Pilot Assignment for Joint Uplink-Downlink Spectral Efficiency Enhancement in Massive MIMO Systems with Spatial Correlation

Tien Hoa Nguyen, Trinh Van Chien, *Member, IEEE*, Hien Quoc Ngo, *Senior Member, IEEE*, Xuan Nam Tran, *Member, IEEE*, and Emil Björnson, *Senior Member, IEEE*

Abstract—This paper proposes a flexible pilot assignment method to jointly optimize the uplink and downlink data transmission in multi-cell Massive multiple input multiple output (MIMO) systems with correlated Rayleigh fading channels. By utilizing a closed-form expression of the ergodic spectral efficiency (SE) achieved with maximum ratio processing, we formulate an optimization problem for maximizing the minimum weighted sum of the uplink and downlink SEs subject to the transmit powers and pilot assignment sets. This combinatorial optimization problem is solved by two sequential algorithms: a heuristic pilot assignment is first proposed to obtain a good pilot reuse set and the data power control is then implemented. Numerical results manifest that the proposed algorithm converges fast to a better minimum sum SE per user than the algorithms in previous works.

Index Terms—Massive MIMO, Max-Min Fairness Optimization, Pilot Assignment.

I. Introduction

Massive multiple input multiple output (MIMO) is a key component of 5G since this technology can improve ergodic spectral efficiency (SE) with the orders of magnitude compared to single-antenna systems. It is achieved by equipping each base station (BS) with a large number of antennas such that the system can spatially multiplex tens of users at the same time and frequency resource [1], [2]. The quasi-orthogonality of all channels allows a simple linear beamforming to yield SE close to the channel capacity for single-cell systems, in which orthogonal pilot signals are assumed to be available for all

users. In cellular Massive MIMO systems, it is an impractical assumption because the pilot overhead is directly proportional to the total number of users [3], while the coherence interval is limited. A small set of orthogonal pilot signals should be reused among the cells, which results in mutually correlated interference known as pilot contamination that downgrades the ergodic SE [4].

In order to mitigate pilot contamination, an observation is that some users will cause more severe contamination to each other when they use the same pilot. This issue should be avoided in the pilot assignment. The system can reuse pilot signals in a way that gives these users a priority to use orthogonal pilots. Nonetheless, an optimal pilot assignment is expensive to find since it is attained by solving a combinatorial problem. Heuristic algorithms with affordable complexity are necessary in practice to eliminate pilot contamination at a reasonable cost. It should be noticed that most previous works only consider the pilot assignment for either the uplink or downlink transmission, see [3], [5] and references therein. The authors [6] were the first to propose a heuristic pilot assignment algorithm taking both uplink and downlink into account, but based on the asymptotic SE from uncorrelated Rayleigh fading that behaves differently from the capacity regime for a limited number of BS antennas and spatially correlated channels. Meanwhile, the combinations of pilot signals to obtain the pilot assignment increases rapidly with the total number of users in all cells making it an intractable solution. Motivated by the fact that uncorrelated channels rarely appear in practice [7], the pilot assignment for spatially correlated channels were considered in [3], [8], but the pilot assignment for jointly enhancing the uplink and downlink SE has not been considered in this context.

In this paper, we assign the pilot signals to jointly maximize the minimum weighted sum of uplink and downlink SE per user with spatially correlated channels. This optimization problem has flexibility since the weights can be used to assign different priorities to the downlink and uplink. The optimization problem is based only on statistical channel information, thus the obtained solution can be utilized as long as the statistics remain the same. Since this optimization problem is combinatorial and NP-hard, we propose a heuristic pilot assignment that works well for systems with a limited number of BS antennas. The obtained pilot assignment solution outperforms the other benchmarks from previous works.

Notation: The upper and lower-bold letters denote vectors

and matrices, respectively. The notation $(\cdot)^H$ is the Hermitian transpose. $\mathbb{E}\{\cdot\}$ is the expectation of a random variable, while $\mathcal{CN}(\cdot,\cdot)$ denotes circularly complex Gaussian distribution. The identity matrix of size $M \times M$ is denoted by $\mathbf{I}_M$. Finally, $\mathrm{tr}(\cdot)$ is the trace operator, $\|\cdot\|$ is Euclidean norm, and $\mathcal{O}(\cdot)$ denotes the big-$\mathcal{O}$ notation.

II. SYSTEM MODEL

We consider a multi-cell Massive MIMO system comprising L cells, each with an M-antenna BS serving K single-antenna users. The system uses a time-division duplexing protocol. Let τ_c be the length of each coherence block whereof τ_p symbols are used for the uplink training and the remaining is used for the data transmission. We denote by γ^{ul} and γ^{dl} the fraction of the $\tau_c - \tau_p$ symbols used for the uplink and downlink data transmission and satisfied $\gamma^{\mathrm{ul}} + \gamma^{\mathrm{dl}} = 1$. The set $\mathcal{S}$ contains all tuple of cell and user indices in the system as

$$\mathcal{S} = \{(i,t) : i \in \{1,\ldots,L\}, t \in \{1,\ldots,K\}\}. \quad (1)$$

