Received January 21, 2018, accepted February 17, 2018, date of publication March 1, 2018, date of current version April 18, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2811040

Sum Rate Analysis and Power Allocation for Massive MIMO Systems With Mismatch Channel

XINSHUI WANG[1,2], YING WANG[1], WEIHENG NI[3], (Student Member, IEEE), RUIJIN SUN[1], AND SACHULA MENG[1]

[1]State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]Institute of Information Science and Engineering, Qufu Normal University, Rizhao 276826, China
[3]Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92037,USA

Corresponding author: Ying Wang (wangying@bupt.edu.cn)

This work was supported by the National Nature Science Foundation of China under Grant 61431003 and Grant 61421061.

ABSTRACT Massive multiple-input multiple-output (MIMO) has been regarded as one of the key technologies of fifth-generation cellular systems due to its excellent performance in spectral and energy efficiency, whose performance has also been widely studied. However, most related works focus only on the impact of the wireless channels. In fact, its performance is affected not only by the wireless channel but also by the transceiver radio frequency (RF) circuits. Random variation of RF gain would lead to a mismatch channel, where the downlink is not the transpose of the uplink in time-division duplex (TDD) mode. Therefore, the impact of the transceiver RF circuits should be considered when we evaluate the performance of the massive MIMO systems. In this paper, we develop a detailed analysis on the downlink sum rate of the massive MIMO systems and derive its closed-form expressions using maximum ratio transmission and zero-forcing (ZF) precoding. The derived results provide some good insight into how the system performance is affected by the RF mismatch parameters. Based on the analytical results, we further investigate the optimal power allocation scheme to maximize the sum rate subject to the total power constraint and lowest rate requirement. For the simplest case of user equipment side mismatch with the ZF precoding, we apply the water-filling solution, while for the other mismatch cases, we conduct a convex relaxation on the non-convex problem through lower bound inequality, variable substitution, and Taylor expansion techniques, before applying some convex optimization solving tools. In the end, we propose an iterative algorithm to successively improve the iterative results for approaching the optimal solution. Simulations demonstrate that, for the massive MIMO systems with RF mismatch, our power allocation schemes achieve significant capacity improvement relative to an equal power scheme, and it performs well for the ZF precoding in the case of the RF mismatch only at the base station.

INDEX TERMS MIMO communication, performance analysis, optimization, MRT/ZF precoding.

I. INTRODUCTION

Recently, massive Multiple-Input Multiple-Output (MIMO) has been widely studied since it can bring substantial improvements in spectral and energy efficiency. Hence, it has been identified as one of the most promising techniques for fifth generation (5G) wireless communication networks [1]–[4]. In massive MIMO systems, the effects of uncorrelated noise and small-scale fading will be eliminated when the number of base station (BS) antennas goes to infinity [1]. Furthermore, with the perfect knowledge of the downlink channel state information (CSI), BSs can use the simple linear precoding/detection approaches, e.g., matched-filter (MF) precoding/detection—to obtain the optimal capacity [2], [3].

The sum rate analysis on different transmission/reception schemes is one of the hottest topics in massive MIMO systems [5]–[12]. By assuming imperfect CSI and adopting conjugate and generalized zero-forcing (ZF) precoding, the system performance is examined in [5] and [6]. The impact of pilot contamination on sum rate is discussed in [7]. Besides, there are numerous results about the bounds on the sum rate of massive MIMO systems, e.g., the upper bound based on maximum-ratio combination (MRC) receivers in [8], the upper and lower bounds based on ZF receivers in [9]. Furthermore, [10] proposes lower bounds on the sum rate with MRC, ZF and minimum mean square error (MMSE) receivers in the cases of perfect CSI and imperfect CSI. Most of the above-mentioned analytical

results are based on the assumption of Rayleigh channels. In order to make study results more practical, researchers start to consider various real-life channel models. By using the same receivers as in [10], the asymptotic sum rate under Rayleigh correlated channel is investigated through random matrix theory in [11]. Assuming Ricean channels, the works in [10] are extended in [13] for providing a more general analysis.

However, all the aforementioned studies take no account of the transceiver radio frequency (RF) circuit gains at both ends of the communication link. In fact, in practical systems, the end-to-end communication channel consists of not only the wireless propagation part but also the transceiver RF circuits [14]. In general, the transceiver RF circuits comprise filters, analog-to-digital (A/D) converters, power amplifiers, etc. Its gain is superimposed to the wireless propagation channel through a complex magnification coefficient [15] and usually changes randomly since it heavily depends on the external environment — e.g., temperature, humidity. It would be impossible for downlink and uplink RF circuits to have the same change; hence, the reciprocity property of the whole communication channel for time-division duplex (TDD) systems is destroyed. Such RF mismatch will severely degrade the various performances of the system including sum rate [16]. Therefore, reciprocity calibration should be performed to compensate RF mismatch for TDD systems. There is a large body of literature investigating the impact of RF mismatch [16]–[19] so that some highly efficient and practical calibration methods can likely be found. Total-least-squares-based calibration has been proposed for MIMO systems [17]. Nevertheless, the calibration information is required to be exchanged between the transmitters and the receivers for that method, which results in heavy feedback overhead and prohibits the use of massive antenna arrays. In [18], a hardware calibration method performed at BS is proposed from the perspective of dual-layer beamforming, while the authors in [19] propose that user equipment (UE) side hardware calibration can also be performed to enhance downlink transmission. However, [16] notes that RF mismatch at the BS side will have a greater impact on system performance than that at the UE side.

Among most prior works in this area, the impact of RF mismatch on system performance has been evaluated only through simulations. The theoretical analysis on the performance of massive MIMO systems with RF mismatch is presented in [20] and [21]. Such analysis is pursued in [20] under regularized ZF and MRT precoding schemes, where the RF circuit gains are treated as known constants during the data transmission interval concerned. As RF gains vary over time, to obtain the average sum rate, it also needs take expectation over time aside from channel realizations. [21] does not consider large-scale fading effects, thus limiting the generality of the analysis.

Building on the prior work of [20] and [21], we pursue a detailed analysis with MRT and ZF precoding considering large-scale fading and RF mismatch at both the BS side and the UE side. Our derived results are insightful in terms of characterizing the impact of the RF mismatch parameters and path-loss on the ergodic sum rate. In addition, we note that most studies on power allocation for massive MIMO systems assume an ideal match channel (i.e., excluding the RF circuit) [22]–[24]. For example, [24] proposes the corresponding uplink and downlink power allocation schemes for massive MIMO systems with ZF detection/precoding by assuming an ideal Rayleigh channel. To our best knowledge, there is no any power allocation scheme proposed for the system considering a non-ideal channel.

For this reason, other than our analytical results, we also investigate the optimal power allocation scheme to maximize the system sum rate subject to the total power constraint and lowest rate requirement. For the case with mismatch only at the UE side (thereafter simply termed UE side mismatch; likewise, the case with mismatch only at the BS side is termed BS side mismatch) with ZF precoding, the optimal scheme is a water-filling solution, while the formulated problem for the other mismatch cases with ZF and MRT precoding is non-convex and intractable. To solve the original problem efficiently, we relax it to an equivalent form based on virtue of lower bound inequality. Using appropriate variable substitution and the first-order Taylor expansion, the equivalent problem can then be transformed into a standard convex optimization problem, which can be solved efficiently by off-the-shelf solvers, e.g., CVX [25]. Since we are optimizing the lower bound on the sum rate, we propose an iterative algorithm that can improve these lower bounds successively and finally approach the optimal solution to the original problem by using the successive convex approximation method (SCAM) [26]. Simulations show that our proposed power allocation schemes, combined with existing hardware calibration methods, can further eliminate the impact of RF mismatch on system performance.

To summarize, the key contributions of this paper are now listed as follows.

- For massive MIMO systems with an RF mismatch channel, we first provide an analytical result on the sum rate with MRT precoding. We also extend the results in [21] considering large-scale fading effects.
- Based on the derived analytical results, we further investigate the optimal power allocation. For the simplest case of UE side mismatch with ZF precoding, we derive the water-filling solution, whereas for the cases of mismatch at both the BS and UE sides with MRT precoding as well as BS side mismatch with ZF precoding, we transform the non-convex problem into a standard convex one by a series of procedures, including lower bound inequality, variable substitution and Taylor expansion, then apply an iterative algorithm to solve it using a constrained concave convex procedure (CCCP).
- Simulations are conducted with different system configurations to verify the accuracy of the analytical results and the effectiveness of our proposed power allocation scheme.

The remainder of the paper is organized as follows. In Section II, we describe the system model. Then, the analytical results on sum rate for mismatch channel with MRT and ZF precodings are developed in Section III. In Section IV, based on the results in Section III, we present the power allocation schemes to maximize system sum rate. Numerical simulation results and discussions are provided in Section V. Finally, the main results of the paper are summarized in Section VI.

Notations-Throughout the paper, boldface lower-case and upper-case letters denote vectors and matrices, respectively. $[\,\cdot\,]_{m,n}$ and $\mathrm{Tr}(\cdot)$ denote the (m,n)th element and the trace of a matrix, respectively, $\mathbf{I}_M$ represents an $M \times M$ identity matrix and $\mathbf{e}_i$ is the ith column vector of $\mathbf{I}_M$. Moreover, the operators $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote the transpose, conjugate and conjugate transpose, respectively. $\|\cdot\|$ represents the Euclidean norm of a vector, while $\|\cdot\|_F$ represents the Frobenius norm of a matrix, $|\cdot|$ refers to the modulus of the operand. $E(x)$ represents the expectation of the random variable x, and $x \sim \mathcal{CN}(\mu, \delta^2)$ denotes that random variable x is circularly symmetric complex Gaussian distribution with mean μ and variance δ^2. The set $\mathcal{K}$ is defined as $\mathcal{K} \triangleq \{1, 2, \cdots, K\}$.

