

Large-Scale Antenna Systems with UL/DL Hardware Mismatch: Achievable Rates Analysis and Calibration

Wence Zhang, *Student Member, IEEE*, Hong Ren, Cunhua Pan *Student Member, IEEE*, Ming Chen, *Member, IEEE* and Rodrigo C. de Lamare, *Senior Member, IEEE*, Bo Du *Student Member, IEEE*, Jianxin Dai, *Member, IEEE*

Abstract— This paper studies the impact of hardware mismatch (HM) between the base station (BS) and the user equipment (UE) in the downlink (DL) of large-scale antenna systems. Analytical expressions to predict the achievable rates are derived for different precoding methods, i.e., matched filter (MF) and regularized zero-forcing (RZF), using large system analysis techniques. Furthermore, the upper bounds on achievable rates of MF and RZF with HM are investigated, which are related to the statistics of the circuit gains of the mismatched hardware. Moreover, we present a study of HM calibration, where we take zero-forcing (ZF) precoding as an example to compare two HM calibration schemes, i.e., Pre-precoding Calibration (Pre-Cal) and Post-precoding Calibration (Post-Cal). The analysis shows that Pre-Cal outperforms Post-Cal schemes. Monte-Carlo simulations are carried out, and numerical results demonstrate the correctness of the analysis.

Index Terms—Large-scale antenna systems, Massive MIMO, hardware mismatch (HM), channel calibration, channel reciprocity.

I. INTRODUCTION

Multiple-Input-Multiple-Output (MIMO) technique plays a key role in modern communications systems to increase the data rates and link reliability [1]. Large system analysis have long been employed to analyze the performance of MIMO systems, e.g., [2]–[5], but the idea of ‘Massive MIMO systems’ or ‘large-scale antenna systems’ with a large number of antennas deployed at the base station (BS) begins to draw more and more research interest recently [6]–[8]. Operating in time-division duplexing (TDD) mode, large-scale antenna systems utilize the channel reciprocity in the uplink (UL) and the downlink (DL), and obtain the channel state information (CSI) through UL pilot training. With an increasing number

of antennas, large-scale antenna systems can exploit excess degrees of freedom. As the degree of orthogonality between channels of different users improves, inter-user interference can be suppressed by simple processing [9]. Large-scale antenna systems exhibits advantages in many aspects, such as data rates, symbol error rates, computational complexity of signal processing and energy efficiency [7], [10], [11].

In frequency-division duplexing (FDD) systems, the DL channel coefficients are estimated at the user equipment (UE) and then sent back to the BS. So it brings heavy training overhead when the number of transmit and receive antennas is large. Although several papers have tried to address this problem, e.g., [12] and references therein, it remains a big challenge for FDD systems with a large number of BS antennas. Therefore, the benefits brought by large-scale antenna systems should be based on channel reciprocity in TDD mode, which considers the channel coefficients to remain the same in the UL and the DL within the channel coherent time. This is generally true for wireless propagation [13], [14], but not exactly when considering the hardware employed in the front-ends of both the BS and the UEs. Some research points out that the I/Q imbalance of the front-end analog processing has negative effect on the channel reciprocity [15]. In fact, many approaches have been proposed to compensate such I/Q imbalance in the digital field, e.g. [16] and references therein. Besides I/Q imbalance, the most important impact comes from the ‘hardware mismatch’ (HM) between the UL and the DL, i.e., the base station (BS) and the UEs employ different devices (e.g., antennas, filters and amplifiers), and consequently the equivalent channel coefficients that take into account the analog circuit gains differ in the UL and the DL. Although HM generally does not affect the performance in the UL (the channel estimation is carried out in the UL), it results in an essential constraint of the DL performance [11].

HM generates ‘uncertainty’ in the DL. One possible way to deal with this uncertainty is through robust design [17] at the cost of extra implementation complexity. However, further research have advocated estimation and calibration of these HM parameters [14], [18]–[25]. The estimation methods of HM parameters can be divided into two categories, according to whether the users are involved or not. When the HM estimation is carried out at the BS without involving a user, the circuit gains of all the BS antennas are adjusted to a reference antenna by either resorting to additional circuitry in the transceivers [18], or by treating an antenna as the

W. Zhang, H. Ren, C. Pan, M. Chen and B. Du are with National Mobile Communications Research Lab. (NCRL), Southeast University, China. Emails: {wencezhang, renhong, cunhuapan, chenming, dubo}@seu.edu.cn.

