

## Pilot Power Allocation Through User Grouping in Multi-Cell Massive MIMO Systems

Liu, P., Jin, S., Jiang, T., Zhang, Q., & Matthaiou, M. (2016). Pilot Power Allocation Through User Grouping in Multi-Cell Massive MIMO Systems. IEEE Transactions on Communications. DOI: 10.1109/TCOMM.2016.2645767

**Published in:**  
IEEE Transactions on Communications

**Document Version:**  
Peer reviewed version

**Queen's University Belfast - Research Portal:**  
[Link to publication record in Queen's University Belfast Research Portal](#)

**Publisher rights**  
Copyright 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

**General rights**  
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**  
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [openaccess@qub.ac.uk](mailto:openaccess@qub.ac.uk).

# Pilot Power Allocation Through User Grouping in Multi-Cell Massive MIMO Systems

Pei Liu, *Student Member, IEEE*, Shi Jin, *Member, IEEE*, Tao Jiang, *Senior Member, IEEE*,  
Qi Zhang, and Michail Matthaiou, *Senior Member, IEEE*

**Abstract**—In this paper, we propose a relative channel estimation error (RCEE) metric, and derive closed-form expressions for its expectation  $\text{Exp}_{\text{rcee}}$  and the achievable uplink rate holding for any number of base station antennas  $M$ , with the least squares (LS) and minimum mean squared error (MMSE) estimation methods. It is found that RCEE and  $\text{Exp}_{\text{rcee}}$  converge to the same constant value when  $M \rightarrow \infty$ , which renders the pilot power allocation (PPA) substantially simplified and a PPA algorithm is proposed to minimize the average  $\text{Exp}_{\text{rcee}}$  per user with a total pilot power budget  $P$  in multi-cell massive multiple-input multiple-output systems. Numerical results show that the PPA algorithm brings considerable gains for the LS estimation compared with equal PPA (EPPA), while the gains are significant only with large frequency reuse factor (FRF) for the MMSE estimation. Moreover, for large FRF and large  $P$ , the performance of the LS approaches to the MMSE. Besides, a scheduling strategy is proposed to allocate pilot power whole system, which can approach the optimal performance. For the achievable uplink rate, the PPA scheme delivers almost the same average achievable uplink rate and improves the minimum achievable uplink rate compared with the EPPA scheme.

**Index Terms**—Massive multiple-input multiple-output (MIMO), pilot power allocation, relative channel estimation error, achievable uplink rate.

## I. INTRODUCTION

Manuscript received March 22, 2016; revised August 22, 2016 and October 30, 2016; accepted December 19, 2016. Date of publication XXX XX, 2017; date of current version XXX XX, 2017. This work was supported in part by National Science Foundation of China (NSFC) for Distinguished Young Scholars with Grant number 61325004, NSFC with Grants 61531011 and 61501193, National High Technology Development 863 Program of China under Grants 2015AA01A710 and 2014AA01A704, and Key Project of Hubei Province in China with Grant 2015BAA074. The associate editor coordinating the review of this paper and approving it for publication was Prof. H. Li. (Corresponding author: Tao Jiang.)

P. Liu is with Wuhan National Laboratory for Optoelectronics, School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China, and also with the Department of Electronic Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, China (e-mail: peil@hust.edu.cn; pliu32-c@my.cityu.edu.hk).

S. Jin is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: jinshi@seu.edu.cn).

T. Jiang is with Wuhan National Laboratory for Optoelectronics, School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: Tao.Jiang@ieee.org).

Q. Zhang is with Jiangsu Key Laboratory of Wireless Communications, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: zhangqiqi\_1212@126.com).

M. Matthaiou is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, BT3 9DT, U.K. (e-mail: m.matthaiou@qub.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier XXX

**M**ULTIPLE-INPUT multiple-output (MIMO) systems have been well integrated in the fourth generation mobile communication technology since they can improve the achievable data rates and suppress channel fading effects. Unlike traditional MIMO, massive MIMO [1–3] technology deploys hundreds of antennas to serve tens of users who share the same time-frequency resources, and has attracted wide attention from academia and industry in recent years. Massive MIMO can achieve considerable spatial multiplexing gains and improve energy efficiency more effectively. Hence, it is considered as one of the key technologies for the fifth generation mobile communication networks [4].

A great amount of research results have been reported on massive MIMO, in the context of linear precoders and detectors [5, 6], achievable sum-rate analysis [3, 7], hardware impairments [8], channel estimation [9, 10], and so on. Among these important topics, based on the seminal work [1], the channel estimation is one of the biggest challenges pertaining to massive MIMO. In this context, the time-division duplex (TDD) mode dominates the massive MIMO literature, since the pilot sequence length is analogous to the number of users [11]. Interestingly, based on the equal pilot power allocation (EPPA) scheme, some works investigate the channel estimation performance via different evaluation metrics. For example, through the symbol error probability [12] and the bit error rate [13, 14] all reveal the channel estimation performance via the number of errors in the data transmission. The mean squared error (MSE) [15, 16] discloses the quality of the estimated channel based on the absolute channel estimation error. In [16], a general MSE expression was given based on the correlation matrix of the user channel. Considering the normalized relative error of the estimated channel, [10, 16, 17] utilized the normalized MSE (NMSE) metric, whilst [9] utilized the normalized channel estimation error metric. However, closed-form expressions for these normalized relative error metrics are barely available in the literature. Very recently, [18] provided an analytical expression for the NMSE, while mainly paying attention on how to extract the desired source data from the received signal; it also proposed a semi-blind channel estimation method. Finally, [19] studied a channel estimation MSE metric, which involves the channel estimation error and the true channel, and compared the proposed soft pilot reuse scheme and the all pilot reuse scheme through simulations. From the above discussion, it becomes apparent that a detailed theoretical analysis of such normalized channel estimation error metrics for massive MIMO is missing from the open literature.

On a parallel avenue, pilot power allocation (PPA) represents another formidable opportunity in massive MIMO systems to improve the system performance. In [20], a resource allocation scheme was proposed to maximize the sum spectral efficiency (SE), which determines the optimal values of the pilot sequence length, the pilot signal power, and the data signal power. A joint pilot and data transmit power control was provided in [21] to minimize the total power consumption of all users with the constraints of per user signal to interference plus noise ratio (SINR) and per user power. In [22], a PPA policy was developed to maximize the minimum asymptotic SINR in each cell, by adopting a pilot scheme where all users in each cell share the same pilot sequence and keep the pilot sequences orthogonal for different cells. In [23], a binary search PPA algorithm was provided to maximize the achievable downlink sum rate with matched filter based precoding in single cell massive MIMO. A pilot power control scheme was provided in [24], which can improve the channel estimation performance. Moreover, in [25], as the number of antennas increased, simulation results show that more power should be allocated for pilot transmission to improve the channel estimation performance. Recently, in [26], a joint optimal pilot and data power allocation was proposed in single cell uplink massive MIMO systems to maximize the minimum SE and maximize the sum SE through geometric programming and an algorithm based on Karush-Kuhn-Tucker (KKT) points, respectively. From the above, it is easy to see that the overwhelming majority of works use the PPA scheme to improve SE and SINR or other equivalent performance metrics.

In our work, we take a substantially different approach to the problem of channel estimation. In particular, we strive to improve channel estimation performance by designing a PPA scheme in a heuristic manner. In this paper, we study the performance of channel estimation and achievable uplink rate in multi-cell massive MIMO systems based on the least squares (LS) and minimum mean squared error (MMSE) estimation methods, respectively, by considering a modified NMSE metric derived from the MSE metric, called the relative channel estimation error (RCEE). Closed-form expressions for the expectation of RCEE ( $\text{Exp}_{\text{rcee}}$ ) and achievable uplink rate are obtained, which depend only on the large-scale fading coefficients. We conclude that when the number of base station (BS) antennas ( $M$ ) grows to infinity, the RCEE and the  $\text{Exp}_{\text{rcee}}$  approach to the some constant value. This important observation, which is due to the channel hardening effect, enables us to design a simple PPA algorithm, which aims at minimizing the average  $\text{Exp}_{\text{rcee}}$  per user in the target cell. Numerical simulations justify the accuracy of our analytical results, and it is found that the proposed PPA algorithm can approach the solution of the general constrained optimization problem very effectively. For the channel estimation performance, compared with the EPPA scheme, the MMSE PPA scheme obtains an increasing gain as the frequency reuse factor grows. However, the gain of the LS PPA scheme remains fixed. Most importantly, the simple LS estimation method can achieve almost the same performance as with the MMSE estimation method, with both higher frequency

reuse factors and larger total pilot power. Further, a joint user-cell grouping (JUCG) scheduling strategy is proposed for pilot power allocation across different cells to achieve nearly optimal performance. For the achievable uplink rate performance, compared with the EPPA scheme, the PPA scheme maintains almost the same average achievable uplink rate and improves considerably the minimum achievable uplink rate. In particular, the main contributions of the paper are outlined as follows.

- We investigate the channel estimation performance of multi-cell massive MIMO through the RCEE metric and  $\text{Exp}_{\text{rcee}}$  for which we derive a closed-form expression. Based on this expression, we also deduce a new and tractable expression for the achievable uplink rate.
- We develop a novel PPA scheme to improve the channel estimation performance through appropriate user grouping. Our analysis accounts for both the LS and MMSE estimation methods.
- We finally design a JUCG scheme to allocate pilot power across different cells. This novel scheme yields near-optimal performance with very low complexity.

The rest of the paper is organized as follows. The channel model and the procedure of the channel estimation and uplink data transmission are described in Section II. The optimization problem, the analysis of the RCEE and the  $\text{Exp}_{\text{rcee}}$ , and the corresponding PPA algorithm are presented in Section III. Section IV presents the simulation results to check the effectiveness of the PPA algorithm and compare the PPA and EPPA schemes on the basis of channel estimation and achievable uplink rate performance. Conclusions are presented in Section V.

*Notation:* Lower-case (underlined lower-case) and upper-case (underlined upper-case) boldface letters denote vectors and matrices, respectively. The  $\mathbb{C}^{M \times N}$  denotes the  $M \times N$  complex space. The notations  $\mathbf{A}^H$ ,  $\mathbf{A}^{-1}$ , and  $\text{tr}(\mathbf{A})$  indicate the Hermitian transpose, the inverse, and the trace of the matrix  $\mathbf{A}$ , respectively. The  $M \times M$  identity matrix is  $\mathbf{I}_M$ . The  $M \times N$  zero matrix is  $\mathbf{0}_{M \times N}$ . The expectation operation is  $\mathbb{E}\{\cdot\}$ . A complex Gaussian random vector  $\mathbf{x}$  is denoted as  $\mathbf{x} \sim \mathcal{CN}(\bar{\mathbf{x}}, \mathbf{\Sigma})$ , where the mean vector is  $\bar{\mathbf{x}}$  and the covariance matrix is  $\mathbf{\Sigma}$ . Finally,  $\|\cdot\|_2$  denotes the 2-norm of vector.

## II. SYSTEM MODEL

### A. Signal and Channel Model

The system architecture is a typical cellular communication system with  $L$  hexagonal cells. Each cell contains  $K$  single-antenna users and one  $M$ -antenna BS. We assume that the BSs and users in the whole system are perfectly synchronized. The whole system operates under a TDD protocol and adopts the orthogonal frequency division multiplexing (OFDM) technique. Moreover, for convenience, the dependency of the user's channel on the sub-carrier index is suppressed. Hence, for each channel use, the received signal vector  $\mathbf{y}_j \in \mathbb{C}^{M \times 1}$  at the BS in cell  $j$  is given by

$$\mathbf{y}_j = \sqrt{p_u} \sum_{l=1}^L \mathbf{H}_{jl} \mathbf{x}_l + \mathbf{n}_j, \quad (1)$$

where  $\mathbf{x}_l \in \mathbb{C}^{K \times 1} \sim \mathcal{CN}(\mathbf{0}_{K \times 1}, \mathbf{I}_K)$  is the transmit signal vector of the users in cell  $l$ ,  $p_u$  denotes the average normalized transmitted power of all users in all cells, and  $\mathbf{n}_j \in \mathbb{C}^{M \times 1}$  is the noise vector containing independent elements  $\sim \mathcal{CN}(0, 1)$  in cell  $j$ . Also,  $\mathbf{H}_{jl} \in \mathbb{C}^{M \times K}$  denotes the channel between all users in cell  $l$  and the BS in cell  $j$ , which is given by

$$\mathbf{H}_{jl} \triangleq [\mathbf{h}_{jl1}, \dots, \mathbf{h}_{jlk}, \dots, \mathbf{h}_{jLK}], \quad (2)$$

where  $\mathbf{h}_{jlk}$  is the uplink channel transmission vector between the user  $k$  in cell  $l$  and the BS in cell  $j$ . The channel  $\mathbf{h}_{jlk}$  is modeled as a combination of small-scale fading and large-scale fading and is written as [1, 5]

$$\mathbf{h}_{jlk} = \underline{\mathbf{h}}_{jlk} \beta_{jlk}^{\frac{1}{2}}, \quad (3)$$

where  $\underline{\mathbf{h}}_{jlk} \sim \mathcal{CN}(\mathbf{0}_{M \times 1}, \mathbf{I}_M)$  is the small-scale fading and  $\beta_{jlk}$  is the large-scale fading coefficient. Here, we make the block fading assumption that the large-scale fading coefficients are kept fixed over lots of coherence time intervals and also assume that large-scale fading coefficients are known at the BS [5], while small-scale fading coefficients remain fixed within a coherence time interval and change between any two adjacent coherence time intervals. At the same time, each user's channel is considered to be independent from other users' channels.