II. SYSTEM MODEL

A. AMPLITUDE AND PHASE MISMATCH OF TRANSCEIVER RF CIRCUITS

We consider a downlink single-cell massive MIMO system operating in TDD mode, in which the M-antenna BS simultaneously serves K single-antenna UEs and usually $M \gg K$ is satisfied. The whole communication channel of the considered system comprises not only the wireless propagation channel, but also the transceiver RF circuits at both ends of the radio link [14], [21]. Each antenna at BS or UE has an independent transceiver RF module, which imposes a random change complex magnification coefficient on the transmission signal. For notational convenience, we place these magnification coefficients into diagonal positions of a diagonal matrix. Let the diagonal elements of $\mathbf{B_t}$ and $\mathbf{B_r}$ denote the coefficients of RF gain in the transmission and reception at the BS side, respectively. Likely, the diagonal elements of $\mathbf{U_t}$ and $\mathbf{U_r}$ denote the coefficients of RF gain in the transmission and reception at the UE side, respectively. All these diagonal matrices are defined as

$$\mathbf{B}_r = \mathrm{diag}\left\{b_{r,1}, b_{r,2}, \cdots, b_{r,M}\right\}, \quad (1)$$
$$\mathbf{B}_t = \mathrm{diag}\left\{b_{t,1}, b_{t,2}, \cdots, b_{t,M}\right\}, \quad (2)$$
$$\mathbf{U}_r = \mathrm{diag}\left\{u_{r,1}, u_{r,1}, \cdots, u_{r,K}\right\}, \quad (3)$$
$$\mathbf{U}_t = \mathrm{diag}\left\{u_{t,1}, u_{t,2}, \cdots, u_{t,M}\right\}, \quad (4)$$

where $b_{r,m}, b_{t,m}\,(m = 1, 2, \cdots, M)$ and $u_{r,k}, u_{t,k}\ (k = 1, 2, \cdots, K)$ are the complex RF circuit gains, which can be written as $b_{r,m} = |b_{r,m}|\, e^{i\varphi^b_{r,m}}$, $b_{t,m} = |b_{t,m}|\, e^{i\varphi^b_{t,m}}$, $u_{r,k} = |u_{r,k}|\, e^{i\varphi^u_{r,k}}$, and $u_{t,k} = |u_{t,k}|\, e^{i\varphi^u_{t,k}}$. The amplitude of the RF gain is assumed to be of log-normal distribution [16], [27], denoted as

$$\ln|b_{r,m}| \sim \mathcal{N}\left(0, \delta^2_{b,r}\right), \forall m, \quad (5)$$

where $\delta^2_{b,r}$ is the mismatch variance of the RF gain in the reception of BS, and the subscripts b, r in $\delta^2_{b,r}$ represent the base reception side. Similarly, we have $\ln|b_{t,m}| \sim \mathcal{N}\left(0, \delta^2_{b,t}\right)$, $\ln|u_{r,k}| \sim \mathcal{N}\left(0, \delta^2_{u,r}\right)$ and $\ln|u_{t,k}| \sim \mathcal{N}\left(0, \delta^2_{u,t}\right)$. The phase of the RF gain is assumed to be of uniform distribution [16], [27], which is denoted as

$$\varphi^b_{r,m} \sim \mathrm{U}\left[-\theta_{b,r}, \theta_{b,r}\right], \forall m, \quad (6)$$

$\theta_{b,r}$ is the maximal mismatch phase of RF gain in the reception of BS. Similarly, we have $\varphi^b_{t,m} \sim \mathrm{U}\left[-\theta_{b,t}, \theta_{b,t}\right]$, $\varphi^u_{t,k} \sim \mathrm{U}\left[-\theta_{u,t}, \theta_{u,t}\right]$, and $\varphi^u_{r,k} \sim \mathrm{U}\left[-\theta_{u,r}, \theta_{u,r}\right]$.

B. CHANNEL MODEL

As mentioned above, we consider a more practical MIMO channel consisting of the path-loss effects depending only on the distance from BS to UEs and the transceiver RF circuit gains. The whole uplink and downlink channels can then be expressed as

$$\mathbf{G}_U = \mathbf{B}_r \tilde{\mathbf{H}}^T \mathbf{D}^{\frac{1}{2}} \mathbf{U}_t, \quad (7)$$
$$\mathbf{G}_D = \mathbf{U}_r \mathbf{D}^{\frac{1}{2}} \tilde{\mathbf{H}} \mathbf{B}_t, \quad (8)$$

where $\tilde{\mathbf{H}} = \left[\underline{\tilde{\mathbf{h}}}_1^T, \underline{\tilde{\mathbf{h}}}_2^T, \cdots, \underline{\tilde{\mathbf{h}}}_K^T\right]^T \in \mathbb{C}^{K\times M}$ contains the independent and identically distributed (i.i.d.) complex Gaussian entries with zero mean and unit variance, $\underline{\tilde{\mathbf{h}}}_k = \left[\tilde{h}_{k,1}, \tilde{h}_{k,2}, \cdots \tilde{h}_{k,M}\right]$ denotes the k-th row entries of $\tilde{\mathbf{H}}$, and the diagonal entries of diagonal matrix $\mathbf{D} = \mathrm{diag}\{\beta_1, \beta_2, \cdots, \beta_K\}$ represent path-loss coefficients. Notice that $\mathbf{D}$, $\mathbf{U}_t$ and $\mathbf{U}_r$ are all diagonal matrices, and (7) and (8) can be written as $\mathbf{G}_U = \mathbf{H}_U \mathbf{D}^{\frac{1}{2}}$ and $\mathbf{G}_D = \mathbf{D}^{\frac{1}{2}} \mathbf{H}_D$, respectively, where $\mathbf{H}_U \triangleq \mathbf{B}_r \tilde{\mathbf{H}}^T \mathbf{U}_t$ and $\mathbf{H}_D \triangleq \mathbf{U}_r \tilde{\mathbf{H}} \mathbf{B}_t$. From (7) and (8), it can be readily observed that the uplink link channel is non-reciprocal with the corresponding downlink channel due to the RF mismatch, i.e., $\mathbf{G}_D \neq \mathbf{G}_U^T$

III. ANALYSIS ON SUM RATE FOR MASSIVE MIMO SYSTEMS WITH MISMATCH CHANNEL

In this section, by exploiting the two schemes of MRT and ZF precoding we provide the results of the sum rate for massive MIMO systems with a mismatch channel. It is known that the CSI of uplink channel is required at the BS for precoding in the downlink. We assume that BS has a perfect uplink CSI, which is a reasonable approximation in environments with low noise or moderate mobility [24], and has also been widely used in the literature [9], [21], [24]. It follows that the corresponding downlink channel by transposing the uplink channel since the BS operates in TDD mode. After precoding, the output vector $\mathbf{x} = [x_1, x_2, \cdots, x_K]^T$ at the BS is given by

$$\mathbf{x} = \alpha \mathbf{W} \mathbf{P}^{\frac{1}{2}} \mathbf{s}, \quad (9)$$

where $\mathbf{W}$ is the precoding matrix exhibiting different forms according to the MRT/ZF schemes, $\mathbf{P}$ is a diagonal matrix with the diagonal entries $[\mathbf{P}]_{k,k} = p_k$ denoting the transmission power allocated to the k-th user, $\mathbf{s} = [s_1, s_2, \cdots, s_K]^T$ is the signal vector transmitted to all the K UEs satisfying $E\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_K$, and α is the scaling factor to satisfy the transmission power constraint $E\left(\|\mathbf{x}\|^2\right) = p$, p is the total transmission power of the BS, i.e., $p \triangleq \sum_{k=1}^{K} p_k$. From the property of norm $\|\mathbf{A}\mathbf{x}\|^2 \le \|\mathbf{A}\|_F^2 \|\mathbf{x}\|^2$, we have $\|\mathbf{x}\|^2 = \alpha^2 \left\|\mathbf{W}\mathbf{P}^{\frac{1}{2}}\mathbf{s}\right\|^2 \le \alpha^2 \|\mathbf{W}\|_F^2 \left\|\mathbf{P}^{\frac{1}{2}}\mathbf{s}\right\|^2$. Therefore, the equality $\alpha^2 E\left(\|\mathbf{W}\|_F^2\right) = 1$ will meet the above power constraint, that is

$$\alpha = \frac{1}{\sqrt{E\left\{\mathrm{Tr}\left(\mathbf{W}\mathbf{W}^H\right)\right\}}}. \tag{10}$$

A. ANALYSIS OF SUM RATE WITH MRT SCHEME

With the ideal uplink channel estimation (i.e., (7) can be accurately estimated), the precoding matrix $\mathbf{W}$ for the MRT scheme can be given as $\mathbf{W} = \left(\mathbf{G}_U^T\right)^H = \mathbf{H}_U^* \mathbf{D}^{\frac{1}{2}}$. Substituting this result into (10), we have

$$\alpha = \frac{1}{\sqrt{E\left[\mathrm{Tr}\left(\tilde{\mathbf{H}}^H E\left(\mathbf{U}_t \mathbf{U}_t^*\right) \mathbf{D} \tilde{\mathbf{H}} E\left(\mathbf{B}_r \mathbf{B}_r^*\right)\right)\right]}}, \tag{11}$$

$$= \frac{1}{\sqrt{e^{2\delta_{u,t}^2 + 2\delta_{b,r}^2} \mathrm{Tr}\left[E\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right)\mathbf{D}\right]}}, \tag{12}$$

$$= \frac{1}{\sqrt{M e^{2\delta_{u,t}^2 + 2\delta_{b,r}^2} \sum_{k'=1}^{K} \beta_{k'}}}. \tag{13}$$

In (11) we use the property of trace operator $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$ and the commutative property of diagonal matrices $\mathbf{D}\mathbf{U}_t = \mathbf{U}_t\mathbf{D}$ as well as the independence between $\mathbf{U}_t$ and $\mathbf{B}_r$, while (12) results from the n-th moment property of the log-normal distribution ($\ln x \sim \mathcal{N}\left(\mu, \delta^2\right)$, $x > 0$), i.e.,

$$E\left(x^n\right) = \exp\left(n\mu + \frac{n^2\sigma^2}{2}\right), n = \cdots -1, 0, 1, 2, \cdots, \tag{14}$$

which means $E\left(\left|u_{t,k}\right|^2\right) = e^{2\delta_{u,t}^2}$ and $E\left(\left|b_{r,k}\right|^2\right) = e^{2\delta_{b,r}^2}$. And (13) is obtained from the property of the Wishart distribution, which is, if $\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H \sim \mathrm{W}_K(M, \mathbf{I}_K)$ is a $K \times K$ complex Wishart central matrix with M ($M > K$) degrees of freedom, $E\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right) = M\mathbf{I}_K$ [28].

Defined $\mathbf{y} = [y_1, y_2, \cdots, y_K]^T$ as the signal vector received by K UEs in the downlink, then

$$\begin{aligned}\mathbf{y} &= \mathbf{G}_D\mathbf{x} + \mathbf{n} \\ &= \alpha \mathbf{D}^{\frac{1}{2}} \mathbf{U}_r \tilde{\mathbf{H}} \mathbf{B}_t \mathbf{B}_r^* \tilde{\mathbf{H}}^H \mathbf{U}_t^* \mathbf{D}^{\frac{1}{2}} \mathbf{P}^{\frac{1}{2}} \mathbf{s} + \mathbf{n},\end{aligned} \tag{15}$$

where $\mathbf{n}$ is a $K \times 1$ noise vector and has the i.i.d. normalized entries distributed according to $\mathcal{CN}(0, 1)$.

