independent, where \(g_{(l,k),j} \sim \mathcal{CN}(0, \beta_{(l,k),j}^2 I_M)\) and \(g_{(l,k’),j} \sim \mathcal{CN}(0, \beta_{(l,k’),j}^2 I_M)\). Hence, the asymptotic orthogonality condition can be expressed as in [3, 9, 19] as

\[
\lim_{M \to \infty} \frac{g_{(l,k),j}^H}{M} g_{(l,k’),j} \to \beta_{(l,k),j}^\dagger,
\]

\[
\lim_{M \to \infty} \frac{g_{(l,k’),j}^H}{M} g_{(l,k),j} \to 0.
\]

Figure 2: Illustration of a frame, which contains two parts: the first one in the left is dedicated to the pilot transmission, whereas the second part is reserved for data transmission.
The given of (6) is the foothold of the M-MIMO technology, where the channel vectors become mutually orthogonal for \( M \rightarrow \infty \). Hence, the LSF can be exploited to describe the channel of each user, instead of using the SSFs, which varies quickly and hard to be exploited.

3. Communication Schemes and the Effect of Pilot Contamination

This section is aimed at analyzing the different processes of the wireless communication between the users and their respective BSs, therefore showing the main influence of the PCP on the fulfilled performance of the M-MIMO. Thus, the widely adopted block-fading channel model, based TDD protocol, is considered [4, 9, 15, 20], where the channel responses remain constant. Thus, three specific phases can be distinguished: channel estimation through uplink training, data detection, and data precoding. Therefore, by adopting the system model presented in Section 2, these three phases are provided in the following subsections.

3.1. Training Phase and Channel Estimation. During the training phase, the users of each cell, synchronously, uplink their pilot sequences of length \( \tau \). Thus, the received pilot signal \( y_j^\tau \in \mathbb{C}^{M \times \tau} \) at the BS of the \( j^{th} \) cell can be expressed as

\[
y_j^\tau = \sqrt{\rho_p} \sum_{k=1}^{K} g_{(j,k)} \phi_k + n_j^\tau,
\]

where \( \rho_p \) represents the transmitted power, which is used by each user to uplink its pilot sequences; \( \phi_k \in \mathbb{C}^{1 \times \tau} \) is the pilot sequences of the \( k^{th} \) user, where the orthogonality condition \( \phi_k \phi_k^H = \delta_{k,K} \) is satisfied [21, 22]; \( n_j^\tau \in \mathbb{C}^{M \times \tau} \) is an additive white Gaussian noise (AWGN) at the antennas BS of the \( j^{th} \) cell. Therefore, by correlating the received pilot signal \( y_j^\tau \), with the local pilot sequences \( \phi_k \), the estimated channel of \( \langle j, k \rangle \) at the BS of the \( j^{th} \) cell is expressed as

\[
\hat{g}_{(j,k)} = \frac{y_j^\tau \phi_k^H}{\sqrt{\rho_p}} = \sum_{l=1}^{L} g_{(j,k)} \phi_k + \sum_{k=1}^{K} \sum_{k'=1 \neq k}^{K} g_{(k',k)} \phi_k \phi_k^H + \frac{n_j^\tau \phi_k^H}{\sqrt{\rho_p}} = g_{(j,k)} + \sum_{l=1}^{L} g_{(j,k)} + \frac{n_j^\tau \phi_k^H}{\sqrt{\rho_p}}.
\]

3.2. Data Detection. Users of each cell uplink their data symbols. Accordingly, the user \( \langle j, k \rangle \) uplink its data symbols \( x_j^\mu \), where \( E \left[ |x_j^\mu|^2 \right] = 1 \). Therefore, the received data signal \( y_j^\mu \) at the BS of the \( j^{th} \) cell can be expressed as

\[
y_j^\mu = \sqrt{\rho_p} \sum_{l=1}^{L} \sum_{k=1}^{K} g_{(j,k)} x_j^\mu + n_j^\mu,
\]

where \( n_j^\mu \) denotes the AWGN at the antennas BS of the \( j^{th} \) cell. Therefore, based on the matched-filter (MF) detector (i.e., maximum ratio combining) [23], the estimated data symbols \( \hat{x}_{(j,k)}^\mu \) of the user \( (j, k) \), at the BS of the \( j^{th} \) cell, can be computed as follows:

\[
\hat{x}_{(j,k)}^\mu = \frac{\hat{g}_{(j,k)}^H y_j^\mu}{\sqrt{\rho_p}} = \sqrt{\rho_p} g_{(j,k)}^H ( \hat{g}_{(j,k)} y_j^\mu + \hat{n}_{(j,k)}^\mu + \hat{e}_{(j,k)}^\mu ),
\]

where \( e_{(j,k)}^\mu \) arranges the interference and the noisy components, and it is expressed as

\[
e_{(j,k)}^\mu = \sqrt{\rho_p} g_{(j,k)}^H \sum_{l=1}^{L} g_{(j,k)}^H x_j^\mu + \sum_{l=1}^{L} g_{(j,k)}^H n_j^\mu + \sum_{l=1}^{L} g_{(j,k)}^H n_j^\mu.
\]

Therefore, the uplink SINR of \( \langle j, k \rangle \) can be expressed as

\[
\text{SINR}_{(j,k)} = \frac{\left| g_{(j,k)} \right|^2}{\sum_{l=1}^{L} \left| g_{(j,k)} \right|^2 + \left| e_{(j,k)}^\mu \right|^2 / \rho_p}.
\]

Hence, the corresponding average uplink achievable rate is then expressed as

\[
C_{(j,k)} = (1 - \mu) E \left[ \log_2 \left( 1 + \text{SINR}_{(j,k)} \right) \right],
\]

where \( 0 < \mu < 1 \) denotes a parameter that evaluates the decrease of the spectral efficiency, which is caused by the problem of pilot overhead. It is well seen from (12) that the SIN \( R_{(j,k)} \) saturates (i.e., tends to constant value however how larger the number of antennas deployed at the BSs) when the number of the antennas BSs tends toward the infinity, which is inconsistent with the main goal of the M-MIMO technology, which is aimed at enhancing the system throughput; this is by using a massive number of antennas at the BSs.