### B. Channel Estimation

At the start of the coherence interval, before the user sends data to the BS, the BS needs to acquire channel state information (CSI) by estimating the channel between the user and itself. We use uplink pilot sequences to perform channel estimation. The pilot sequence sent by the user  $k$  in cell  $l$  is

$$\mathbf{s}_{lk} = \sqrt{\rho_{lk}} \underline{\mathbf{s}}_{lk} \in \mathbb{C}^{\tau \times 1}, \quad (4)$$

where  $\underline{\mathbf{s}}_{lk} \in \mathbb{C}^{\tau \times 1}$  is the pilot sequence with length  $\tau$  sent by the user  $k$  in cell  $l$ , and  $\rho_{lk}$  is the pilot power of the user  $k$  in cell  $l$ .<sup>1</sup> Moreover, to ensure the orthogonality of users' pilot sequences within one cell, we set the pilot sequence length to be larger than the number of users, i.e.,  $\tau \geq K$ , and

$$\underline{\mathbf{s}}_{lk}^H \underline{\mathbf{s}}_{lk} = 1 \text{ and } \underline{\mathbf{s}}_{lk_1}^H \underline{\mathbf{s}}_{lk_2} = 0, \forall k_1 \neq k_2. \quad (5)$$

From the perspective of pilot contamination (PC), the worst case choice is to reuse  $\underline{\mathbf{s}}_{lk}$  in all  $L$  cells [1, 6, 9] for user  $k$ .<sup>2</sup> That is,  $\forall i, j$ ,

$$\underline{\mathbf{s}}_{ik} = \underline{\mathbf{s}}_{jk}. \quad (6)$$

Hence, in the phase of channel estimation, the BS in cell  $j$  receives the signal matrix

$$\mathbf{Y}_j = \sum_{l=1}^L \sum_{n=1}^K \sqrt{\rho_{ln}} \mathbf{h}_{jln} \underline{\mathbf{s}}_{ln}^H + \mathbf{N}_j, \quad (7)$$

<sup>1</sup>Since we elaborate on the channel estimation performance, we allow the pilot power in (4) to be variable for different users in each cell, while the transmit data power in (1) is kept the same for all users in all cells. This assumption also offers analytical tractability.

<sup>2</sup>Since we want to study only how the pilot power allocation affects the system performance, we make this worst-case assumption to ensure that each user in each cell is prone to the same number of interference sources.

where  $\mathbf{N}_j$  represents the  $M \times \tau$  additive white Gaussian noise matrix with independent and identically distributed zero-mean and unit-variance elements.

Then, in order to estimate the channel between the user  $k$  in cell  $j$  and the BS in cell  $j$ , while ignoring the terms in the right-hand side of (7) as noise except the term  $\sqrt{\rho_{jk}} \mathbf{h}_{jjk} \underline{\mathbf{s}}_{jk}^H$ , through the LS estimation method [27], we can get

$$\hat{\mathbf{h}}_{jjk}^{\text{LS}} \triangleq \frac{1}{\sqrt{\rho_{jk}}} \mathbf{Y}_j \underline{\mathbf{s}}_{jk}, \quad (8)$$

where  $\hat{\mathbf{h}}_{jjk}^{\text{LS}}$  is the estimated vector of channel  $\mathbf{h}_{jjk}$  based on the LS estimation method. Therefore, substituting (5)-(7) into (8), we obtain

$$\hat{\mathbf{h}}_{jjk}^{\text{LS}} = \mathbf{h}_{jjk} + \sum_{l \neq j}^L \frac{\sqrt{\rho_{lk}}}{\sqrt{\rho_{jk}}} \mathbf{h}_{jlk} + \frac{1}{\sqrt{\rho_{jk}}} \mathbf{N}_j \underline{\mathbf{s}}_{jk}. \quad (9)$$

Moreover, based on the distribution of  $\mathbf{h}_{jlk} (\forall j, l, k)$  and  $\mathbf{N}_j$ , the MMSE estimation method, whose target is to minimize the MSE of the estimated parameter, can be used to estimate  $\mathbf{h}_{jjk}$  with the help of  $\hat{\mathbf{h}}_{jjk}^{\text{LS}}$  [27]. We define

$$\hat{\mathbf{h}}_{jjk}^{\text{MMSE}} \triangleq \mathbb{E} \left\{ \mathbf{h}_{jjk} (\hat{\mathbf{h}}_{jjk}^{\text{LS}})^H \right\} \left( \mathbb{E} \left\{ \hat{\mathbf{h}}_{jjk}^{\text{LS}} (\hat{\mathbf{h}}_{jjk}^{\text{LS}})^H \right\} \right)^{-1} \hat{\mathbf{h}}_{jjk}^{\text{LS}}, \quad (10)$$

to denote the estimator of the channel  $\mathbf{h}_{jjk}$  based on the MMSE estimation method. Then, it is easy to prove that  $\mathbf{N}_j \underline{\mathbf{s}}_{jk} \sim \mathcal{CN}(\mathbf{0}_{M \times 1}, \mathbf{I}_M)$  in (9).<sup>3</sup> By substituting (3) and (9) into (10), after some manipulations, we get

$$\hat{\mathbf{h}}_{jjk}^{\text{MMSE}} = \frac{\rho_{jk} \beta_{jjk}}{\sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1} \hat{\mathbf{h}}_{jjk}^{\text{LS}}. \quad (11)$$

In the following, we denote the estimator of  $\mathbf{h}_{jjk}$  as  $\hat{\mathbf{h}}_{jjk}$  for both the LS and MMSE estimation methods, except otherwise denoted.

### C. Uplink Data Transmission

After channel estimation, the BS in each cell uses the obtained CSI to detect the received signal of (1). We consider the standard linear detector maximal-ratio combining (MRC) [5]. Hence, for the BS in cell  $j$ , the received signal  $\mathbf{y}_j$  in (1) is separated into  $K$  streams by multiplying it with the MRC detector, that is,

$$\mathbf{r}_j = \hat{\mathbf{H}}_{jj}^H \mathbf{y}_j \in \mathbb{C}^{K \times 1}, \quad (12)$$

where

$$\hat{\mathbf{H}}_{jj} \triangleq [\hat{\mathbf{h}}_{jj1}, \dots, \hat{\mathbf{h}}_{jjk}, \dots, \hat{\mathbf{h}}_{jjK}] \in \mathbb{C}^{M \times K}. \quad (13)$$

<sup>3</sup>Since we consider the block fading assumption in this paper, the operation  $\mathbb{E}\{\cdot\}$  regards the variables, such as small-scale fading coefficient and noise, which change over time, as the random variables and the large-scale fading coefficient as a constant.

Then, expanding  $\mathbf{y}_j$  and denoting the  $n$ th component value of  $\mathbf{x}_l$  as  $x_{ln}$ , the  $k$ th entry  $r_{jk}$  of  $\mathbf{r}_j$  can be written as

$$r_{jk} = \underbrace{\sqrt{\rho_u} \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjk} x_{jk}}_{\text{Desired signal}} + \underbrace{\sqrt{\rho_u} \sum_{n \neq k} \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjn} x_{jn}}_{W_1: \text{Intra-cell interference}} + \underbrace{\sqrt{\rho_u} \sum_{l \neq j} \sum_{n=1}^K \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jln} x_{ln}}_{W_2: \text{Inter-cell interference}} + \underbrace{\hat{\mathbf{h}}_{jjk}^H \mathbf{n}_j}_{W_3: \text{Noise}}. \quad (14)$$

Since we want to investigate the achievable uplink rate of the  $k$ th user in cell  $j$ , using the worst-case technique of Gaussian noise from [6, 7], we assume the term  $\mathbb{E} \left\{ \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjk} \right\}$  is perfectly known at the BS  $j$ . Hence,  $r_{jk}$  is written as

$$r_{jk} = \underbrace{\sqrt{\rho_u} \mathbb{E} \left\{ \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjk} \right\} x_{jk}}_{\text{Effective signal}} + \underbrace{\tilde{n}_{jk}}_{\text{Equivalent noise}}, \quad (15)$$

where

$$\tilde{n}_{jk} \triangleq \sqrt{\rho_u} \left\{ \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjk} - \mathbb{E} \left\{ \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjk} \right\} \right\} x_{jk} + W_1 + W_2 + W_3. \quad (16)$$

From (14)-(16), we can infer that the ‘‘Effective signal’’ is uncorrelated with the ‘‘Equivalent noise’’. Hence, considering the effects of the pilot overhead, total bandwidth, frequency reuse factors, and the overhead of the cyclic prefix as in [1, Eq. (14)], the achievable uplink rate of user  $k$  in cell  $j$ , in units of bits/sec, is given by

$$R_{jk} = \left( \frac{B}{\Gamma} \right) \left( \frac{T_s - T_p}{T_s} \right) \left( \frac{T_u}{T_o} \right) \log_2 (1 + \text{SINR}_{jk}), \quad (17)$$

where  $\Gamma$  is the frequency reuse factor,  $B$  is the total bandwidth,  $T_s$  is the slot length,  $T_p$  is the time spent transmitting pilot sequences,  $T_u$  is the useful symbol duration, and  $T_o$  is the OFDM symbol interval. Also, using the definition of the effective SINR in multi-cell massive MIMO systems as in [28, Eq. (12)] and [29, Eq. (20)],  $\text{SINR}_{jk}$  is defined as

$$\text{SINR}_{jk} \triangleq \frac{\rho_u \left| \mathbb{E} \left\{ \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjk} \right\} \right|^2}{\rho_u \sum_{l=1}^L \sum_{n=1}^K \mathbb{E} \left\{ \left| \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jln} \right|^2 \right\} - \rho_u \left| \mathbb{E} \left\{ \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjk} \right\} \right|^2 + \mathbb{E} \left\{ \left\| \hat{\mathbf{h}}_{jjk} \right\|_2^2 \right\}}. \quad (18)$$

The following theorem presents a closed-form expression for  $\text{SINR}_{jk}$  for both the LS and MMSE estimation methods.

*Theorem 1:* The exact  $\text{SINR}_{jk}$ , for both the LS and MMSE estimation methods, can be analytically evaluated as

$$\text{SINR}_{jk} = \frac{M \rho_{jk} \beta_{jjk}^2}{M \sum_{l \neq j} \rho_{lk} \beta_{jlk}^2 + \left( \sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1 \right) \left( \frac{1}{\rho_u} + \sum_{l=1}^L \sum_{n=1}^K \beta_{jln} \right)}. \quad (19)$$

*Proof:* See Appendix A. ■

Note that the exact expression in *Theorem 1* can be easily evaluated since it involves only the pilot power, data power,

and large-scale fading coefficients, as well as,  $M$ . Interestingly, from (19), the achievable uplink rate for both the LS and MMSE estimation methods are identical, and this result can be easily obtained by substituting (11) into (18) for the MMSE case and then divide both the numerator and the denominator of (18) by the common factor  $(\rho_{jk} \beta_{jjk})^2 / (\sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1)^2$ .