The received signal of the k-th UE can be denoted as

$$\begin{aligned}y_k = {} & \alpha \sqrt{p_k} \beta_k u_{r,k} u_{t,k}^* \underline{\tilde{\mathbf{h}}}_k \mathbf{B}_t \mathbf{B}_r^* \underline{\tilde{\mathbf{h}}}_k^H s_k \\ & + \alpha \sqrt{\beta_k} \sum_{j=1, j\neq k}^{K} \sqrt{\beta_j p_j} u_{r,k} u_{t,j}^* \underline{\tilde{\mathbf{h}}}_k \mathbf{B}_t \mathbf{B}_r^* \underline{\tilde{\mathbf{h}}}_j^H s_j + n_k.\end{aligned} \tag{16}$$

The received signal-to-interference-plus-noise ratio (SINR) of the k-th UE for MRT precoding, denoted by γ_k^{MRT}, is given by

$$\gamma_k^{\mathrm{MRT}} = \frac{p_k \beta_k^2 \left|u_{r,k}\right|^2 \left|u_{t,k}\right|^2 \left|\underline{\tilde{\mathbf{h}}}_k \mathbf{B}_t \mathbf{B}_r^* \underline{\tilde{\mathbf{h}}}_k^H\right|^2}{\beta_k \sum_{j=1, j\neq k}^{K} \beta_j p_j \left|u_{r,k}\right|^2 \left|u_{t,j}\right|^2 \left|\underline{\tilde{\mathbf{h}}}_k \mathbf{B}_t \mathbf{B}_r^* \underline{\tilde{\mathbf{h}}}_j^H\right|^2 + \alpha^{-2}}. \tag{17}$$

The ergodic sum rate for the considered system is then given by

$$R^{\mathrm{MRT}} = E\left[\sum_{k=1}^{K} \log_2\left(1 + \gamma_k^{\mathrm{MRT}}\right)\right]. \tag{18}$$

The expectation operation is taken over all the channel realizations of $\tilde{\mathbf{H}}$ as well as mismatch variables at UE side and BS side. To evaluate the ergodic sum rate, we apply [13, Lemma 1] to (18), which can be translated as $E\left[\log_2\left(1 + \frac{x}{y}\right)\right] \approx \log_2\left(1 + \frac{E(x)}{E(y)}\right)$, where $x = \sum_{i=1}^{t_1} x_i$ and $y = \sum_{i=1}^{t_2} y_i$ are sums of nonnegative random variables x_is and y_is respectively, and x and y are not necessarily independent.

A closed-form expression to approximate the achievable downlink rate with the MRT scheme is given in Theorem 1. This result holds for any finite number of antennas and it can also reveal explicitly the impact of the path-loss and the RF mismatch variances at two ends of the link on the performance.

Theorem 1: Using MRT precoding, the achievable downlink sum rate of the K UEs for the considered massive MIMO system with the mismatch channel can be approximated as

$$\begin{aligned}R^{\mathrm{MRT}} &\approx \tilde{R}^{\mathrm{MRT}} \\ &= \sum_{k=1}^{K} \log_2\left(1 + \frac{p_k \beta_k^2 \left[3 + (M-1)\, e^{-\left(\delta_{b,r}^2 + \delta_{b,t}^2\right)}\right]}{\beta_k \sum\limits_{j=1, j\neq k}^{K} \beta_j p_j + e^{-2\left(\delta_{u,r}^2 + \delta_{b,t}^2\right)} \sum\limits_{k'=1}^{K} \beta_{k'}}\right),\end{aligned} \tag{19}$$

where the notation $\tilde{R}^{\mathrm{MRT}}$ denotes the approximate sum rate. From (19), we can see that RF mismatch phases at both the UE side and BS side has no impact on the downlink sum rate since (19) is not related to the mismatch phases, and downlink sum rate does not depend on the RF transmission mismatch variance—i.e., $\delta_{u,t}^2$ at the UE side.

Proof: See Appendix. ■

Corollary 1: For the case of RF mismatch only at the UE side—i.e., $\delta_{u,r}^2 \neq 0$, $\delta_{u,t}^2 \neq 0$, $\delta_{b,r}^2 = \delta_{b,t}^2 = 0$—the

approximate sum rate for the system is reduced to

$$\tilde{R}_{\text{Umis}}^{\text{MRT}} \approx \sum_{k=1}^{K} \log_2 \left(1 + \frac{p_k \beta_k^2 e^{2\delta_{u,r}^2} (M+2)}{\beta_k e^{2\delta_{u,r}^2} \sum\limits_{j=1, j\neq k}^{K} \beta_j p_j + \sum\limits_{k'=1}^{K} \beta_{k'}} \right). \tag{20}$$

Proof: The proof can be completed by substituting $\delta_{b,t}^2 = 0$ into (19). ■

Corollary 2: For the case of RF mismatch only at the BS side—i.e., $\delta_{b,r}^2 \neq 0$, $\delta_{b,t}^2 \neq 0$, $\delta_{u,r}^2 = \delta_{u,t}^2 = 0$—the approximate sum rate for the system is reduced to

$$\tilde{R}_{\text{Bmis}}^{\text{MRT}} \approx \sum_{k=1}^{K} \log_2 \left(1 + \frac{p_k \beta_k^2 \left[3 + (M-1) e^{-\left(\delta_{b,r}^2 + \delta_{b,t}^2\right)} \right]}{\beta_k \sum\limits_{j=1, j\neq k}^{K} \beta_j p_j + e^{-2\delta_{b,t}^2} \sum\limits_{k'=1}^{K} \beta_{k'}} \right). \tag{21}$$

Proof: The proof is straightforward by setting $\delta_{u,r}^2 = 0$ in (19). ■

B. ANALYSIS OF SUM RATE WITH ZF SCHEME

The precoding matrix for the ZF precoding scheme based on the ideal uplink channel estimation, $\mathbf{W}$ can be written as $\mathbf{W} = \left(\mathbf{G}_U^T\right)^H \left[\mathbf{G}_U^T \left(\mathbf{G}_U^T\right)^H\right]^{-1} = \mathbf{H}_U^* \left(\mathbf{H}_U^T \mathbf{H}_U^*\right)^{-1} \mathbf{D}^{-\frac{1}{2}}$. Similar to the process of solving the power constraint coefficient of the MRT scheme, substituting this result into (10) and exploiting the independence among the random variables, we have

$$\alpha = \frac{1}{\sqrt{E\left\{ \text{Tr}\left[\left(\tilde{\mathbf{H}} \mathbf{B}_r \mathbf{B}_r^* \tilde{\mathbf{H}}^H \right)^{-1} E\left[\left(\mathbf{U}_t \mathbf{U}_t^* \right)^{-1} \right] \mathbf{D}^{-1} \right] \right\}}}.$$

Note that

$$\begin{aligned} E\left[\left(\mathbf{U}_t \mathbf{U}_t^* \right)^{-1} \right] &= E\left(diag\left\{ \left|u_{t,1}\right|^{-2}, \left|u_{t,2}\right|^{-2}, \cdots, \left|u_{t,K}\right|^{-2} \right\} \right) \\ &= e^{2\delta_{u,t}^2} \mathbf{I}_K. \end{aligned} \tag{22}$$

Plugging (22) into the above equation yields

$$\alpha = \frac{1}{\sqrt{e^{2\delta_{u,t}^2} E\left\{ \text{Tr}\left[\left(\tilde{\mathbf{H}} \mathbf{B}_r \mathbf{B}_r^* \tilde{\mathbf{H}}^H \right)^{-1} \mathbf{D}^{-1} \right] \right\}}}. \tag{23}$$

Combining $\tilde{\mathbf{H}} \mathbf{B}_r \mathbf{B}_r^* \tilde{\mathbf{H}}^H = \sum_{m=1}^{M} \left|b_{r,m}\right|^2 \tilde{\mathbf{h}}_m \tilde{\mathbf{h}}_m^H$ ($\tilde{\mathbf{h}}_m$ is the m-th column vector of $\tilde{\mathbf{H}}$) with $E\left\{ \left|b_{r,m}\right|^2 \right\} = e^{\delta_{b,r}^2}, \forall m$ and using the characteristic of Wishart matrices [29], we have $\tilde{\mathbf{H}} \mathbf{B}_r \mathbf{B}_r^* \tilde{\mathbf{H}}^H \sim \text{W}_K\left(M, e^{\delta_{b,r}^2} \mathbf{I}_K \right)$, then $\left(\tilde{\mathbf{H}} \mathbf{B}_r \mathbf{B}_r^* \tilde{\mathbf{H}}^H \right)^{-1}$ has an inverse Wishart distribution $\left(\tilde{\mathbf{H}} \mathbf{B}_r \mathbf{B}_r^* \tilde{\mathbf{H}}^H \right)^{-1} \sim \text{W}_K^{-1}\left(M, e^{-\delta_{b,r}^2} \mathbf{I}_K \right)$. According to [29, eq. (39)], we obtain $E\left\{ \left(\tilde{\mathbf{H}} \mathbf{B}_r \mathbf{B}_r^* \tilde{\mathbf{H}}^H \right)^{-1} \right\} = (M-K)^{-1} e^{-\delta_{b,r}^2} \mathbf{I}_K$. Using this result, (23) can be simplified as

$$\alpha = \sqrt{\frac{M-K}{e^{2\left(\delta_{u,t}^2 - \delta_{b,r}^2\right)} \sum_{k'=1}^{K} \beta_{k'}^{-1}}}. \tag{24}$$

Similarly, the signal vector received by K UEs with ZF precoding is given by

$$\mathbf{y} = \alpha \mathbf{D}^{\frac{1}{2}} \mathbf{U}_r \mathbf{F} \mathbf{U}_t^{-1} \mathbf{D}^{-\frac{1}{2}} \mathbf{P}^{\frac{1}{2}} \mathbf{s} + \mathbf{n}, \tag{25}$$

where $\mathbf{F} \triangleq \tilde{\mathbf{H}} \mathbf{B}_t \mathbf{B}_r^* \tilde{\mathbf{H}}^H \left(\tilde{\mathbf{H}} \mathbf{B}_r \mathbf{B}_r^* \tilde{\mathbf{H}}^H \right)^{-1}$. The signal received by the k-th UE is then given by

$$\begin{aligned} y_k = \alpha \sqrt{p_k \beta_k} \frac{u_{r,k}}{u_{t,k}} [\mathbf{F}]_{k,k} s_k + \alpha \beta_k \\ \times \sum_{i=1, i\neq k}^{K} \sqrt{\frac{p_i}{\beta_i}} \frac{u_{r,k}}{u_{t,i}} [\mathbf{F}]_{k,i} s_i + n_k. \end{aligned} \tag{26}$$

Hence, the SINR for the k-th UE with ZF precoding can be written as

$$\gamma_k^{ZF} = \frac{p_k \beta_k \left|u_{r,k}\right|^2 \left|u_{t,k}\right|^{-2} \left|[\mathbf{F}]_{k,k}\right|^2}{\beta_k \sum\limits_{i=1, i\neq k}^{K} \frac{p_i}{\beta_i} \left|u_{r,k}\right|^2 \left|u_{t,i}\right|^{-2} \left|[\mathbf{F}]_{k,i}\right|^2 + \alpha^{-2}}. \tag{27}$$

The ergodic sum rate of the system is then given by

$$R^{\text{ZF}} = E\left[\sum_{k=1}^{K} \log_2\left(1 + \gamma_k^{\text{ZF}} \right) \right]. \tag{28}$$

For the general case of having RF-mismatch at both the BS and the UE sides simultaneously, the sum rate is very arduous to be obtained due to random RF circuit gains and the inverse of the ZF precoding matrix, and it has not yet been addressed in the literature. Consequently, in the following, we investigate separately the ergodic sum rate of the system with ZF precoding for analytical convenience.