R. C. de Lamare is with Dept. of Electronics, the University of York, UK and CETUC, PUC-Rio, Brazil. Email:rcdl500@york.ac.uk.

J. Dai is with School of Science, Nanjing University of Posts and Telecommunications, China. Email:daijx@njupt.edu.cn.

The work of W. Zhang, H. Ren, C. Pan, M. Chen and B. Du was supported in part by China Scholarship Council (CSC), National Nature Science Foundation of China (No. 61372106 & 61172077 & 61223001), National 863 High Technology Development Project (No. 2014AA01A701), Scientific Research Foundation of Graduate School of Southeast University (No. YBJJ1433). The work of R. C. de Lamare was supported by CNPq, FAPERJ, PUC-Rio and the University of York. The work of J. Dai was supported by Natural Science Foundation of Nanjing University of Posts and Telecommunications (No. NY2014090).

Part of this work is accepted for publication in IEEE GlobeCom 2014.

reference and exchanging pilot signals between the reference antenna and all the other antennas at the BS [21], [22], [24]. This scheme is usually termed as ‘self calibration’ [14], or ‘relative calibration’ [21]. In this way, the circuit gains of all the antennas at the BS are calibrated to a constant value and will not affect the the channel reciprocity. The self calibration scheme can also be used to users with multiple antennas, but the performance degrades due to the low cost amplifiers used in users [26]. Another approach to obtain the HM parameters demands the assistance from the users and is often termed as ‘over-the-air calibration’ because pilot signals are transmitted between the BS and the users through the air interface [14], [19], [20], [23], [25]. The main limitations of this approach are the occupation of system resources and the selection of supporting users [20].

Once the HM parameters are reliably obtained, channel calibration is carried out to retrieve reciprocity between the UL and the DL. Some works proposed to only calibrate the HM parameters at the BS (we refer to these schemes as ‘partial calibration’), because they have the dominant impact on the level of inter-user interference in the DL [18], [20]–[22]. However, the quality of service (QoS) of each user is difficult to guarantee without knowledge of the user’s HM parameters. This problem can be handled either by employing additional filters at the users [18], or by HM calibration [14], [19], [23], [25]. We refer to these calibration approaches as ‘full calibration’.

As most research work on HM calibration, to the best of the authors’ knowledge, there are few results of the HM’s impact on the DL performance except for [26], in which the problem of different HM calibration schemes is studied for Block-Diagonalization (BD) and Zero-Forcing (ZF) precoding. However, few analytical results can be found and the impact of HM on DL performance remains unclear for large-scale antenna systems.

In this paper, the impact of HM on different DL precoding methods, i.e., matched filter (MF) and regularized zero-forcing (RZF) is investigated. MF is a potential technique for DL transmission in large-scale antenna systems due to its simplicity, low complexity and scalability [6], [7], [21]. In contrast, RZF has higher computational complexity and inferior scalability, but it offers much higher data rates. We give the analytical expression of the achievable rates for both MF and RZF. Prior related works on large system analysis of MIMO systems, e.g., [2]–[5] and references therein, always assume the same antenna correlation profile for all the users. Therefore, the combined DL channel matrix of all the users can be represented by a matrix with independent entries multiplied by two correlation matrices on both sides. In this work, however, we model the channel in a more generalized way in which different users can have different channel correlation matrices, as in [27], [28]. Unlike [27], [28], which analyze the DL performance of MF and RZF in large-scale antenna systems without considering HM between the UL and the DL, in this work we take HM into consideration, and the problem becomes very different especially for RZF. The correlated wireless channel matrix is further corrupted by two diagonal matrices comprising the HM parameters of the system, which

makes the problem more complicated. By utilizing random matrix theory and some useful lemmas, we derive analytical expressions for the achievable rates of MF and RZF with HM in the UL and the DL for large-scale antenna systems.

We also present a study of HM calibration using the same downlink scenario. As briefly described in [22], the HM calibration can be carried out either before or after the precoding. These two scenarios are referred to as Pre-precoding Calibration (Pre-Cal) and Post-precoding Calibration (Post-Cal). Because zero-forcing (ZF) precoding has a simpler expression than RZF and shows more insights on HM’s impact than MF, in this work we take ZF as an example to compare the performance of Pre-Cal and Post-Cal under both partial calibration and full calibration.