Note that this common factor stems from the fact that (11) of the LS estimator is proportional to the MMSE estimator. The underlying reason behind (11) is that the Rayleigh fading model with no correlation between the antennas in the BS is considered, that is,  $\mathbb{E} \{ \mathbf{h}_{jlk} \mathbf{h}_{jlk}^H \} = \beta_{jlk} \mathbf{I}_M (\forall j, l, k)$ . The fact that LS and MMSE estimators can achieve the same uplink rate with an MRC detector was first observed in [30] for massive MIMO systems with no correlation between the antennas. However, the authors in [30] just gave a qualitative explanation, while we also give the quantitative analysis in *Theorem 1*. As the next result shows, the exact  $\text{SINR}_{jk}$  admits further simplifications in the large antenna regime.

*Corollary 1:* When  $M \rightarrow \infty$ , the exact analytical expression in (19) approaches to

$$\lim_{M \rightarrow \infty} \text{SINR}_{jk} = \frac{\rho_{jk} \beta_{jjk}^2}{\sum_{l \neq j} \rho_{lk} \beta_{jlk}^2}. \quad (20)$$

*Proof:* The proof is completed by calculating the limit of (19) when  $M \rightarrow \infty$ . ■

It is important to note from (20) that in the high  $M$  regime, the PC interference from the other cells’ users, which use the same pilot sequence as the user in the target cell, is the only limit for the achievable uplink rate performance. Moreover, when  $\rho_{jk} = \rho_{lk}, \forall j, l$ , the  $\text{SINR}_{jk}$  is equal to the signal-to-interference ratio (SIR) based on LS estimation in [1, Eq. (13)] and the SIR based on MMSE estimation in [31, Eq. (32)], respectively. Hence, the conclusion of *Corollary 1* gives a universal formula for  $\text{SINR}_{jk}$  when  $M \rightarrow \infty$  for any pilot power setting.

### III. PILOT POWER ALLOCATION

In this section, we introduce the RCEE metric and, thereafter, aim to find a PPA scheme to minimize the average  $\text{Exp}_{\text{rcee}}$  per user. Based on the closed-form expressions for  $\text{Exp}_{\text{rcee}}$ , we propose a PPA algorithm to solve this optimization problem.

#### A. RCEE

We can now define the RCEE, which is basically a modified NMSE metric derived from the MSE metric, of user  $k$  in cell  $j$  as follows

$$\Lambda_{jk} \triangleq \frac{\| \mathbf{h}_{jjk} - \hat{\mathbf{h}}_{jjk} \|^2}{\| \mathbf{h}_{jjk} \|^2}, \quad (21)$$

and also, its expectation or  $\text{Exp}_{\text{rcee}}$  given by

$$\Delta_{jk} \triangleq \mathbb{E} \{ \Lambda_{jk} \}. \quad (22)$$

The former metric indicates the instantaneous relative change between the channel estimation error and the true channel in any coherence interval, while the latter metric indicates this relative change over many coherence intervals. The following theorem gives a closed-form expression for  $\text{Exp}_{\text{rcee}}$ .

*Theorem 2:* The term  $\Delta_{jk}$  can be analytically evaluated as

$$\Delta_{jk} = \begin{cases} \infty, & \text{LS\&MMSE, } M=1, \\ \frac{M \left( \sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1 \right)}{(M-1) \rho_{jk} \beta_{jjk}}, & \text{LS, } M \geq 2, \\ \frac{\left( \sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1 \right) \left( \sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1 + \frac{M \rho_{jk} \beta_{jjk}}{M-1} \right)}{\left( \sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1 \right)^2}, & \text{MMSE, } M \geq 2. \end{cases} \quad (23)$$

*Proof:* See Appendix B. ■

Note that the formulas in (23) are only meaningful when  $M \geq 2$ . Moreover, the exact analytical expressions in *Theorem 2* can be easily evaluated as they involve only the number of BS antennas, pilot power, as well as, the large-scaling fading coefficients. Compared with  $\Delta_{jk}^{\text{LS}}$ ,  $\Delta_{jk}^{\text{MMSE}}$  is more complicated.<sup>4</sup> Hence, the following corollary gives an upper bound for  $\Delta_{jk}^{\text{MMSE}}$  in the  $M \geq 2$  regime.

*Corollary 2:* In the  $M \geq 2$  regime,  $\Delta_{jk}^{\text{MMSE}}$  satisfies

$$\Delta_{jk}^{\text{MMSE}} < \tilde{\Delta}_{jk}^{\text{MMSE}} = \frac{M \left( \sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1 \right)}{(M-1) \left( \sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1 \right)}. \quad (24)$$

*Proof:* The proof is trivial and thus omitted. ■

It is important to note that  $\tilde{\Delta}_{jk}^{\text{MMSE}}$  in *Corollary 2* will be particularly useful for the PPA problem in subsection III.C.

*Corollary 3:* When  $M \rightarrow \infty$ , the exact analytical expressions in (23) approach to

$$\lim_{M \rightarrow \infty} \Delta_{jk} = \bar{\Delta}_{jk} = \begin{cases} \frac{\sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1}{\rho_{jk} \beta_{jjk}}, & \text{LS,} \\ \frac{\sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1}{\sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1}, & \text{MMSE,} \end{cases} \quad (25)$$

whilst  $\tilde{\Delta}_{jk}^{\text{MMSE}}$  also approaches to the same limit as  $\Delta_{jk}^{\text{MMSE}}$ .

*Proof:* The proof is completed by calculating the limit of (23) and  $\tilde{\Delta}_{jk}^{\text{MMSE}}$  when  $M \rightarrow \infty$ . ■

It is interesting to note from *Corollary 3* that as  $M$  grows,  $\text{Exp}_{\text{rcee}}$  will decrease and approach a constant value for both the LS and MMSE estimation methods. Moreover, from (23) and (25), we see that MMSE performs better than LS since the

MMSE estimation method utilizes the additional second-order statistical information.

*Corollary 4:* When  $M \rightarrow \infty$ , the relationship between the RCEE and the  $\text{Exp}_{\text{rcee}}$  is given by

$$\Lambda_{jk} \xrightarrow{a.s.} \bar{\Delta}_{jk}, \quad (26)$$

where  $\xrightarrow{a.s.}$  denotes almost sure convergence.

*Proof:* By utilizing the law of large numbers [5, Eq. (4)], we can obtain

$$\frac{1}{M} (\mathbf{h}_{jjk} - \hat{\mathbf{h}}_{jjk})^H (\mathbf{h}_{jjk} - \hat{\mathbf{h}}_{jjk}) \xrightarrow{a.s.} \frac{1}{M} \mathbb{E} \{ \|\mathbf{h}_{jjk} - \hat{\mathbf{h}}_{jjk}\|_2^2 \}, \quad (27)$$

$$\frac{1}{M} \mathbf{h}_{jjk}^H \mathbf{h}_{jjk} \xrightarrow{a.s.} \frac{1}{M} \mathbb{E} \{ \|\mathbf{h}_{jjk}\|_2^2 \}. \quad (28)$$

The proof is completed by substituting (27) and (28) into (21) and utilizing the Appendix B. ■

It is important to note that the Rayleigh fading model with no correlation between the antennas in the BS is adopted in this paper. Hence, the analysis of the relationship between the  $\Lambda_{jk}$  and  $\text{Exp}_{\text{rcee}}$  is proper when  $M \rightarrow \infty$ . The case of spatial correlation is outside the scope of this paper and left to future work. Interestingly, from *Corollary 3* and *Corollary 4*, RCEE converge to its statistical value ( $\text{Exp}_{\text{rcee}}$ ) when  $M$  is large. In other words, the stochastic nature of the RCEE disappears when  $M$  is big enough. Note that (26) reflects the channel hardening effect via the channel estimation performance, though in a slightly different way than in [32].

## B. Constrained Optimization Problem

Now, to avoid an iterative non-stationary optimization problem, we focus on the PPA in one cell, while the pilot power in other cells is kept fixed, and we call it as the PPA scheme. It is clear that the RCEE which is obtained from one time channel estimation is not able to represent the channel estimation performance in a period of time. For the massive MIMO setup consideration, we can avail of the channel hardening effect and work exclusively with the  $\text{Exp}_{\text{rcee}}$  instead of RCEE. Moreover, considering the fairness of all users in the target cell, we choose the average  $\text{Exp}_{\text{rcee}}$  per user as our objective function to evaluate the system channel estimation performance. We can now formulate the following constrained PPA optimization problem

$$\begin{aligned} & \text{minimize} \quad \frac{1}{K} \sum_{k=1}^K \Delta_{jk}, \\ & \text{subject to} \quad \sum_{k=1}^K \rho_{jk} \leq P, \\ & \quad \quad \quad \rho_{\min} \leq \rho_{jk} \leq \rho_{\max}, \forall k = 1, \dots, K, \end{aligned} \quad (29)$$

where  $P$  is the total pilot power budget of the users in each cell,  $\rho_{\min}$  and  $\rho_{\max}$  denote the lower and upper bounds of the

<sup>4</sup>Here, for convenience, we use the notation  $\Delta_{jk}^{\text{LS}}$  and  $\Delta_{jk}^{\text{MMSE}}$  to replace  $\Delta_{jk}$  based on LS and MMSE estimation methods, respectively.

variation range of  $\rho_{jk}$ .<sup>5</sup> Once the problem (29) is solved, it means that the pilot power  $\rho_{jk}$  is determined.

We now turn our attention to the special case, EPPA scheme, where all cell users have the same pilot power setting ( $\rho_{jk} = P/K, \forall j, k$ ), and obtain the following corollary.

*Corollary 5:* When  $M \rightarrow \infty$  and for the EPPA scheme, the exact analytical expressions in (23) approach to

$$\lim_{M \rightarrow \infty} \Delta_{jk} = \bar{\Delta}_{jk} = \begin{cases} \frac{\sum_{l \neq j}^L \beta_{jlk} + \frac{K}{P}}{\beta_{jjk}}, & \text{LS,} \\ \frac{\sum_{l \neq j}^L \beta_{jlk} + \frac{K}{P}}{\sum_{l=1}^L \beta_{jlk} + \frac{K}{P}}, & \text{MMSE.} \end{cases} \quad (30)$$

*Proof:* The proof is completed by calculating the limit of (23) when considering the EPPA scheme and  $M \rightarrow \infty$ . ■

### C. Approximate Unconstrained Solution

According to *Theorem 2*, we have already obtained the asymptotic analytical expression for the objective function in (29). However,  $\Delta_{jk}^{\text{MMSE}}$  is complicated and the second constraint in (29) makes it challenging to solve. Hence, we now replace  $\Delta_{jk}^{\text{MMSE}}$  with  $\tilde{\Delta}_{jk}^{\text{MMSE}}$  and release the constraint  $\rho_{\min} \leq \rho_{jk} \leq \rho_{\max}$  in (29). The optimization problem can be rewritten as

$$\begin{aligned} & \text{minimize} \quad \frac{1}{K} \sum_{k=1}^K \tilde{\Delta}_{jk}, \\ & \text{subject to} \quad \sum_{k=1}^K \rho_{jk} \leq P, \end{aligned} \quad (31)$$

where

$$\tilde{\Delta}_{jk} \triangleq \begin{cases} \frac{M \left( \sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1 \right)}{(M-1) \rho_{jk} \beta_{jjk}}, & \text{LS,} \\ \frac{M \left( \sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1 \right)}{(M-1) \left( \sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1 \right)}, & \text{MMSE.} \end{cases} \quad (32)$$

The following theorem presents closed-form expressions for the optimal solution of (31).

