IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 19, 2017, accepted January 18, 2018, date of publication February 5, 2018, date of current version April 4, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2797874

Optimal Power Control and Load Balancing for Uplink Cell-Free Multi-User Massive MIMO

TIEN HOA NGUYEN, THE KHAI NGUYEN, HUY DUNG HAN, AND VAN DUC NGUYEN
School of Electronics and Telecommunications, Hanoi University of Science and Technology, Hanoi 100000, Vietnam

Corresponding author: Tien Hoa Nguyen (hoa.nguyentien@hust.edu.vn)

This work was supported by the National Foundation for Science and Technology Development, Vietnam, under Grant DFG.102.2014.02.

ABSTRACT In this paper, we consider the problem of power control and load balancing in the uplink of the cell-free (CF) multi-user (MU) massive multiple-input multiple-output system. The power control problem is solved using three different criteria: 1) power minimization; 2) maximize min quality of service (QoS); and 3) maximize sum spectral efficiency (SE) under imperfect channel state information. While power minimization and min-QoS maximization problems can be solved in polynomial time, sum SE maximization is NP-hard. Hence, we apply a successive approach to convert it into a geometric program and achieve near optimum solution. As the number of connections in CF-MU system is large, we minimize the number of base stations (BSs) serving each user while maintaining the optimized power consumption and QoS from the previous stage. We propose an iterative elimination (IE) algorithm to remove ineffective BSs for each user. The system analysis is performed under two common BS receivers: maximum ratio combining and zero forcing. The simulation shows that our method is better than both maximum signal-to-noise ratio association and full-set joint transmission, especially in the high QoS regime. Another notable result is that the selection of the linear receiver becomes more influential than the association method.

INDEX TERMS Massive MIMO, power allocation, user association.

I. INTRODUCTION

Multiple-Input and Multiple-Output (MIMO) has become a key technique of various wireless communication specifications such as Long Term Evolution (LTE), 802 family standards and Digital Video Broadcasting-Terrestrial (DVB-T) [1]. In multi-user MIMO systems, a base station (BS) is equipped with several antennas to serve several mobile users using the same frequency resource. Meanwhile, the number of BS antennas in Massive MIMO could be higher, resulting in the availability of linear processing scheme that makes the complexity at the mobile terminals dropped. The deployment of linear processing scheme can be considered as nearly optimal thanks to massive antennas array effect [2]. However, the multiple users (MU) in the system generate inter-user interference which has to be cancelled by doing resource allocation such as power control and load balancing for the downlink channel and uplink one respectively.

Massive MIMO (an extension of MIMO) is a potential 5G wireless access technology. This technique is expected to offer many advantages in enhancing the spectrum and energy efficiency in the case that the data rate rapidly increases [3]–[5]. A Massive MIMO based network is typically equipped with hundreds of antennas [6]. There have been several existing works focusing on large MIMO systems in various research directions such as information theoretical capacity and resource allocation [7]–[9]. Similar to the most of previous works, the time division duplex (TDD) mode is used while the channel state information (CSI) at the transmitter is obtained with the use of channel reciprocity via an uplink pilot training step [10], [11]. Because of the importance of CSI in Massive MIMO systems, the effect of the perfect and imperfect characteristics of CSI has been widely investigated in [4], [5], and [12].

In this work, we consider a cell-free (CF) multi-users (MU) massive MIMO system where a set of BSs equipped with hundreds antennas serve a group of autonomous users distributed in the network and a backhaul network where signal from all the BSs is centered for baseband processing. Cell-Free implies no cell or cell boundaries in the network and all the BSs can connect to a user at the same time [13]–[16]. The CF-MU Massive MIMO system with single antenna access points (APs), can bring almost 5-fold improvement in spectral efficiency (SE) compared to small-cell schemes [16]. However, it is not necessary for an user to associate with all BSs in the network. Van Chien *et al.* [17] show that in downlink of Massive

2169-3536 © 2018 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

MIMO, although a user can connect to multiple BSs, only one BS serves that user in most of the cases. There has not been any works investigating the optimal number of BSs to serve each user in UL of CF-MU Massive MIMO. In order to investigate the joint power control and load balancing in UL CF-MU Massive MIMO, we develop optimization problems which integrates 2 types of linear receiver: zero forcing (ZF) and maximum ratio combining (MRC) receivers under imperfect CSI schemes.

Despite the fact that low-complexity linear detectors/precoders with the use of MRC or ZF, are implemented to reduce the complexity of signal processing [18]–[20], the joint power optimization and load balancing under multiple constraints is still a computationally complex and combinatorial problem. Therefore, the current works have been focused on only association aiming to enhance the user's performance while the load balance across different tiers of BSs is kept unchanged [21]–[27]. When the association strategies between BS and users are applied, the power control and load balancing conditions significantly impact the system performance. The cell-users association problem has been investigated in [28] with regards to minimizing the maximum per BS load, subjected to the user throughput target. Nevertheless, this load balancing problem has been solved for only a single cell. In [17] and [26], the user cell association problem is investigated for only the downlink channel aiming to minimize the total transmit power under fixed spectral efficiency constraints.

It is important to note that the aforementioned works only considered the downlink channel of a Massive MIMO system. In few studies on the uplink one, the transmitting power of the end-user devices is assumed to be constant. However, this assumption is not optimal due to the inter-users interference [29]. The previous solutions of the power control and cell users association have their drawback of only dealing with the combinatorial problem [30]. Motivated by the aforementioned analysis, we investigate the joint transmission power control and load balancing problem for the uplink channel of the CF-MU massive MIMO. The performance metrics are power consumption and system's throughput under the effect of large-scale fading and interference among users, which is commonly analyzed in a wide range of related works in Massive MIMO [7], [16], [26], [31]. We propose an iterative elimination process (IE) to optimize the number of BSs that serve a mobile user with a reasonable power budget. The main contributions of this paper can be listed as follows:

- We propose a new ergodic spectral efficiency (SE) expression for uplink CF-MU Massive MIMO when using MRC and ZF combining vectors. The signal of an arbitrary user will be treated as the sum signal received at the back-haul network of all streams from that user to all BSs in the system under imperfect CSI condition.
- In order to minimize the total transmission power, we consider a lower bound for ergodic SE requirement at each mobile user with a predicted maximum level of transmit power at the BSs. Furthermore, to fully assess the performance of our model, we consider max min-QoS and max sum-SE problem. To solve the optimization problem with NP-hardness, we apply an efficient method for finding Karush-Kuhn-Tucker (KKT) local maximum points. The novelty of our method is based on solving a set of geometric programs (GPs) which is approximated from the original problem via an successive approach of convex optimization.
- We develop an algorithm to optimize the load balancing while the total transmission power is minimized. With the simulation results, we show how the optimal BSs are selected to serve each user. We also discuss how the association is affected by the large scale fading, channel estimation quality as well as signal to interference and noise ratio (SINR).
- Numerical results and discussion based on Rayleigh fading channel model demonstrate the effectiveness of our solution compared to conventional association methods. We indicate that the power allocation, array gain, and user association are 3 vital factors to reduce the power consumption in the network. In addition, it can be that the type of linear receiver has a more dominant impact to transmit power than compared to association methods.

This paper is organized in 5 sections: Section 2 describes the system model for the CF-MU Massive MIMO uplink system. Section 3 formulates the joint power control and user association problem under the imperfect CSI condition. After that, we show algorithmic solutions for this optimization problem. Numerical results are presented in Section 4. Finally, the conclusions are given in Section 5.

Notations: Vectors are written in boldface using lowercase. We use boldface with capital letters to represent matrices. The transpose and its conjugate are presented by the superscripts T and H, respectively. $\mathbf{I}_K$ is the $K \times K$ identity matrix. The operator $\mathbb{E}\{.\}$ denotes the expectation of a random variable. The notation $||.||$ stands for the Euclidean norm and tr(.) represents the trace of a matrix. $n \in CN(0; C)$ means that $\mathbf{n}$ is the zero-mean complex Gaussian vector with covariance matrix $\mathbf{C}$.