1) RF MISMATCH ONLY AT UE SIDE

In this case, $\delta_{b,t}^2 = \delta_{b,r}^2 = 0$, or equivalently $\mathbf{B}_t = \mathbf{B}_r = \mathbf{I}_M$. Hence, $\mathbf{F}$ is reduced to a unit matrix $\mathbf{I}_K$, and (27) can be simplified to

$$\gamma_k^{\text{ZF}} = \frac{p_k \beta_k \alpha^2 \left|u_{r,k}\right|^2}{\left|u_{t,k}\right|^2}. \tag{29}$$

Substituting (29) into (28) and then using Jensen's inequality $E\left[\log_2\left(1+e^x\right)\right] \geq \log_2\left(1+e^{E(x)}\right)$, we can easily obtain the lower bound on sum rate $R^{\text{ZF}} \geq \underline{R}_{\text{Umis}}^{\text{ZF}}$ with

$$\underline{R}_{\text{Umis}}^{\text{ZF}} = \sum_{k=1}^{K} \log_2\left(1 + e^{\ln(p_k \beta_k \alpha^2) + 2E\left[\ln\left(\left|u_{r,k}\right|\right) - \ln\left(\left|u_{t,k}\right|\right)\right]} \right).$$

By setting $\delta_{b,r}^2 = 0$ in (24) and together with the property of log-normal distribution $\ln\left|u_{r,k}\right| \sim \mathcal{N}\left(0, \delta_{u,r}^2\right)$ and

$\ln|u_{t,k}| \sim \mathcal{N}(0, \delta_{u,t}^2)$, we obtain the following result for sum rate, which is given in the following theorem.

Theorem 2: When ZF precoding is used, the ergodic downlink sum rate of the system for RF mismatch only at the UE side can be lower bounded by

$$\underline{R}_{\text{Umis}}^{\text{ZF}} = \sum_{k=1}^{K} \log_2\left(1 + \frac{(M-K)\,p_k\beta_k}{e^{2\delta_{u,t}^2}\sum_{k'=1}^{K}\beta_{k'}^{-1}}\right). \tag{30}$$

From (30), we can observe that RF reception mismatch at the UE side has no impact on the sum rate of the system. Clearly, compared (30) with the sum rate for ideal channel $R_{\text{Ideal}}^{\text{ZF}} = \sum_{k=1}^{K}\log_2\left(1+\frac{(M-K)p_k\beta_k}{\sum_{k'=1}^{K}\beta_{k'}^{-1}}\right)$, which is obtained by setting $\delta_{u,t}^2 = 0$ in (30), RF mismatch at the UE side excluding phase will deteriorate the sum-rate performance of the system.

2) RF MISMATCH ONLY AT BS SIDE

In this case, $\delta_{u,t}^2 = \delta_{u,r}^2 = 0$. Like before, (27) and (24) can be respectively simplified to

$$\gamma_k^{\text{ZF}} = \frac{p_k\beta_k\left|[\mathbf{F}]_{k,k}\right|^2}{\beta_k\sum\limits_{i=1,i\neq k}^{K}\frac{p_i}{\beta_i}\left|[\mathbf{F}]_{k,i}\right|^2 + \alpha^{-2}} \tag{31}$$

and

$$\alpha = \sqrt{\frac{(M-K)e^{2\delta_{b,r}^2}}{\sum_{k'=1}^{K}\beta_{k'}^{-1}}}. \tag{32}$$

In the following, we represent $[\mathbf{F}]_{k,k}$ in another form for evaluation convenience.

$$\begin{aligned}[\mathbf{F}]_{k,k} &= \mathbf{e}_k^T\tilde{\mathbf{H}}\mathbf{B}_t\mathbf{B}_r^*\tilde{\mathbf{H}}^H\left(\tilde{\mathbf{H}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)^{-1}\mathbf{e}_k \\ &= \underline{\mathbf{a}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\frac{adj\left(\tilde{\mathbf{H}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)\mathbf{e}_k}{\det\left(\tilde{\mathbf{H}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)},\end{aligned} \tag{33}$$

where $\underline{\mathbf{a}} \triangleq \mathbf{e}_k^T\tilde{\mathbf{H}}\mathbf{B}_t = \left[\frac{\tilde{h}_{k,1}b_{t,1}}{b_{r,1}}, \frac{\tilde{h}_{k,2}b_{t,2}}{b_{r,2}}, \cdots, \frac{\tilde{h}_{k,M}b_{t,M}}{b_{r,M}}\right]$ and $adj(\cdot)$ denotes an adjoint matrix. Let $\mathbf{C} \triangleq \mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H = [\mathbf{c}_1, \mathbf{c}_2\cdots, \mathbf{c}_K] \in \mathbb{C}^{M\times K}$; then from (33), we have

$$adj\left(\tilde{\mathbf{H}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)\mathbf{e}_k = \left[\mathbf{e}_k^T\left((-1)^{p+q}\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)^{(pq)}\right)_{K\times K}\right]^T,$$

where $\left(\tilde{\mathbf{H}}\mathbf{C}\right)^{(pq)}$ is a $(K-1)\times(K-1)$ matrix with the same elements as $\tilde{\mathbf{H}}\mathbf{C}$ but with the p-th row and q-th column removed, $(\cdot)_{K\times K}$ denotes a $K\times K$ matrix, and $p = 1, 2, \cdots, K, q = 1, 2, \cdots, K$. Since $\mathbf{e}_k^T\mathbf{A}$ extracts the k-th row of $\mathbf{A}$, $adj\left(\tilde{\mathbf{H}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)\mathbf{e}_k$ can be rewritten as

$$\begin{aligned}&adj\left(\tilde{\mathbf{H}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)\mathbf{e}_k \\ &= \left[(-1)^{k+1}\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)^{(k1)}, (-1)^{k+2}\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)^{(k2)},\right. \\ &\qquad \left.\cdots, (-1)^{k+K}\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)^{(kK)}\right]^T.\end{aligned} \tag{34}$$

Substituting (34) into (33) gives

$$\begin{aligned}[\mathbf{F}]_{k,k} &= \frac{\sum_{j=1}^{K}(-1)^{k+j}\underline{\mathbf{a}}\mathbf{c}_j\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)^{(kj)}}{\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)} \\ &= \frac{\sum_{j=1}^{K}(-1)^{k+j}\underline{\mathbf{a}}\mathbf{c}_j\det\left(\tilde{\mathbf{H}}_{(k)}\mathbf{C}^{(j)}\right)}{\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)},\end{aligned} \tag{35}$$

where $\tilde{\mathbf{H}}_{(k)}$ and $\mathbf{C}^{(j)}$ correspond respectively to $\tilde{\mathbf{H}}$ with the k-th row removed and $\mathbf{C}$ with the j-th column removed. We further define a new matrix $\tilde{\mathbf{H}}_{k\underline{\mathbf{a}}}$, which is obtained by replacing the k-th row of $\tilde{\mathbf{H}}$ with $\underline{\mathbf{a}}$. Note that $\tilde{\mathbf{H}}_{k\underline{\mathbf{a}}}$ with the k-th row removed is the same as $\tilde{\mathbf{H}}$ with the k-th row removed—i.e., $\tilde{\mathbf{H}}_{k\underline{\mathbf{a}}(k)} = \tilde{\mathbf{H}}_{(k)}$. Consequently, (35) can be rewritten as [21]

$$\begin{aligned}[\mathbf{F}]_{k,k} &= \frac{\sum_{j=1}^{K}(-1)^{k+j}\left[\tilde{\mathbf{H}}_{k\underline{\mathbf{a}}}\mathbf{C}\right]_{k,j}\det\left(\tilde{\mathbf{H}}_{k\underline{\mathbf{a}}(k)}\mathbf{C}^{(j)}\right)}{\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)} \\ &= \frac{\det\left(\tilde{\mathbf{H}}_{k\underline{\mathbf{a}}}\mathbf{C}\right)}{\det\left(\tilde{\mathbf{H}}\mathbf{C}\right)} = \frac{\det\left(\tilde{\mathbf{H}}_{k\underline{\mathbf{a}}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)}{\det\left(\tilde{\mathbf{H}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)},\end{aligned} \tag{36}$$

where the second equality of (36) follows from the property of expanding the determinant along a row. Following the similar lines as in the derivation of (36), we can obtain [21]

$$[\mathbf{F}]_{k,i} = \frac{\det\left(\tilde{\mathbf{H}}_{i\underline{\mathbf{a}}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)}{\det\left(\tilde{\mathbf{H}}\mathbf{B}_r\mathbf{B}_r^*\tilde{\mathbf{H}}^H\right)}, \tag{37}$$

where $\tilde{\mathbf{H}}_{i\underline{\mathbf{a}}}$ is obtained by replacing the i-th row of $\tilde{\mathbf{H}}$ with $\underline{\mathbf{a}}$.

When the number of BS antennas is large enough, combining (36) with (37) and exploiting [21, Ths. 1 and 2], we can obtain the approximate rate of the system, denoted by $\tilde{R}_{\text{Bmis}}^{\text{ZF}}$, which is given in the theorem below.

Theorem 3: When the number of BS antennas is large enough and ZF precoding is used, the ergodic downlink sum rate of the K UEs for RF mismatch only at the BS side can be approximated as

$$\tilde{R}_{\text{Bmis}}^{\text{ZF}} = \sum_{k=1}^{K}\log_2\left(1 + \frac{(M-K)\,p_k\beta_k e^{\delta_{b,t}^2-\delta_{b,r}^2}\upsilon^2}{\zeta\beta_k\sum\limits_{i=1,i\neq k}^{K}p_i\beta_i^{-1} + \sum\limits_{k'=1}^{K}\beta_{k'}^{-1}}\right), \tag{38}$$

where $\upsilon \triangleq \text{sinc}(\theta_{b,t})\,\text{sinc}(\theta_{b,r})$ and $\zeta \triangleq \frac{M-K}{M}\left[e^{2\delta_{b,t}^2} + e^{2\delta_{b,r}^2} - 2e^{\frac{\delta_{b,t}^2+\delta_{b,r}^2}{2}}\upsilon\right]$. It can be seen from (38) that the performance is related not only to RF mismatch variances at the BS side but also to RF mismatch phases.