The channel between user t in cell i and BS l is denoted as $\mathbf{h}_{i,t}^l \in \mathbb{C}^M$ and is assumed to feature correlated Rayleigh fading: $\mathbf{h}_{i,t}^l \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{i,t}^l)$, where $\mathbf{R}_{i,t}^l \in \mathbb{C}^{M \times M}$ is the channel correlation matrix. All BSs know the channel statistics, but need to estimate the realizations in each coherence block.[1]

A. Uplink training phase

Let us introduce a set $\mathcal{P}$ of τ_p mutually orthogonal pilot signals, $K \leq \tau_p \leq KL$, reused among the users. The pilot signal assigned to user t in cell l is denoted as $\boldsymbol{\psi}_{i,t}$. In the uplink training phase, the received baseband pilot signal $\mathbf{Y}_{p,l} \in \mathbb{C}^{M \times \tau_p}$ at BS l is

$$\mathbf{Y}_{p,l} = \sum_{(i,t)\in\mathcal{S}} \mathbf{h}_{i,t}^l \boldsymbol{\psi}_{i,t}^H + \mathbf{N}_{p,l}, \quad (2)$$

where $\mathbf{N}_{p,l}$ is an $M \times \tau_p$ noise matrix with independent elements distributed as $\mathcal{CN}(0, \sigma_{\mathrm{UL}}^2)$ and σ_{ul}^2 being the noise variance in the uplink. The channel estimate $\hat{\mathbf{h}}_{l,k}^l$ of $\mathbf{h}_{l,k}^l$ is obtained from (2) by MMSE estimation [4] as

$$\hat{\mathbf{h}}_{l,k}^l = \|\boldsymbol{\psi}_{l,k}\|^2 \mathbf{R}_{l,k}^l (\mathbf{F}_{l,k}^l)^{-1} \mathbf{Y}_{p,l} \boldsymbol{\psi}_{l,k}, \quad (3)$$

where $\mathbf{F}_{l,k}$ is given as

$$\mathbf{F}_{l,k} = \sum_{(i,t)\in\mathcal{S}} \mathbf{R}_{i,t}^l |\boldsymbol{\psi}_{i,t}^H \boldsymbol{\psi}_{l,k}|^2 + \sigma_{\mathrm{UL}}^2 \|\boldsymbol{\psi}_{l,k}\|^2 \mathbf{I}_M. \quad (4)$$

For all l, k, the channel estimates are distributed as

$$\hat{\mathbf{h}}_{l,k}^l \sim \mathcal{CN}\left(\mathbf{0}, \|\boldsymbol{\psi}_{l,k}\|^4 \mathbf{R}_{l,k}^l \mathbf{F}_{l,k}^{-1} \mathbf{R}_{l,k}^l\right). \quad (5)$$

The channel estimate in (3) together with its statistical information in (5) are used to formulate the linear processing vectors for the data transmission and compute closed-form ergodic SE expressions for each user.

[1]For sake of the simplicity, we assume that the channel correlation matrices are known. In practical systems, we can easily estimate the channel correlation matrices by averaging over many different instantaneous channel realizations.

B. Data transmission

In the uplink data transmission, the K users in each cell simultaneously send data to the serving BS. Specifically, user k in cell l sends a complex data symbol $s_{l,k}$ with $\mathbb{E}\{|s_{l,k}|^2\} = 1$. The received signal $\mathbf{y}_l \in \mathbb{C}^M$ at BS l is

$$\mathbf{y}_l = \sum_{(i,t)\in\mathcal{S}} \sqrt{p_{i,t}^{\mathrm{ul}}} \mathbf{h}_{i,t}^l s_{i,t} + \mathbf{n}_l, \quad (6)$$

where $p_{i,t}^{\mathrm{ul}}$ is the transmit data power and $\mathbf{n}_l \in \mathbb{C}^M$ is the uplink Gaussian noise with $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathrm{ul}}^2 \mathbf{I}_M)$. We assume BS l detects the desired signal from its user k by utilizing a maximum-ratio combining vector as

$$\mathbf{v}_{l,k} = \hat{\mathbf{h}}_{l,k}^l. \quad (7)$$

The desired signal is then obtained from

$$\mathbf{v}_{l,k}^H \mathbf{y}_l = \sum_{(i,t)\in\mathcal{S}} \sqrt{p_{i,t}^{\mathrm{ul}}} \hat{\mathbf{h}}_{l,k}^{l,H} \mathbf{h}_{i,t}^l s_{i,t} + \hat{\mathbf{h}}_{l,k}^{l,H} \mathbf{n}_l. \quad (8)$$