II. SYSTEM MODEL

We consider an uplink CF-MU Massive MIMO with L BS, where each BS is equipped with M antenna elements. In our system, K distributed single-antenna users independently transmit their signal on the same time-frequency slot to BSs. It is assumed that the system works in the TDD mode. During a symbol coherence interval τ_c, the channels are assumed to be constant and frequency-flat. We assume that a user can associate and be served by more than one BS. In such system there is no cell boundary. Hence, a predefinition of cell indices is unnecessary. We can label users from 1 to K. The received signal at the l-th BS can be given as

$$\boldsymbol{y}_l = \sum_{k=1}^{K} \boldsymbol{h}_{l,k} \sqrt{p_k} x_k + \boldsymbol{n}_l, \tag{1}$$

where the transmit power is specified by p_k, which is assigned to the normalized data signal x_k of the k-th user with $\mathbb{E}\left\{|x_k|^2\right\} = 1$. $\boldsymbol{n_l}$ indicates the additive noise at the l-th BS, which is specified as $\boldsymbol{n}_l \sim \mathcal{CN}\left(0, \sigma_{UL}^2 \mathbf{I}_M\right)$. The channel coefficients are assumed to be uncorrelated Rayleigh fading vectors. It means the channel realizations between users, BS antennas and coherence intervals are independent. In other words, the channel vector from the k-th user to the l-th BS , mathematically denoted by $\boldsymbol{h}_{l,k}$, is a realization of the circularly symmetric complex Gaussian distribution with $\boldsymbol{h}_{l,k} \sim \mathcal{CN}\big(0, \beta_{l,k}\mathbf{I}_M\big)$. $\beta_{l,k}$ is the large-scale fading representing for the attenuation due to diffraction, path-loss and shadowing between the l-th BS and the k-th user.

We define the channel matrix $H_l = [\boldsymbol{h}_{l,1}, \ldots, \boldsymbol{h}_{l,K}] \in \mathbb{C}^{M\times K}$, the diagonal matrix $\mathbf{P} = \operatorname{diag}(p_1, \ldots, p_K) \in \mathbb{C}^{K\times K}$, and the useful signal vector $\boldsymbol{x} = [x_1, \ldots, x_K]^T \in \mathbb{C}^K$. Thus, the UL signal can be modeled as

$$\boldsymbol{y}_l = \mathbf{H}_l \mathbf{P}^{1/2} \boldsymbol{x} + \boldsymbol{n}_l. \tag{2}$$

It is very common to assume that the BS is able to estimate the CSI between itself and the users based on the pilot signal. After that, the estimated CSI is used to generate the precoding vectors for DL as well as the receiver combining vectors for UL. In the TDD mode, the length of pilot sequence, specified as τ_p, is often chosen to be more than or equal to the number of users in the network. Consequently, the pilot signal received at l-th BS is

$$\mathbf{Y}_l = \mathbf{H}_l \tilde{\mathbf{P}}^{1/2} \boldsymbol{\Phi} + \mathbf{N}_l, \tag{3}$$

where $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_K]$ are mutual orthogonal pilot sequences, which can be mathematically expressed as $\boldsymbol{\Phi}^H \boldsymbol{\Phi} = \tau_p \mathbf{I}_K$. $\mathbf{N}_l \in \mathbb{C}^{M\times\tau_p}$ denotes the Gaussian noise. The pilot power is represented by the matrix $\tilde{\mathbf{P}} = \operatorname{diag}(\rho_1, \ldots, \rho_K) \in \mathbb{C}^{K\times K}$.

For receiving data correctly, the channel must be estimated. We recall the following Lemma [32].

Lemma 1 (Uplink Channel Estimation): By using the standard MMSE estimation of Gaussian random variables, we acquire the estimated value of $\mathbf{h}_{l,k}$ *(mentioned as* $\hat{\mathbf{h}}_{l,k}$*) and error covariance (as* $\boldsymbol{e}_{l,k}$*)*

$$\hat{\boldsymbol{h}}_{l,k} \sim \mathcal{CN}\left(0, \frac{\rho_k \tau_p \beta_{l,k}^2}{\rho_k \tau_p \beta_{l,k} + \sigma_{UL}^2} \mathbf{I}_M\right), \tag{4}$$

$$\boldsymbol{e}_{l,k} \sim \mathcal{CN}\left(0, \frac{\beta_{l,k}^2 \sigma_{UL}^2}{\rho_k \tau_p \beta_{l,k} + \sigma_{UL}^2} \mathbf{I}_M\right), \tag{5}$$

with $\boldsymbol{e}_{l,k} = \boldsymbol{h}_{l,k} - \hat{\boldsymbol{h}}_{l,k}$.

To minimize the error of channel estimation, we fix the value of pilot power equal to the maximum UL transmit power. Therefore, only data power is considered in the optimization problems of this paper.

A. UPLINK PAYLOAD TRANSMISSION MODEL

In the UL, all K users simultaneously transmit data to all L BSs in the network. It means that an arbitrary BS will receive the signal from all users and then detects its useful data. The signal received at the l-th BS is given as (1). To detect the signal transmitted from the t-th user, the l-th BS multiples the received signal with corresponding combining vector $\boldsymbol{v}_{l,t}$, which has been obtained in the channel estimation step. $\boldsymbol{v}_{l,t}$ clearly depends on the used linear receiver. The decoded signal at the l-th BS can be read as

$$s_{l,t} = \boldsymbol{v}_{l,t}^H \boldsymbol{y}_l = \sum_{k=1}^{K} \boldsymbol{v}_{l,t}^H \boldsymbol{h}_{l,k} \sqrt{p_k} x_k + \boldsymbol{v}_{l,t}^H \boldsymbol{n}_l. \tag{6}$$

Then the desired data x_t will be detected from $s_{l,k}$ as

$$s_{l,t} = \underbrace{\boldsymbol{v}_{l,t}^H \boldsymbol{h}_{l,t} \sqrt{p_t} x_t}_{\text{Desired signal}} + \underbrace{\sum_{k=1,k\neq t}^{K} \boldsymbol{v}_{l,t}^H \boldsymbol{h}_{l,k} \sqrt{p_k} x_k}_{\text{Interference}} + \underbrace{\boldsymbol{v}_{l,t}^H \boldsymbol{n}_l}_{\text{Noise}}. \tag{7}$$

The first component in (7) refers to the t-th user's desired signal. The second one represents the multi-user interference, which degrades the quality of the signal detection. The last component in (7) denotes the additive white Gaussian noise (AWGN) with $\boldsymbol{n}_l \sim \mathcal{CN}\left(0, \sigma_{UL}^2 \mathbf{I}_M\right)$.

Under the assumption that BSs receive desired signals beaming from all the users, we formulate the sum rate capacity of a user in the system. After being multiplied with the corresponding combining vectors, the signal from each BS is sent to a backhaul network. As a result, the signal used to detect the UL data of user t-th is the sum of all $s_{l,t}$. In other words, the received signal at that contains the uplink data of t-th user in the back-haul network is

$$r_t = \sum_{l=1}^{L} s_{l,t} = \sum_{l=1}^{L} \sum_{t=1}^{K} \boldsymbol{v}_{l,t} \boldsymbol{h}_{l,t} \sqrt{p_t} x_t + \sum_{l=1}^{L} \boldsymbol{v}_{l,t}^H \boldsymbol{n}_l. \tag{8}$$

From this r_t, the UL data x_t of the t-th user will be detected. From this point of view, when only channel information are available BS, a lower bound for UL ergodic SE of the t-th user is introduced in *Theorem 1*.

Theorem 1: The lower bound on UL ergodic sum spectral efficiency (SE) of the t-th user is

$$R_t \geq \left(1 - \frac{\tau_p}{\tau_c}\right) \log_2 \left(1 + \mathrm{SINR}_t\right), \tag{9}$$

where SINR_t *has been specified as in* (20).

Proof: For the sake of channel estimation, in every τ_c symbol-length coherent interval, we spend τ_p symbols for pilot signaling, which means there will be $(\tau_c - \tau_p)$ data symbols per coherent interval. As a result, we have the ratio between the number of data symbols to the total length of a coherent interval be $\frac{\tau_c - \tau_p}{\tau_c}$. In other words, the sum SE of t-th user will be proportional with $\left(1 - \frac{\tau_p}{\tau_c}\right)$ as in (9). Using the method mentioned in [33], a lower bound on UL ergodic SE that the l-th BS serves an arbitrary user can be found via utilizing the definition of the mutual information between the original base-band signal x_t and the sum signal.