IV. POWER ALLOCATION TO MAXIMIZE THE ERGODIC SUM RATE

We note that for the mismatch channel, most related works focused on improving the system performance by a hardware calibration method. In this work, we attempt to improve the system performance from another point of view of resource allocation. Specifically, we optimize the power allocation to adapt the mismatch channel such that the performance can be further improved. Based on the derived ergodic sum rate expressions, we discuss the optimization problems in the above-mentioned cases. To approach the solution of the non-convex optimization problem, whose objective function has a difference of convex (d.c.) structure—we will make use of the following lower bound inequality [26]:

$$\log_2(1+z) \geq \xi \log_2 z + \mu, \tag{39}$$

which is tight at $z = z_0$ when the approximation constants are chosen as

$$\xi = \frac{z_0}{1+z_0} \tag{40a}$$

$$\mu = \log_2(1+z_0) - \frac{z_0}{1+z_0}\log_2 z_0. \tag{40b}$$

A. POWER ALLOCATION FOR MRT PRECODING

The optimal power allocation policy to maximize sum rate subject to power constraints at BS can be obtained by solving

$$\mathcal{P}1: \ \mathbf{p}^* = \underset{\{p_1,p_2,\cdots,p_K\}}{\arg max} \ \tilde{R}^{\text{MRT}} \tag{41a}$$

$$s.t. \ \sum_{i=1}^{K} p_i \leq p \tag{41b}$$

$$p_i \geq p_{i,\min}, \quad \forall i \in \mathcal{K}, \tag{41c}$$

where $p_{i,\min}$ is the minimal power allocated to the i-th user such that it can meet the lowest rate requirement of the i-th user. Clearly, the objective function has a d.c. structure and hence it is non-convex. It is of importance to transform this problem into a tractable form. For ease of exposition, we define $b \triangleq 3 + (M-1)\,e^{-\left(\delta_{u,r}^2+\delta_{b,t}^2\right)}$, $c \triangleq e^{-2\left(\delta_{u,r}^2+\delta_{b,t}^2\right)}\sum_{k'=1}^{K}\beta_{k'}$ and

$$z_k \triangleq \frac{bp_k\beta_k^2}{\beta_k\sum_{j=1,j\neq k}^{K}\beta_j p_j + c}. \tag{42}$$

Applying (39) and transformation $\tilde{p}_i = \ln p_i$ [30] to $\mathcal{P}1$, the original problem can be relaxed to the equivalent problem $\mathcal{P}2$ below,

$$\mathcal{P}2: \ \tilde{\mathbf{p}}^* = \underset{\{\tilde{p}_1,\tilde{p}_2,\cdots,\tilde{p}_K\}}{\arg max} \ \tilde{R}_{\text{LB}}^{\text{MRT}} \tag{43a}$$

$$s.t. \ \sum_{i=1}^{K} e^{\tilde{p}_i} \leq p \tag{43b}$$

$$e^{\tilde{p}_i} \geq p_{i,\min}, \quad \forall i \in \mathcal{K}, \tag{43c}$$

where

$$\begin{aligned}\tilde{R}_{\text{LB}}^{\text{MRT}} &= \sum_{k=1}^{K}\left[\xi_k \log_2 z_k + \mu_k\right] \\ &= \sum_{k=1}^{K}\xi_k\left[\tilde{p}_k - \log_2\left(\beta_k\sum_{j=1,j\neq k}^{K}\beta_j e^{\tilde{p}_j} + c\right)\right] \\ &\quad + \sum_{k=1}^{K}\left[\xi_k\log_2\left(b\beta_k^2\right) + \mu_k\right],\end{aligned} \tag{44}$$

ξ_k and μ_k are given by (40a) and (40b), respectively. Given ξ_k and μ_k, (44) is concave with respect to the variables $\tilde{\mathbf{p}} = [\tilde{p}_1, \tilde{p}_2, \cdots, \tilde{p}_K]$ since it is the sum of linear and concave terms within the square brackets (note that log-sum-exp is convex [**??**, Sec. 3.1.5]).

However, $\mathcal{P}2$ is still non-convex due to the constraint (43c). We use the CCCP method, which is widely adopted for solving the d.c. problem [33], to approximate constraint (43c) in the n-th iteration by its first-order Taylor expansion around the current point $\tilde{p}_i^{n-1}$ obtained from the $(n-1)$-th iteration. According to [32], using the first-order Taylor expansion of (43c), $\mathcal{P}2$ can be transformed into

$$\mathcal{P}3: \ \tilde{\mathbf{p}}^* = \underset{\{\tilde{p}_1,\tilde{p}_2,\cdots,\tilde{p}_K\}}{\arg max} \ \tilde{R}_{\text{LB}}^{\text{MRT}} \tag{45a}$$

$$s.t. \ \sum_{i=1}^{K} e^{\tilde{p}_i} \leq p \tag{45b}$$

$$e^{\tilde{p}_i^{n-1}} + e^{\tilde{p}_i^{n-1}}\left(\tilde{p}_i - \tilde{p}_i^{n-1}\right) \geq p_{i,\min}, \quad \forall i \in \mathcal{K}, \tag{45c}$$

Obviously, after the above transformation, problem $\mathcal{P}3$ becomes a standard concave maximization problem, which can be solved efficiently by off-the-shelf solvers—e.g., CVX [25]. Note that the last term in (44) is not related to $\tilde{\mathbf{p}} = [\tilde{p}_1, \tilde{p}_2, \cdots, \tilde{p}_K]$ and hence does not impact the optimization problem and can be omitted.

Once a solution $\tilde{\mathbf{p}}$ is obtained, we use the inverse transformation $p_i = e^{\tilde{p}_i}$ to achieve the solution $\mathbf{p}$ to the original problem. Since we are optimizing a lower bound on the achievable sum rate, it becomes natural to approach the optimal solution by using SCAM. Therefore, we propose the following iterative algorithm to solve this convex relaxation problem. This algorithm does not need a brute-force or heuristic search of any kind and hence has a low computational complexity.

B. POWER ALLOCATION FOR ZF PRECODING

1) POWER ALLOCATION FOR MISMATCH ONLY AT UE SIDE

In this case, similar to establishing the optimization problem of MRT precoding, according to (30), the optimization

Algorithm 1 The Iterative Algorithm to Solve $\mathcal{P}3$

1: Initialize $p_1^0, p_2^0, \cdots, p_K^0$, and compute $z_1^0, z_2^0, \cdots, z_K^0$ according to (42) and $p_1^0, p_2^0, \cdots, p_K^0$, and then compute $\xi_1^0, \xi_2^0, \cdots, \xi_K^0$ according to $\xi_k = \frac{z_k}{1+z_k}$; set error tolerance $\varepsilon = 10^{-6}$ and iteration counter $n := 1$;
2: **do**
3: Solving the transformation problem
With the initial parameters, invoke CVX to solve $\mathcal{P}3$ to obtain the solution $\tilde{p}_1^n, \tilde{p}_2^n, \cdots, \tilde{p}_K^n$;
4: Solving the original problem
According to the results in step 3 and $p_k = e^{\tilde{p}_k}$, compute the solution $p_1^n, p_2^n, \cdots, p_K^n$ to original problem $\mathcal{P}1$;
5: Update parameters
Update parameters $z_1^n, z_2^n, \cdots, z_K^n$ according to (42) and $p_1^n, p_2^n, \cdots, p_K^n$ in step 4, then update $\xi_1^n, \xi_2^n, \cdots, \xi_K^n$ according to $\xi_k = \frac{z_k}{1+z_k}$;
6: Increment n := n + 1;
7: Until convergence—i.e., $\left\|p^n - p^{n-1}\right\|^2 < \varepsilon$
8: Output the optimal results $p_1^{*,n}, p_2^{*,n}, \cdots, p_K^{*,n}$

problem can be formulated as

$$\mathcal{P}4: \ \mathbf{p}^* = \underset{\{p_1,p_2,\cdots,p_K\}}{\arg\max} \ \underline{R}_{\text{Umis}}^{\text{ZF}} \tag{46a}$$

$$s.t. \ \sum_{i=1}^{K} p_i \le p \tag{46b}$$

$$p_i \ge p_{i,\min}, \quad \forall i \in \mathcal{K}, \tag{46c}$$

The following theorem gives the solution to this problem.

Theorem 4: The water-filling solution to the downlink power allocation problem $\mathcal{P}4$ is

$$p_k^* = \left(\lambda - \eta_{\text{Umis},k}^{\text{ZF}} - p_{i,\min}\right)^+ + p_{i,\min}, \forall k \tag{47}$$

where $\eta_{\text{Umis},k}^{\text{ZF}} \triangleq \frac{e^{2\delta_{u,t}^2} \sum_{k'=1}^{K} \beta_{k'}^{-1}}{(M-K)\beta_k}$ and λ is chosen to satisfy $\sum_{k=1}^{K}\left(\lambda - \eta_{\text{Umis},k}^{\text{ZF}} - p_{i,\min}\right)^+ = p - Kp_{i,\min}$.

Proof: Let $p_i' = p_i - p_{i,\min}$ for $\forall i \in \mathcal{K}$, $\mathcal{P}4$ can be written as

$$\mathcal{P}4.1: \ \mathbf{p}'^* = \underset{\{p_1',p_2',\cdots,p_K'\}}{\arg\max} \ \underline{R}_{\text{Umis}}^{\text{ZF}} \tag{48a}$$

$$s.t. \ \sum_{i=1}^{K} p_i' \le p - Kp_{i,\min} \tag{48b}$$

$$p_i' \ge 0, \quad \forall i \in \mathcal{K}, \tag{48c}$$

where $\underline{R}_{\text{Umis}}^{\text{ZF}}$ in (48a) can be written as

$$\underline{R}_{\text{Umis}}^{\text{ZF}} = \sum_{k=1}^{K} \log_2 \frac{\eta_{\text{Umis},k}^{\text{ZF}} + p_{k,\min}}{\eta_{\text{Umis},k}^{\text{ZF}}} + \sum_{k=1}^{K} \log_2\left(1 + \frac{p_k'}{\eta_{\text{Umis},k}^{\text{ZF}} + p_{k,\min}}\right) \tag{49}$$

According to (49), the water-filling solution to $\mathcal{P}4.1$ is given by [31]

$$p_k' = \left(\lambda - \eta_{\text{Umis},k}^{\text{ZF}} - p_{k,\min}\right)^+. \tag{50}$$

The substitution of $p_k' = p_k - p_{k,\min}$ into (50) yields the desired result (47). ■