*Theorem 3:* The optimal solution  $\rho_{jk}^*$  of (31) is formulated as

$$\rho_{jk}^* = \begin{cases} \frac{1}{\lambda_{\text{LS}}} \left( \frac{v_{jk}}{\beta_{jjk}} \right)^{\frac{1}{2}}, & \text{LS,} \\ \frac{1}{\lambda_{\text{MMSE}}} \left( \frac{v_{jk}}{\beta_{jjk}} \right)^{\frac{1}{2}} - \frac{v_{jk}}{\beta_{jjk}}, & \text{MMSE,} \end{cases} \quad (33)$$

<sup>5</sup>Here, to ensure the user in each cell has at least the minimum power for pilot transmission, we assume the minimum pilot power is  $\rho_{\min} = P/2K$ . Then, to ensure that a user does not use too much pilot power that decreases other users' pilot power substantially, we assume the maximum pilot power is  $\rho_{\max} = \mu P/K$ . To ensure that at least one user in the target cell can be allocated the maximum power to sent pilots, we let  $K \geq 2$  and  $\rho_{\min}(K-1) + \rho_{\max} \leq P$ , that is,  $\mu \leq (K+1)/2$ . Finally, to ensure that the channel estimation performance of the user who is allocated the maximum power can be improved effectively, we assume  $\mu \geq 3/2$ . Hence,  $\mu \in [3/2, (K+1)/2]$ .

where

$$v_{jk} \triangleq \sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1, \quad (34)$$

$$\lambda_{\text{LS}} \triangleq \frac{1}{P} \sum_{k=1}^K \left( \frac{v_{jk}}{\beta_{jjk}} \right)^{\frac{1}{2}}, \quad (35)$$

$$\lambda_{\text{MMSE}} \triangleq \frac{\sum_{k=1}^K \left( \frac{v_{jk}}{\beta_{jjk}} \right)^{\frac{1}{2}}}{P + \sum_{k=1}^K \frac{v_{jk}}{\beta_{jjk}}}. \quad (36)$$

*Proof:* The proof is completed by using the Lagrange multiplier method [33]. ■

Note that from (33)-(36), we see that the optimal solution of (31) is independent of the number of BS antennas. Most importantly, the optimal solution in (31) can serve as a very effective starting point for solving the constrained optimization problem (29) approximately. This is one of the main ideas behind the PPA algorithm. The second idea is to partition the users in cell  $j$  into three groups. Users in group 1 have the same pilot power  $\rho_{\min}$ . Users in group 2 have the same pilot power  $\rho_{\max}$ . Group 3 users' pilot power is determined by solving (31) with the remaining available total pilot power budget. We firstly consider the optimal solution in (31). Next, if the optimal solution in (31) satisfies the power constraint in (29), then all users are put into the group 3 and the whole process stops. Otherwise, we select the user whose pilot power is far lower than  $\rho_{\min}$  or far more than  $\rho_{\max}$ . Finally, we allocate the selected user into the group 1 or group 2, and return to carry out the above process for the rest of users and the remaining available total pilot power budget in the target cell until all users' pilot power, in the target cell, meets the power constraint in (29). Inspired by this, we propose a PPA algorithm (**Algorithm 1**), based on a simple decision criterion for both the LS and MMSE estimation methods, to obtain an approximate solution of the optimization problem (29). In **Algorithm 1**,  $\mathcal{K}_{P,\min}$  and  $\mathcal{K}_{P,\max}$  denote the user groups where users' pilot powers are the minimum pilot power  $\rho_{\min}$  and the maximum pilot power  $\rho_{\max}$ , respectively. Also, after removing the users in  $\mathcal{K}_{P,\min}$  or  $\mathcal{K}_{P,\max}$ , the remaining users belong to the user group  $\mathcal{K}_P$ . The users' pilot powers have the following relationship

$$\sum_{k \in \mathcal{K}_{P,\min}} \rho_{\min} + \sum_{k \in \mathcal{K}_{P,\max}} \rho_{\max} + \sum_{k \in \mathcal{K}_P} \rho_{jk} = P. \quad (37)$$

These groups are initialized in **Algorithm 1**. The cardinalities of these groups are denoted as  $K_{P,\min}$ ,  $K_{P,\max}$  and  $K_P$ , respectively, which satisfy

$$K_P + K_{P,\min} + K_{P,\max} = K. \quad (38)$$

Generally speaking, the core idea of the PPA algorithm is to give an approximately optimal partition of the users in the target cell through the quality of their channels. Note that the entire procedure of the PPA algorithm is depicted in Fig. 1. Given the fact that in the massive MIMO regime, the stochastic nature of RCEE vanishes, large-scale fading coefficients are

**Algorithm 1** PPA algorithm

---

```

1: procedure PPA
2:   input:  $P, K, \beta_{jlk}, \rho_{lk} (l \neq j), \rho_{\min}, \rho_{\max},$ 
3:     and Flag (Flag = 0 or Flag = 1).
4:   initialization:  $\mathcal{K}_P = \{1, \dots, K\}, \mathcal{K}_{P,\min} = \emptyset,$ 
5:     and  $\mathcal{K}_{P,\max} = \emptyset.$ 
6:   if Flag == 0
7:     Calculate the first expression in (33).
8:   else
9:     Calculate the second expression in (33).
10:  end if
11:  for  $i \leftarrow 1, \dots, K$  do
12:    if  $\rho_{jk} \geq \rho_{\min} \ \& \ \rho_{jk} \leq \rho_{\max}, \forall k \in \mathcal{K}_P$ 
13:      break
14:    end if
15:    Obtain  $\{k_i \mid \rho_{jk_i} < \rho_{\min}, k_i \in \mathcal{K}_P\},$ 
16:      and  $\{t_i \mid \rho_{jt_i} > \rho_{\max}, t_i \in \mathcal{K}_P\}.$ 
17:    Obtain  $k_i^* = \arg \max_{k_i^* \in \{k_i\}} |\rho_{jk_i} - \rho_{\min}|,$ 
18:      and  $t_i^* = \arg \max_{t_i^* \in \{t_i\}} |\rho_{jt_i} - \rho_{\max}|.$ 
19:    if  $(k_i^* \& t_i^* \& |\rho_{jk_i^*} - \rho_{\min}| \geq |\rho_{jt_i^*} - \rho_{\max}|) \parallel \{t_i\} = \emptyset$ 
20:       $\rho_{jk_i^*} \leftarrow \rho_{\min}, P \leftarrow P - \rho_{jk_i^*}, \mathcal{K}_P \leftarrow \mathcal{K}_P \setminus \{k_i^*\},$ 
21:      and  $\mathcal{K}_{P,\min} \leftarrow \mathcal{K}_{P,\min} \cup \{k_i^*\}.$ 
22:    if Flag == 0
23:      Replace  $\sum_{k=1}^K$  to  $\sum_{k \in \mathcal{K}_P}$  in (35) and calculate (35).
24:    for  $k \in \mathcal{K}_P$  do
25:      Calculate the first expression in (33).
26:    end for
27:    else
28:      Replace  $\sum_{k=1}^K$  to  $\sum_{k \in \mathcal{K}_P}$  in (36) and calculate (36).
29:    for  $k \in \mathcal{K}_P$  do
30:      Calculate the second expression in (33).
31:    end for
32:    end if
33:  else
34:     $\rho_{jt_i^*} \leftarrow \rho_{\max}, P \leftarrow P - \rho_{jt_i^*}, \mathcal{K}_P \leftarrow \mathcal{K}_P \setminus \{t_i^*\},$ 
35:    and  $\mathcal{K}_{P,\max} \leftarrow \mathcal{K}_{P,\max} \cup \{t_i^*\}.$ 
36:    Execute the procedure of lines 22-32.
37:  end if
38: end for
39: output:  $\rho_{jk}, \mathcal{K}_P, \mathcal{K}_{P,\min},$  and  $\mathcal{K}_{P,\max}.$ 
40: end procedure

```

---

the main parameters in our objective function in (29). This observation simplifies the problem of PPA substantially, since we can simply put the large-scale fading coefficients and other parameters into the PPA algorithm and determine the pilot power of the users in the target cell. Moreover, we have considered the block fading assumption such that the large-scale fading coefficients remain fixed over lots of coherence time intervals. Hence, once we start the PPA algorithm and obtain the pilot power, regardless of the small-scale variations between two consecutive coherence intervals, our pilot power keeps fixed within these coherence time intervals.

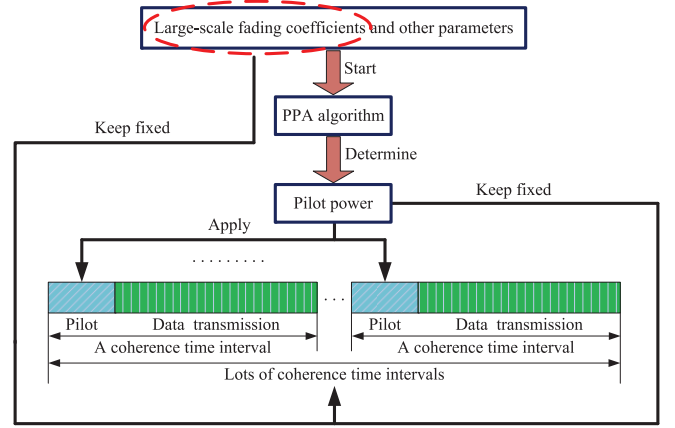


Fig. 1: The procedure of the PPA algorithm.

Since we propose a PPA algorithm to allocate pilot power in the target cell to minimize the average  $\text{Exp}_{\text{rcee}}$  per user for a given  $P$ , the limiting performance will also be obtained with infinitely high  $P$ . Hence, the following theorem gives the asymptotic expression for  $\text{Exp}_{\text{rcee}}$  when both  $P$  and  $M$  tend to infinity.<sup>6</sup>

*Theorem 4:* When  $M \rightarrow \infty$  and  $P \rightarrow \infty$ , and with the help of (25),  $\text{Exp}_{\text{rcee}}$  is asymptotically approximated by

$$\lim_{P \rightarrow \infty} \bar{\Delta}_{jk} = \check{\Delta}_{jk}^{\text{LS}} \approx \begin{cases} \frac{\sum_{l \neq j}^L \delta_{lk} \beta_{jlk}}{\alpha \beta_{jjk}}, & k \in \dot{\mathcal{K}}_{\min}, \\ \frac{\sum_{l \neq j}^L \delta_{lk} \beta_{jlk}}{\mu \beta_{jjk}}, & k \in \dot{\mathcal{K}}_{\max}, \\ \frac{\phi}{\varphi \psi}, & k \in \dot{\mathcal{K}}, \end{cases} \quad (39)$$

with LS estimation, and

$$\lim_{P \rightarrow \infty} \bar{\Delta}_{jk} = \check{\Delta}_{jk}^{\text{MMSE}} \approx \begin{cases} \frac{\sum_{l \neq j}^L \delta_{lk} \beta_{jlk}}{\sum_{l \neq j}^L \delta_{lk} \beta_{jlk} + \alpha \beta_{jjk}}, & k \in \dot{\mathcal{K}}_{\min}, \\ \frac{\sum_{l \neq j}^L \delta_{lk} \beta_{jlk}}{\sum_{l \neq j}^L \delta_{lk} \beta_{jlk} + \mu \beta_{jjk}}, & k \in \dot{\mathcal{K}}_{\max}, \\ \frac{\phi}{(\varphi + \omega) \psi}, & k \in \dot{\mathcal{K}}, \end{cases} \quad (40)$$

with MMSE estimation, where  $\delta_{lk} \in (0, 1)$  is the scale factor which satisfies  $\delta_{lk} = \rho_{lk}/P$  for  $l \neq j$ ,  $\alpha = \rho_{\min}K/P = 0.5$ ,

<sup>6</sup>Note that the relationship between the RCEE and  $\text{Exp}_{\text{rcee}}$  has already been given in *Corollary 4* when  $M \rightarrow \infty$ . Hence, for convenience, we just study the performance of  $\text{Exp}_{\text{rcee}}$  in the subsequent parts.



and

$$\phi \triangleq \sum_{l \neq j}^L \delta_{lk} \beta_{jlk} \sum_{k \in \mathcal{K}} \left( \frac{1}{\beta_{jjk}} \sum_{l \neq j}^L \delta_{lk} \beta_{jlk} \right)^{\frac{1}{2}}, \quad (41)$$

$$\varphi \triangleq 1 - \frac{\alpha \dot{K}_{\min} + \mu \dot{K}_{\max}}{K}, \quad (42)$$

$$\psi \triangleq \left( \beta_{jjk} \sum_{l \neq j}^L \delta_{lk} \beta_{jlk} \right)^{\frac{1}{2}}, \quad (43)$$

$$\varpi \triangleq \sum_{k \in \mathcal{K}} \frac{\sum_{l \neq j}^L \delta_{lk} \beta_{jlk}}{\beta_{jjk}}. \quad (44)$$

Also,  $\dot{K}_{\min}$ ,  $\dot{K}_{\max}$ ,  $\dot{K}$ ,  $\dot{K}_{\min}$ ,  $\dot{K}_{\max}$ , and  $\dot{K}$  are defined in Appendix C.<sup>7</sup>

*Proof:* See Appendix C. ■

It is important to note that, as  $M \rightarrow \infty$  and  $P \rightarrow \infty$ ,  $\text{Exp}_{\text{rcee}}$  is approximately controlled by the large-scale fading coefficients and the ratio of the user's pilot power to the total pilot power budget. Also, by comparing (39) and (40), for the user groups  $\dot{K}_{\min}$  and  $\dot{K}_{\max}$ , respectively, the denominator of (40) contains the impact of the users' large-scale fading coefficients in the other cells compared with the denominator of (39). The main reason of this phenomenon is that the MMSE estimation method accounts for inter-cell interference (the large-scale fading from other cells), whereas the LS estimation method treats the interference from other cells as noise. Interestingly, for the group  $\dot{K}$ , the main difference between the third expression for (39) and (40) is the term  $\varpi$  since the MMSE estimation method considers the effect of  $\beta_{jlk}$  ( $l \neq j$ ).