The lower bound on UL ergodic SE is calculated at the back-haul network as

$$R_t \geq (1-\frac{\tau_p}{\tau_c}) I\left(x_t; r_t, \hat{\mathcal{H}}\right), \tag{10}$$

where $\hat{\mathcal{H}}$ denote the availability of channel estimation at the BSs. With $x_t \sim \mathcal{CN}(0,1)$, the mutual information can be equivalently expressed as

$$\begin{aligned} I\left(x_t; r_t, \hat{\mathcal{H}}\right) &= h(x_t) - h\left(x_t | r_t, \hat{\mathcal{H}}\right) \\ &= \log_2(\pi e) - h\left(x_t | r_t, \hat{\mathcal{H}}\right), \end{aligned} \tag{11}$$

where $h(x_t)$ is the differential entropy and $h\left(x_t|r_t, \hat{\mathcal{H}}\right)$ is the conditional entropy function. Using the property that subtracting a known variable does not change the entropy, we can bound $h\left(x_t|r_t, \hat{\mathcal{H}}\right)$ from above as

$$\begin{aligned} h\left(x_t|r_t, \hat{\mathcal{H}}\right) &= h\left(x_t - \alpha r_t | r_t, \hat{\mathcal{H}}\right) \\ &\leq h(x_t - \alpha r_t) \\ &\leq \log_2\left(\pi e \mathbb{E}\left\{|x_t - \alpha r_t|^2\right\}\right), \end{aligned} \tag{12}$$

where α is some deterministic scalar. In order to find the best upper bound for $h\left(x_t|r_t, \hat{\mathcal{H}}\right)$, we minimize the expectation in (12) with respect to α. The expectation in (12) is

$$\begin{aligned} &\mathbb{E}\left\{|x_t - \alpha r_t|^2\right\} \\ &= \mathbb{E}\left\{\left|x_t - \alpha \sum_{l=1}^{L} v_{l,t}^H \left(\sum_{k=1}^{K} h_{l,k}\sqrt{p_k}x_k + n_l\right)\right|^2\right\} \\ &= 1 - 2\alpha\mathbb{E}\left\{\sum_{l=1}^{L} v_{l,t}^H h_{l,t}\right\}\sqrt{p_t} \\ &\quad + \alpha^2\left[\sum_{k=1}^{K} p_k \mathbb{E}\left\{\left|\sum_{L=1}^{L} v_{l,t}^H h_{l,k}\right|^2\right\} + \sigma_{UL}^2 \mathbb{E}\left\{\sum_{l=1}^{L} ||v_{l,t}||^2\right\}\right]. \end{aligned} \tag{13}$$

As can be seen in (13), the expectation is in the form of a quadratic function with respect to α. It is easy to calculate its minimum value. Therefore, we have:

$$\mathbb{E}\left\{|x_t - \alpha r_t|^2\right\} \geq \frac{1}{1 + \mathrm{SINR}_t}. \tag{14}$$

We select this value for calculating the lower bound of $I\left(x_t; r_t, \hat{\mathcal{H}}\right)$. Plugging the result from (11) to (14) into (10) we obtain the lower bound for SE that the l-th BS can serve the t-th user. Adding up the SE from all BSs in the system, we yield the result as in *Theorem 1*.

Intuitively, since BSs do not have the knowledge of the actual CSI in the system, we are unable to use the sum of $\boldsymbol{v}_{l,t}^H \boldsymbol{h}_{l,t}\sqrt{p_t}x_t$ from all L BSs as the useful signal to apply Shannon theorem. However, the channel information is assumed to be available. Hence, the received signal at the back-haul network for the t-th user can be reformulated as (19), as shown at the top of the next page, by add-and-subtract the term $D_t = \sum_{l=1}^{L} \mathbb{E}\left\{\boldsymbol{v}_{l,t}^H \boldsymbol{h}_{l,t}\right\}\sqrt{p_t}x_t$ which is the expectation of the desired signal. The second term denoted by U_t is the uncorrelated deviation $\sum_{l=1}^{L}\left(\boldsymbol{v}_{l,t}^H \boldsymbol{h}_{l,t} - \mathbb{E}\left\{\boldsymbol{v}_{l,t}^H \boldsymbol{h}_{l,t}\right\}\right)\sqrt{p_t}x_t$ which originates from the difference between the expectation and the actual value of the desired signal. The term $I_t = \sum_{l=1}^{L}\sum_{k=1,k\neq t}^{K} \boldsymbol{v}_{l,t}^H \boldsymbol{h}_{l,k}\sqrt{p_k}x_k$ is the inter-users interference while $N_t = \sum_{l=1}^{L} \boldsymbol{v}_{l,t}^H \boldsymbol{n}_l$ is AWGN. The SINR is therefore expressed as

$$\mathrm{SINR}_t = \frac{\mathbb{E}\left\{|D_t|^2\right\}}{\mathbb{E}\left\{|U_t|^2\right\} + \mathbb{E}\left\{|I_t|^2\right\} + \mathbb{E}\left\{|N_t|^2\right\}}. \tag{15}$$

With each component specified above, we have the SINR expression as in *Theorem 1* intuitively explained. ■

The SINR expression in (20) will be used for QoS constraint in the power optimization problem throughout this paper. Close-form formula for either MRC or ZF deployment will be given in the next section.

B. OBTAINABLE SPECTRAL EFFICIENCY COVERED BY RAYLEIGH FADING

We model the network under the assumption that the BSs use either MRC or ZF to precode the payload data before transmission. The combining vectors can be expressed as

$$\boldsymbol{v}_{l,t} = \begin{cases} \hat{\boldsymbol{h}}_{l,t} & \text{for MRC} \\ \hat{\mathbf{H}}_l \boldsymbol{u}_{l,t} & \text{for ZF,} \end{cases} \tag{16}$$

with $\boldsymbol{u}_{l,t}$ represents the t-th column of the matrix $\left(\hat{\mathbf{H}}_l^H \hat{\mathbf{H}}_l\right)^{-1}$. Let $\gamma_{l,t}$ denote $\frac{\rho_t \tau_p \beta_{l,t}^2}{\rho_t \tau_p \beta_{l,t} + \sigma_{UL}^2}$, the lower bound on the ergodic SE in *Theorem 1* can be obtained in closed-forms for MRC and ZF combining as shown in *Corollaries 1.1* and *1.2*, respectively.

Corollary 1: Let M denote the number of BS antennas. The lower bound on the UL ergodic SE in Theorem 1 utilized by MRC combining is given as

$$R_t^{\mathrm{MRC}} \geq (1 - \frac{\tau_p}{\tau_c})\log_2\left(1 + \mathrm{SINR}_t^{\mathrm{MRC}}\right), \tag{17}$$

where

$$\mathrm{SINR}_t^{\mathrm{MRC}} = \frac{Mp_t\left(\sum_{l=1}^{L}\gamma_{l,t}\right)^2}{\sum_{l=1}^{L}\sum_{k=1}^{K} p_k \beta_{l,k}\gamma_{l,t} + \sigma_{UL}^2 \sum_{l=1}^{L}\gamma_{l,t}}. \tag{18}$$

Proof: We will first calculate every component of the Gaussian random variables in (20). As shown in (16), $\boldsymbol{v}_{l,t} = \hat{\boldsymbol{h}}_{l,t}$ for the MRC case. The expected squared norm of the Rayleigh distribution channel between the l-th BS and the t-th user can be represented as

$$\begin{aligned} \mathbb{E}\left\{||\boldsymbol{v}_{l,t}||^2\right\} &= \mathbb{E}\left\{||\hat{\boldsymbol{h}}_{l,t}||^2\right\} \\ &= \frac{Mp_t \tau_p \beta_{l,t}^2}{p_t \tau_p \beta_{l,t} + \sigma^2} \\ &= \gamma_{l,t} M. \end{aligned} \tag{21}$$

$$r_t = \underbrace{\mathbb{E}\left\{\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,t}\right\}\sqrt{p_t}x_t}_{D_t} + \underbrace{\sum_{l=1}^{L}\left(\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,t} - \mathbb{E}\left\{\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,t}\right\}\right)\sqrt{p_t}x_t}_{U_t} + \underbrace{\sum_{l=1}^{L}\sum_{k=1,k\neq t}^{K}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}\sqrt{p_k}x_k}_{I_t} + \underbrace{\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{n}_l}_{N_t}. \tag{19}$$

$$\mathrm{SINR}_t = \frac{p_t\left|\mathbb{E}\left\{\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,t}\right\}\right|^2}{\sum_{k=1}^{K}p_k\mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}\right|^2\right\} - p_t\left|\mathbb{E}\left\{\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,t}\right\}\right|^2 + \sigma_{UL}^2\mathbb{E}\left\{\sum_{l=1}^{L}\|\boldsymbol{v}_{l,t}\|^2\right\}}. \tag{20}$$

Because the error of estimation is free from the corresponding estimate, we consider each term in the numerator of (20), as shown at the top of the this page, as follows

$$\begin{aligned}\mathbb{E}\left\{\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,t}\right\} &= \mathbb{E}\left\{\hat{\boldsymbol{h}}_{l,t}^H(\hat{\boldsymbol{h}}_{l,t}+\boldsymbol{e}_{l,t})\right\}\\ &= \mathbb{E}\left\{\|\hat{\boldsymbol{h}}_{l,t}\|^2\right\}\\ &= \gamma_{l,t}M,\end{aligned} \tag{22}$$

and the expectation $\mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}\right|^2\right\}$ in case of $k \neq t$ is

$$\begin{aligned}\mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}\right|^2\right\} &= \sum_{l=1}^{L}\mathbb{E}\left\{\left|\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}\right|^2\right\}\\ &\quad + \sum_{l=1}^{L}\sum_{\substack{i=1,\\ i\neq l}}^{L}\mathbb{E}\left\{\left|\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}\left(\boldsymbol{v}_{i,t}^H\boldsymbol{h}_{i,k}\right)\right|\right\}.\end{aligned} \tag{23}$$