3.3. Data Precoding. The data symbols of the users are precoded at the BSs before being transmitted toward the users. Accordingly, based on the MF precoder, the data symbols \( x_j^d \) (\( E \left[ |x_j^d|^2 \right] = 1 \) of \( \langle j, k \rangle \) are precoded at the BS of the \( j^{th} \) cell. This is based on the following precoder:

\[
n_{(j,k)} = \frac{\hat{g}_{(j,k)}^H}{\sqrt{\rho_p}}.
\]
where $\lambda_{(j,k)} = \overline{g}_{(j,k);i}^T \overline{g}_{(j,k);i}^*$ is a normalization factor. Therefore, the received data signal $y_{(j,k)}^d$ at $(j, k)$ can be expressed as

\[
y_{(j,k)}^d = \sqrt{P_d} \sum_{l=1}^{L} g_{(j,k);l}^* \lambda_{(j,k)} x_{(j,k)} + n_j^d
\]

\[
= \sqrt{P_d} g_{(j,k);j}^T \frac{g_{(j,k);j}^*}{\sqrt{\lambda_{(j,k)}}} x_{(j,k)}^d + n_j^d
\]

\[
= \sqrt{P_d} g_{(j,k);j}^T \frac{g_{(j,k);j}^*}{\sqrt{\lambda_{(j,k)}}} x_{(j,k)}^d + \sqrt{P_d} \sum_{l=j}^{L} g_{(j,k);l}^T \frac{g_{(j,k);l}^*}{\sqrt{\lambda_{(j,k)}}} x_{(j,k)}^d + n_j^d
\]

\[
= \sqrt{P_d} g_{(j,k);j}^T \frac{g_{(j,k);j}^*}{\sqrt{\lambda_{(j,k)}}} x_{(j,k)}^d + \sqrt{P_d} \sum_{l=j}^{L} g_{(j,k);l}^T \frac{g_{(j,k);l}^*}{\sqrt{\lambda_{(j,k)}}} x_{(j,k)}^d + \frac{\sqrt{P_d}}{\sqrt{\lambda_{(j,k)}}} \left( \frac{n_j^d \phi_k^H}{\sqrt{P_P}} \right)^* x_{(j,k)}^d + \sqrt{P_d} \sum_{l=j}^{L} g_{(j,k);l}^T \frac{g_{(j,k);l}^*}{\sqrt{\lambda_{(j,k)}}} x_{(j,k)}^d
\]

Thus, the downlink SINR of $(j, k)$ can be expressed as follows:

\[
\text{SINR}_{(j,k)}^d = \frac{\left| g_{(j,k);j}^T g_{(j,k);j}^* \right|^2 / \lambda_{(j,k)}}{\sum_{l=j}^{L} \left| g_{(j,k);l}^T g_{(j,k);l}^* \right|^2 / \lambda_{(j,k)} + \frac{\left| \epsilon_{(j,k)}^d \right|^2}{P_d} + \sum_{l=j}^{L} \beta_{(j,k);l}^2 / \lambda_{(j,k)}}.
\]

Therefore, the average downlink achievable rate can be expressed similarly to (13) as follows:

\[
C_{j,k}^d = (1 - \mu) E \left\{ \log_2 \left( 1 + \text{SINR}_{(j,k)}^d \right) \right\}.
\]

It is well seen from (12) and (17) that the fulfilled SINR is saturated. Specifically, by increasing the number of antennas $M$ toward the infinity, the fulfilled SINR remains constant; hence, it is said to be saturated in a fixed value, which is inconsistent with the main target of M-MIMO. Thus, to benefit from the M-MIMO, the increases of $M$ should be accompanied by the increases of the fulfilled SINR. It should be noticed that the saturation of the SINR, in the asymptotic regime (i.e., $M \rightarrow \infty$), is mainly caused by the PCP.

4. Proposed S3SR

Since the LSF coefficients are not the same for different users, and the QoS of a user depends on its LSF coefficient (e.g., (12)); the users of each cell benefit from an unfair QoS. Accordingly, the users of each cell are separated into two groups, which are called the center and edge users; therefore, to control the use of the OPSs while mitigating the PCP among the users of each group, a BS controller is exploited to compute the SR, which is employed to make better use of the OPSs while making the decision on which strategy of the two strategies (i.e., SPR-MBDP or the WGC-PA) is desired to be used to mitigate the PCP upon the users of each group (i.e., center and edge users). This process is described in detail in the following subsections.

4.1. User Classification. The users of each cell are randomly distributed and have different LSF coefficients. Thus, by exploiting their respective LSF coefficients, the $K$ users of the $j$th cell can be classified at the aim of the following threshold:

\[
\rho_j = \lambda \sum_{k=1}^{K} \beta_{(j,k);j}^2 \quad \forall j = 1, \cdots, L,
\]

where $0 < \lambda < 1$ denotes a weighted parameter, which is exploited to adjust $\rho_j$ regarding the system configuration. Thus, for $\lambda = 0$, the $K$ users are viewed as center users, while they are considered as edge users for $\lambda = 1$. Therefore, the user $(j, k)$ is classified based on its LSF coefficient as follows:

\[
\beta_{(j,k);j}^2 > \rho_j \quad \rightarrow \quad \left\{ \begin{array}{ll}
\text{Yes, Isacenteruser,} & \\
\text{No, Isanedgeuser.} &
\end{array} \right.
\]