We now consider the special case of the EPPA scheme. The following corollary evaluates RCEE and  $\text{Exp}_{\text{rcee}}$  with EPPA scheme, when  $M$  and  $P$  are asymptotically large.

*Corollary 6:* When  $M \rightarrow \infty$  and  $P \rightarrow \infty$ , for the EPPA scheme,

$$\lim_{P \rightarrow \infty} \bar{\Delta}_{jk} = \begin{cases} \frac{\sum_{l \neq j}^L \beta_{jlk}}{\beta_{jjk}}, & \text{LS,} \\ \frac{\sum_{l \neq j}^L \beta_{jlk}}{\sum_{l=1}^L \beta_{jlk}}, & \text{MMSE.} \end{cases} \quad (45)$$

*Proof:* The proof is completed by calculating the limit of (30) when  $P \rightarrow \infty$ . ■

Note that in the EPPA scheme, the expressions for  $\text{Exp}_{\text{rcee}}$  with infinite  $M$  and  $P$  are simpler than that in PPA scheme. Moreover, (45) provides physical insights into the channel estimation performance for the LS and MMSE estimation methods, respectively, to reflect the PC effect; interestingly, these expressions are inversely analogous to the SIR in [1, Eq. (13)].

<sup>7</sup>In fact, the user groups for the LS and MMSE estimation methods may be different. However, for convenience, we still use these groups  $\dot{K}_{\min}$ ,  $\dot{K}_{\max}$ , and  $\dot{K}$  to represent both the LS and MMSE estimation methods.

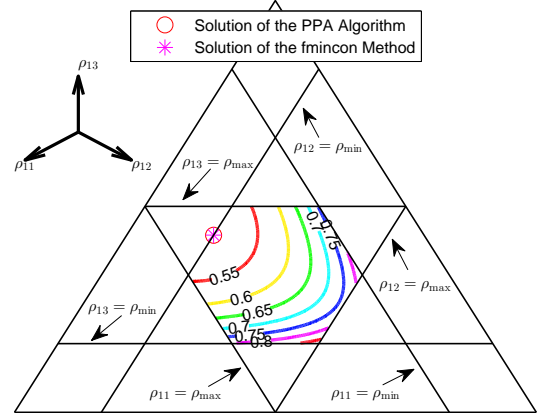


Fig. 2: The objective function values in (29) and the solutions of PPA algorithm and fmincon method are shown in the form of contour lines.

#### IV. NUMERICAL RESULTS

In this section, we consider a hexagonal cellular network with a set of  $L$  cells and radius  $r$  meters where users are distributed uniformly in each cell. Also, the large-scale fading coefficients, which account for geometric attenuation and shadow fading, are set as [34]

$$\beta_{jlk} = \frac{z_{jlk}}{1 + (r_{jlk}/r_{\min})^\gamma}, \quad (46)$$

where  $z_{jlk}$  is a log-normal random variable with standard deviation  $\sigma$ ,  $\gamma$  is the path loss exponent,  $r_{jlk}$  is the distance between the user  $k$  in cell  $l$  and the BS in cell  $j$ , and  $r_{\min}$  is the reference distance. In our simulations, we choose  $L = 7$ ,  $r = 500\text{m}$ ,  $\sigma = 8\text{dB}$ ,  $\gamma = 3.8$ , and  $r_{\min} = 200\text{m}$ , which also follow the methodology of [34].

##### A. Verification of the PPA Algorithm

In this subsection, we will check the PPA algorithm for the LS method based on *Theorem 3*, while the evaluation based on the MMSE estimation method can be omitted due to similarity. In order to give a visual display of our PPA algorithm to solve (29), the number of users  $K$  in each cell is assumed to be 3. Since we assume that the noise variance is 1,  $P$  can be interpreted as the transmit signal to noise ratio (SNR) and, thus, can be expressed in dB. For convenience, we set the average per user pilot power  $P/K$  to 30 dB. Also, we set the pilot sequence length to  $\tau = K$ ,  $\mu = 1.5$ , and  $M = 200$ .

We choose the center cell of the 7-cell hexagonal cellular network as our objective and call it cell 1, which means we choose  $j = 1$ . The large-scale fading coefficients of the channel between the users in cell 1-7 and the BS in cell 1 are randomly generated and given in Table Ia.

We use the PPA algorithm to allocate the three users' pilot power in cell 1. In order to validate the effectiveness of the PPA algorithm, we also compute the approximate solution of the constrained optimization problem in (29) through the

TABLE I

(a) List of the large-scale coefficients			
$\beta_{1lk}$ \begin{smallmatrix} User $k$ \\ Cell $l$ \end{smallmatrix>	1	2	3
1	0.0304	1.2899	0.0655
2	0.0006	0.0290	0.0389
3	0.0045	0.0024	0.0070
4	0.0080	0.0039	0.0045
5	0.0008	0.0842	0.0028
6	0.0078	0.0003	0.0026
7	0.0011	0.0177	0.0014

(b) List of the average runtime values for solving (29)		
$K$	Average runtime $\times 10^{-5}$ (s)	
	Method	
	PPA	fmincon
2	2.2	1239.2
3	4.0	1354.6
4	6.4	1405.4
5	9.1	1471.3
6	12.3	1573.4
7	16.2	1702.1
8	20.8	1898.0
9	25.6	1966.2
10	31.4	2018.7

MATLAB function “fmincon”. Note that we use the solution of (31) as the start point for fmincon. Based on the specific structure and the monotonicity of the objective function, the KKT conditions, as well as, the constraints in (29), it is obvious that the global optimal solution point of (29) is at the hyperplane

$$\sum_{k=1}^K \rho_{1k} = P, \rho_{\min} \leq \rho_{1k} \leq \rho_{\max}, \forall k = 1, \dots, K. \quad (47)$$

Therefore, considering that the objective function in (29) now has three variables, we use the method of coordinate transform to transform (47) into the two-dimensional plane when  $K = 3$ . Hence, the form of contour lines can be used to show the objective function values in (29).

Fig. 2 gives the objective function values in (29) in the form of contour lines. Also, the solution points of the PPA algorithm and fmincon method are marked by a red circle and a black asterisk, respectively. This graph shows that the geometrical centers of these two mark symbols overlap almost completely, which implies that the solution obtained from the PPA algorithm is almost the same as the one obtained from the numerically evaluated method.

We now turn our attention to the issue of computational complexity by computing the average complexity of both techniques. The average runtime of each algorithm is obtained by averaging over 1000 independent large-scale coefficient realizations. The average runtime values for both PPA algorithm and fmincon method, as well as, different number of users in cell 1, are presented in Table Ib. We can see that, when the number of users  $K$  gets larger, the average runtime increases for all cases under consideration. Most importantly, compared with the fmincon method, the PPA algorithm’s average runtime is significantly reduced, (e.g., the average runtime of the proposed PPA algorithm decreases about 600 times for  $K = 2$  and 70 times for  $K = 10$ ). Hence, our PPA algorithm offers the important advantage of low computational complexity.

### B. Channel Estimation Performance Comparison

In this subsection, we compare the channel estimation performance of the PPA scheme and the EPPA scheme. In this cellular system, we consider  $K = 10$  users in each cell and set  $\mu = 3$ . The other parameters are the same as in Section IV-A.

Moreover, different frequency reuse factors are also considered as in [1].

1) *One-Cell Performance:* Here, for convenience, the objective cell is again the cell 1. For each analytical result, 100 independent large-scale coefficient realizations are generated. The results of the PPA scheme and the EPPA scheme are obtained by averaging over 100 independent small-scale fading channels for each large-scale channel realization.

Fig. 3 gives the analytical and Monte-Carlo simulated average  $\text{Exp}_{\text{rcee}}$  per user with the LS and MMSE estimation methods, respectively. Results are shown for different frequency reuse factors, and  $P = 40$  dB. Again, we see that in all cases the analytical curves (obtained from (23)) match precisely with the simulated curves (obtained from (22)), which proves the validity of *Theorem 2*. When  $M \rightarrow \infty$ , we also see that all results tend to different constants, which match the asymptotic expressions, derived in *Corollary 3* and *Corollary 5*, respectively. Moreover, it is found that the PPA scheme performs systematically better than the EPPA scheme, which justifies the effectiveness of our PPA scheme. Also, from Fig. 3a, as the frequency reuse factor increases, the gap between the EPPA scheme and the PPA scheme is almost fixed for LS. From Fig. 3b, when the frequency reuse factor increases, the gap between the EPPA scheme and the PPA scheme is gradually increased for MMSE estimation. When  $M \rightarrow \infty$ , compared with the LS EPPA scheme, the LS PPA scheme reduces the average  $\text{Exp}_{\text{rcee}}$  per user 50.4%, 47.8%, and 45.3% for  $\Gamma = 1, 3$ , and 7, respectively, while the MMSE case remains nearly similar.

Fig. 4a presents the cumulative distribution function (CDF) of the analytical average  $\text{Exp}_{\text{rcee}}$  per user with  $P = 40$  dB and frequency reuse factors  $\Gamma = 1, 3$ , and 7, for the LS and MMSE estimation methods, respectively, under the setting of asymptotically large antenna number. This figure shows that using the PPA algorithm to allocate pilot power improves the average  $\text{Exp}_{\text{rcee}}$  per user in the whole probability distribution range compared with the EPPA scheme for all cases. We also see that by increasing the frequency reuse factor, the channel estimation performances of the LS and MMSE methods get closer for both the EPPA and PPA schemes, respectively. To be more specific, when  $\Gamma = 1$ , the co-channel interference is still stronger than the channel power between the user in cell 1 and the BS in cell 1, so the average  $\text{Exp}_{\text{rcee}}$  per user of the

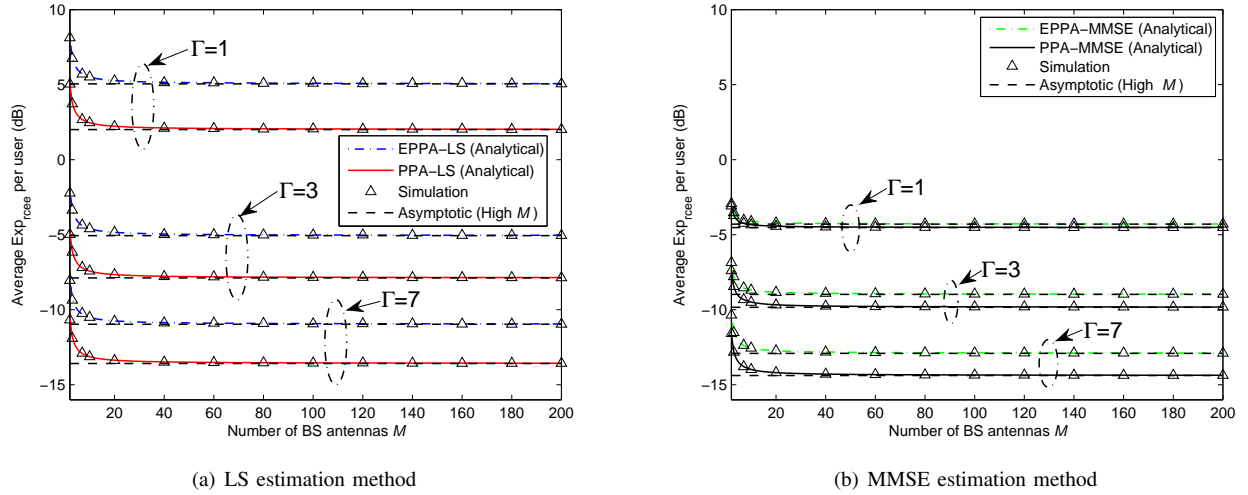


Fig. 3: The average  $\text{Exp}_{\text{rcee}}$  per user with asymptotic expressions (25) for PPA scheme and (30) for the EPPA scheme as a function of the number of BS antennas with  $P = 40$  dB.