2) POWER ALLOCATION FOR MISMATCH ONLY AT BS SIDE

Likewise, according to (38) the optimization problem can be formulated as

$$\mathcal{P}5: \ \mathbf{p}^* = \underset{\{p_1,p_2,\cdots,p_K\}}{\arg\max} \ \tilde{R}_{\text{Bmis}}^{\text{ZF}} \tag{51a}$$

$$s.t. \ \sum_{i=1}^{K} p_i \le p \tag{51b}$$

$$p_i \ge p_{\min}, \quad \forall i \in \mathcal{K}, \tag{51c}$$

Since (51a) still has a d.c. structure, following the similar lines as in the derivation of $\mathcal{P}3$, $\mathcal{P}5$ can be transformed into the following problem $\mathcal{P}6$,

$$\mathcal{P}6: \ \tilde{\mathbf{p}}^* = \underset{\{\tilde{p}_1,\tilde{p}_2,\cdots,\tilde{p}_K\}}{\arg\max} \ \tilde{R}_{\text{Bmis,LB}}^{\text{ZF}} \tag{52a}$$

$$s.t. \ \sum_{i=1}^{K} e^{\tilde{p}_i} \le p \tag{52b}$$

$$e^{\tilde{p}_i^{n-1}} + e^{\tilde{p}_i^{n-1}}\left(\tilde{p}_i - \tilde{p}_i^{n-1}\right) \ge p_{i,\min}, \quad \forall i \in \mathcal{K}, \tag{52c}$$

where

$$\tilde{R}_{\text{Bmis,LB}}^{\text{ZF}} = \sum_{k=1}^{K} \xi_k \left[\tilde{p}_k - \log_2\left(\beta_k \zeta \sum_{i=1,i\neq k}^{K} e^{\tilde{p}_i}\beta_i^{-1} + \widehat{c}\right)\right] + \sum_{k=1}^{K}\left[\xi_k \log_2\left(\widehat{b}\beta_k\right) + \mu_k\right]. \tag{53}$$

The associated parameters are defined as $\widehat{b} \triangleq (M-K)e^{\delta_{b,t}^2 - \delta_{b,r}^2}\upsilon^2$, $\widehat{c} \triangleq \sum_{k'=1}^{K}\beta_{k'}^{-1}$ and

$$z_k = \frac{\widehat{b}p_k\beta_k}{\beta_k\zeta\sum_{i=1,i\neq k}^{K} p_i\beta_i^{-1} + \widehat{c}}. \tag{54}$$

Therefore, $\mathcal{P}6$ can be solved by invoking Algorithm 1 by computing z_k^n via (54).

V. NUMERICAL RESULTS

Numerical simulations are now conducted to validate the accuracy of the derived sum rate results and the effectiveness of our proposed power allocation scheme. Consider a massive MIMO system with a cell radius of 500 meters, in which the M-antenna BS is located in the center of the cell and concurrently serves $K = 5$ users in the same time-frequency resource. Without loss of generality, the large-scale shading coefficients of UEs are assumed to be $[\beta_1, \beta_2, \beta_3, \beta_4, \beta_5] = [0.1, 0.2, 0.3, 0.5, 0.8]$, and the minimal power required for the lowest rate is set to be $p_{k,min} = 0.1$ dB, $\forall k$.

First, we validate the results with MRT precoding. Fig. 1 illustrates the advantage of our proposed power allocation

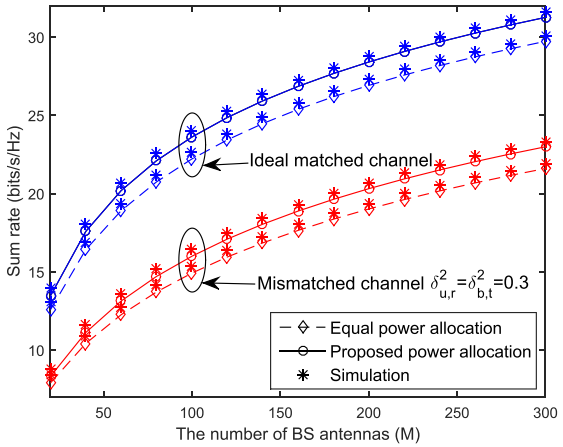


FIGURE 1. Sum rate vs. the number of antennas at BS for MRT precoding ($p = 40$dB).

scheme over the equal power allocation in the cases of the mismatch channel and of the ideal channel $\left(\delta^2_{b,t} = \delta^2_{u,r} = 0\right)$ respectively. From Fig. 1, we can observe that the analytical sum rate can match well with the Monte-Carlo result in the entire antenna range considered, and the sum rate for the case of the ideal channel outperforms that for the case of the mismatch channel in the entire antenna range, which shows that the mismatch channel will deteriorate the sum rate of the system. Regardless of whether the channel is ideal, our proposed power allocation scheme is better than the equal power scheme. In addition, the sum rate increment of our proposed scheme relative to equal power scheme becomes more obvious as the number of BS antennas grows large.

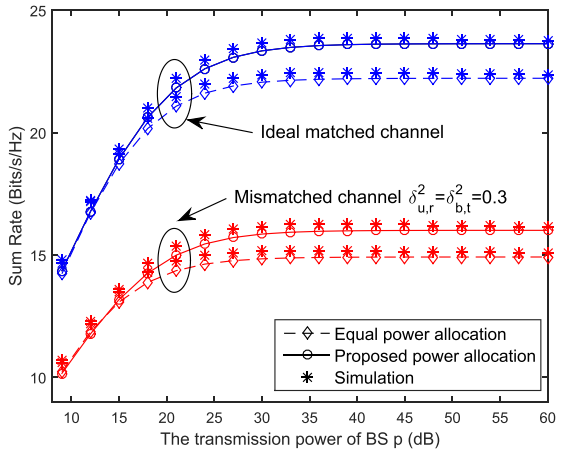


FIGURE 2. Sum rate vs. transmission power of BS for MRT precoding ($M = 100$).

In Fig. 2, we investigate how the sum rate of the system changes over the transmission power of BS with two kinds of power allocation schemes. The comparison between the mismatched channel and the ideal channel is also considered. The sum rate for the case of ideal channel still outperforms that for the case of mismatch channel in the entire transmission power range. For both cases of ideal channel and mismatch channel, our proposed scheme outperforms the equal power scheme. The advantage of our proposed scheme gradually becomes more obvious with increasing BS transmission power. However, when BS operates in the high SINR regime (i.e., $\rho > 15$ dB), the sum rate of the system will trend to convergence and the advantage will disappear. The reason is that when the transmission power becomes large enough, the mulit-user interference resulting from MRT precoding becomes large, which in turn restricts the increment of sum rate and the advantage of our proposed scheme.

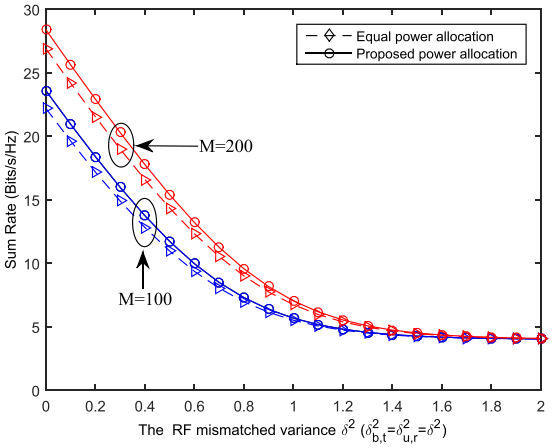


FIGURE 3. Sum rate vs. RF mismatch variance for MRT precoding ($p = 40$dB).

Fig. 3 depicts how sum rates for two different antenna configurations ($M = 100, 200$) are influenced by the RF mismatch variance. It can be observed that the sum rate will deteriorate when the mismatch variance becomes large. This has further verified that the mismatch channel will deteriorate the sum rate of the system. This is because the larger mismatch variance will bring more uncertainty to the channel, which leads to less effectiveness of MRT precoding. This ultimately results in the lower performance of the system. In addition, we can also see that the sum rate with our proposed scheme outperforms that with the equal power scheme for different RF variance in the cases of $M = 100$ and $M = 200$. However, the advantage of our proposed scheme, in terms of the sum rate increment, compared to the equal power scheme tends to vanish as the variance becomes large. Furthermore, the system with $M = 100$ antennas can achieve a very close performance to that with $M = 200$ antennas when the variance is large enough, which means the improvement on the sum rate obtained through equipping more antennas at BS will disappear when the variance is large enough. This indicates that it does not help to improve the sum rate of the system via increasing the number of antennas or optimizing users' power when the channel has high uncertainty.

Fig. 4 illustrates the impact of RF mismatch variance on the relative gain of sum rate for different antenna configurations. The relative gain of the sum rate is defined as $\frac{R^*-R}{R}$, where R^* and R denote the sum rates obtained by our proposed scheme and by the equal power scheme, respectively. It can

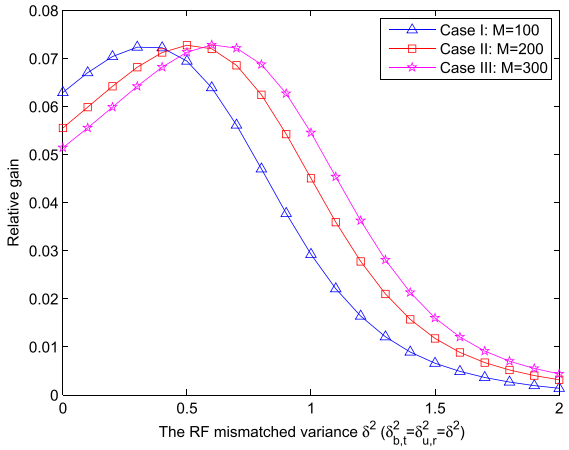


FIGURE 4. Sum rate relative gain vs. mismatch variance for MRT precoding ($p = 40$dB).

be observed from Fig. 4 that the relative gain first increases and then decreases as the RF mismatch variance becomes larger. Note that the relative gain approaches zero when the RF mismatch variance is large enough (e.g., $\delta^2 = 2$). This shows that our proposed scheme is no longer better than the equal power allocation scheme as the channel becomes highly uncertain, which means that a simple equal power allocation scheme may be a good choice in this case. Another interesting observation from Fig. 4 is that the larger the number of antennas, the greater the mismatch variance corresponding to optimal gain.

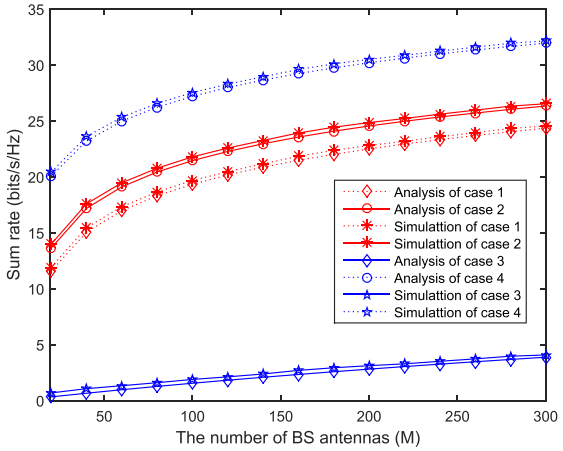


FIGURE 5. Sum rate vs. the number of antennas at BS for ZF precoding ($p = 40$dB). Case 1 and case 2 with $\delta^2_{u,r} = 1$; case 3 and case 4 with $\delta^2_{b,r} = \delta^2_{b,t} = 0.5$, $\theta_{b,r} = \theta_{b,t} = 0$.