4.2. Supervision Rule. In order to construct the SR, the number of edge and center users is, separately, computed in each cell.
Thus, the number of edge $k_{j,e}$ and center $k_{j,c}$ users of the $j$th cell is computed based on (19) and (20) as follows:

$$
\begin{align*}
{k_{j,e}} &= \text{card} \left\{ k, \beta^2_{j(k,j)} > \rho \right\}, \quad k = 1, 2, \ldots, K, \\
{k_{j,c}} &= \text{card} \left\{ k, \beta^2_{j(k,j)} \leq \rho \right\}, \quad k = 1, 2, \ldots, K,
\end{align*}
$$

and so on for the $L$ cells (i.e., $\forall j = 1, \ldots, L$). Therefore, each BS compute and transmit its number of center and edge users to the BS controller, which is in charge of computing the total number of edge users $K^{T_{\text{tot},e}}$ and center users $K^{T_{\text{tot},c}}$, which are wandering in the $L$ cells, that is as follows:

$$
\begin{align*}
K^{T_{\text{tot},e}} &= \sum_{j=1}^{L} k_{j,e}, \\
K^{T_{\text{tot},c}} &= \sum_{j=1}^{L} k_{j,c}.
\end{align*}
$$

Hence, the derived SR at the BS controller can be expressed as follows:

$$
K^{T_{\text{tot},e}} > K^{T_{\text{tot},c}} \quad \rightarrow \quad \begin{cases} 
\text{Yes, WGC-PA to edge users, while SPR-MBDP to center users}, \\
\text{No, WGC-PA to center users, while SPR-MBDP to edge users.}
\end{cases}
$$

The number of OPs required in the conventional SPR-MBDP is computed as follows:

$$
K_{\text{SPR}} = K^{T_{\text{tot},e}} + \max \left\{ k_{j,e}, j = 1, 2, \ldots, L \right\}.
$$

To clarify the benefits of the proposed S3SR strategy, let us analyze the example given in Figure 3, which depicts a set of three hexagonal cells (i.e., $L = 3$). Each cell is equipped with eight, randomly distributed, users (i.e., $K = 8$). Thus, Figure 3(a) shows the conventional SPR-MBDP, where the edge users are assigned with $K^{T_{\text{tot},e}} = 11$ OPs, while the center users are obliged to reuse $\max \left\{ k_{j,e}, j = 1, 2, 3 \right\} = 5$ OPs. Accordingly, $K_{\text{SPR}} = 16$ is the number of the OPs, which is required by the SPR-MBDP to overcome the PCP. However, from Figure 3(a), it is well recognized that the PCP remains upon the center users. Figure 3(b) depicts the conventional WGC-PA strategy, which is aimed at reducing the PCP upon the users that employ the same OPs; however, for a large number of users per cell, the WGC-PA tends to perform like a random pilot assignment strategy.
The two sides of the S3SR are shown in Figures 3(c) and 3(d). Specifically, in Figure 3(c) $K_{Tot,c} = 14 > K_{Tot,e} = 10$, consequently, the S3SR strategy applies, simultaneously, the SPR-MBDP strategy to the edge users, whereas the WGC-PA is dedicated to the center users. Therefore, the number of the OPs, which is required in this case based on (24) is $K_{sw} = 10 + 5 = 15 = K_{SPR}$. The second side of S3SR is depicted in Figure 3(d), where $K_{Tot,e} = 17 > K_{Tot,c} = 7$, and according to (23), the S3SR strategy applies simultaneously the SPR-MBDP strategy to the center users, while the WGC-PA strategy is applied to the edge users. Thus, the number of the required OPs is only $K_{sw} = 7 + 6 = 13$ OPs. Still referring to Figure 3(d), the use of the conventional SPR-MBDP will need the use of $K_{SPR} = 17 + 3 = 20$ OPs $\gg K_{sw} = 13$. Thus, the S3SR sacrifices with less OPs compared to the SPR-MBDP.

Since the S3SR supports two cases (23), the same analysis can be, symmetrically, done for both cases. The following analysis is restricted to one side of the S3SR, where $K_{Tot,c} > K_{Tot,e}$ (e.g., Figure 3(d)).

4.4. The Allocation of the OPs. Let $\Phi_{sw} \in \mathbb{C}^{K_{sw} \times L}$ be the set of the OPs (i.e., $\Phi_{sw} \Phi_{sw}^H = I_{K_{sw}}$), which is required by the S3SR strategy, where $K_{sw} = K_{Tot,c} + \max \{k_{j,c} : j = 1, \ldots, L\}$. Therefore, $\Phi_{sw}$ can be expressed as

$$\Phi_{sw} = \begin{bmatrix} \Phi_{c1}^T & \Phi_{c2}^T & \cdots & \Phi_{cL}^T \end{bmatrix}^T,$$

(26)

where $\Phi_{sw,c} \in \mathbb{C}^{K_{sw,c} \times L}$ is the set of the OPs, which is allocated to the $K_{Tot,c}$ users, while $\Phi_{sw,e} \in \mathbb{C}^{K_{sw,e} \times L}$ denotes the OPs allocated to the edge users. Since SPR-MBDP is selected to be applied to the center users, the set $\Phi_{sw,c}$ must be shared by the $K_{Tot,c}$ users of the $L$ cells. Therefore, it can be split as follows:

$$\Phi_{sw,c} = \begin{bmatrix} \Phi_{c1}^T & \Phi_{c2}^T & \cdots & \Phi_{cL}^T \end{bmatrix}^T,$$

(27)

where $\Phi_{c,j} \in \mathbb{C}^{K_{sw,c} \times L}$ denotes the subset of the OPs, which are allocated to the $k_{j,c}$ center users of the $j$th cell, and it can be expressed as follows:

$$\Phi_{c,j} = \begin{bmatrix} \Phi_{j1}^T & \Phi_{j2}^T & \cdots & \Phi_{jL}^T \end{bmatrix}^T,$$

(28)

where $\Phi_{j,k} \in \mathbb{C}^{r \times 1}$ denotes the pilot sequences of the $k$th center user of the $j$th cell. On the other hand, the set of the OPs $\Phi_{sw,e}$, which is reused by the edge users (i.e., $k_{e,j}, \forall j \in \{1, \cdots, L\}$), can be expressed as

$$\Phi_{sw,e} = \begin{bmatrix} \Phi_{e1}^T & \Phi_{e2}^T & \cdots & \Phi_{eL}^T \end{bmatrix}^T,$$

(29)

where $\Phi_{e,j} \in \mathbb{C}^{r \times 1}$ denotes the pilot sequences assigned to the $k$th edge user of the $j$th cell.