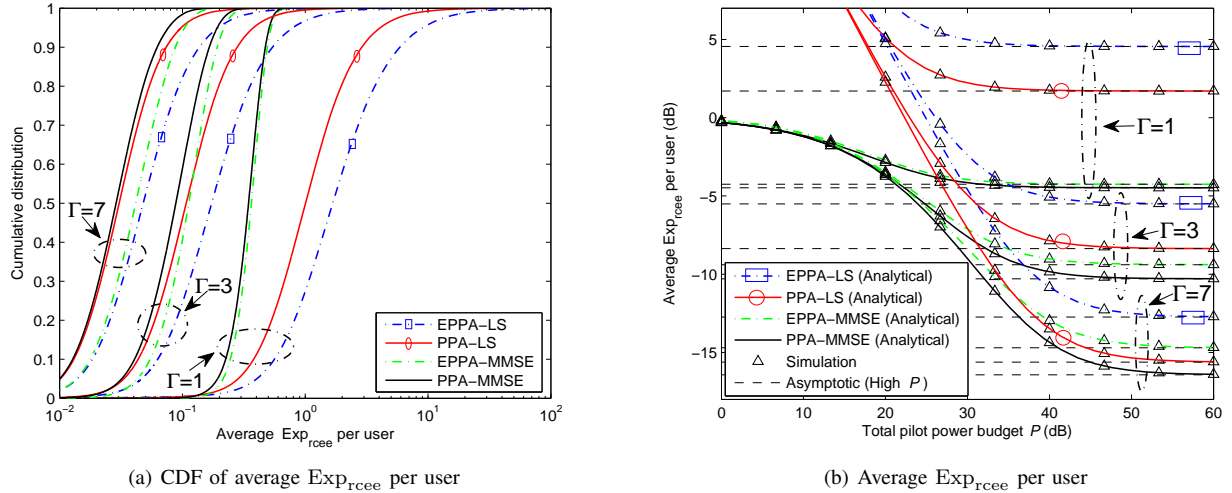


Fig. 4: The CDF of the average  $\text{Exp}_{\text{rcee}}$  per user for asymptotically large antenna numbers with  $P = 40$  dB, as well as, the average  $\text{Exp}_{\text{rcee}}$  per user as  $P$  increases with  $M = 200$ . The asymptotic expressions are based on (39-40) for the PPA scheme and (45) for the EPPA scheme.

LS estimation method is much worse than the MMSE, since the former treats co-channel interference as noise and fails to deliver precise estimates. On the contrary, when  $\Gamma = 7$ , the co-channel interference becomes small compared with the thermal noise at the receiver, such that the performance of the LS PPA scheme approaches the performance of the MMSE PPA scheme.

Fig. 4b investigates the impact of the total pilot power budget  $P$  on the average  $\text{Exp}_{\text{rcee}}$  per user performance. In this figure, the number of BS antennas is set to 200. It shows that the analytical values and simulation values are almost indistinguishable for both EPPA and PPA schemes, regardless of the value of  $P$ . When  $P \rightarrow \infty$ , the average  $\text{Exp}_{\text{rcee}}$  per user approaches to different constant values, which match the asymptotic expressions (obtained by calculating the limit of (23) when  $P \rightarrow \infty$  and combining with *Theorem 4*

and *Corollary 6*) well, respectively. This result showcases the beneficial impact of larger frequency reuse factor on the channel estimation performance, since the average  $\text{Exp}_{\text{rcee}}$  per user is lower for larger  $\Gamma$ . Moreover, as the frequency reuse factor and  $P$  increase, the channel estimation performance of LS and MMSE estimation methods are close to each other for both the EPPA and PPA schemes. In other words, by extending the distance between cells which use the same frequency benefits the channel estimation performance, and the simple estimation method (LS) approaches the more sophisticated estimation method (MMSE) when  $P$  is big enough.

2) *Multi-cell Performance With Scheduling Strategy*: Now, we consider to allocate all the users' pilot power in multi-cell systems to minimize the average  $\text{Exp}_{\text{rcee}}$  per cell per user. Based on the user grouping concept of the proposed PPA scheme and inspired by the cell grouping concept in

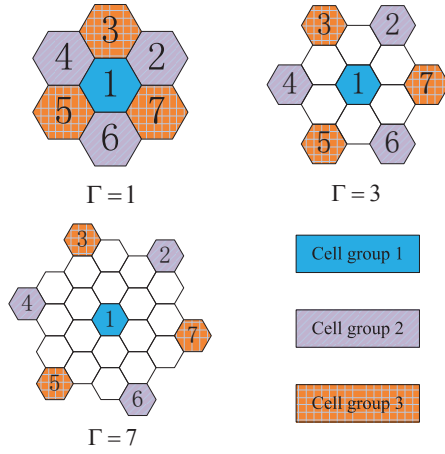


Fig. 5: The JUCG scheduling strategy for  $\Gamma = 1, 3$ , and  $7$ , respectively.

TABLE II: List of the average  $\text{Exp}_{\text{rce}}$  per cell per user

$\Gamma$		1	3	7
Method	EPPA	2.4757dB	-7.6347dB	-12.8171dB
	JUCG	-0.1959dB	-10.1606dB	-15.1958dB
	fmincon	-1.0478dB	-10.5540dB	-15.2979dB
MMSE	EPPA	-5.4686dB	-10.5643dB	-14.1414dB
	JUCG	-5.7410dB	-11.5722dB	-15.6628dB
	fmincon	-5.8442dB	-11.7061dB	-15.6986dB

[35], we propose and adopt the JUCG scheduling strategy in Fig. 5. First, we assume that all users in the whole system utilize the EPPA scheme, that is,  $\rho_{jk} = P/K, \forall j, k$ . Next, the users in cell group 1 (Cell 1) adjust their pilot power simultaneously based on the PPA scheme, while the users in cell group 2 (Cell 2, 4, and 6) and cell group 3 (Cell 3, 5, and 7) keep their pilot power fixed. The scheduling strategy stops when the similar action is taken for the cell group 2 and cell group 3, respectively. As a benchmark, we compare with the optimal performance provided by the fmincon method across all cells, and with the EPPA scheme based on  $M = 200$  and the analytical expressions (23) for both the LS and MMSE estimation methods and different  $\Gamma$  in Table II. The results are obtained by averaging over 100 independent large-scale coefficient realizations.

From Table II, we can see that, for all cases, the performance of the average  $\text{Exp}_{\text{rce}}$  per cell per user increases when  $\Gamma$  increases. For both the LS and MMSE estimation methods, not only the performance of the JUCG scheduling strategy is better than the EPPA scheme, but also the gap between the JUCG and fmincon method is negligible for all  $\Gamma$  except to the little gap which exists for the LS case when  $\Gamma = 1$ . However the high complexity of the fmincon method makes its implementation a big challenge. Therefore, the JUCG is a very attractive solution. In other words, our PPA scheme is still very useful for practical applications and yields nearly optimal performance.

### C. Achievable Uplink Rate Comparison

In this subsection, we compare the achievable uplink rate performance of the PPA scheme and the EPPA scheme. The

relative parameters are the same as subsection B except for  $\rho_u = 20$  dB and  $P = 40$  dB. Also, by considering the methodology of [1], we set  $T_o = 71.4$  ms,  $T_u = 66.7$  ms,  $B = 20$  MHz, and  $(T_s - T_p)/T_s = 3/7$ . For comparison, we define two metrics called “Minimum achievable uplink rate” and “Average achievable uplink rate” in target cell 1, which are given as

$$R_{\min} = \min\{R_{11}, \dots, R_{1k}, \dots, R_{1K}\}, \quad (48)$$

and

$$R_{\text{av}} = \frac{1}{K} \sum_{k=1}^K R_{1k}, \quad (49)$$

respectively.

Fig. 6a gives the minimum achievable uplink rate based on (17), (19), and (48) for both the EPPA and PPA schemes, as well as, the fmincon method.<sup>8</sup> For all cases, when  $M$  increases,  $R_{\min}$  increases and the PPA scheme performance is almost the same as the fmincon performance. To be more specific, when  $\Gamma = 1$ , the PPA scheme performance is better than the EPPA scheme. Moreover, the LS PPA scheme is better than the MMSE PPA scheme since (19) is an increasing function of the target user’s pilot power and the PPA algorithm allocates more power to the user whose relative channel estimation performance is bad; yet, note that for a given user’s pilot power in the target cell, the LS estimation performance is worse than the MMSE. Compared with  $\Gamma = 1$ , although  $\Gamma = 3$  makes the available bandwidth  $B/3$ , the effective SINR grows larger since the interference from other cells’ users drops drastically. Hence,  $\Gamma = 3$  is better than  $\Gamma = 1$ . Also, compared with the EPPA scheme, the PPA scheme performance improves considerably. Note that the performance of the MMSE PPA scheme approaches that of the LS PPA scheme. When  $\Gamma = 7$ , compared with  $\Gamma = 3$ , the available bandwidth becomes substantially smaller, while the effective SINR grows marginally. Hence, regarding  $R_{\min}$ ,  $\Gamma = 7$  performs worse than  $\Gamma = 3$ . Therefore,  $\Gamma = 3$  yields the best performance and we can infer that the channel estimation performance is not proportional to  $\Gamma$ , due to the trade off between the available bandwidth and the effective SINR. Note also that the MMSE performance is almost the same with that of the LS. Moreover, compared with the EPPA scheme, for both the LS and MMSE estimation methods, our PPA scheme improves substantially the performance of the user with the minimum achievable uplink rate.

Fig. 6b gives the CDF of the average achievable uplink rate with asymptotically large antenna numbers based on (17), (20), and (49) for both the EPPA and PPA schemes, as well as, the fmincon method. For all cases, the performance difference between the PPA scheme and the fmincon is almost negligible. For the case of the LS estimation method, when  $\Gamma = 1$ , the LS PPA scheme performance is slightly worse than that of the EPPA scheme since the proposed PPA algorithm puts more pilot power to the user who suffers strong co-channel interference;

<sup>8</sup>Based on Theorem 1, for convenience, we only need to use the case EPPA scheme to replace the cases of the LS EPPA and MMSE EPPA schemes in Fig. 6.

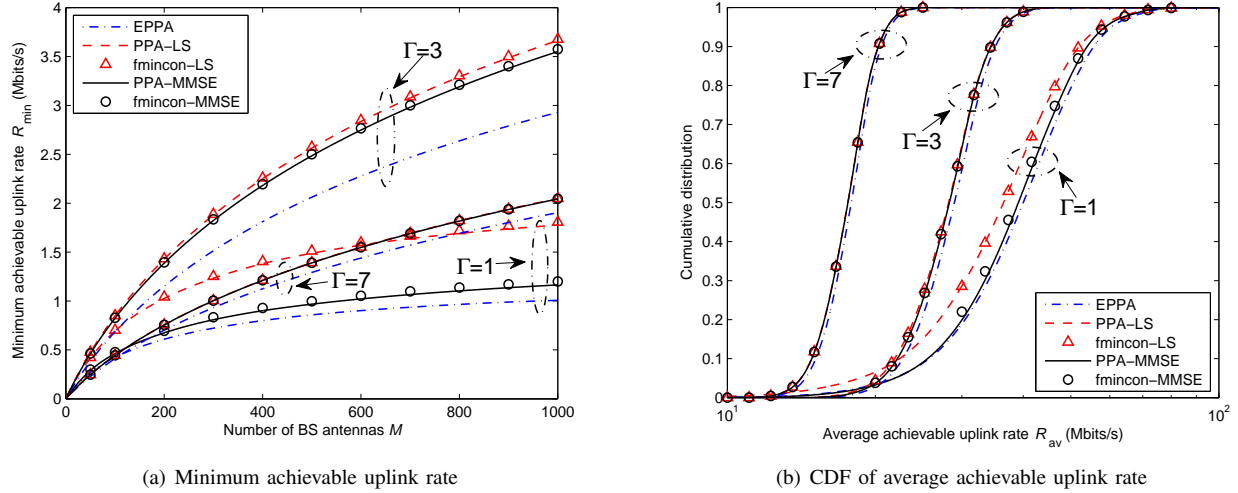


Fig. 6: The minimum achievable uplink rate as a function of the number of BS antennas with  $P = 40$  dB and  $\rho_u = 20$  dB, as well as, the CDF of the average achievable uplink rate with asymptotically large antenna numbers and  $P = 40$  dB.

this affects negatively the rate performance of the users who suffer small co-channel interference. As  $\Gamma = 3$  and  $\Gamma = 7$ , the difference of co-channel interference level between the users in the cell becomes small; hence, the difference between the PPA scheme and the EPPA scheme can be ignored. For the case of the MMSE estimation method, for all frequency reuse factors, the PPA scheme is almost the same as the EPPA scheme. Moreover, with an increasing frequency reuse factor, the performance of the LS approaches the performance of the MMSE. In summary, compared with the EPPA scheme, for both the LS and MMSE estimation methods, our PPA scheme offers almost the same average achievable uplink rate.