In the downlink data transmission, BS l transmits a signal $\mathbf{x}_l \in \mathbb{C}^M$ to its K users, which is formulated as

$$\mathbf{x}_l = \sum_{t=1}^{K} \sqrt{p_{l,t}^{\mathrm{dl}}} \mathbf{w}_{l,t} q_{l,t}, \quad (9)$$

where $p_{l,t}^{\mathrm{dl}}$ is the power allocated to the data symbol $q_{l,t}$ with $\mathbb{E}\{|q_{l,t}|^2\} = 1$. The maximum ratio (MR) precoding vector

$$\mathbf{w}_{l,t} = \frac{\hat{\mathbf{h}}_{l,t}^l}{\sqrt{\mathbb{E}\left\{\|\hat{\mathbf{h}}_{l,t}^l\|^2\right\}}} = \frac{\hat{\mathbf{h}}_{l,t}^l}{\sqrt{\|\boldsymbol{\psi}_{l,k}\|^4 \mathrm{tr}(\mathbf{R}_{l,k}^l \mathbf{F}_{l,k}^{-1} \mathbf{R}_{l,k}^l)}}, \quad (10)$$

is used. The received signal at user k in cell l is a superposition of the signals from all L BSs as

$$r_{l,k} = \sum_{(i,t)\in\mathcal{S}} \sqrt{p_{i,t}^{\mathrm{dl}}} \left(\mathbf{h}_{l,k}^i\right)^H \mathbf{w}_{i,t} q_{i,t} + n_{l,k}, \quad (11)$$

where $n_{l,k}$ denotes the additive noise which is distributed as $n_{l,k} \sim \mathcal{CN}(0, \sigma_{\mathrm{dl}}^2)$ and σ_{dl}^2 is the noise variance in the downlink. By applying the standard Massive MIMO methodology [4] to (8) and (11), the closed-form expression of the ergodic uplink and downlink SEs in Lemma 1 are attained.

Lemma 1. *Closed-form expression of the uplink and downlink ergodic SEs of user k in cell l are respectively*

$$R_{l,k}^{\mathrm{ul}} = \gamma^{\mathrm{ul}}\left(1 - \frac{\tau_p}{\tau_c}\right)\log_2\left(1 + \mathrm{SINR}_{l,k}^{\mathrm{ul}}\right), \quad (12)$$

$$R_{l,k}^{\mathrm{dl}} = \gamma^{\mathrm{dl}}\left(1 - \frac{\tau_p}{\tau_c}\right)\log_2\left(1 + \mathrm{SINR}_{l,k}^{\mathrm{dl}}\right), \quad (13)$$

where the effective SINR values are given in (14) and (15).

Proof. The proof follows along the lines of Corollaries 4.5 and 4.9 in [4] except for the different notations and the fact that pilot reuse pattern is arbitrary and not defined in advance. $\square$

In the SINR expressions, the numerator represents an array gain as the trace of the covariance matrix is proportional to the number of BS antennas. The first term of the denominator represents coherent interference originating from pilot

$$\text{SINR}_{l,k}^{\text{ul}} = \frac{p_{l,k}^{\text{ul}}\|\boldsymbol{\psi}_{l,k}\|^4 \text{tr}(\mathbf{R}_{l,k}^l \mathbf{F}_{l,k}^{-1}\mathbf{R}_{l,k}^l)}{\sum_{(i,t)\in\mathcal{S}\setminus\{(l,k)\}} p_{i,t}^{\text{ul}}|\boldsymbol{\psi}_{l,k}^H\boldsymbol{\psi}_{i,t}|^2 \frac{|\text{tr}(\mathbf{R}_{i,t}^l \mathbf{F}_{l,k}^{-1}\mathbf{R}_{l,k}^l)|^2}{\text{tr}(\mathbf{R}_{l,k}^l \mathbf{F}_{l,k}^{-1}\mathbf{R}_{l,k}^l)} + \sum_{(i,t)\in\mathcal{S}} p_{i,t}^{\text{ul}}\frac{\text{tr}(\mathbf{R}_{i,t}^l \mathbf{R}_{l,k}^l \mathbf{F}_{l,k}^{-1}\mathbf{R}_{l,k}^l)}{\text{tr}(\mathbf{R}_{l,k}^l \mathbf{F}_{l,k}^{-1}\mathbf{R}_{l,k}^l)} + \sigma_{\text{ul}}^2} \tag{14}$$

$$\text{SINR}_{l,k}^{\text{dl}} = \frac{p_{l,k}^{\text{dl}}\|\boldsymbol{\psi}_{l,k}\|^4 \text{tr}(\mathbf{R}_{l,k}^l \mathbf{F}_{l,k}^{-1}\mathbf{R}_{l,k}^l)}{\sum_{(i,t)\in\mathcal{S}\setminus\{(l,k)\}} p_{i,t}^{\text{dl}}|\boldsymbol{\psi}_{l,k}^H\boldsymbol{\psi}_{i,t}|^2 \frac{|\text{tr}(\mathbf{R}_{i,t}^i \mathbf{F}_{i,t}^{-1}\mathbf{R}_{i,t}^i)|^2}{\text{tr}(\mathbf{R}_{i,t}^i \mathbf{F}_{i,t}^{-1}\mathbf{R}_{i,t}^i)} + \sum_{(i,t)\in\mathcal{S}} p_{i,t}^{\text{dl}}\frac{\text{tr}(\mathbf{R}_{i,t}^i \mathbf{F}_{i,t}^{-1}\mathbf{R}_{i,t}^i \mathbf{R}_{i,t}^i)}{\text{tr}(\mathbf{R}_{i,t}^i \mathbf{F}_{i,t}^{-1}\mathbf{R}_{i,t}^i)} + \sigma_{\text{dl}}^2} \tag{15}$$
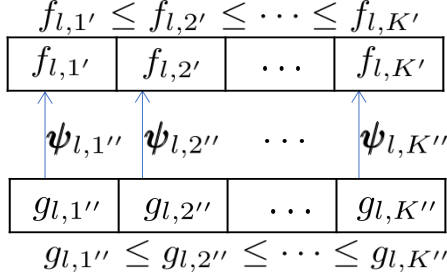