Due to the fact that $\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}$ $(\forall l \in 1,\ldots,L)$ are independent of each other, therefore the second term in (23) can be mitigated, which leads to

$$\begin{aligned}\mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}\right|^2\right\} &= \sum_{l=1}^{L}\mathbb{E}\left\{\left|\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k}\right|^2\right\}\\ &= \sum_{l=1}^{L}\mathbb{E}\left\{\left\|\hat{\boldsymbol{h}}_{l,t}\right\|^2\right\}\mathbb{E}\left\{\left\|\boldsymbol{h}_{l,k}\right\|^2\right\}\\ &= \sum_{L=1}^{L}\mathbb{E}\left\{\left\|\hat{\boldsymbol{h}}_{l,t}\right\|^2\right\}\beta_{l,k}\\ &= \sum_{l=1}^{L}\beta_{l,k}\gamma_{l,t}M.\end{aligned} \tag{24}$$

In case of $k = t$, we have the first term of the denominator in (20), which has been given in [34] as

$$\begin{aligned}&\mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,t}\right|^2\right\}\\ &= \mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\left(\hat{\boldsymbol{h}}_{l,t}+\boldsymbol{e}_{l,t}\right)\right|^2\right\}\\ &= \mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\hat{\boldsymbol{h}}_{l,t}\right|^2\right\} + \mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^H\boldsymbol{e}_{l,t}\right|^2\right\}\\ &= \sum_{l=1}^{L}\gamma_{l,t}^2(M+M^2) + \sum_{l=1}^{L}\gamma_{l,t}\left(\beta_{l,t}-\gamma_{l,t}\right)M\\ &\leq \left(\sum_{l=1}^{L}\gamma_{l,t}M\right)^2 + \sum_{l=1}^{L}\beta_{l,t}\gamma_{l,t}M.\end{aligned} \tag{25}$$

Substituting (21), (22), (24) and (25) to (20), we obtain the SINR as in *Corollary 1.1.* ■

The advantage of using MRC is clearly shown in (18) that the desired signal power increases proportionally to the number of BS antennas. In addition, the interference is not affected by the number of BS antennas.

Corollary 2: The lower bound on the UL ergodic SE in Theorem 1 utilized by ZF combining is given as

$$R_t^{\mathrm{ZF}} \geq (1-\frac{\tau_p}{\tau_c})\log_2\left(1+\ \mathrm{SINR}_t^{\mathrm{ZF}}\right), \tag{26}$$

where

$$\mathrm{SINR}_t^{\mathrm{ZF}} = \frac{(M-K)p_tL^2}{\sum_{l=1}^{L}\sum_{k=1}^{K}p_k\frac{\beta_{l,k}-\gamma_{l,k}}{\gamma_{l,t}} + \sigma_{UL}^2\sum_{l=1}^{L}\frac{1}{\gamma_{l,t}}}. \tag{27}$$

Proof: As shown in (16), we have $\boldsymbol{v}_{l,k} = \hat{\mathbf{H}}_l\boldsymbol{u}_{l,k}$ in the ZF case. Using the result in Lemma 2.10 given in [34] and the properties of $K \times K$ central complex Wishart matrix, in which M is the degrees of freedom and satisfies $M \geq K+1$, we have

$$\begin{aligned}\mathbb{E}\left\{\|\boldsymbol{v}_{l,t}\|^2\right\} &= \mathbb{E}\left\{\left[\left(\hat{\mathbf{H}}_l^H\hat{\mathbf{H}}_l\right)^{-1}\right]_{t,t}\right\}\\ &= \frac{1}{\gamma_{l,t}\left(M-K\right)},\end{aligned} \tag{28}$$

where $\left[\left(\hat{\mathbf{H}}_l^H\hat{\mathbf{H}}_l\right)^{-1}\right]_{t,t}$ represents the element at the t-th row and the t-th column of the matrix $\left[\left(\hat{\mathbf{H}}_l^H\hat{\mathbf{H}}_l\right)^{-1}\right]$. ZF combining has to ensure the following property

$$\boldsymbol{v}_{l,t}^H\boldsymbol{h}_{l,k} = \begin{cases}1, & k = t\\ 0, & k \neq t.\end{cases} \tag{29}$$

Therefore, the numerator of (20) can be written as

$$\begin{aligned}\mathbb{E}\left\{\boldsymbol{v}_{l,t}^{H}\boldsymbol{h}_{l,t}\right\} &= \mathbb{E}\left\{\boldsymbol{v}_{l,t}^{H}(\hat{\boldsymbol{h}}_{l,t}+\boldsymbol{e}_{l,t})\right\}\\ &= \mathbb{E}\left\{1\right\}+\mathbb{E}\left\{\boldsymbol{v}_{l,t}^{H}\boldsymbol{e}_{l,t}\right\}\\ &= 1. \end{aligned} \tag{30}$$

The first term in the denominator of (20) in the case of $k \neq 1$ is given as

$$\begin{aligned}\mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^{H}\boldsymbol{h}_{l,k}\right|^2\right\} &= \mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^{H}\left(\hat{\boldsymbol{h}}_{l,k}+\boldsymbol{e}_{l,k}\right)\right|^2\right\}\\ &= \mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^{H}\boldsymbol{e}_{l,k}\right|^2\right\}\\ &= \sum_{l=1}^{L}\mathbb{E}\left\{\left\|\hat{\boldsymbol{h}}_{l,t}\right\|^2\right\}\left(\beta_{l,k}-\gamma_{l,k}\right)\\ &= \sum_{l=1}^{L}\frac{\left(\beta_{l,k}-\gamma_{l,k}\right)}{\gamma_{l,t}\left(M-K\right)}. \end{aligned} \tag{31}$$

Similarly, in case of $t = k$, we have

$$\begin{aligned}&\mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^{H}\boldsymbol{h}_{l,t}\right|^2\right\}\\ &= \mathbb{E}\left\{\left|\sum_{l=1}^{L}\boldsymbol{v}_{l,t}^{H}\left(\hat{\boldsymbol{h}}_{l,t}+\boldsymbol{e}_{l,t}\right)\right|^2\right\}\\ &= \sum_{l=1}^{L}\mathbb{E}\left\{1\right\}+\sum_{l=1}^{L}\mathbb{E}\left\{\left\|\hat{\boldsymbol{h}}_{l,t}\right\|^2\right\}\left(\beta_{i,k}-\gamma_{l,t}\right)\\ &\leq L^2+\sum_{l=1}^{L}\frac{\left(\beta_{l,t}-\gamma_{l,t}\right)}{\gamma_{l,t}\left(M-K\right)}. \end{aligned} \tag{32}$$

Substituting (28), (30), (31) and (32) into (20), we obtain the SINR for ZF combining as in *Corollary 1.2*. ■

It can be seen that ZF offers a lower array gain than MRC. As a result, MRC works better in systems with high noise level at BSs, because MRC aims to improve the SNR by multiplying the desired signal power with the numbers of BS antennas while the noise and interference remains unchanged. Meanwhile, ZF is suitable for systems with a system having a high number of users. In this case, interference between users outweighs noise, but can be effectively mitigated as seen in (27). That is the strategy how the optimal transmit power p_t is selected.

III. JOINT POWER CONTROL FOR UPLINK CELL-FREE MULTI-USERS MASSIVE MIMO

In this section, we formulate the power control problem in three approaches: 1) optimal power delivery, 2) max-min fairness and 3) SE maximization. Finally, we also offer solutions for each problem.

A. POWER MINIMIZATION PROBLEM WITH LINEAR PROGRAM

We investigate the joint power control for an uplink massive MIMO system, in which the main target is to minimize the total transmit power of users while each user is ensured with a predefined QoS. The problem is formulated as

$$\begin{aligned}&\underset{p_t}{\mathit{minimize}}\ \sum_{t=1}^{K}p_t\\ &\text{subject to } R_t \geq \xi_t, \quad \forall t \in [1,\ldots,K]\\ &\qquad p_t \leq p_{max}, \quad \forall t \in [1,\ldots,K],\\ &\qquad p_t \geq 0, \quad \forall t \in [1,\ldots,K]. \end{aligned} \tag{33}$$

where ξ_t depicts the target QoS of the t-th user. Due to the requirement of energy saving on mobile devices, it is necessary to limit the transmit power of each user by a maximum value of P_{max}. Following the result in (9), the QoS constraint in (33) can be considered as SINR and rewritten as

$$1+\mathrm{SINR}_t \geq 2^{\frac{\xi_t\tau_c}{\tau_c-\tau_p}}. \tag{34}$$

Plugging (18) and (27) into (33), we obtain the optimization solution for MR and ZF linear receiver in Lemma 2.