4.5. The WGC-PA Applied to the Edge Users. A set of max $\{k_{j,e} : j = 1, 2, \cdots, L\}$ OPs is, fully, reused by the edge users of the $L$ cells. Hence, WGC-PA [15] is employed to manage the assignment of the OPs to the edge users, therefore reducing the strength of the PCP upon these edge users. Specifically, the strength of the PCP between two edge users $(j, k)$ and $(l, k')$ (i.e., where $k \neq k'$ and $j \neq l$) that employ the same pilot sequence, i.e., $\Phi_{e,j} = \Phi_{e,l}$, can be expressed as

$$\zeta_{(j,k),(l,k')} = \frac{\beta^2(j,k,j)}{\beta^2(j,k,j)} + \frac{\beta^2(j,k,j)}{\beta^2(j,k',j)},$$

(30)

Notice that as $\zeta_{(j,k),(l,k')}$ increases, the strength of the PCP between the users $(j, k)$ and $(l, k')$ increases too and vice versa. Therefore, the WGC-PA is aimed at reusing the same OPs for the edge users of different cells, where $\zeta_{(j,k),(l,k')} \forall \{k \neq k', j \neq l\}$ is minimized. To apply the WGC-PA strategy, it is desired first to define an interference graph $Gr = (E, V)$, where $V$ denotes the vertexes of Gr, while $E$ denotes the severity of the PCP upon the edge users. Hence, $Gr$ can be expressed similarly to [15] as follows:

$$Gr = (V, E): \begin{cases} V = \{(j, k)\}, \\ E = \{(j,k),(l,k'), j \neq l\}, \end{cases}$$

(31)

where $(j, l) = \{(1, 2, \cdots, L), k = 1, 2, \cdots, k_{j,e} \}$ and $k' = 1, 2, \cdots, k_{j,e}$.

4.6. Channel Estimation. During the training phase, the received pilot signal $Y_j^p$ at the BS of the $j$th cell can be expressed as

$$Y_j^p = \sqrt{p_j} \left( \sum_{l=1}^{L} G_{j,l}^c \Phi_{sw,c} + \sum_{l=1}^{L} G_{j,l}^e \Phi_{sw,e} \right) + N_j^p,$$

(32)

where $G_{j,l}^c \in \mathbb{C}^{M_{x} \times \max \{k_{c,j}\}}$ and $G_{j,l}^e \in \mathbb{C}^{M_{x} \times K_{sw,e}}$ denote, respectively, the channel matrices of the edge and center users, while $N_j^p \in \mathbb{C}^{L_{x} \times 1}$ is the AWGN at the antennas BS of the $j$th cell. It should be noticed that, only, the first $k_{j,e}$ rows of $\Phi_{sw,e}$ are considered when $k_{j,e} < \max \{k_{j,e} : i = 1, \cdots, L\}$.

Since the SPR-MBDP strategy serves the center users, their corresponding channels will be, perfectly, estimated. On the other hand, the WGC-PA strategy is dedicated to the edge users of the overall cells. Thus, the strength of the PCP is minimized, upon the edge users. Therefore, the estimated channels at the BS of the $j$th cell can be
expressed as follows:

\[
\begin{align*}
\mathcal{G}_{e,j}^c &= \frac{1}{\sqrt{P_p}} Y_p^c \varphi_{e,j}^c = \mathcal{G}_{e,j}^c + \frac{N_p}{\sqrt{P_p}} \varphi_{e,j}^c, \\
\bar{G}_{e,j}^c &= \frac{1}{\sqrt{P_p}} Y_p^c \varphi_{e,j}^c = \mathcal{G}_{e,j}^c + \sum_{l} G_{e,j}^c + \frac{N_p}{\sqrt{P_p}} \varphi_{e,j}^c.
\end{align*}
\]

(33)

Therefore, based on (4), the noisy components in (33), i.e., \( (N_p^p/\sqrt{P_p})\varphi_{e,j}^c \) and \( (N_p^p/\sqrt{P_p})\varphi_{e,j}^c \), can be canceled. This is by increasing the number of the antennas at the BSs (i.e., \( M \rightarrow \infty \)). Notice that only the first \( k_{e,j} \) column of \( \varphi_{e,j}^c \) are considered (33) if \( k_{e,j} < \max \{ k_{e,i}, i = 1, \ldots, L \} \).

4.7. Data Detection. Simultaneously, the \( K \) users of each cell uplink their data symbols. Thus, the received data signals at the BS of the \( j \)th cell can be expressed as follows:

\[
Y_j^u = \sqrt{P_u} \sum_{l=1}^{L} \left( \mathcal{G}_{e,j}^c x_{j}^{u,c} + G_{e,j}^c x_{j}^{u,e} \right) + N_j^u,
\]

(34)

where \( P_u \) denotes the transmit power during the uplink phase, while \( x_{j}^{u,c} = [x_{j}^{u,c}(1), x_{j}^{u,c}(2), \ldots, x_{j}^{u,c}(L)]^T \) and \( x_{j}^{u,e} = [x_{j}^{u,e}(1), x_{j}^{u,e}(2), \ldots, x_{j}^{u,e}(L)]^T \) represent, respectively, the data symbols of the \( k_{e,j} \) center users and those of the \( k_{e,j} \) edge users of the \( j \)th cell. The AWGN at the antennas BS of the \( j \)th cell is expressed by \( N_j^u \).