Fig. 1. The proposed pilot assignment for the K users in cell l.

contamination caused by the pilot reuse and it grows with the number of BS antennas. The last terms of the denominator are noncoherent interference and noise. While the uplink SINR expression of each user only depends on the channel estimate of this own user, a superposition of the channel estimation quality from all the users are observed in the downlink SINR expression. The denominators of the SINR expressions in (14) and (15) indicate the different contributions of a pilot reuse pattern to the uplink and downlink data transmission. The coupled nature motivates a pilot assignment for jointly optimizing the both SEs per user instead of either the uplink or downlink SE as in most previous works.

III. MAX-MIN FAIRNESS OPTIMIZATION

This section studies the pilot assignment for the weighted max-min sum SE per user fairness problem with uplink and downlink transmit power constraints. Due to the inherent non-convexity, we propose a heuristic algorithm to obtain a good local solution with tolerable computational complexity.

A. Problem formulation

By introducing the weights $\{w_{l,k}^{\text{ul}}, w_{l,k}^{\text{dl}}\}$ that prioritize the uplink and downlink transmission of arbitrary user k in cell l, the optimization problem is formulated for a given set of orthogonal pilot signals as

$$\underset{\substack{\{p_{l,k}^{\text{ul}}\geq 0\},\{p_{l,k}^{\text{dl}}\geq 0\},\\ \{\boldsymbol{\psi}_{l,k}\in\mathcal{P}\}}}{\text{maximize}} \quad \underset{(l,k)}{\min} \quad w_{l,k}^{\text{ul}}R_{l,k}^{\text{ul}} + w_{l,k}^{\text{dl}}R_{l,k}^{\text{dl}} \tag{16a}$$

$$\text{subject to} \quad p_{l,k}^{\text{ul}} \leq P_{\max,l,k}^{\text{ul}}, \forall l,k, \tag{16b}$$

$$\sum_{k=1}^{K} p_{l,k}^{\text{dl}} \leq P_{\max,l}^{\text{dl}}, \forall l, \tag{16c}$$

where $P_{\max,l,k}^{\text{ul}}, P_{\max,l}^{\text{dl}}$ is the maximum power that each user and BS can allocate to in the uplink and downlink, respectively. Problem (16) is combinatorial and its optimal solution

for the pilots is obtained by exhaustive search over possible pilot assignments. For the pilot length of $\tau_p = K$, there are $(K!)^{L-1}$ different pilot assignments, which is impossible to perform in a large-scale system [3]. We notice that by introducing weights and considering spatially correlated fading, problem (16) is a generalization of previous works which only focus on uncorrelated Rayleigh channel for either the uplink or downlink transmission [9].

B. Heuristic pilot assignment with fixed data powers

A low-complexity heuristic pilot assignment algorithm is proposed, in which the user having the lowest weight sum SE is prioritized. For a given set of transmit power coefficients, problem (16) becomes

$$\underset{\{\boldsymbol{\psi}_{l,k}\}\in\mathcal{P}}{\text{maximize}} \quad \underset{(l,k)}{\min} \quad w_{l,k}^{\text{ul}}R_{l,k}^{\text{ul}} + w_{l,k}^{\text{dl}}R_{l,k}^{\text{dl}}. \tag{17}$$

We assume that all KL users first randomly select the pilot signals such that there is no pilot contamination inside a cell. After that, cell l reallocates the pilot signals to its K users with the availability of pilot assignment information from other cells. If we define the weighted sum SE of user k in cell l as

$$f_{l,k} = w_{l,k}^{\text{ul}}R_{l,k}^{\text{ul}} + w_{l,k}^{\text{dl}}R_{l,k}^{\text{dl}}, \tag{18}$$

then BS l can sort all K users in the ascending order as

$$f_{l,1'} \leq f_{l,2'} \leq \ldots \leq f_{l,K'}, \tag{19}$$

where $\{1', 2', \ldots, K'\}$ is a permutation of the set $\{1, 2, \ldots, K\}$. From these notations, user k' in cell l has the weighted sum SE $f_{l,k'}, \forall k' = 1, \ldots, K$.