## V. CONCLUSION

In this paper, we studied the performance of channel estimation and achievable uplink rate for multi-cell massive MIMO systems. We provided a modified NMSE metric derived from the MSE metric called RCEE and deduced new, tractable, closed-form expressions for  $\text{Exp}_{\text{rcee}}$  and achievable uplink rate for the LS and MMSE estimation methods, respectively. We found that RCEE and  $\text{Exp}_{\text{rcee}}$  tend to the same constant when  $M \rightarrow \infty$ , which reflects the channel hardening effect. Due to the obtained closed-form expressions for  $\text{Exp}_{\text{rcee}}$ , we proposed a PPA algorithm to minimize the average  $\text{Exp}_{\text{rcee}}$  per user with a total pilot power budget. The channel estimation performance of the PPA scheme and EPPA scheme were compared for both the LS and MMSE estimation methods. Regarding the differences between the PPA scheme and EPPA scheme, the LS gap remains almost fixed irrespective of the frequency reuse factor, whilst the MMSE gap increases. Moreover, when co-channel interference is small and the pilot power is high enough, the LS scheme channel estimation performance approaches the MMSE scheme performance. Hence, from a design point of view, the simple LS estimation method is a very viable choice. Also, to allocate pilot power in the whole systems, a JUCG scheduling strategy was proposed, which shows it can obtain almost the optimal performance. Finally,

the average achievable uplink rate of the PPA scheme is almost the same as the EPPA scheme, while, compared with the EPPA scheme, the PPA scheme improves significantly the minimum achievable uplink rate.

## APPENDIX A PROOF OF THEOREM 1

To evaluate the  $\text{SINR}_{jk}$  in (18), we define four terms

$$\mathfrak{A} \triangleq \left| \mathbb{E} \left\{ \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jjk} \right\} \right|^2, \quad (50)$$

$$\mathfrak{B} \triangleq \mathbb{E} \left\{ \left| \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jln} \right|^2 \right\}, \quad (n \neq k), \quad (51)$$

$$\mathfrak{C} \triangleq \mathbb{E} \left\{ \left| \hat{\mathbf{h}}_{jjk}^H \mathbf{h}_{jlk} \right|^2 \right\}, \quad (52)$$

$$\mathfrak{D} \triangleq \mathbb{E} \left\{ \left\| \hat{\mathbf{h}}_{jjk} \right\|_2^2 \right\}. \quad (53)$$

Although  $\hat{\mathbf{h}}_{jjk}$  has different expressions for the LS and MMSE estimation methods, the corresponding proofs for  $\text{SINR}_{jk}$  are similar. Hence, it is convenient to only study the case of the LS estimation.

- Calculate  $\mathfrak{A}$ : Substituting (3) and (9) into (50), we get

$$\mathfrak{A} = \left| \mathbb{E} \left\{ \left\| \mathbf{h}_{jjk} \right\|_2^2 \right\} \right|^2 = M^2 \beta_{jjk}^2. \quad (54)$$

- Calculate  $\mathfrak{B}$ : By considering the  $\hat{\mathbf{h}}_{jjk}^{\text{LS}}$  is uncorrelated with  $\mathbf{h}_{jln}$  when  $n \neq k$ , we can obtain

$$\mathfrak{B} = \mathbb{E} \left\{ (\hat{\mathbf{h}}_{jjk}^{\text{LS}})^H \mathbb{E} \left\{ \mathbf{h}_{jln} \mathbf{h}_{jln}^H \right\} \hat{\mathbf{h}}_{jjk}^{\text{LS}} \right\}. \quad (55)$$

Then, substituting (3) and (9) into (55), yields

$$\mathfrak{B} = \frac{M \beta_{jln}}{\rho_{jk}} \left( \sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1 \right). \quad (56)$$

- Calculate  $\mathfrak{C}$ : Since  $\hat{\mathbf{h}}_{jjk}^{\text{LS}}$  is correlated with  $\mathbf{h}_{jlk}$ , substituting (9) into (52), we can get

$$\mathfrak{C} = \frac{1}{\rho_{jk}} \left( \underbrace{\mathbb{E} \left\{ \left\| \left( \sum_{c \neq l}^L \sqrt{\rho_{ck}} \mathbf{h}_{jck} + \mathbf{N}_j \mathbf{s}_{jk} \right)^H \mathbf{h}_{jlk} \right\|^2 \right\}}_{C_1} + \underbrace{\rho_{lk} \mathbb{E} \{ \|\mathbf{h}_{jlk}\|_2^4 \}}_{C_2} \right), \quad (57)$$

where the closed-form expression for  $C_1$  can be obtained based on the similar way for obtaining (51), that is,

$$C_1 = M\beta_{jlk} \left( \sum_{c \neq l}^L \rho_{ck} \beta_{jck} + 1 \right), \quad (58)$$

whilst for the term  $C_2$ , using the properties of Wishart matrices [36, Lemma 2.9],  $C_2$  is given by

$$C_2 = M(M+1)\beta_{jlk}^2. \quad (59)$$

Therefore, substituting (58) and (59) into (57) and simplifying, we can write  $\mathfrak{C}$  as follows

$$\mathfrak{C} = \frac{1}{\rho_{jk}} \left( M\beta_{jlk} \left( \sum_{c=1}^L \rho_{ck} \beta_{jck} + 1 \right) + M^2 \rho_{lk} \beta_{jlk} \right). \quad (60)$$

- Calculate  $\mathfrak{D}$ : With the help of (3) and (9),  $\mathfrak{D}$  is given by

$$\mathfrak{D} = \frac{M}{\rho_{jk}} \left( \sum_{l=1}^L \rho_{lk} \beta_{jlk} + 1 \right). \quad (61)$$

Finally, substituting (54), (56), (60), and (61) into (18) and simplifying, the closed-form expression for  $\text{SINR}_{jk}$  based on LS estimation method is obtained. ■

#### APPENDIX B PROOF OF THEOREM 2

Although  $\Delta_{jk}$  has different closed-form expressions for LS and MMSE estimation methods, the corresponding proofs are similar. Hence, it is convenient to only study the case of the LS estimation. Then, (22) is rewritten as

$$\Delta_{jk}^{\text{LS}} = \mathbb{E} \left\{ \underbrace{\|\mathbf{h}_{jjk} - \hat{\mathbf{h}}_{jjk}^{\text{LS}}\|_2^2}_{I_1} \times \underbrace{\frac{1}{\|\mathbf{h}_{jjk}\|_2^2}}_{I_2} \right\}. \quad (62)$$

Then, substituting (9) into  $I_1$ , we obtain

$$I_1 = \left\| \sum_{l \neq j}^L \frac{\sqrt{\rho_{lk}}}{\sqrt{\rho_{jk}}} \mathbf{h}_{jlk} + \frac{1}{\sqrt{\rho_{jk}}} \mathbf{N}_j \mathbf{s}_{jk} \right\|_2^2. \quad (63)$$

It is obvious that  $I_1$  does not have the term  $\mathbf{h}_{jjk}$ , which means that  $\Delta_{jk}^{\text{LS}} = \mathbb{E} \{I_1\} \times \mathbb{E} \{I_2\}$ . Moreover, based on the properties of the expectation operator and the trace operator and the assumption of the user's channel, we evaluate (63) to yield

$$\mathbb{E} \{I_1\} = \text{tr} \left( \mathbb{E} \left\{ \sum_{l \neq j}^L \frac{\rho_{lk}}{\rho_{jk}} \mathbf{h}_{jlk} \mathbf{h}_{jlk}^H + \frac{1}{\rho_{jk}} \mathbf{N}_j \mathbf{s}_{jk} (\mathbf{N}_j \mathbf{s}_{jk})^H \right\} \right). \quad (64)$$

Then, substituting (3) into (64), we have the following result

$$\mathbb{E} \{I_1\} = \frac{M}{\rho_{jk}} \left( \sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1 \right). \quad (65)$$

To evaluate  $I_2$  in (62), we firstly consider the case of  $M = 1$ . Hence, we denote  $I_2$  as  $I_2^{M=1}$ . Moreover,  $h_{jjk}$  is now a scalar stochastic quantity  $\sim \mathcal{CN}(0, \beta_{jlk})$ . Based on the special structure of  $I_2^{M=1}$ , after some manipulations, it is easy to obtain

$$\mathbb{E} \{I_2^{M=1}\} \rightarrow \infty. \quad (66)$$

When  $M \geq 2$ , we denote  $I_2$  as  $I_2^{M \geq 2}$ . Using the properties of Wishart matrices [36, Lemma 2.10], thus,  $\mathbb{E} \{I_2^{M \geq 2}\}$  is reduced to

$$\mathbb{E} \{I_2^{M \geq 2}\} = \frac{1}{\beta_{jlk}(M-1)}. \quad (67)$$

Hence, by substituting (65)-(67) into (62) and simplifying, the closed-form expression for  $\Delta_{jk}$  based on LS estimation method is obtained. ■

#### APPENDIX C PROOF OF THEOREM 4

In this proof, the main challenge is to obtain the user groups  $\mathcal{K}_{P,\min}$ ,  $\mathcal{K}_{P,\max}$ , and  $\mathcal{K}_P$  in **Algorithm 1** when  $P \rightarrow \infty$ . However,  $P \rightarrow \infty$  means that the thermal noise in the BS can be ignored. Hence, inspired by (31), we formulate a new optimization problem by replacing the  $\Delta_{jk}$  in (29) with  $\hat{\Delta}_{jk}$  and its subproblem by replacing the  $\tilde{\Delta}_{jk}$  in (31) with  $\hat{\Delta}_{jk}$ , where

$$\hat{\Delta}_{jk} \triangleq \begin{cases} \frac{M \sum_{l \neq j}^L \hat{\rho}_{lk} \beta_{jlk}}{(M-1) \hat{\rho}_{jk} \beta_{jlk}}, & \text{LS,} \\ \frac{M \sum_{l \neq j}^L \hat{\rho}_{lk} \beta_{jlk}}{(M-1) \sum_{l=1}^L \hat{\rho}_{lk} \beta_{jlk}}, & \text{MMSE.} \end{cases} \quad (68)$$

Using the Lagrange multiplier method [33], we can determine the optimal solution  $\hat{\rho}_{jk}^*$  of this new optimization problem, which is similar with the format of (33)-(36) that we omit it. Moreover, we find that  $\hat{\rho}_{jk}^*$  is proportional to  $P$  for the LS and MMSE estimation methods, respectively. Therefore, since the PPA algorithm solves (29) with the help of (31), we also use the core idea of PPA algorithm to solve this new optimization problem with the help of its subproblem. It shows that the output user groups  $\hat{\mathcal{K}}_{P,\min}$ ,  $\hat{\mathcal{K}}_{P,\max}$ , and  $\hat{\mathcal{K}}_P$ ,<sup>9</sup> whose cardinalities are  $\hat{K}_{P,\min}$ ,  $\hat{K}_{P,\max}$ , and  $\hat{K}_P$ , respectively, will not depend on  $P$ . That is,  $\forall P_1, P_2 > 0$ , we have  $\hat{\mathcal{K}}_{P_1,\min} = \hat{\mathcal{K}}_{P_2,\min}$ ,  $\hat{\mathcal{K}}_{P_1,\max} = \hat{\mathcal{K}}_{P_2,\max}$ , and  $\hat{\mathcal{K}}_{P_1} = \hat{\mathcal{K}}_{P_2}$ . Hence, we use  $\hat{\mathcal{K}}_{\min}$ ,  $\hat{\mathcal{K}}_{\max}$ , and  $\hat{\mathcal{K}}$  to replace  $\hat{\mathcal{K}}_{P,\min}$ ,  $\hat{\mathcal{K}}_{P,\max}$ , and  $\hat{\mathcal{K}}_P$ , respectively. Also, the  $\hat{K}_{\min}$ ,  $\hat{K}_{\max}$ , and  $\hat{K}$  are obtained for the same reason. Then, we use  $\hat{\mathcal{K}}_{\min}$ ,  $\hat{\mathcal{K}}_{\max}$ , and  $\hat{\mathcal{K}}$  to approximate the user groups  $\mathcal{K}_{P,\min}$ ,  $\mathcal{K}_{P,\max}$ , and  $\mathcal{K}_P$ , respectively, when  $P \rightarrow \infty$ . Later, considering *Corollary 3*, we hold  $M \rightarrow \infty$

<sup>9</sup>To avoid confusion, we redefine these three user groups when using the core idea of PPA algorithm to solve this new optimization problem.



and study the channel estimation performance when  $P \rightarrow \infty$ . Moreover, it is obvious that the first and second expressions in both (39) and (40) are easy to obtain with the help of the limit theory. Therefore, this proof focuses on the third expressions in both (39) and (40).