The main contributions of this paper are summarized as follows:

- a) Analytical expressions for the achievable rates per user are given for RZF and extended to MF with consideration of UL and DL HM, through the evaluation of expectations and asymptotic deterministic equivalents (ADE) of a series of random variables;
- b) We prove that there are upper bounds for achievable rates of MF and RZF in the high signal to noise (SNR) region for a special case. The upper bounds are only related to the statistics of the HM parameters and the ratio of the number of users and the number of transmit antennas.
- c) We also develop a study of HM calibration and show that the Pre-Cal schemes outperform the Post-Cal schemes.

The rest of the paper is organized as follows. The system model is described in section II. In section III, the set of achievable rates for both RZF is derived and then extended to MF. The study of different HM calibration schemes is carried out in section IV. The simulations and numerical results are shown in section V, and conclusions are drawn in section VI.

Notation: $(\cdot)^*$ is the complex conjugate operation; $(\cdot)^T$ and $(\cdot)^H$ denotes the transpose and Hermitian transpose of a vector or a matrix, respectively; $\mathbb{E}\{\cdot\}$ denotes the expectation operation; $[v_{nk}]_{N \times K}$ denotes an N by K matrix with the (n, k) -th element being v_{nk} ; \mathbf{I}_N denotes an N by N identity matrix; $\mathcal{CN}(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ denotes circularly symmetric complex Gaussian distribution with mean $\boldsymbol{\theta}$ and covariance $\boldsymbol{\Sigma}$; $\text{diag}\{a_1, \dots, a_N\}$ denotes an N by N diagonal matrix with diagonal entries given by a_1, \dots, a_N .

II. SYSTEM MODEL

Consider the DL of a single cell multi-user large-scale antenna system with N transmit antennas at BS and K single antenna users, where N is very large (tens or hundreds). The system operates in TDD mode, so that the channel reciprocity in the UL and the DL can be exploited. An illustration of the system is given in Fig. 1.

Let $\mathbf{h}_k = [h_{1k}, \dots, h_{Nk}]^T$, $\mathbf{h}'_k = [h'_{1k}, \dots, h'_{Nk}]^T$ be the UL and DL channel vectors of user k , respectively, where h_{nk} is the channel gain from user k to the n -th antenna at BS in the UL, and h'_{nk} is that in the DL. As shown in Fig. 2, h_{nk} and its DL counterpart h'_{nk} are modeled, respectively, as [21],

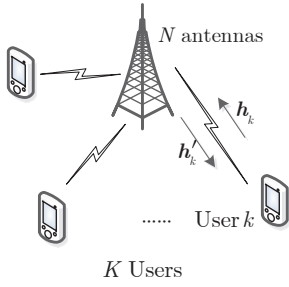


Fig. 1. Single cell MU Massive MIMO systems

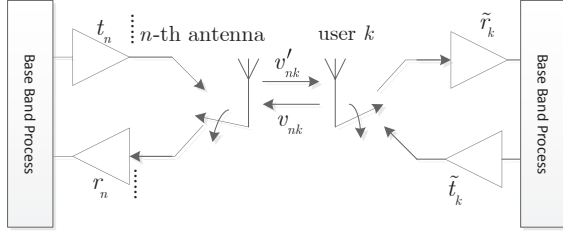


Fig. 2. Hardware Mismatch of UL and DL transmissions

[22]

$$\begin{aligned} h_{nk} &= r_n v_{nk} \tilde{t}_k, \\ h'_{nk} &= t_n v'_{nk} \tilde{r}_k, \end{aligned} \quad (1)$$

where v_{nk} and v'_{nk} are the corresponding UL and DL wireless channel gains, respectively; r_n , t_n are the equivalent receive and transmit circuit gains of the n -th antenna at BS, and \tilde{r}_k, \tilde{t}_k are the equivalent receive and transmit circuit gains of user k , respectively.

In fact, h_k and h'_k are not exactly the same due to HM between UL and DL. In TDD mode, because the electromagnetic waves in UL and DL undergo the same physical propagation environment, i.e., reflections, refractions, scatterers, etc [14], the wireless channel gains in UL and DL transmission are considered to be unchanged within a channel coherent period, i.e., $v_{nk} = v'_{nk}$. However, the circuit gains are related to the hardware configurations, e.g., transmit and receive filters and amplifiers. Because the BS and the UEs usually have different hardware implementation in the UL and DL, so that in general $r_i \neq t_i$, $\tilde{r}_i \neq \tilde{t}_i$, $r_i \neq \tilde{r}_i$ and $t_i \neq \tilde{t}_i$. These circuit gains may change with the working conditions, e.g., temperature [14]. But they vary slowly compared with the wireless channel, so it is reasonable to treat them as known constants during the data transmission interval concerned.