We now compute the normalized mean square error (NMSE) of user k in cell l as

$$g_{l,k} = \frac{\mathbb{E}\{\|\mathbf{e}_{l,k}^l\|^2\}}{\mathbb{E}\{\|\mathbf{h}_{l,k}^l\|^2\}} = 1 - \frac{\text{tr}(\mathbf{R}_{l,k}^l \mathbf{F}_{l,k}^{-1}\mathbf{R}_{l,k}^l)}{\text{tr}(\mathbf{R}_{l,k}^l)}, \tag{20}$$

then the channel estimation quality of the K users in cell l is sorted in the ascending order as

$$g_{l,1''} \leq g_{l,2''} \leq \ldots \leq g_{l,K''}, \tag{21}$$

where $\{1'', 2'', \ldots, K''\}$ is a permutation of the set $\{1, 2, \ldots, K\}$. The pilot signal $\boldsymbol{\psi}_{l,k''}$ currently used by user k'' is reassigned to user k'. The intuition is to dedicate pilots signal subject to less pilot contamination to users with worse conditions, which is viewed as one with a smaller $f_{l,k'}$. Our proposed pilot assignment for users in cell l is illustrated in Fig. 1. This process is implemented cell-by-cell in an iterative algorithm (please see Algorithm 1). Concerning on the per-cell-based pilot assignment, a new pilot reuse set may harm

the minimum weighted sum SE in the entire system. To avoid this issue, we introduce a backtracking condition to assign the pilot signals at iteration n as in Lemma 2.

Lemma 2. *If the pilot signals are only assigned to the K users in cell l when the objective function in* (17) *does not increase, then the proposed iterative pilot assignment approach converges to a fixed point.*

Proof. Let us denote $h_l^{*,(n)}$ and $h_l^{*,(n-1)}$ as the minimum weighted sum SE per user before and after BS l reassigns the pilot signals, i.e.,

$$h_l^{*,(n)} = \min_{(l',k)} \quad w_{l',k}^{\mathrm{ul}} R_{l',k}^{\mathrm{ul},(n)} + w_{l',k}^{\mathrm{dl}} R_{l',k}^{\mathrm{dl},(n)}, \tag{22}$$

$$h_l^{*,(n-1)} = \min_{(l',k)} \quad w_{l',k}^{\mathrm{ul}} R_{l',k}^{\mathrm{ul},(n-1)} + w_{l',k}^{\mathrm{dl}} R_{l',k}^{\mathrm{dl},(n-1)}, \tag{23}$$

where $R_{l',k}^{\mathrm{ul},(n)}$, $R_{l',k}^{\mathrm{dl},(n)}$, $R_{l',k}^{\mathrm{ul},(n-1)}$, and $R_{l',k}^{\mathrm{dl},(n-1)}$ are the uplink and downlink SEs at iteration n and $n-1$, respectively. A criterion is then used to approve if the reassignment is valid by checking the backtracking condition

$$h_l^{*,(n)} \geq h_l^{*,(n-1)}, \tag{24}$$

which ensures a nondecreasing objective function in problem (17). We stress that the condition (24) needs to be implemented in each iteration due to the non-convexity of (22) and (23). Moreover the limited power budgets in (16b) and (16c) ensure that this objective function is bounded from above for any set of pilot and data power coefficients in the feasible domain. Consequently, problem (17) converges to a fixed point and we conclude the proof. $\square$

During assigning the pilot signals to the users over cells, the proposed approach will be stopped when, for example, a small variation of two consecutive iterations, which is computed for all the L cells as

$$\sum_{l=1}^{L} \left| h_l^{*,(n)} - h_l^{*,(n-1)} \right| \leq \epsilon \tag{25}$$

where $\epsilon \geq 0$ is a given accuracy. The proposed pilot assignment approach is applied to all the cells as in Algorithm 1.

C. Data power control

For a given pilot assignment, problem (16) now reduces to the data power control problem. To obtain a low-complexity, the uplink and downlink data power controls can be separately optimized. We therefore present a framework which is applied for both using the nominal parameters: Let us denote $\alpha \in \{\mathrm{ul}, \mathrm{dl}\}$, the max-min fairness problem is now formulated as

$$\underset{\{p_{l,k}^{\alpha} \geq 0\}}{\text{maximize}} \quad \underset{(l,k)}{\min} \quad w_{l,k}^{\alpha} R_{l,k}^{\alpha} \tag{26}$$

$$\text{subject to} \quad \text{Constraints in (27)},$$

where the power budget constraints are

$$\begin{cases} p_{l,k}^{\mathrm{ul}} \leq P_{\max,l,k}^{\mathrm{ul}}, \forall l,k, & \text{for the uplink,} \\ \sum_{k=1}^{K} p_{l,k}^{\mathrm{dl}} \leq P_{\max,l}^{\mathrm{dl}}, \forall l, & \text{for the downlink.} \end{cases} \tag{27}$$
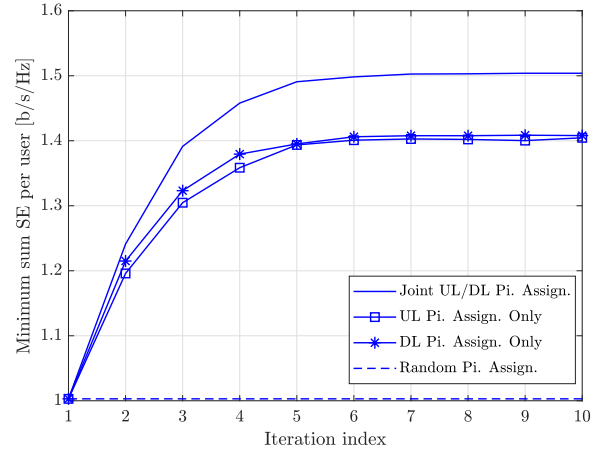

Fig. 2. The convergence of the proposed pilot assignment for a network with 4 users per cell.