The power minimization problem for MR and ZF linear receiver is expressed as

$$\begin{aligned}&\underset{p_t}{\mathit{minimize}}\ \sum_{t=1}^{K}p_t\\ &\text{subject to } 1+\frac{p_t g_t}{\sum_{l=1}^{L}\sum_{k=1}^{K}p_k z_{l,k}w_{l,t}+\sigma_{UL}^2\sum_{l=1}^{L}w_{l,t}}\\ &\qquad \geq 2^{\frac{\xi_t\tau_c}{\tau_c-\tau_p}}, \quad \forall t \in [1,\ldots,K],\\ &\qquad p_t \leq P_{max}, \quad \forall t \in [1,\ldots,K],\\ &\qquad p_t \geq 0, \quad \forall t \in [1,\ldots,K]. \end{aligned} \tag{35}$$

where

$$g_t = \begin{cases} M\left(\sum_{l=1}^{L}\gamma_{l,t}\right)^2 & \text{for MRC}\\ (M-K)L^2 & \text{for ZF,} \end{cases} \tag{36}$$

$$z_{l,t} = \begin{cases} \beta_{l,t} & \text{for MRC}\\ \beta_{l,t}-\gamma_{l,t} & \text{for ZF,} \end{cases} \tag{37}$$

and

$$w_{l,t} = \begin{cases} \gamma_{l,t} & \text{for MRC}\\ \dfrac{1}{\gamma_{l,t}} & \text{for ZF.} \end{cases} \tag{38}$$

The problems in (35) is linear programing and thus, can be solved effectively to find the globally optimal solution in polynomial time. The optimal value for UL transmit power will be utilized to solve user association problem which will be discuss in section III-D of this paper.

B. MAX-MIN QoS FAIRNESS OPTIMIZATION WITH MRC

QoS fairness is a vital criteria of a wireless network to serve users in the system with an equal data rate. In fact, the power control problem with a predefined QoS value is not always feasible. It depends on various factors such as large-scale fading, interference, noise at BSs and quality of channel estimation. Therefore, in this section we consider the min-QoS optimization problem which targets to find the greatest value of QoS that our model can equally serve every user. With the QoS of the t-th being equivalently represented by SINR, the problem is formulated as

$$\begin{aligned} &\underset{p_t}{maximize} \quad \min_{t=1,\ldots,K} R_t \\ &\text{subject to } p_t \leq p_{max}, \quad \forall t \in [1,\ldots,K], \\ &\qquad\qquad p_t \geq 0, \quad \forall t \in [1,\ldots,K]. \end{aligned} \tag{39}$$

Because R_t, which is in the form of $\log(1+x)$, is an ascending function with respect to x, we can rewrite the problem in (39) into epigraph form

$$\begin{aligned} &\underset{p_t,\lambda}{maximize}\ \lambda \\ &\text{subject to } \frac{p_t g_t}{\sum_{l=1}^{L}\sum_{k=1}^{K} p_k z_{l,k} w_{l,t} + \sigma_{UL}^2 \sum_{l=1}^{L} w_{l,t}} \geq \lambda, \\ &\qquad \forall t \in [1,\ldots,K], \\ &\qquad p_t \leq p_{max}, \quad \forall t \in [1,\ldots,K], \\ &\qquad p_t \geq 0, \quad \forall t \in [1,\ldots,K]. \end{aligned} \tag{40}$$

By multiplying both sides of each SINR constrain to the corresponding denominator, we can see that the problem in (40) is a Geometric Program (GP) which can be converted into a convex optimization.

C. SUM SPECTRAL EFFICIENCY OPTIMIZATION

In contrast to the max-min fairness problem, the optimal sum SE aim to maximize the throughput of the entire system. The max sum-SE problem can be presented as

$$\begin{aligned} &\underset{p_t}{maximize} \ \sum_{t=1}^{K} R_t \\ &\text{subject to } p_t \leq p_{max}, \quad \forall t \in [1,\ldots,K], \\ &\qquad\qquad p_t \geq 0, \quad \forall t \in [1,\ldots,K]. \end{aligned} \tag{41}$$

Solving (41), we obtain the optimum power control for max sum-SE strategy as well as the maximal value of sum SE of the network. Unfortunately, it has been confirmed in [35] that the power control optimization problem of this type is NP hard, hence it can not be effectively solved. NP-hard problem can only be solved in polynomial-time with no-deterministic computation. For close-optimal solution, we apply the method proposed in [31] to translate the original problem into a deterministic polynomial time one. The proposed method makes use of a successive solution based on convex optimization. As can be seen in (41), the non-convexity of the problem derived from the QoS constraint. Therefore, we will approximate the QoS constraint into the form of GP constraint. With this conversion, the problem can be solved effectively using CVX [36]. First, we rewrite the problem in (41) using epigraph form

$$\begin{aligned} &\underset{p_t,\lambda_t}{maximize} \ \prod_{t=1}^{K} \lambda_t \\ &\text{subject to } 1 + \frac{p_t g_t}{\sum_{l=1}^{L}\sum_{k=1}^{K} p_k z_{l,k} w_{l,t} + \sigma_{UL}^2 \sum_{l=1}^{L} w_{l,t}} \\ &\qquad \geq \lambda_t, \quad \forall t \in [1,\ldots,K], \\ &\qquad p_t \leq p_{max}, \quad \forall t \in [1,\ldots,K], \\ &\qquad p_t \geq 0, \quad \forall t \in [1,\ldots,K]. \end{aligned} \tag{42}$$

To utilize GP, we approximate $f\left(p_t\right) = 1 + \text{SINR}_t$ to condense its numerator to a monomial, which satisfies the top-level rules of GP. Every QoS constraint has to be approximated, where all new generated QoS functions have to satisfy the conditions given in the following Lemma that has been presented in [37].

Lemma 2: Constructing functions satisfy the following properties. Let i denote the order of the iteration.

1) $f\left(p_t\right) \leq f_i\left(p_t\right), \forall t \in [1,\ldots,K]$ *in the feasible set*
2) $f\left(p_t^{i-1}\right) = f_i\left(p_t^{i-1}\right)$, *where* p_t^{i-1} *is the result from previous iteration*
3) $\nabla f\left(p_t^{i-1}\right) = \nabla f_i\left(p_t^{i-1}\right)$.

The constructed optimal problem is obtained by replacing $f\left(p_t\right)$ by $f_i\left(p_t\right)$. After that, these series of the solutions are transformed to a KKT point of the original problem. Based on Lemma 3, we approximate the SINR constraint in (42), as following Lemma 4.

Lemma 3: For any polynomial $y(x) = m_i(x)$, *it is valid for any* α_i *that*

$$y(x) \geq \tilde{y}(x) = \prod_i \left(\frac{m_i(x)}{\alpha_i}\right)^{\alpha_i}. \tag{43}$$

By choosing $\alpha_i = m_i\left(x_o\right)/y\left(x_o\right)$, the approximate conditions will be satisfied. Note that the SINR constraint in (42) has the form of

$$f_t\left(p_k\right)/g_t\left(p_k\right) \geq \lambda_t, \tag{44}$$

with

$$f_t(p_k) = p_t g_t + \sum_{l=1}^{L}\sum_{k=1}^{K} p_k z_{l,k} w_{l,t} + \sigma_{UL}^2 \sum_{l=1}^{L} w_{l,t}, \tag{45}$$

and

$$g_t(p_k) = \sum_{l=1}^{L}\sum_{k=1}^{K} p_k z_{l,k} w_{l,t} + \sigma_{UL}^2 \sum_{l=1}^{L} w_{l,t}, \tag{46}$$

which is not a valid GP form because $f_t\left(p_k\right)$ and $g_t\left(p_k\right)$ are both polynomials. By transforming each numerator $f_t\left(p_k\right)$ to a monomial $\tilde{f}_t\left(p_k\right)$, we convert the SINR constraint to $\tilde{f}_t\left(p_k\right) \geq \lambda_t g_t\left(p_k\right)$. Here, the left side is a monomial and the right side is a polynomial, which satisfies the top-level rules of GP given. This conversion will be repeatedly conducted for all K SINR constraints to ensure that the SINR constraint satisfies the GP condition. The optimization problem can be then solved by using CVX [36]. This process will be iteratively carried out to finally get a optimal KKT local optimum points of (42). The iteration is stopped when the approximated problem is converged as in *Lemma 3*. The procedure of solving the optimization problem in (33) is represented in Algorithm 1.