In the DL, the received signal y_k at user k is given by

$$y_k = \mathbf{x} \sqrt{\mathbf{P}\mathbf{W}} \mathbf{h}'_k + n_k, \quad \forall k = 1, \dots, K, \quad (2)$$

where $\mathbf{x} = [x_1, \dots, x_K]$ is the transmit signal vector consisting of independent, zero mean and unit energy symbols for respective users, so that $\mathbb{E}\{\mathbf{x}^H \mathbf{x}\} = \mathbf{I}_K$; \mathbf{P} is the power loading matrix, $\mathbf{P} = \text{diag}\{p_1, \dots, p_K\}$ with p_k being the power loading factor for user k ; $\mathbf{W} \in \mathbb{C}^{K \times N}$ is the transmit precoding matrix; n_k is the Gaussian noise at the receiver of user k , i.e., $n_k \sim \mathcal{CN}(0, \sigma_n^2)$.

Let the UL combined channel matrix of all users be $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K]$, and $\mathbf{H} = \mathbf{R}\mathbf{V}\tilde{\mathbf{T}}$, in which $\mathbf{R} =$

$\text{diag}(r_1, \dots, r_N)$ is the receive circuit gain matrix of BS, $\tilde{\mathbf{T}} = \text{diag}(\tilde{t}_1, \dots, \tilde{t}_K)$ is the transmit circuit gain matrix of all the users, and $\mathbf{V} = [v_{nk}]_{N \times K}$ is the wireless channel coefficient matrix. Denote the k -th column of \mathbf{V} as \mathbf{v}_k , and \mathbf{v}_k 's are independent complex Gaussian vectors with correlated entries, i.e., $\mathbf{v}_k \sim \mathcal{CN}(\mathbf{0}, \frac{1}{N} \Phi_k)$, $\forall k = 1, \dots, K$. Note that the channel gain is normalized to $\frac{1}{N}$ for analysis convenience.

The DL channel matrix $\tilde{\mathbf{H}} = [\mathbf{h}'_1, \dots, \mathbf{h}'_K]$ can be represented as $\tilde{\mathbf{H}} = \mathbf{T}\mathbf{V}\tilde{\mathbf{R}}$, where $\mathbf{T} = \text{diag}(t_1, \dots, t_N)$ is the transmit circuit gain matrix of BS, and $\tilde{\mathbf{R}} = \text{diag}(\tilde{r}_1, \dots, \tilde{r}_K)$ is the receive circuit gain matrix of all the users.

Let $\mathbf{y} = [y_1, \dots, y_K]$, $\mathbf{n} = [n_1, \dots, n_K]$, and (2) can be reformulated as

$$\mathbf{y} = \mathbf{x} \sqrt{\mathbf{P}\mathbf{W}} \tilde{\mathbf{H}} + \mathbf{n}. \quad (3)$$

Remark: The wireless channel covariance matrix of user k is $\frac{1}{N} \Phi_k$. This model is very general and encompasses a wide range of scenarios. For example, when $\Phi_k = \mathbf{I}_N$, it means no correlation among the transmit antennas and the path loss of user k reduces to 1. When distributed antennas are considered at BS, the covariance matrix of user k would be of form $\Phi_k = \text{diag}\{a_1, a_2, \dots, a_N\}$, where a_n represents the path loss from user k to the n -th distributed transmit antenna.

III. ACHIEVABLE RATES OF LINEAR PRECODING SCHEMES WITH HM

In this section, we first give an introduction to a set of achievable rates which are upper bounded by the capacity and are very useful for performance analysis [28], [29]. Then we study the effect of HM on DL performance in terms of these achievable rates, and derive analytical expressions for both MF and RZF.

A. The Set of Achievable Rates

Similar to [2]–[5], [28], [29], in order to study the performance of large-scale antenna systems with HM, we first introduce a set of achievable rates as shown in Lemma 1.