By adopting the epigraph representation [10], (26) is equivalent to as

$$\begin{aligned} \underset{\{p_{l,k}^{\alpha} \geq 0\}, \xi}{\text{maximize}} \quad & \xi \\ \text{subject to} \quad & \mathrm{SINR}_{l,k}^{\alpha} \geq \xi, \forall l,k, \\ & \text{Constraints in (27)}, \end{aligned} \tag{28}$$

where the new optimization variable $\xi \in \{\xi^{\mathrm{ul}}, \xi^{\mathrm{dl}}\}$ is the minimum SINR per user. In (28), the objective function and uplink power constraints aligns with monomials. The SINR expressions and downlink power constraints can be recast as posynomials. Consequently, (28) is a geometric program whose global optimum is able to attain by a general purpose optimization toolbox [3].[2]

Our proposal to obtain a local solution to (16) is summarized in Algorithm 1. The computational complexity of the pilot assignment is from sorting the pilot signals and from computing KL inverse matrices. The matrix inversion is more expensive since each BS has many antennas. The computational complexity of the pilot assignment is in the order of $\mathcal{O}\left(\nu N_1 L^2 K M^3\right)$, where N_1 is the number of iterations requires to reach the fixed point and the constant value ν stands for the effectiveness of computing matrix inverse [11]. Next, the data power control by the interior-point method consumes the computational complexity of the order of $\mathcal{O}\left(2N_2^{\alpha} \max\left\{2L^3 K^3, F_1\right\}\right)$, where F_1 is the first and second derivative estimation cost of computing the SINR constraints in (28). Consequently, the total computational complexity of Algorithm 1 is in the order of $\mathcal{O}\left(\nu N_1 L^2 K M^3 + 2N_2^{\alpha} \max\left\{2L^3 K^3, F_1\right\}\right)$. By conditioned on the dominated computational complexity from the matrix inverses $\mathbf{F}_{l,k}^{-1}, \forall l,k$, and the computational complexity of the pilot assignment and either the uplink or downlink data power control is $\mathcal{O}\left(\nu N_1 L^2 K M^3 + N_2^{\alpha} \max\left\{2L^3 K^3, F_1\right\}\right)$.

IV. NUMERICAL RESULTS

A network is considered with 4 square cells covering the area 0.5 km^2 utilizing the wrap-around technique at the edges

---

[2]We have implemented the pilot assignment and data power control iteratively, but no further improvement is observed.

Algorithm 1 An approach finding a fixed point to (16)

1: **Input** Set $\{P_{\max,l,k}^{\text{ul}}, P_{\max,l,k}^{\text{dl}}\}, \epsilon, h_l^{*,(0)} = 0, h_l^{*,(1)} = 1, \forall l$, and preliminary pilot assignment $\{\boldsymbol{\psi}_{l,k}^*\}$ obtained by randomization; Select initial transmit powers $\{p_{l,k}^{\text{ul}}, p_{l,k}^{\text{dl}}\}$. Set $n = 0$.

2: **while** (25) unsatisfied **do**

3:     Set $n = n + 1$

4:     **for** $l = 1, \ldots, L$ **do**

5:         BS l computes (19) and (21), then assigns the pilot signals as in Fig. 1

6:         BS l verifies the backtracking condition $h_l^{*,(n)} \geq h_l^{*,(n-1)}$ **If** this is satisfied, **then** update $h_l^{*,(n)}$ and broadcast the new pilot assignment $\{\boldsymbol{\psi}_{l,k}^*\}_{k=1}^K$. **Otherwise** keep the previous one.

7:         BS l checks the stopping condition. **If** not satisfied, **then** continue by setting $l = l + 1$. **Otherwise** go to step 9.

8:     **end for**

9:     Update the cost $\sum_{l=1}^L \left| h_l^{*,(n)} - h_l^{*,(n-1)} \right|$ for the stopping criterion (25).

10: **end while**

11:     Solve problem (28) to obtain the optimal data powers $\{p_{l,k}^{*,\text{ul}}, p_{l,k}^{*,\text{dl}}\}, \forall l, k$.

12: **Output** $\{p_{l,k}^{*,\text{ul}}, p_{l,k}^{*,\text{dl}}\}$, and $\{\boldsymbol{\psi}_{l,k}^*\}, \forall l, k$.

Fig. 3. The convergence of the proposed pilot assignment for a network with 6 users per cell.