Algorithm 1 Successive Approach Process for Obtaining Convex Optimization Problem

```
Require: choosing $p_t^0$ equal to the maximum power of an
    user (which is feasible for (42))
Ensure: all $p_t^i$
    $i \leftarrow 1$
    while unconvergence do
        The $i$-th approximated problem of (42) is obtained by
        approximating every SINR constraints using lemma 3
        Solving the $i$th approximated problem to get $p_t^i$ of $t$-th
        user
        $i \leftarrow i + 1$
    end while
    return $p_t^i$
```

D. BS-USERS ASSOCIATION METHOD

In this section, we propose a BS-user association method that exploits the optimal transmit power closed-form. Although an user can connect to every BS in the system, there are still some BSs with negligible data being transmitted from the users, which causes noise and interference to the sum signal rather than useful data. Hence, these weak connections negatively affect the performance optimization process of the system. Furthermore, it is obvious that a high number of connections from the users to a BS cause large burden on load balancing of that BS. Therefore, the proposed method effectively eliminates the weak connections from a set serving BS with a user. The detail of the proposed approach is presented in Algorithm 2. According to simulation results, we observed that although an arbitrary user can be served by all BSs in the system, there are some BSs providing more interference than desired signal. We define the group of indices of BSs that serve t-th user as S_t and try to remove inefficient BSs from it by an iterative method. At each loop, if the l-th BS serves t-th user with SINR lower than $R_{threshold}$, which is chosen to be the lowest SINR that a BS must supply to a user, it will be added to a removed BSs set A_t. After that, we will solve (33) with respect to the new serving BSs set of $S_t \setminus A_t$. The change in the optimal power consumption of the system specified as δ must not surpass a predetermined value of $\delta_{threshold}$ and the serving BSs set can now be updated as $S_t = S_t \setminus A_t$. Otherwise, S_t remains unchanged. The same procedure must be carried out for all users in the system to acquired the most effective association.

Algorithm 2 BS Removing Algorithm

```
Require: $S_t \in \{1, \ldots, L\}$ defines the group of indices of BSs
    that serve $t$-th user
    $i = 1; \delta = 0; A_t = \varnothing$
    $P^*_{prev} = P^*$ where $P^*$ is the optimal solution for (33) with
    respect to the original BSs set
    while $\delta < \delta_{threshold}$ and $i \leq L$ do
        Solve (33) with respect to the BS subset whose indices
        in $S_k$, the optimal solution is saved as $P^*_{new}$
        if $R_{i,t} \leq R_{threshold}$ then
            $A_t = A_t \cup \{i\}; S_t = S_t \setminus A_t$
        end if
        $\delta = P^*_{new} - P^*_{prev};$   $P^*_{prev} = P^*_{new};$   $i \leftarrow i + 1$
    end while
    return $S_t$
```

IV. NUMERICAL RESULTS AND DISCUSSION

In this section, we will analyze the performance of our proposed joint power control and balancing load approach for uplink CF-MU massive MIMO system. We first describe the specification of our system used in this experiment. Next, the performance is evaluated with different aspects: the change in per-user transmit power when adjusting the number of antennas or the user average transmit power to obtain the various target QoS levels, QoS fairness and total SE in the network. Numerical results manifest the competitiveness of our proposed IE method versus the max-SNR association (where users connect to a single BS with highest SNR) and full-set joint reception, which means that a user can connect to all BSs.

A. SYSTEM SPECIFICATION

We consider the uplink channel of a CF-MU Massive MIMO system which has 4 fixed base stations serving all users. These users is uniformly and randomly distributed over the active area of the 4 BSs. The upper bound of UL transmit power is set at 23 dBm. The system bandwidth is 20 MHz. The coherence interval contains 200 symbols, in which the length of UL pilot is equal to the number of users in the system. To reduce the complexity of the simulation, The user pilot power are fixed equally to the maximum UL power. The large scale fading coefficients are modeled according to the 3GPP LTE standard [38]. The simulation parameters are presented in Table 1. From the values in Table 1, the large scale fading is $\beta_{l,k} = -131 - 42.8\log 10d + z_{l,k}$ dB where $z_{l,k}$ is shadowing standard deviation. The noise of 5 dB leads to a noise variance of 96 dBm.

B. SIMULATION RESULTS

The change of per-user transmitting power versus the number of antennas, when the QoS is given as 2 bits/HZ/sec, is described in Figure 1. To have a fair comparison, the results have been averaged over 100 Monte Carlo runs. It can be seen from Figure 1 that there is a remarkable gap in transmit power

TABLE 1. **Simulation parameters.**

Parameter	Values
Peak UL radio transmit power	23 dBm
Number of BSs	4
Coherence interval length	200 symbols
Shadowing standard deviation	10 dB
Penetration loss (indoor users)	20 dB
Noise figure	5 dB
Pathloss	$131 + 42.8\log10d$

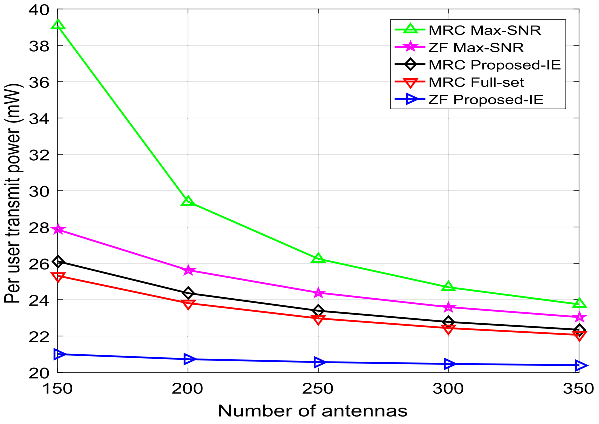

FIGURE 1. **Per user transmit power versus number of antenna with QoS = 2bits/Hz/s.**

TABLE 2. **Power gap between proposed IE method and full-set joint-reception with QoS=2 bits/Hz/s ($P_{IE} - P_{Fullset}$).**

Number of antennas	150	200	250	300	350
MRC gap (in mW)	0.7929	0.5498	0.4196	0.3396	0.2852
ZF gap (in mW)	-1361	-449.1	-276.2	-200	-156.9

between our proposed IE method and that one based on max SNR. The comparison between our proposed IE method using MRC and max SNR using MRC shows also the difference of around 13 mW with 150 antennas. With the same number of antennas, when system utilizes ZF combining, the gap between max-SNR and our proposed method is 7 mW. Meanwhile, the gap in power consumption between our proposed method and the full-set joint transmission one for both ZF and MRC case is show in Table 2.

We can also see from Figure 1 that the per-user transmitting power for max-SNR association significantly drops with the number increasing of the transmit antenna thanks to the array gain from linear receivers. Furthermore, the power gap between MRC and ZF is reduced by the number of BS antennas, since interference is mitigated more efficiently. Meanwhile, with the growing number of antennas from 150 to 350, the change in per-user transmit power of our proposed method witnesses just a slight decrease of around 2 mW, which shows the stability of our method despite significant change in the number of antennas.

TABLE 3. **Per user transmit power fof ZF full-set joint reception with fixed QoS=2 bits/Hz/s.**

Number of antennas	150	200	250	300	350
Per user transmit power (in mW)	1382.2	469.8	296.8	220.5	177.3

According to Table 2, it can be seen that the power consumption of our proposed IE method is always just about 0.8 mW greater than Full-set's figure for MRC case. By contrast, our proposed IE method shows a superior power efficiency compared to full-set joint reception, whose figures specified in Table 3. This shows that the joint reception with full BSs set is not always the optimal association strategy. Because the multiple BS cooperation increases not only the array gain but also the mutual interference that appear in the numerator and denominator of the SINR expression, respectively.

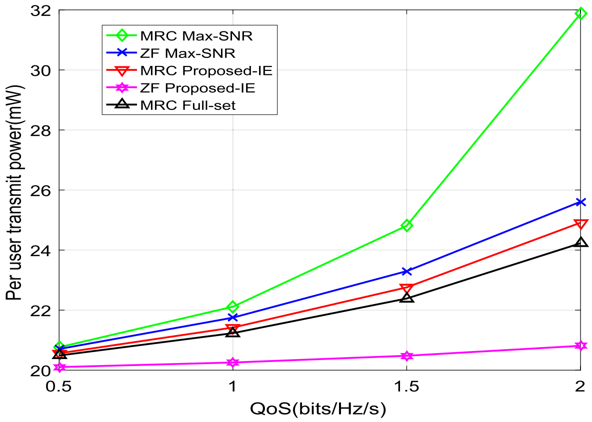

FIGURE 2. **Per-user transmit power $\sum_{i=1}^{N} P_i/N$ versus the target QoS at the users with 200 BS antennas.**

Figure 2 shows the user average transmit power to obtain the different target QoS levels at the users. The number of antenna is set at 200. Clearly, when the required QoS increases, the transmit power also goes up. As mentioned above, the power consumption of MRC is lower in the low QoS target, where the noise power dominates the signal ones. Meanwhile, ZF combining consumes less power in high QoS case, when multi-user interference is more significant. In the low QoS regime, the gap between ZF and MRC combining power consumption is quite small compared to the transmit power. The same trend are observed for the max-SNR association method. However, the gap between the max-SNR and proposed IE association methods is quite significant. It implies that our proposed method are far more better than max-SNR and asymptotic to the optimal association for MRC case. With ZF, joint reception with full BSs set has a worse result compare to max-SNR and our propose IE-method. With the limit of transmit power, the power minimization problem is infeasible in most of the time when QoS is greater than 2 bits/Hz/s. While removing this upper bound, the performance of joint-reception with full BSs set is shown in Table 4.