Next, we validate the advantage of our proposed power allocation scheme with ZF precoding. For convenience of showing legends in the following figures, we define the following four cases. Cases 1-4 denote UE side mismatch and equal power allocation scheme, UE side mismatch and our proposed power allocation scheme, BS side mismatch and equal power allocation scheme, and BS mismatch and our proposed power allocation scheme, respectively. In Fig. 5, we plot the changes in the sum rate of the system with the number of antennas at BS for for the above four cases. It is observed from Fig. 5 that our proposed analytical results on sum rate can still match well with the Monte-Carlo results in the whole antenna range for all the four cases. Similar to MRT precoding, the use of large-scale antennas can also bring significant improvement in spectral efficiency for ZF precoding. Although we set $\delta^2_{b,r} = \delta^2_{b,t} = 0.5 < \delta^2_{u,r} = 1$, the sum rate obtained for BS side mismatch is much smaller than that for UE side mismatch, which has further verified the conclusion that the BS side mismatch has a more important impact on system performance than UE side mismatch. However, after allocating the BS's power to the UEs with our proposed algorithm, the sum rate for the BS side mismatch can be improved significantly.

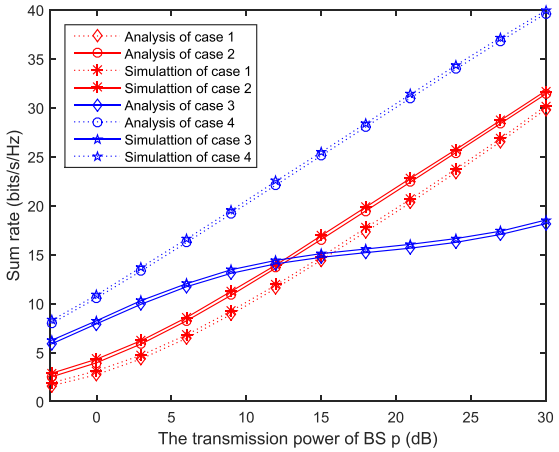


FIGURE 6. Sum rate vs. transmission power of BS for ZF precoding ($M = 100$). Case 1 and case 2 with $\delta^2_{u,r} = 1$; case 3 and case 4 with $\delta^2_{b,r} = \delta^2_{b,t} = 0.01$.

Fig. 6 shows the lines that the sum rate of the system changes with the transmission power of BS for the four cases. Similar to Fig. 5, there is a precise match between our analytical results of the sum rate and the Monte-Carlo results for all the cases. Moreover, for the cases of UE side mismatch and BS side mismatch, enhancing the transmission power of BS both can help raise the sum rate of the system. In addition, the sum rate for the case of UE side mismatch can acquire nearly linear growth as the transmission power of BS increases, while the growth of the sum rate for the case of BS side mismatch gets slower when the transmission power is high. However, by using our proposed scheme, the sum rate for the case of BS side mismatch can be greatly improved in the high transmission power regime.

Fig. 7 illustrates the impact of RF mismatch variance on sum rate. It can be seen that the sum rate for the case of UE side mismatch decreases linearly as the mismatch variance gradually grows from zero, whereas the sum rate for BS side mismatch experiences a sharp decline and rapidly converges to zero, which once more verifies the conclusion that the BS side mismatch has a greater impact on system performance than the UE side mismatch. However, our proposed power allocation scheme can prevent this degradation trend to a

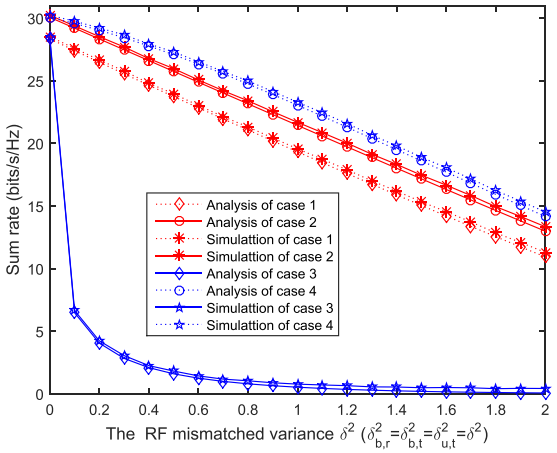


FIGURE 7. Sum rate vs. RF mismatch variance for ZF precoding ($p = 40$dB, $M = 100$, $\theta_{b,r} = \theta_{b,t} = \theta_{u,t} = 0$).

great extent. Besides, we can further see that the sum rates of the system for the two cases of BS side match and UE side match (i.e., $\delta^2_{b,r} = \delta^2_{b,t} = \delta^2_{u,t} = 0$, $\theta_{b,r} = \theta_{b,t} = \theta_{u,t} = 0$) are the same when both adopt an equal power allocation scheme. This is because the two case are essentially equivalent to each other under the same ideal channel, so the sum rates for both cases are expected to be equal.

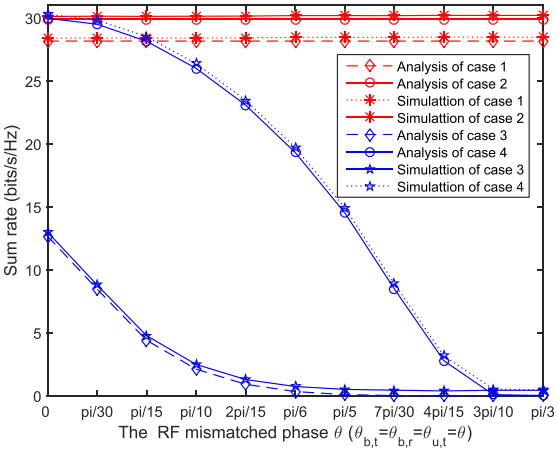


FIGURE 8. Sum rate vs. RF mismatch phase for ZF precoding ($p = 40$dB, $M = 100$, $\delta^2_{b,r} = \delta^2_{b,t} = \delta^2_{u,t} = 0.5$).

In Fig. 8, we investigate the impact of RF mismatch phase on the performance of the system. It is observed from Fig. 8 that RF mismatch phase at the BS side still has a significant impact on the sum rate of the system for fixed variances ($\delta^2_{b,r} = \delta^2_{b,t} = 0.5$). When the mismatch phase at BS increases from zero, the performance of the system will also suffer significant deterioration. However, the sum rate for the case of BS side phase mismatch is fundamentally different from the case of the UE side phase mismatch. Specifically, the sum rate holds constant for any mismatch phase at UE side since the sum rate expression (30) does not include this parameter. Therefore, it is necessary to concentrate only on performing calibration at the BS side rather than the UE side.

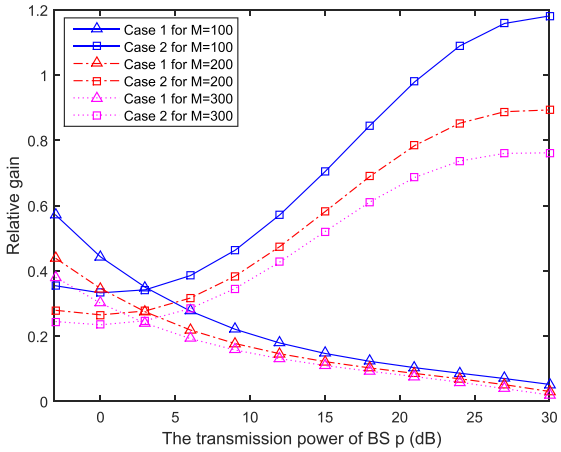


FIGURE 9. Sum rate relative gain vs. transmission power of BS ($\delta^2_{b,r} = \delta^2_{b,t} = 0.01$, $\delta^2_{u,t} = 1$, $\theta_{b,r} = \theta_{b,t} = \theta_{u,t} = 0$).

To conclude this section, in Fig. 9 we study how the relative gain of sum rate is affected by the transmission power of BS. The definition of relative gain is the same as in Fig. 4, and cases 1 and 2 in the legend of Fig. 9 represent the case of UE side mismatch and the case of BS side mismatch, respectively. As shown in Fig. 9, the relative gains under three different antenna configurations (i.e., $M = 100, 200, 300$) always drop for case 1 as the transmission power of BS increases. The main reason is that from Fig. 5 and Fig. 6, the rate difference of the two schemes does not significantly increase with increasing transmission power. However, for the sum rate in case 2, the relative gains under three configurations first decrease and then increase with increasing transmission power. This phenomenon can be explained by virtue of the lines of sum rate for BS side mismatch in Fig. 6. When increasing transmission power, the growth of sum rates of the equal power scheme shows a marked increase in the range of low power; however, it slows down in the range of high power, and at the same time, the sum rate can be greatly improved after adopting our proposed scheme. According to the definition of relative gain, it is not difficult to understand such a phenomenon.

VI. CONCLUSIONS

In this paper, we have investigated downlink massive MIMO systems considering both large-scale and small-scale fading effects as well as transceiver RF mismatch effects. Closed-form expressions on the ergodic sum rate using MRT and ZF precoding have been derived. The results can offer useful insights into the impact of the RF mismatch parameters and path-loss on the sum rate. Based on the derived results, we have further investigated the power allocation scheme to maximize the system sum rate. The water-filling solution has been given to solve the simplest case of UE side mismatch with ZF precoding, while for other cases, we have proposed an iterative algorithm to obtain near-optimal power allocation performance. Extensive simulation results have validated the accuracy of the derived analytical

results and the effectiveness of our proposed power allocation scheme. Furthermore, simulations have also demonstrated that, for ZF precoding, the case of RF mismatch only at BS side can reap significant improvement in spectrum frequency over that only at UE side when our proposed scheme is adopted.

Appendix
Proof of Theorem 1

Obviously, the expectation of the numerator and denominator of (17) must be taken when applying [31, Lemma 1] to (18). To this end, we first calculate the expectation of the numerator, denoted by Nu, in (17).

$$\begin{aligned} E\left(Nu\right) &= p_k\beta_k^2 E\left(\left|u_{r,k}\right|^2\right) E\left(\left|u_{t,k}\right|^2\right) E\left[\left|\underline{\tilde{\mathbf{h}}}_k\mathbf{B}_t\mathbf{B}_r^*\underline{\tilde{\mathbf{h}}}_k^H\right|^2\right] \\ &= p_k\beta_k^2 e^{2\delta_{u,r}^2+2\delta_{u,t}^2} E\left[\left|\underline{\tilde{\mathbf{h}}}_k\mathbf{B}_t\mathbf{B}_r^*\underline{\tilde{\mathbf{h}}}_k^H\right|^2\right], \end{aligned} \tag{55}$$

where (55) is the direct result after using (14).