to avoid boundary effects, and therefore one BS has eight neighbors. In each cell, a BS with 200 antennas is at the center and serving K uniformly distributed users with a minimum distance to the serving BS being 35 m. There are K orthogonal pilot signals, while the maximum power is $P_{\max,l,k}^{\text{ul}} = 200$ mW and $P_{\max,l}^{\text{dl}} = 200K$ mW for an equal total power budget of the uplink and downlink data transmission thanks to the uplink-downlink duality [12]. The noise variance is -96 dBm corresponding to the noise figure 5 dB. The large-scale fading coefficients are

$$\beta_{l,k}^j [\text{dB}] = -148.1 - 37.6 \log_{10}(d_{l,k}^j/1\text{km}) + z_{l,k}^j, \quad (29)$$

where $d_{l,k}^j$ in km is the distance between user k in cell l and BS j [3]. The shadow fading $z_{l,k}^j$ follows a log-normal distribution with standard deviation 7 dB. The covariance matrix of the channel from user k in cell l and BS j is defined by the exponential correlation model, which models a uniform linear array as

$$\mathbf{R}_{l,k}^j = \beta_{l,k}^j \begin{bmatrix} 1 & r_{l,k}^{j,*} & \cdots & \left(r_{l,k}^{j,*}\right)^{M-1} \\ r_{l,k}^j & 1 & \cdots & \left(r_{l,k}^{j,*}\right)^{M-2} \\ \vdots & \vdots & \ddots & \vdots \\ \left(r_{l,k}^j\right)^{M-1} & \left(r_{l,k}^j\right)^{M-2} & \cdots & 1 \end{bmatrix}, \quad (30)$$

where the spatial correlation $r_{l,k}^j = \mu e^{j\theta_{l,k}^j}$ with the correlation magnitude μ in the range $[0, 1]$ and the user incidence angle to the array boresight being $\theta_{l,k}^j$. By setting the weights $w_{l,k}^{\text{ul}}, w_{l,k}^{\text{dl}} \in \{0, 1\}, \forall l, k$, (i.e., $w_{l,k}^{\text{dl}} = 0$ if only considering the uplink data transmission; $w_{l,k}^{\text{ul}} = 0$ if only considering the downlink data transmission; $w_{l,k}^{\text{ul}} = w_{l,k}^{\text{dl}} = 1$ if both the uplink
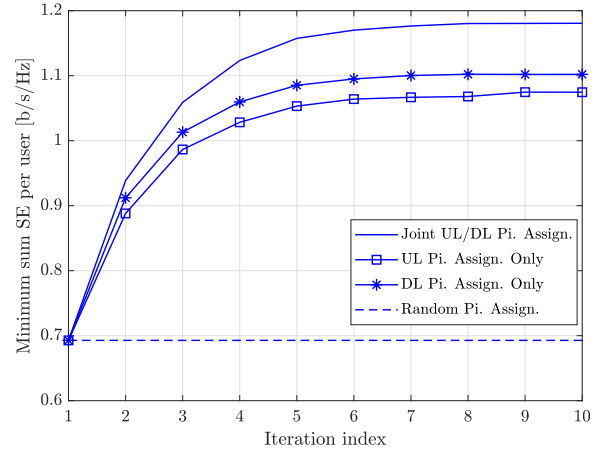
and downlink data transmissions are considered) the following benchmarks are used for comparison:[3]

1) *Random pilot assignment (Denoted as "Ran. Pi. Assign." in the figures):* The pilot signals are randomly assigned to all users, which was used in, for example [3].

2) *Greedy pilot assignment (Denoted as "Gre. Pi. Assign." in the figures):* The pilot signals are assigned based on the similarity between the covariance matrices, which was proposed by [8].

3) *Uplink pilot assignment (Denoted as "UL. Pi. Only" in the figures):* The pilot signals are assigned based on the uplink SE only.

4) *Downlink pilot assignment (Denoted as "DL. Pi. Only" in the figures):* The pilot signals are assigned based on the downlink SE only.

5) *Pilot assignment for the joint UL/DL SE enhancement (Denoted as "Joint UL/DL Pi. Assign." in the figures):* The pilot signals are assigned by the weighted sum SE per user.

Figs. 2 and 3 display the convergence of proposed pilot assignments for a network with 4 users and 6 users per cell, respectively. The convergence is obtained after less than 8 iterations for all the considered scenarios. When each BS serves 4 users, the sum SE per user converges to about 1.4 b/s/Hz when utilizing either the uplink or downlink SE as the utility metric to assign the pilot signals. However, relying on the downlink SE to assign the pilot signals yields 2% better the sum SE per user than utilizing the uplink SE when increasing the number of users per cell to 6 users.

Fig. 4 shows the cumulative distribution function (CDF) of the minimum weighted SE per user for a network where each cell has 4 users. The greedy pilot gives $1.2\times$ SE better than the random assignment. While assigning the pilot signals based on either the uplink or downlink SE gives almost equivalent performance, but it provides better performance than the random pilot assignment by $1.5\times$. Meanwhile, the improvement of