TABLE 4. **Per user transmit power for ZF full-set joint reception with 200 BS antennas according to QoS level.**

QoS Level	0.5	1	1.5	2
Per user transmit power (in mW)	47.9	97.7	199.9	501.7

The cumulative distribution function (CDF) of the max sum SE optimized QoS level is depicted in Figure 3. It can be seen that the QoS of our proposed IE association method can reach to just under 5 bits/Hz/s when using MRC and eventually 9 bits/Hz/s using ZF combining. That means ZF delivers almost double the capacity comparing with MRC. The reason is that when the transmitted power of users is high, the interference between users are also increased. Hence, ZF combinator is more effective to cancel the interference term. The same trends can be seen in max-SNR association method. From Figure 3, we observe that max-SNR has around 0.5 bits/Hz/s lower than our proposed method. It is worth to mention that the ZF with full BSs-set joint reception has the worst performance. This is due to the fact that the gain of useful signal can not surpass the enormous amount of noise and interference added up from all BSs in the network. It can also be seen that ZF provides about 50% higher QoS than MRC. From this result, we conclude that the impact of linear receiver type is more significant than association method's.

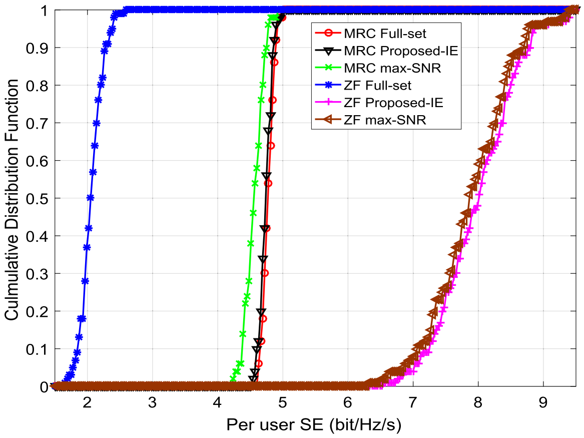

FIGURE 3. **The cumulative distribution function of per user SE with 200 BS antennas.**

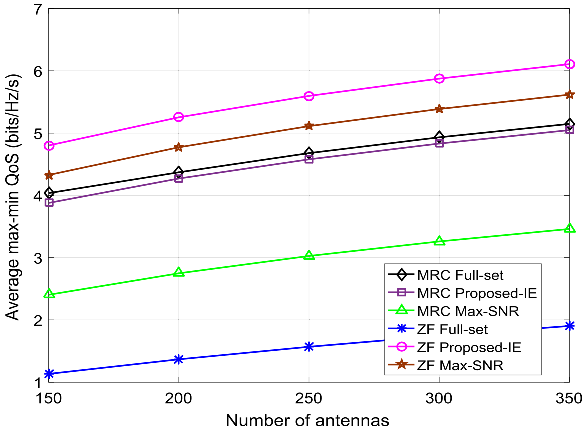

FIGURE 4. **The average QoS level versus the number of BS antennas.**

Figure 4 illustrates the QoS levels that the system can equally provide to the all users. Our method provides up to 1.5 bits/Hz/s higher than the max-SNR association for MRC and 0.5 bits/Hz/s for ZF. Especially, ZF combining with max-SNR association provides better average QoS than both MRC with full-set joint reception and our proposed method. This result can be explain by the same way as we did with Figure 3 that the performance of ZF combining outweighs MRC in high QoS regime. The proposed IE method remove the ineffective BSs, which have less than a specified capacity. Therefore, with the proposed IE method, the complexity of the receiver is greatly reduced compared to the full-set joint transmission while the performance of the former is just a little bit worse than that of the later in MRC case and eventually better with ZF. In other words, the proposed method provide a good performance - complexity trade off compared to the full-set joint transmission method. The decrease in performance of IE over optimal methods is around 0.1 bit/Hz/s using MRC, meanwhile with ZF, our method is far better than full-set joint reception. It is notable that with ZF, full-set joint reception average QoS is lower than 2 bits/Hz/s, which shows the reason why the minimization problem for it is infeasible for most of the cases.

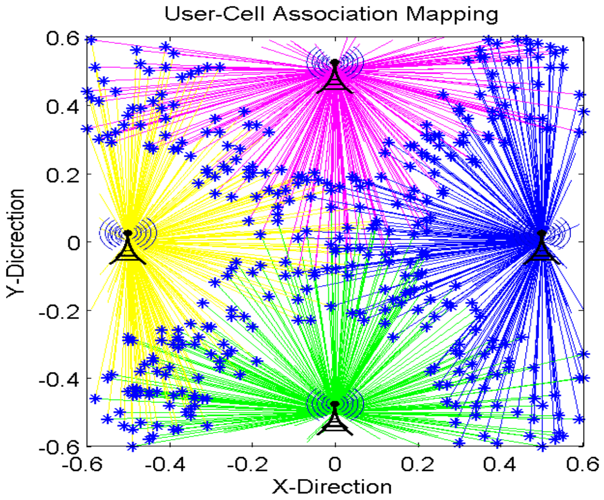

FIGURE 5. **Association mapping with randomly distributed 100 users.**

In Figure 5, we further illustrate the association map in our simulation model with 100 users. The stars denote the users which are served by multi BSs. From this figure, we observe that an user tends to associate with the BSs whose smaller

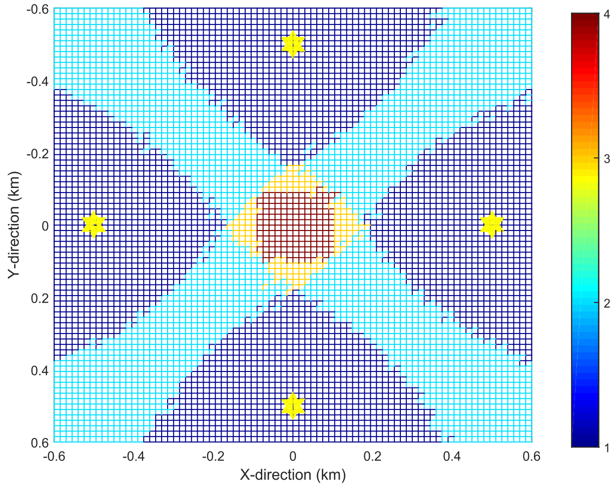

FIGURE 6. Number of BSs are serving an user for MRC.

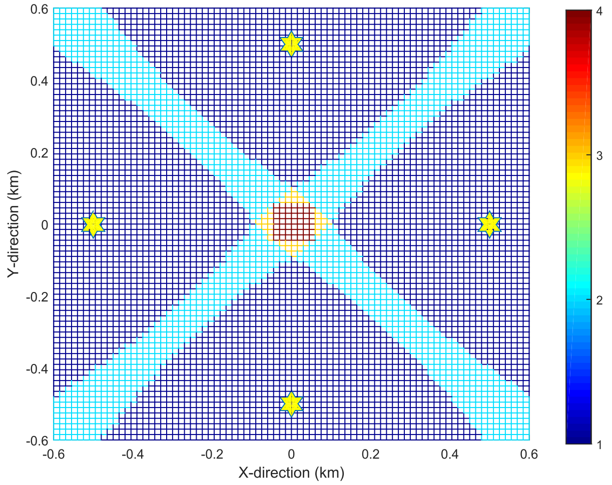

FIGURE 7. Number of BSs are serving an user for ZF.

distance it. This phenomenon is resulted from that smaller distance guarantees lower attenuation. In other words, an user's signal received at the nearer BS will be stronger and that connection can supply a higher data rate compared to a further one. For example, we consider the BS at the coordinate (0.5,0). The majority of users in the radius of 0.5 kilometers from this BS tends to connect to it. Remarkably, some other users in further distance also connect to this BS as the result of several factors such as interference from other users, channel estimation or shadow fading. Beside that, users in the co-coverage between BSs usually connect with multiple users. The dependence of the number of BSs serving one users on user's location is shown in Figure 6 and 7. It is noticed that the users in the center of the map can be served by all 4 BSs, which brings higher efficiency in transmit power but more complexity in load balancing.