The expectation term in (55) is calculated as follows.

$$\begin{aligned} &E\left[\left|\underline{\tilde{\mathbf{h}}}_k\mathbf{B}_t\mathbf{B}_r^*\underline{\tilde{\mathbf{h}}}_k^H\right|^2\right] \\ &= \sum_{m=1}^{M} E\left(\left|\tilde{h}_{k,m}\right|^4\left|b_{t,m}\right|^2\left|b_{r,m}\right|^2\right) \\ &\quad + \sum_{m=1}^{M}\sum_{m'\neq m}^{M} E\left(\left|\tilde{h}_{k,m}\right|^2\left|\tilde{h}_{k,m'}\right|^2 b_{t,m}b_{t,m'}b_{r,m}^*b_{r,m'}^*\right) \\ &= 3Me^{2\delta_{b,t}^2+2\delta_{b,r}^2} + M\left(M-1\right)e^{\delta_{b,t}^2+\delta_{b,r}^2} \\ &= Me^{\delta_{b,t}^2+\delta_{b,r}^2}\left(3e^{\delta_{b,t}^2+\delta_{b,r}^2}+M-1\right). \end{aligned} \tag{56}$$

Note the independence among the random variables; the second equality in (56) results from (14) and the fourth moment property of normal distribution $E\left(x^4\right)=\mu^4+6\sigma^2\mu^2+3\sigma^4$ if $x \sim N\left(\mu,\delta^2\right)$ as well as the properties $E\left(b_{t,m}\right) = E\left(\left|b_{t,m}\right|\right)e^{iE\left(\varphi_{b,m}^t\right)} = e^{\frac{1}{2}\delta_{b,t}^2}$, $E\left(b_{t,m}\right) = E\left(b_{t,m'}\right) = e^{\frac{1}{2}\delta_{b,t}^2}$, $E\left(b_{r,m}^*\right) = E\left(b_{r,m'}^*\right) = e^{\frac{1}{2}\delta_{b,r}^2}$. Plugging (56) into (55) follows as

$$E\left(Nu\right) = Mp_k\beta_k^2 e^{2\delta_{u,r}^2+2\delta_{u,t}^2+\delta_{b,r}^2+\delta_{b,t}^2}\left(3e^{\delta_{b,r}^2+\delta_{b,t}^2}+M-1\right). \tag{57}$$

Next, we begin to compute the expectation of the denominator in (17). Let F denote the first term of the denominator; the expectation of F is derived as follows:

$$\begin{aligned} E(F) &= \beta_k\sum_{j=1,j\neq k}^{K}\beta_j p_j E\left(\left|u_{r,k}\right|^2\left|u_{t,j}\right|^2\left|\underline{\tilde{\mathbf{h}}}_k\mathbf{B}_t\mathbf{B}_r^*\underline{\tilde{\mathbf{h}}}_j^H\right|^2\right), \\ &= \beta_k e^{2\delta_{u,r}^2+2\delta_{u,t}^2}\sum_{j=1,j\neq k}^{K}\beta_j p_j\left[\sum_{m=1}^{M} E\left(\left|\tilde{h}_{k,m}b_{t,m}b_{r,m}^*\tilde{h}_{j,m}^*\right|^2\right)\right. \\ &\quad \left. + \sum_{m=1}^{M}\sum_{m'=m}^{M} E\left(\tilde{h}_{k,m}b_{t,m}b_{r,m}^*\tilde{h}_{j,m}^*\tilde{h}_{k,m'}b_{t,m}b_{r,m}^*\tilde{h}_{j,m'}^*\right)\right] \\ &= \beta_k e^{2\delta_{u,r}^2+2\delta_{u,t}^2}\sum_{j=1,j\neq k}^{K}\beta_j p_j\sum_{m=1}^{M} E\left(\left|\tilde{h}_{k,m}b_{t,m}b_{r,m}^*\tilde{h}_{j,m}^*\right|^2\right) \\ &= \beta_k e^{2\delta_{u,r}^2+2\delta_{u,t}^2}\sum_{j=1,j\neq k}^{K}\beta_j p_j\sum_{m=1}^{M} E\left(\left|b_{t,m}\right|^2\right)E\left(\left|b_{r,m}^*\right|^2\right), \end{aligned} \tag{58}$$

where the third equality in (58) is obtained from the independence among random variables and the zero mean property $E\left(\tilde{h}_{k,m}\right) = E\left(\tilde{h}_{j,m}^*\right) = 0$, whereas the last equality in (58) follows from $E\left(\left|\tilde{h}_{k,m}\right|^2\right) = E\left(\left|\tilde{h}_{j,m}^*\right|^2\right) = 1$. After using the property of the log-normal distribution, (58) can be simplified as

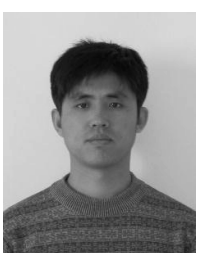
$$E\left(F\right) = M\beta_k e^{2\left(\delta_{u,r}^2+\delta_{u,t}^2+\delta_{b,r}^2+\delta_{b,t}^2\right)}\sum_{j=1,j\neq k}^{K}\beta_j p_j. \tag{59}$$

Substituting (57), (59) and (13) into (18) and applying some algebraic manipulations, we can obtain the desired result (19).

REFERENCES

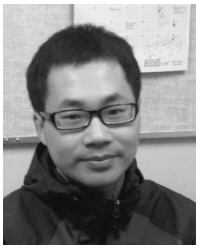
[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
[2] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–46, Jan. 2013.
[3] 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.
[4] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[5] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.
[6] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, "Channel estimation and linear precoding in multiuser multiple-antenna TDD systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2102–2116, Jun. 2011.
[7] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
[8] X. Wang, Y. Wang, and S. Ma, "Upper bound on uplink sum rate for large-scale multiuser MIMO systems with MRC receivers," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2154–2157, Dec. 2015.
[9] M. Matthaiou, C. Zhong, M. R. McKay, and T. Ratnarajah, "Sum rate analysis of ZF receivers in distributed MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 180–191, Feb. 2013.
[10] 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.
[11] 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.
[12] X. Wang, Y. Wang, and R. Sun, "Approximate sum rate for massive multiple-input multiple-output two-way relay with Ricean fading," *IET Commun.*, vol. 10, no. 12, pp. 1493–1500, 2016.
[13] 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.
[14] A. Bourdoux, B. Come, and N. Khaled, "Non-reciprocal transceivers in OFDM/SDMA systems: Impact and mitigation," in *Proc. IEEE Radio Wireless Conf.*, Boston, MA, USA, Aug 2003, pp. 183–186.
[15] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems," in *Proc. IEEE Future Netw. Mobile Summit*, Florence, Italy, Jun. 2010, pp. 1–10.

[16] F. Huang, Y. Wang, J. Geng, and D. Yang, "Antenna mismatch and calibration problem in coordinated multi-point transmission system," *IET Commun.*, vol. 6, no. 3, pp. 289–299, 2012.
[17] H. Wei, D. Wang, and X. You, "ICD reciprocity calibration for distributed large-scale MIMO systems with BD precoding," in *Proc. IEEE/CIC Int. Conf. Commun. China—Workshops (CIC/ICCC)*, Shenzhen, China, Nov. 2015, pp. 132–136.
[18] *Hardware Calibration Requirement for Dual Layer Beamforming*, document R1-092359, 3GPP TSG RAN WG1 Meeting 57, Huawei, Los Angeles, CA, USA, 2009.
[19] *On the Need for UE Calibration for Enhanced Downlink Transmission*, document R1-092016, 3GPP TSG RAN WG1 Meeting 57, Ericsson, San Francisco, CA, USA, 2009.
[20] W. Zhang *et al.*, "Large-scale antenna systems with UL/DL hardware mismatch: Achievable rates analysis and calibration," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1216–1229, Apr. 2015.
[21] H. Wei, D. Wang, J. Wang, and X. You, "Impact of RF mismatches on the performance of massive MIMO systems with ZF precoding," *Sci. China Inf. Sci.*, vol. 59, no. 2, pp. 1–14, 2016.
[22] 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 Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 5184–5190.
[23] J. Choi, "Massive MIMO with joint power control," *IEEE Wireless Commun. Lett.*, vol. 3, no. 4, pp. 329–332, Aug. 2014.
[24] Q. Zhang, S. Jin, M. 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.
[25] M. Grant and S. Boyd. (Aug. 2012). *CVX: MATLAB Software for Disciplined Convex Programming, Version 1.22*. [Online]. Available: http://cvxr.com/cvx
[26] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009.
[27] J. Geng *et al.*, "Antenna gain mismatch calibration for cooperative base stations," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, San Francisco, CA, USA, Sep. 2011, pp. 1–5.
[28] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Breda, The Netherlands: Now Publishers, 2004.
[29] D. Maiwald and D. Kraus, "Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices," *IEE Proc.-Radar Sonar Navigat.*, vol. 147, no. 4, pp. 162–168, 2000.
[30] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.
[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[32] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*. Palo Alto, CA, USA: Meboo Publishing, 2010.
[33] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.



XINSHUI WANG received the B.S. and M.S. degrees in communication and information system from Lanzhou University, China, in 2001 and 2004, respectively. He is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunication. From 2004 to 2013, he was a Lecturer with the Institute of Information Science and Engineering, Qufu Normal University. His research interests include massive MIMO and energy-efficient communications.

YING WANG received the Ph.D. degree in circuits and systems from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2003. In 2004, she was a Visiting Researcher with the Communications Research Laboratory (renamed NICT from 2004), Yokosuka, Japan. She was a Research Associate with The University of Hong Kong, Hong Kong, in 2005. She is currently a Professor with BUPT and the Director of the Radio Resource Management Laboratory, Wireless Technology Innovation Institute, BUPT. She has authored over 100 papers in international journals and conferences proceedings. Her research interests include cooperative and cognitive systems, radio resource management, and mobility management in 5G systems. She is active in standardization activities of 3GPP and ITU. She took part in performance evaluation work of the Chinese Evaluation Group, as a Representative of BUPT. She was a recipient of first prizes of the Scientific and Technological Progress Award by the China Institute of Communications in 2006 and 2009, respectively, and a second prize of the National Scientific and Technological Progress Award in 2008. She was also selected in the New Star Program of the Beijing Science and Technology Committee and the New Century Excellent Talents in University, Ministry of Education, in 2007 and 2009, respectively.



WEIHENG NI (S'14) received the B.E. degree in communication engineering from the Beijing University of Posts and Telecommunications, China, in 2012, and the M.A.Sc degree in electrical and computer engineering from the University of Victoria, Canada, in 2015. He is currently pursuing the Ph.D. degree with the University of California at San Diego, USA. From 2015 to 2016, he was with Fortinet Technologies, Burnaby, BC, Canada, where he was involved in the development of the FortiCloud system. His research interests include cognitive radio, energy harvesting, massive multiple antenna systems, millimeter-wave techniques, and some machine learning-based designs.

RUIJIN SUN received the B.S. degree in communications engineering from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2013. She is currently pursuing the Ph.D. degree with the Wireless Technology Innovation Institute, Beijing University of Posts and Telecommunications, China. Her research interests include MIMO, cooperative communications, and energy harvesting communications in future wireless networks.

SACHULA MENG received the M.S. degree in electronics engineering from Inner Mongolia University, China, in 2014. She is currently pursuing the Ph.D. degree in telecommunication and information system with the Wireless Technology Innovation Institute, Beijing University of Posts and Telecommunications. Her research interests include mobile cloud computing and resource management in wireless networks.