For data detection, let us adopt a low-complexity MF detector (i.e., \( D_{e,j}^c \in \mathcal{C}^{K \times M} \) for the center users and \( D_{e,j}^e \in \mathcal{C}^{K \times M} \) for the edge users), therefore, the detected data symbols can be expressed, separately, for the two groups as follows:

4.7.1. For the Center Users.

\[
\bar{x}_{j}^{u,ce} = D_{e,j}^c Y_j^u = \left( \mathcal{G}_{e,j}^c + \frac{N_p}{\sqrt{P_p}} \varphi_{e,j}^c \right) \left( \sqrt{P_u} \sum_{l=1}^{L} G_{e,j}^c x_{j}^{u,e} + N_j^u \right)
\]

\[
= \left( \mathcal{G}_{e,j}^c \right)^H G_{e,j}^c x_{j}^{u,e} + \text{Inter}_{e,j}^{u,ce},
\]

where \( \text{Inter}_{e,j}^{u,ce} \) denotes the interference terms, and it can be expressed as

\[
\text{Inter}_{e,j}^{u,ce} = \left( \mathcal{G}_{e,j}^c \right)^H \left( \sqrt{P_u} \sum_{l=1}^{L} G_{e,j}^c x_{j}^{u,e} + N_j^u \right) + \frac{N_p^p}{\sqrt{P_p}} \varphi_{e,j}^c.
\]

(36)

Therefore, the detected data symbols \( \bar{x}_{e,j}^{u,ce} \) of the \( k \)th center user wandering in the \( j \)th cell can be expressed as follows:

\[
\bar{x}_{e,j}^{u,ce} = \sqrt{P_u} \left( g_{e,j}^c \right)^H g_{e,j}^c x_{e,j}^{u,ce} + \text{Inter}_{e,j}^{u,ce},
\]

(37)

where \( \text{Inter}_{e,j}^{u,ce} \) refers to the corresponding interference of the \( k \)th center user in the \( j \)th cell. Therefore, the uplink SINR of the \( k \)th center user in the \( j \)th cell is

\[
\text{SINR}_{e,j}^{u,ce} = \frac{\rho_j \left( g_{e,j}^c \right)^H g_{e,j}^c x_{e,j}^{u,ce} + \text{Inter}_{e,j}^{u,ce}}{\left| \text{Inter}_{e,j}^{u,ce} \right|^2} \xrightarrow{\rho_j \rightarrow 0} \infty.
\]

(38)

Based on (6), the denominator of (38) tends toward zero. Consequently, the SINR_{e,j}^{u,ce} is significantly boosted, compared to the fulfilled SINR in the conventional (i.e., (12)). Therefore, the average uplink achievable rate of the \( k \)th center user of the \( j \)th cell, can be expressed similarly to [9] as follows:

\[
C_{e,j}^{u,ce} = \frac{1}{U} K \sqrt{\frac{N_p}{P_p}} \left( \log_2 \left( 1 + \text{SINR}_{e,j}^{u,ce} \right) \right).
\]

(39)

4.7.2. For the Edge Users.

\[
\bar{x}_{j}^{u,ce} = D_{e,j}^e Y_j^u = \left( G_{e,j}^e + \frac{N_p}{\sqrt{P_p}} \varphi_{e,j}^e \right) \left( \sqrt{P_u} \sum_{l=1}^{L} G_{e,j}^e x_{j}^{u,e} + N_j^u \right)
\]

\[
= \left( G_{e,j}^e \right)^H G_{e,j}^e x_{j}^{u,e} + \text{Inter}_{j}^{u,ce}.
\]

(40)

Similar to (37), \( \text{Inter}_{j}^{u,ce} \) denotes the corresponding interferences. Then, the detected data symbol of the \( k \)th edge user wandering in the \( j \)th cell is expressed as follows:

\[
\bar{x}_{e,j}^{u,ce} = \sqrt{P_u} \left( g_{e,j}^e \right)^H g_{e,j}^e x_{e,j}^{u,ce} + \text{Inter}_{e,j}^{u,ce},
\]

(41)

where \( \text{Inter}_{e,j}^{u,ce} \) denotes the interference term, which corresponds to the \( k \)th edge user of the \( j \)th cell. Therefore,
the uplink SINR of the $k^{th}$ edge user in the $j^{th}$ cell is expressed as follows:

$$\text{SINR}^{u,c}_{(j,k)} = \frac{p_u \left( g^c_{(j,k),j} \right)^H \mathbf{g}^c_{(j,k),j} }{ p_u \sum_{l \neq j} \left( g^c_{(j,k),j} \right)^H \mathbf{g}^c_{(j,k),j} } + \left| \text{int}^{u,c}_{(j,k)} \right|^2 \sum_{l \neq j} \left( \mathbf{g}^c_{(l,k),j} \right)^2 .$$ (42)

Even though (42) looks like the conventional strategy (12), they are not similar. Due to the use of the WGC-PA, which is aimed at reducing the interference components in the denominator of (42), hence, the fulfilled SINR is enhanced for the edge users. Therefore, the average uplink achievable rate can be expressed similarly to (39) by replacing $\text{SINR}^{u,c}_{(j,k)}$ by $\text{SINR}^{u,c}_{(j,k)}$.

4.8. Data Precoding. Before transmitting the data symbols toward the target user, each BS precodes its local data symbols. This is based on a constructed precoder (e.g., MF precursor which is considered herein). Thus, the center users will receive their respective data symbols based on SPR-MBDP, while the WGC-PA is used for edge users.

4.8.1. “For Center Users. From the first line of (33), it is well recognized that the PCP upon the center users is completely canceled. Accordingly, the BSs become able to partially estimate the channel matrices of the center users of the adjacent cells, which are then exploited to find out the subspace spanned by the intercell interference. Therefore, to project data symbols through the null space of the subspace spanned by the intercell interference. This process is explained in more detail in the following.