---

[3]Exhaustive research is not included for comparison due to its extremely heavy complexity. One realization of user locations needs to evaluate the SE of $(K!)^{L-1} = 373, 248, 000$ combinations of pilot signals.
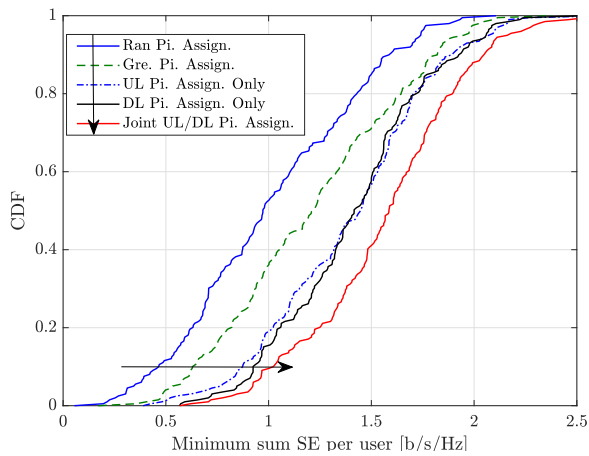
Fig. 4. The cumulative distribution function of the minimum sum SE per user without data power control.
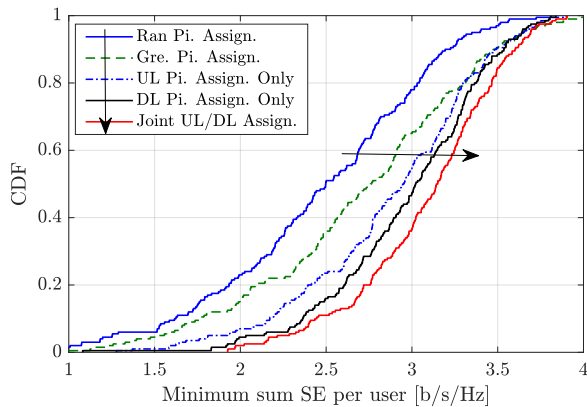

Fig. 5. The cumulative distribution function of the minimum sum SE per user with data power control.

joint pilot assignment is up to 39% in weighted minimum sum SE per user compared with the second best and it approves the locality of Algorithm 1. Finally, Fig. 5 manifests the benefits of data power control based on the proposed pilot assignment over the other benchmarks. We observe that the greedy pilot assignment only outperforms the random pilot assignment $1.1\times$. Meanwhile, an improvement up to $2.29\times$ better weighted sum SE per user than random pilot assignment is obtained. In addition, the data power control improves the sum SE per user up to about $2\times$ compared with the fixed data power allocation.

Fig. 6 plots the minimum sum SE per user versus the number of users per cell. Specifically, the minimum sum SE per user decreases when there are more users in the coverage area that generate more mutual interference. For instance, The joint pilot assignment reduce the minimum sum SE per user $1.7\times$ as the number of user per cell increases from 2 to 8 users. We also observe the benefits of combining the joint pilot assignment and data power control that results in a superior SE improvement up to $1.4\times$ compared with the random assignment.
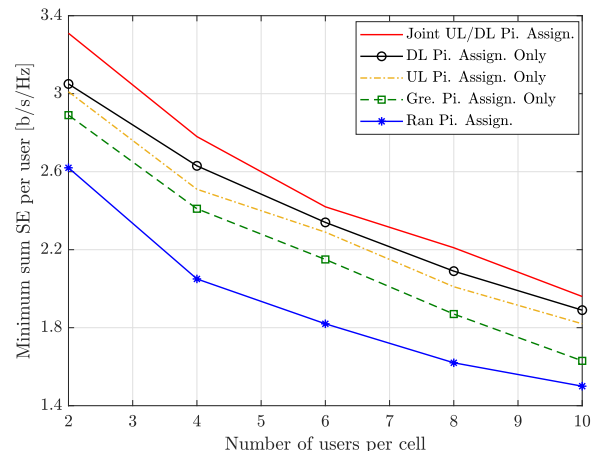

Fig. 6. The minimum sum SE per user versus the number of users per cell with data power control.

V. CONCLUSION

This paper has formulated and solved a max-min sum SE per user optimization problem considering both the pilot assignment and data power control for cellular Massive MIMO systems with correlated Rayleigh channels. We observed significant improvements of pilot assignment to the minimum sum SE per user compared with the other related works. Interestingly, only deploying the uplink or downlink SE as side information to assign pilot still yields good sum SE to weak users if the max-min fairness optimization is considered.

REFERENCES

[1] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
[2] T. A. Le, T. V. Chien, M. R. Nakhai, and T. Le-Ngoc, "Pareto-optimal pilot design for cellular Massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13 206–13 215, 2020.
[3] T. V. Chien, E. Björnson, and E. G. Larsson, "Joint pilot design and uplink power allocation in multi-cell Massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2000 – 2015, 2018.
[4] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154 – 655, 2017.
[5] S. Jin, M. Li, Y. Huang, Y. Du, and X. Gao, "Pilot scheduling schemes for multi-cell massive multiple-input-multiple-output transmission," *IET Commu.*, vol. 9, no. 5, pp. 689–700, 2015.
[6] J. C. Marinello Filho, C. M. Panazio, and T. Abrão, "Joint uplink and downlink optimization of pilot assignment for massive mimo with power control," *Trans. Emerging Telecommun. Techno.*, vol. 28, no. 12, p. e3250, 2017.
[7] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, 2015.
[8] L. You, X. Gao, X. G. Xia, N. Ma, and Y. Peng, "Pilot reuse for Massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, 2015.
[9] X. Zhu, Z. Wang, L. Dai, and C. Qian, "Smart pilot assignment for Massive MIMO," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1644 – 1647, 2015.
[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
[11] A. Krishnamoorthy and D. Menon, "Matrix inversion using Cholesky decomposition," in *Proc. IEEE SPA*, 2013.
[12] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," in *Proc. IEEE VTC-Fall*, 2002, pp. 87–91.