V. CONCLUSIONS

In this paper, we consider the Uplink Cell-Free Multi-user Massive MIMO system that comprises a distributed number of base stations (BS) equipped with a large number of antennas. We have jointly investigated the optimization of power control for the uplink channel and load balancing problem under both perfect and imperfect CSI schemes for uplink cell-free multi-user massive MIMO systems. The UL coherent joint transmission was designed to minimize the total transmit power consumption while satisfying QoS. Because the joint optimization under multiple constraints is a combinatorial problem, and thereby is computational complex, we proposed a low complexity algorithm for removing inefficient BS, therefore the load balancing problem can be solved. For tackling the power control optimization, which is NP-Hard, we used successive approach to approximate the constraints into the form of GP. We have proved that the use of maximum signal to noise ratio (max-SNR) association in Massive MIMO systems is not effective, especially in the high QoS regime. This makes the chosen linear receiver become more important than the BS-Users association method.

REFERENCES

[1] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[2] 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.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] 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.

[5] P. S. Parihar, R. Saraswat, and S. Maheshwari, "Energy and spectral efficiency of very large multiuser MIMO systems," *Int. J. Comput. Appl.*, vol. 111, no. 5, pp. 1–3, 2015.

[6] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[7] 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.

[8] F. Boccardi, R. W. Heath, 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.

[9] X. Chen and C. Yuen, "Efficient resource allocation in a rateless-coded MU-MIMO cognitive radio network with QoS provisioning and limited feedback," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 395–399, Jan. 2013.

[10] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[11] E. Lähetkangas *et al.*, "Achieving low latency and energy consumption by 5G TDD mode optimization," in *Proc. IEEE Int. Conf. Commun. Workshops*, Jun. 2014, pp. 1–6.

[12] 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.

[13] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and H. Yang, "Cell-free massive MIMO systems," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2015, pp. 695–699.

[14] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun.*, Jun./Jul. 2015, pp. 201–205.

[15] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy efficiency in cell-free massive MIMO with zero-forcing precoding design," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1871–1874, Aug. 2017.

[16] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[17] T. Van Chien, E. Björnson, and E. G. Larsson, "Joint power allocation and user association optimization for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6384–6399, Sep. 2016.

[18] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[19] S. Wang, Y. Li, and J. Wang, "Multiuser detection in massive spatial modulation MIMO with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2156–2168, Apr. 2015.

[20] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo, "Low-complexity iterative detection for large-scale multiuser MIMO-OFDM systems using approximate message passing," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 902–915, Oct. 2014.

[21] D. Bethanabhotla, O. Y. Bursalioglu, H. C. Papadopoulos, and G. Caire, "Optimal user-cell association for massive MIMO wireless networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1835–1850, Mar. 2016.

[22] R. Chai, M. Chen, Q. Chen, and Y. Gao, "An optimal joint user association and power allocation algorithm for secrecy information transmission in heterogeneous networks," *Wireless Commun. Mobile Comput.*, vol. 2017, Mar. 2017, Art. no. 5120538.

[23] D. Liu, L. Wang, Y. Chen, T. Zhang, K. Chai, and M. Elkashlan, "Distributed energy efficient fair user association in massive MIMO enabled HetNets," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1770–1773, Oct. 2015.

[24] L. Liu and R. Zhang, "Downlink SINR balancing in C-RAN under limited fronthaul capacity," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2016, pp. 3506–3510.

[25] G. Su, B. Chen, X. Lin, H. Wang, and L. Li, "A submodular optimization framework for outage-aware cell association in heterogeneous cellular networks," *Math. Problems Eng.*, vol. 2016, Feb. 2016, Art. no. 4567625.

[26] T. Van Chien and E. Björnson, and E. G. Larsson, "Downlink power control for massive MIMO cellular systems with optimal user association," in *Proc. IEEE Int. Conf. Commun.*, May 2016, pp. 1–6.

[27] N. Wang, E. Hossain, and V. K. Bhargava, "Joint downlink cell association and bandwidth allocation for wireless backhauling in two-tier HetNets with large-scale antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3251–3268, May 2016.

[28] S. Zhang, R. Zhang, and T. J. Lim, "Massive MIMO with per-antenna power constraint," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Dec. 2014, pp. 642–646.

[29] D. I. Kim, L. B. Le, and E. Hossain, "Joint rate and power allocation for cognitive radios in dynamic spectrum access environment," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5517–5527, Dec. 2008.

[30] K. Elkhalil, A. Kammoun, T. Y. Al-Naffouri, and M.-S. Alouini, "A blind antenna selection scheme for single-cell uplink massive MIMO," in *Proc. IEEE Globecom Workshops*, Dec. 2016, pp. 1–6.

[31] 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. Int. Symp. Wireless Commun. Syst.*, Aug. 2015, pp. 396–400.

[32] S. K. Sengijpta, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.

[33] T. Van Chien and E. Björnson, "Massive MIMO communications," in *5G Mobile Communications*. Cham, Switzerland: Springer, 2017, pp. 77–116.

[34] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[35] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[36] M. Grant, S. Boyd, and Y. Ye. (2015). *CVX: Matlab Software for Disciplined Convex Programming*. [Online]. Available: http://www.stanford.edu/boyd/cvx

[37] B. R. Marks and G. P. Wright, "Technical note—A general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, pp. 681–683, 1978.

[38] F. Rezaei, M. Hempel, and H. Sharif, "LTE PHY performance analysis under 3GPP standards parameters," in *Proc. IEEE 16th Int. Workshop Comput. Aided Modeling Design Commun. Links Netw.*, Jun. 2011, pp. 102–106.

TIEN HOA NGUYEN received the Dip.-Ing. degree in electronics and communication engineering from Hanover University and the Ph.D. degree from the Department Wireless Communication Technique, Hanoi University of Science and Technology, in 2016. He was with the Research and Development Department in Image Processing and the Development of SDR-Based Drivers, Bosch. He has three years of experiments with the Research and Development Team, MIMOon, to develop embedded signal processing and radio modules for 4G and 5G mobile networks. He is currently a Lecture with the Hanoi University of Science and Technology. His area of research interests are resource allocation in cognitive radios, massive MIMO, and vehicular communication systems.

THE KHAI NGUYEN received the degree (Hons.) from the School of Electronics and Telecommunications, Hanoi University of Science and Technology, Vietnam. He has two years of experience at the Signal Processing Laboratory, as a Senior Member. His research interests include microprocessors, convex optimization, and future 5G mobile networks. His current research interests include power allocation and user association for massive MIMO.

HUY DUNG HAN received the B.S. degree from the Faculty of Electronics and Telecommunications, Hanoi University of Science and Technology, Hanoi, Vietnam, in 2001, the M.Sc. degree from the Technical Faculty, University of Kiel, Germany, in 2005, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA, USA, in 2012. He is currently with the Department of Electronics and Computer Engineering, School of Electronics and Telecommunications, Hanoi University of Science and Technology. His research interests are in the area of wireless communications and signal processing, with current emphasis on blind and semi-blind channel equalization for single and multi-carrier communication systems and convex optimization. Dr. Han has been serving on the Technical Programs of the IEEE International Conference on Communications and Electronics.

VAN DUC NGUYEN was born in Thanh Hoa, Vietnam in 1973. He received the B.E. and M.E. degrees in electronics and communications from the Hanoi University of Technology, Vietnam, in 1995 and 1997, respectively, and the Dr.-Eng. degree in communications engineering from the University of Hannover, Germany, in 2003. From 1995 to 1998, he was with the Technical University of Hanoi as an Assistant Researcher. In 1996, he participated in the Student Exchange Program between the Technical University of Hanoi and the Munich University of Applied Sciences for one term. From 1998 to 2003, he was with the Institute of Communications Engineering, University of Hannover, first as a DAAD Scholarship Holder and then as a member of the scientific staff. From 2003 to 2004, he was with the University of Agder at Grimstad, Norway, as a Postdoctoral Researcher. He was with the International University of Bremen as a Postdoctoral Fellow. In 2007, he spent two months with the Sungkyungkwan University, South Korea, as a Research Professor. From 2008 to 2009, he was a Visiting Researcher with the Alpen-Adria-Universität Klagenfurt, Austria, and the University of Agder, Norway, respectively. His current research interests include mobile radio communications, especially MIMO-OFDM systems, radio resource management, and channel coding for cellular and ad-hoc networks.

• • •