Let $\mathbf{A}_j \in \mathbb{C}^{(K^{\text{int}}_j-k_{j,r}) \times M}$ be the matrix that arranges the estimated channels of intercell interferences, which are generated by the $(K^{\text{int}}_j - k_{j,r})$ center users of the $L$ cells to the BS of the $j^{th}$ cell. Thus, the matrix $\mathbf{A}_j$ can be expressed as follows:

$$\mathbf{A}_j = \left[ \mathbf{C}^c_{(j,1)}, \cdots, \mathbf{C}^c_{(j,J-1)}, \mathbf{C}^c_{(j,J+1)}, \cdots, \mathbf{C}^c_{(j,K)} \right]^T .$$ (43)

Therefore, to find out the directions of the null subspace spanned by $\mathbf{A}_j$, the SVD tool will be exploited. Thus, the matrix of intercell interferences can be expressed as

$$\mathbf{A}_j = U_j \Sigma_j V_j^H ,$$ (44)

where $U_j \in \mathbb{C}^{(K^{\text{int}}_j-k_{j,r}) \times (K^{\text{int}}_j-k_{j,r})}$ and $V_j \in \mathbb{C}^{M \times M}$ denote, respectively, the left and the right singular vector, while

$$\Sigma_j \in \mathbb{C}^{(K^{\text{int}}_j-k_{j,r}) \times M}$$

refers to the matrix that comprises the singular values and it can be expressed as

$$\Sigma_j = \begin{bmatrix} \mathbf{\Sigma}_j, & \mathbf{O}_{r_j \times (M-r_j)} \\ \mathbf{O}_{(K^{\text{int}}_j-k_{j,r}) \times r_j}, & \mathbf{O}_{(K^{\text{int}}_j-k_{j,r}) \times (M-r_j)} \end{bmatrix} ,$$ (45)

where $r_j = \text{rank}(\mathbf{A}_j)$ denotes the rank of $\mathbf{A}_j$, and $\mathbf{\Sigma}_j = \text{diag}\{ \sigma_{j,1}, \sigma_{j,2}, \cdots, \sigma_{j,\beta} \}$ is an $r_j \times r_j$ diagonal matrix of the singular values, which are arranged as follows:

$$\sigma_{j,1} \geq \sigma_{j,2} \geq \cdots \geq \sigma_{j,\beta} > 0 .$$ (46)

Since $M$ is huge compared to the number of the users, the condition $M > r_j$ is necessarily satisfied. Hence, the null subspace of $\mathbf{A}_j$ is spanned by the column of a matrix $\mathbf{B}_j \in \mathbb{C}^{M \times (M-r_j)}$, which represents the submatrix of $V_j$ and it can be expressed as

$$\mathbf{B}_j = \left[ v_{(j,r_1+1)}, v_{(j,r_2+1)}, \cdots, v_{(j,M)} \right] .$$ (47)

Thanks to the SVD tool, the null space of interference is arranged in $\mathbf{B}_j$, which can be exploited to construct the projection operator $\mathbf{P}^u_{\mathbf{B}_j}$ as follows:

$$\mathbf{P}^u_{\mathbf{B}_j} = \mathbf{B}_j \mathbf{B}_j^H ,$$ (48)

where $\mathbf{B}_j \in \text{Null}(\mathbf{A}_j)$; therefore, by projecting the conventional MF precoder (14) onto the null subspace of $\mathbf{A}_j$, the new precoder can be expressed as follows:

$$\mathbf{A}^{\text{SRR}}_{(j,k)} = \mathbf{P}^u_{\mathbf{B}_j} \mathbf{G}^c_{(j,k)} \mathbf{P}^H_{\mathbf{B}_j} .$$ (49)

Therefore, $\forall (j \neq l) = 1, \cdots, L$, we have

$$\mathbf{A}^{\text{SRR}}_{(j,k)} = 0 \longrightarrow \mathbf{G}^c_{(j,k)} \mathbf{A}^{\text{SRR}}_{(j,k)} = 0 .$$ (50)
Hence, based on the enhanced MF precoder (49), the received data signal $y_{d_j}^{dc}$ at the $k_{th}$ center users wandering in the $j^{th}$ cell can be expressed as follows:

$$y_{d_j}^{dc} = \sqrt{P_d} \sum_{l=1}^{L} (G_{c(j),l})^T \Theta_j^{S3SR} x_{d_j}^{dc} + N_{d_j}^{dc},$$

$$= \sqrt{P_d} (G_{c(j),l})^T P_{B_j} (G_{c(j),l})^* x_{d_j}^{dc} + N_{d_j}^{dc},$$

(51)

where $x_{d_j}^{dc}$ refers to the data symbols that will be transmitted to the $k_{th}$ users, while $N_{d_j}^{dc}$ denotes the AWGN at the antennas of the those $k_{th}$ users. Therefore, the received data signal $y_{d(j,k)}^{dc}$ at the $k^{th}$ center user of the $j^{th}$ cell is expressed as follows:

$$y_{d(j,k)}^{dc} = \sqrt{P_d} \frac{(g_{c(j,k),j})^T B_j^T (g_{c(j,k),j})^*}{\| P_{B_j} G_{c(j,k)}^T \|_F} x_{d(j,k)}^{dc}$$

$$+ \sqrt{P_d} \frac{(g_{c(j,k),j})^T B_j^T n_{d(j,k)}^{dc,*}}{\| P_{B_j} G_{c(j,k)}^T \|_F} x_{d(j,k)}^{dc} + N_{d(j,k)}^{dc},$$

(52)

where $x_{d(j,k)}^{dc}$ and $N_{d(j,k)}^{dc}$ denote, respectively, the respective data symbols of the user $(j,k)$ and the AWGN at its antenna. Thereafter, the downlink SINR of the $k^{th}$ center user in the $j^{th}$ cell is expressed as

$$\text{SINR}_{d(j,k)}^{dc} = \frac{P_d \left\| \frac{(g_{c(j,k),j})^T B_j^T (g_{c(j,k),j})^*}{\|P_{B_j} G_{c(j,k)}^T\|_F} \right\|^2}{\| \Theta_{d(j,k)} \|^2 + \| N_{d(j,k)}^{dc} \|^2},$$