Although, when  $M \rightarrow \infty$  and  $P \rightarrow \infty$ , the limit value of  $\Delta_{jk}$  has different closed-form expressions for LS and MMSE estimation methods, the corresponding proofs are similar. Hence, it is convenient to only study the case of the LS estimation. We now substitute the first line of (33) into the first line of (25) and considering that the user groups  $\mathcal{K}_P$ ,  $\mathcal{K}_{P,\min}$ , and  $\mathcal{K}_{P,\max}$  are obtained by the PPA algorithm; hence, for the users in groups  $\mathcal{K}_P$ , we replace  $\sum_{k=1}^K$  with  $\sum_{k \in \mathcal{K}_P}$  and  $P$  with  $P - \rho_{\min} K_{P,\min} - \rho_{\max} K_{P,\max}$  and set  $\alpha = \rho_{\min} K/P$ , as well as, we consider the definition of  $\mu$ , to obtain

$$\frac{\sum_{l \neq j}^L \rho_{lk} \beta_{jlk} + 1}{\rho_{jk} \beta_{jjk}} = \frac{\sum_{l \neq j}^L \delta_{lk} P \beta_{jlk} + 1}{P - K_{P,\min} \frac{\alpha P}{K} - K_{P,\max} \frac{\mu P}{K}} \times \sum_{k \in \mathcal{K}_P} \left( \frac{1}{\beta_{jjk}} \left( \sum_{l \neq j}^L \delta_{lk} P \beta_{jlk} + 1 \right) \right)^{\frac{1}{2}} \times \left( \beta_{jjk} \left( \sum_{l \neq j}^L \delta_{lk} P \beta_{jlk} + 1 \right) \right)^{-\frac{1}{2}}. \quad (69)$$

Finally, using the limit theory when  $P \rightarrow \infty$ , and simplifying, we obtain the third expression in (39). ■

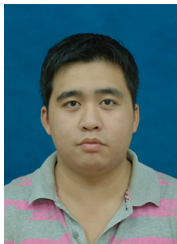
#### ACKNOWLEDGMENT

The authors would like to thank Dr. Emil Björnson, from Linköping University, for his generous help in improving this paper. The authors would also like to thank the associate editor and the anonymous reviewers for their constructive suggestions and comments during the review process which had helped to improve the quality of the paper.

#### REFERENCES

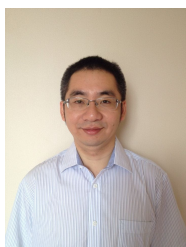
- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.
- [2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.
- [3] Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966-981, Oct. 2014.
- [4] J. Boccardi, R. W. Heath Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74-80, Feb. 2014.
- [5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436-1449, Apr. 2013.
- [6] 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.
- [7] 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.
- [8] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112-7139, Nov. 2014.
- [9] 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.
- [10] 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.
- [11] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Tech. J.*, vol. 20, pp. 11-22, 2015.
- [12] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Proc. IEEE ICASSP*, Mar. 2012, pp. 3249-3252.
- [13] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773-786, Oct. 2014.
- [14] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3111-3124, Jun. 2014.
- [15] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356-1368, Mar. 2015.
- [16] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 815-830, Oct. 2014.
- [17] D. Kong, D. Qu, K. Luo, and T. Jiang, "Channel estimation under staggered frame structure for massive MIMO system," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1469-1479, Feb. 2016.
- [18] D. Hu, L. He, and X. Wang, "Semi-blind pilot decontamination for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 525-536, Jan. 2016.
- [19] X. Zhu, Z. Wang, C. Qian, L. Dai, J. Chen, S. Chen, and L. Hanzo, "Soft pilot reuse and multi-cell block diagonalization precoding for massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3285-3298, May 2016.
- [20] H. Q. Ngo, M. Matthaiou, and E. G. Larsson, "Massive MIMO with optimal power and training duration allocation," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 605-608, Dec. 2014.
- [21] K. Guo, Y. Guo, G. Fodor, and G. Ascheid, "Uplink power control with MMSE receiver in multi-cell MU-massive-MIMO systems," in *Proc. IEEE ICC*, Jun. 2014, pp. 5184-5190.
- [22] Z. Xiang, M. Tao, and X. Wang, "Massive MIMO multicasting in noncooperative cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1180-1193, Jun. 2014.
- [23] X. Zheng, H. Zhang, W. Xu, and X. You, "Optimized pilot power allocation for heterogeneous users in massive MIMO downlinks," in *Proc. IEEE ICTC*, Oct. 2014, pp. 246-250.
- [24] V. Saxena, G. Fodor, and E. Karipidis, "Mitigating pilot contamination by pilot reuse and power control schemes for massive MIMO systems," in *Proc. IEEE VTC Spring*, May 2015.
- [25] G. Fodor and M. Telek, "On the pilot-data power trade off in single input multiple output systems," in *Proc. EW*, May 2014, pp. 485-492.
- [26] H. V. Cheng, E. Björnson, and E. G. Larsson, "Uplink pilot and data power control for single cell massive MIMO systems with MRC," in *Proc. IEEE ISWCS*, Aug. 2015, pp. 396-400.
- [27] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1993.
- [28] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1293-1308, Feb. 2016.
- [29] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353-4368, Aug. 2015.
- [30] D. Neumann, M. Joham, and W. Utschick, "Channel estimation in massive MIMO systems," Mar. 2015. [Online]. Available: <http://arxiv.org/abs/1503.08691>
- [31] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350-2361, Jun. 2013.

- [32] T. L. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 847-860, Oct. 2014.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [34] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721-1737, Sept. 2014.
- [35] Q. Zhang, S. Jin, M. R. McKay, D. Morales-Jimenez, and H. Zhu, "Power allocation schemes for multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 5941-5955, Nov. 2015.
- [36] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1-182, Jun. 2004.



**Pei Liu** (S'17) received the B.S. degree in Telecommunication Engineering from Huazhong University of Science and Technology, China, in 2014. He is currently working toward the Ph.D. degree in Information and Telecommunication Engineering at Wuhan National Laboratory for Optoelectronics, School of Electrical Information and Communications, Huazhong University of Science and Technology, and the Ph.D. degree in Electronic Engineering at the Department of Electronic Engineering, City University of Hong Kong. His research interests include Massive MIMO Systems, Large-Scale Distributed Antennas Systems, Random Matrix Theory, and Convex Optimization Theory.

Mr. Liu served as the Technical Program Committee Member of 10th International Conference on Communications and Networking in China (CHINACOM 2015) and serves as the reviewer for the IEEE TRANSACTIONS ON COMMUNICATIONS. He is the Student Member of IEEE, The Optical Society (OSA), and Chinese Institute of Electronics (CIE).



**Shi Jin** (S'06-M'07) received the B.S. degree in communications engineering from Guilin University of Electronic Technology, Guilin, China, in 1996, the M.S. degree from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003, and the Ph.D. degree in communications and information systems from the Southeast University, Nanjing, in 2007. From June 2007 to October 2009, he was a Research Fellow with the Adastral Park Research Campus, University College London, London, U.K. He is currently with the faculty of the National

Mobile Communications Research Laboratory, Southeast University. His research interests include space time wireless communications, random matrix theory, and information theory. He serves as an Associate Editor for the IEEE Transactions on Wireless Communications, and IEEE Communications Letters, and IET Communications. Dr. Jin and his co-authors have been awarded the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory and a 2010 Young Author Best Paper Award by the IEEE Signal Processing Society.



**Tao Jiang** (M'06-SM'10) is currently a Distinguished Professor with the School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan, P. R. China. He received the B.S. and M.S. degrees in applied geophysics from China University of Geosciences, Wuhan, P. R. China, in 1997 and 2000, respectively, and the Ph.D. degree in information and communication engineering from Huazhong University of Science and Technology, Wuhan, P. R. China, in April 2004. From Aug. 2004 to Dec. 2007, he

worked in some universities, such as Brunel University and University of Michigan-Dearborn, respectively. He has authored or co-authored over 200 technical papers in major journals and conferences and 8 books/chapters in the areas of communications and networks. He served or is serving as symposium technical program committee membership of some major IEEE conferences, including INFOCOM, GLOBECOM, and ICC, etc.. He was invited to serve as TPC Symposium Chair for the IEEE GLOBECOM 2013, IEEE WCNC 2013 and ICC 2013. He has served or serving as associate editor of some technical journals in communications, including in IEEE Transactions on Signal Processing, IEEE Communications Surveys and Tutorials, IEEE Transactions on Vehicular Technology, and IEEE Internet of Things Journal, etc.. He is a recipient of the NSFC for Distinguished Young Scholars Award in 2013, and he is also a recipient of the Young and Middle-Aged Leading Scientists, Engineers and Innovators by the Ministry of Science and Technology of China in 2014. He was awarded as the Most Cited Chinese Researchers announced by Elsevier in 2014 and 2015. He is a Senior Member of IEEE.



**Qi Zhang** received the B.S. and Ph.D. degree in Communication & Information Engineering from Nanjing University of Posts & Telecommunications, Nanjing, China, in 2010 and 2015, respectively. She is currently working as a Postdoc Research Fellow at the Singapore University of Technology & Design, Singapore. Her research interests include massive MIMO systems, stochastic geometry and space-time wireless communications.



**Michail Matthaiou** (S'05-M'08-SM'13) was born in Thessaloniki, Greece in 1981. He obtained the Diploma degree (5 years) in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece in 2004. He then received the M.Sc. (with distinction) in Communication Systems and Signal Processing from the University of Bristol, U.K. and Ph.D. degrees from the University of Edinburgh, U.K. in 2005 and 2008, respectively. From September 2008 through May 2010, he was with the Institute for Circuit Theory and Signal

Processing, Munich University of Technology (TUM), Germany working as a Postdoctoral Research Associate. He is currently a Senior Lecturer at Queen's University Belfast, U.K. after holding an Assistant Professor position at Chalmers University of Technology, Sweden. His research interests span signal processing for wireless communications, massive MIMO, hardware-constrained communications, and performance analysis of fading channels.

Dr. Matthaiou was the recipient of the 2011 IEEE ComSoc Best Young Researcher Award for the Europe, Middle East and Africa Region and a co-recipient of the 2006 IEEE Communications Chapter Project Prize for the best M.Sc. dissertation in the area of communications. He was co-recipient of the Best Paper Award at the 2014 IEEE International Conference on Communications (ICC) and was an Exemplary Reviewer for IEEE COMMUNICATIONS LETTERS for 2010. In 2014, he received the Research Fund for International Young Scientists from the National Natural Science Foundation of China. He currently serves as Senior Editor for IEEE COMMUNICATIONS LETTERS, an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, and was the Lead Guest Editor of the special issue on "Large-scale multiple antenna wireless systems" of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is the chair of the Wireless Communications Symposium (WCS) at IEEE GLOBECOM 2016. He was an associate member of the IEEE Signal Processing Society SPCOM and SAM technical committees.