(53)

where $\Theta_{d(j,k)} = \sqrt{P_d} \| (g_{c(j,k),j})^T B_j^T (n_{d(j,k)}^{dc,*}) /\|P_{B_j} G_{c(j,k)}^T\|_F \|$ arranges the interference components. Notice that $E\{ |x_{d(j,k)}^{dc}|^2 \} = 1$ still hold. Therefore, based on (6), the denominator of (53) can be forced towards zero, which proves that the SINR is significantly boosted, and the average downlink achievable rate, in this case, is expressed as

$$C_{d(j,k)}^{dc} = \left( 1 - \frac{K_w}{K} \right) E \{ \log_2 \left( 1 + \text{SINR}_{d(j,k)}^{dc} \right) \},$$

(54)

4.8.2. *For the Edge Users.* Regarding the fact that the estimated channels of the edge users (i.e., second line of (33)) appear like those derived in the conventional strategy (8), hence, the same analysis given in Subsection 3.3 can be similarly done for the edge users. Thus, the downlink SINR of the $k^{th}$ edge user in the $j^{th}$ cell is then:

$$\text{SINR}_{d(j,k)}^{dc} = \frac{\left\| \frac{(g_{c(j,k),j})^T B_j^T (g_{c(j,k),j})^*}{\|P_{B_j} G_{c(j,k)}^T\|_F} \right\|^2}{\sum_{l=1}^{L} \left\| \frac{(g_{c(j,k),j})^T B_j^T n_{d(j,k)}^{dc,*}}{\|P_{B_j} G_{c(j,k)}^T\|_F} \right\|^2} / \rho_d.$$

(55)

Therefore, the average downlink achievable rate for the $k^{th}$ edge user in the $j^{th}$ cell is then expressed as

$$C_{d(j,k)}^{dc} = \left( 1 - \frac{K_w}{K} \right) E \{ \log_2 \left( 1 + \text{SINR}_{d(j,k)}^{dc} \right) \},$$

(56)

where $0 < \mu < 1$ denotes a parameter that evaluates the decrease of the spectral efficiency, which is caused by the problem of pilot overhead.

5. Simulation Results

This section is aimed at evaluating the performance of the proposed S3SR. That is through a set of Monte Carlo simulation applied to the typical multicell structure, which is described in Section 2, where the $K$ users are randomly distributed within each cell. The specific parameters used in this work are listed in Table 1.

Table 1: System settings.

| Number of the cells | $L = 7$ |
|---------------------|---------|
| Number of the antennas BSs | $32 \leq M \leq 512$ |
| Number of the users per cell | $K = 10$ |
| Number of OPs for the WGC-PA [15] | $S = 10$ |
| Adjustment parameter | $0 \leq \lambda \leq 1$ |
| Cell radius | $R = 500 m$ |
| Inner radius | $r = 30 m$ |
| Transmit power $P_0$, $P_a$ | $12$ dBm |
| Transmit power $\rho_d$ | $10$ dBm |
| Path loss exponent | $\alpha = 3$ |
| log-normal shadowing fading | $\sigma_{\text{shadow}} = 8$ dB |
| Pilot overhead parameter | $0 < \mu < 1$, fixed $\mu = 0.01$ |

Figure 4 shows the average uplink achievable rate versus the number of antennas deployed in the BSs. As can be recognized from Figure 4, the superiority of the proposed S3SR is well seen compared to the other strategies (i.e., conventional, SPR-MBDP, and WGC-PA), especially for a large number of antennas $M \geq 256$, where the gap is increased. On the other hand, the fulfilled performance by the SPR-MBDP is too close to those achieved by the WGC-PA, owing to the problem of pilot overhead (Table 2). Furthermore, based on Table 2 where $K_{\text{SPR}} = K_{\text{WGC}} = 41$, we notice that the S3SR performs better than the SPR-MBDP, even if they employ the same number of the OPs, that proves the effectiveness of the proposed S3SR over the conventional strategies.
The average downlink achievable rate versus $M$ is depicted in Figure 5, where the superiority of the S3SR compared to the WGC-PA is started from a large number of antennas; specifically, from $M \geq 256$, the S3SR performs better compared to the other strategies.

In order to show the robustness of the S3SR versus different distribution of the users within cells, which is controlled by the value of $\lambda$, Figure 6 depicts the mean square error (MSE) of the estimated channels versus $\lambda$. The MSE can be expressed as

$$
\text{MSE} = E\left\{ \frac{1}{K} \sum_{j=1}^{L} \sum_{k=1}^{K} \left| \hat{g}_{(j,k),j} - g_{(j,k),j} \right|^2 \right\}.
$$

Thus, from Figure 6, and for different values of $\lambda$, the accuracy of the S3SR is well seen, which is not the case for the SPR-MBDP that loses its accuracy for $\lambda < 0.6$ and performs better for $\lambda > 0.6$. Specifically, the accuracy of the SPR-MBDP is enhanced, only, when the users are regarded as edges (i.e., $\lambda \rightarrow 1$). Indeed, the S3SR keeps approximately the same accuracy for different values of $\lambda$, which proves its robustness over different distributions of the users within cells.

5.1. Computational Complexity. Unlike the previous decontaminating strategies, where the computational complexity is fixed, the complexity of the S3SR depends on the SR in (23). Hence, the complexity of the S3SR is summarized in Table 3, where $k_{\text{max}}^c = \max \{ k_{j,1}, j = 1, \ldots, L \}$ and $k_{\text{max}}^e = \max \{ k_{j,1}, j = 1, \ldots, L \}$ denote, respectively, the maximum number of the center and edge users in the $L$ cells.

6. Conclusion

In this paper, we have proposed a novel strategy to mitigate the PCP within multicell M-MIMO systems. Unlike the previous studies, which use a specific decontaminating strategy to resolve the PCP upon the users within cells, our proposed
Table 3: Computational complexity of the S3SR, SPR-MBDP [9], and WGC-PA [15].

| Method          | Complexity                        |
|-----------------|-----------------------------------|
| S3SR if the first line of (24) is considered | $O\left(M\left(K_{\text{Tot},c}\right)^2 + S\left(k_{\text{max},c}L\right)^3\right)$ |
| S3SR if the second line of (24) is considered | $O\left(M\left(K_{\text{Tot},c}\right)^2 + S\left(k_{\text{max},c}L\right)^3\right)$ |
| SPR-MBDP        | $O\left(M\left(K_{\text{Tot},c}\right)^2 + M\left(k_{\text{max},c}\right)^2\right)$ |
| WGC-PA          | $O\left(S(KL)^3\right)$          |

S3SR is continually updated and adapted to a different distribution of the users within cells. This process is done based on a SR, which is aimed at guaranteeing a trade-off between the use of the precious OPSs and mitigating the PCP. In other words, the S3SR is, mainly, based on the consolidation of two powerful decontaminating strategies (i.e., WGC-PA and SPR-MBDP) under a SR. Specifically, the LSP coefficients are exploited, to separate the users of each cell into two groups, which are the center and the edge users; therefore, a derived SR is used by the proposed S3SR to switch between the WGC-PA and the SPR-MBDP. This switching leads to make better use of the OPS resources during the training phase. The superiority and the robustness of the S3SR, compared to the SPR-MBDP, the WGC-PA, and the conventional strategy are proved through simulations.

Data Availability
The overall simulation results included in this paper are generated based on Matlab.

Conflicts of Interest
The authors declare that there is no conflict of interest.

References
[1] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Transactions on Wireless Communications, vol. 9, no. 11, pp. 3590–3600, 2010.
[2] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, Fundamentals of Massive MIMO, Cambridge University Press, 2016.
[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” IEEE Transactions on Communications, vol. 61, no. 4, pp. 1436–1449, 2013.
[4] L. Sanguinetti, E. Bjornson, and J. Hoydis, “Toward massive MIMO 2.0: understanding spatial correlation, interference suppression, and pilot contamination,” IEEE Transactions on Communications, vol. 68, no. 1, pp. 232–257, 2020.
[5] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, “A comprehensive survey of pilot contamination in massive MIMO-5G system,” IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 905–923, 2016.
[6] W. A. W. M. Mahyiddin, P. A. Martin, and P. J. Smith, “Pilot contamination reduction using time-shifted pilots in finite massive MIMO systems,” in 2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall), pp. 1–5, Vancouver, BC, 2014.
[7] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, “Inter-cell interference in noncooperative TDD large scale antenna systems,” IEEE Journal on Selected Areas in Communications, vol. 31, no. 2, pp. 192–201, 2013.
[8] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels,” IEEE Transactions on Signal Processing, vol. 52, no. 2, pp. 461–471, 2004.
[9] X. Zhu, Z. Wang, C. Qian et al., “Soft pilot reuse and multicell block diagonalization precoding for massive MIMO systems,” IEEE Transactions on Vehicular Technology, vol. 65, no. 5, pp. 3285–3298, 2016.
[10] L. Ku, J. Fan, and J. Deng, “Adaptive pilot reuse scheme for nonzero neighborhood structured downlink channel in massive MIMO,” International Journal of Electronics and Communications (AEU), vol. 99, pp. 48–58, 2019.
[11] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, “A coordinated approach to channel estimation in large-scale multiple-antenna systems,” IEEE Journal on Selected Areas in Communications, vol. 31, no. 2, pp. 264–273, 2013.
[12] L. Ku, J. Fan, and J. Deng, “Low complexity pilot allocation in massive MIMO systems,” in 2016 8th IEEE International Conference on Communication Software and Networks (ICCSN), pp. 402–406, 2016.
[13] J. Chen, C. Wen, S. Jin, and K. Wong, “A low complexity pilot scheduling algorithm for massive MIMO,” IEEE Wireless Communications Letters, vol. 6, no. 1, pp. 18–21, 2017.
[14] J. Fan, W. Li, and Y. Zhang, “Pilot contamination mitigation by fractional pilot reuse with threshold optimization in massive MIMO systems,” Digital Signal Processing, vol. 78, pp. 197–204, 2018.
[15] X. Zhu, L. Dai, Z. Wang, and X. Wang, “Weighted-graph-coloring-based pilot decontamination for multicell massive MIMO systems,” IEEE Transactions on Vehicular Technology, vol. 66, no. 3, pp. 2829–2834, 2017.
[16] W. Yuan, X. Yang, and R. Xu, “A novel pilot decontamination scheme for uplink massive MIMO systems,” Procedia Computer Science, vol. 131, pp. 72–79, 2018.
[17] W. Chang, H. Chan, and Y. Hua, “Weighted graph coloring based softer pilot reuse for TDD massive MIMO systems,” IEEE Transactions on Vehicular Technology, vol. 67, no. 7, pp. 6272–6285, 2018.
[18] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: benefits and challenges,” IEEE Journal of Selected Topics in Signal Processing, vol. 8, no. 5, pp. 742–758, 2014.
[19] J. C. Marinello and T. Abrao, “Pilot distribution optimization in multi-cellular large scale MIMO systems,” AEU - International Journal of Electronics and Communications, vol. 70, no. 8, pp. 1094–1103, 2016.
[20] V.-K. Dinh, M.-T. Le, V.-D. Ngo, and C.-H. Ta, “PCA-aided linear precoding in massive MIMO systems with imperfect CSI,” Wireless Communications and Mobile Computing, vol. 2020, Article ID 3425952, 9 pages, 2020.
[21] S. M. Sanaie and H. K. Bizaki, “Performance analysis of multicell massive MIMO THP with pilot contamination,” Wiley Transactions on Emerging Telecommunications Technologies, vol. 30, no. 5, p. e3574, 2019.
[22] M. Boulouird, A. Riadi, and M. M. Hassani, “Pilot contamination in multi-cell massive-MIMO systems in 5G wireless communications,” in 2017 International Conference on Electrical and Information Technologies (ICEIT), pp. 15–18, Rabat, Morocco, 2017.

[23] F. Rusek, D. Persson, Buon Kiong Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, “Scaling up MIMO: opportunities and challenges with very large arrays,” IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 40–60, 2013.