Lemma 1: Consider a point-to-point communication link, the signal model of which is given by

$$\begin{aligned} y_k &= g_{k,k} \sqrt{p_k} x_k + \sum_{i=1, i \neq k}^K g_{k,i} \sqrt{p_i} x_i + n_k \\ &= \mathbb{E}\{g_{k,k}\} \sqrt{p_k} x_k + (g_{k,k} - \mathbb{E}\{g_{k,k}\}) \sqrt{p_k} x_k \\ &\quad + \sum_{i=1, i \neq k}^K g_{k,i} \sqrt{p_i} x_i + n_k, \end{aligned} \quad (4)$$

where $k = 1, \dots, K$, $\mathbf{x} = [x_1, \dots, x_K]$ and n_k are as described in (2); $g_{k,i}$ is the equivalent channel gain, and $g_{k,i}$'s are correlated. Then the following set of rates is achievable

$$R_k = \log_2(1 + \gamma_k), \quad \forall k = 1, \dots, K, \quad (5)$$

where

$$\gamma_k = \frac{p_k |\mathbb{E}\{g_{k,k}\}|^2}{\sigma_n^2 + p_k \text{var}\{g_{k,k}\} + \sum_{i=1, i \neq k}^K p_i \mathbb{E}\{|g_{k,i}|^2\}}. \quad (6)$$

Proof: See Theorem 1 in [29]. ■

By assuming that the users only have the mathematical expectations of the equivalent channel gains, and treating the variations of signals and all the correlated interferences as independent noise, this set of achievable rates for each user are actually upper bounded by the capacity, and thus *achievable*. For more study on this set of achievable rates, please see [29].

In order to derive the closed-form expressions of the achievable rates of MF and RZF, we first formulate the signal model as in (4), and then evaluate the mathematical expectations involved in (6).

B. Achievable Rates of Linear Precoding Schemes

In this subsection, we derive the achievable rates for RZF and MF. Since RZF has a generalized expression and MF can be treated as special cases of it, the achievable rates of RZF is first derived and then the results are extended to MF, which has a much simpler expression. Because the expression of achievable rates of ZF is similar to that of RZF and it can not be simplified, we omitted it here.

Using RZF for DL transmission, the precoding matrix is given by $\mathbf{W}_{\text{RZF}} = \sqrt{\lambda_{\text{RZF}}}[\mathbf{H}^H\mathbf{H} + \rho\mathbf{I}_K]^{-1}\mathbf{H}^H$, where $\rho = \frac{K\sigma_n^2}{P_T}$ for the minimum mean square error (MMSE) design [30], and

$$\lambda_{\text{RZF}} = \frac{P_T}{\mathbb{E}\text{Tr}\left\{\mathbf{P}[\mathbf{H}^H\mathbf{H} + \rho\mathbf{I}_K]^{-1}\mathbf{H}^H\mathbf{H}[\mathbf{H}^H\mathbf{H} + \rho\mathbf{I}_K]^{-1}\right\}}. \quad (7)$$

where P_T is the total transmit power.

The received signal vector $\mathbf{y} \in \mathbb{C}^{1 \times K}$ for the K users is

$$\begin{aligned} \mathbf{y} &= \sqrt{\lambda_{\text{RZF}}}\mathbf{x}\sqrt{\mathbf{P}}[\mathbf{H}^H\mathbf{H} + \rho\mathbf{I}_K]^{-1}\mathbf{H}^H\tilde{\mathbf{H}} + \mathbf{n} \\ &= \sqrt{\lambda_{\text{RZF}}}\mathbf{x}\sqrt{\mathbf{P}}\mathbf{H}^H[\mathbf{H}\mathbf{H}^H + \rho\mathbf{I}_N]^{-1}\tilde{\mathbf{H}} + \mathbf{n} \\ &= \sqrt{\lambda_{\text{RZF}}}\mathbf{x}\sqrt{\mathbf{P}}\tilde{\mathbf{T}}^H\mathbf{V}^H\mathbf{R}^H\mathbf{Q}\mathbf{T}\mathbf{V}\tilde{\mathbf{R}} + \mathbf{n}, \end{aligned} \quad (8)$$

where $\mathbf{Q} = [\mathbf{R}\mathbf{V}\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H\mathbf{V}^H\mathbf{R}^H + \rho\mathbf{I}_N]^{-1}$.

By simple algebraic manipulation, the received signal for user k is given by

$$\begin{aligned} y_k &= \sqrt{\lambda_{\text{RZF}}}\tilde{r}_k \sum_{i=1}^K \sqrt{p_i}x_i\tilde{t}_i^* \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q} \mathbf{T} \mathbf{v}_k + n_k \\ &= g_{k,k}\sqrt{p_k}x_k + \sum_{i=1, i \neq k}^K g_{k,i}\sqrt{p_i}x_i + n_k, \end{aligned} \quad (9)$$

where $g_{k,i} = \sqrt{\lambda_{\text{RZF}}}\tilde{r}_k\tilde{t}_i^* \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q} \mathbf{T} \mathbf{v}_k$. The achievable rates with RZF are obtained using Lemma 1.

To deal with the mathematical expectations in (6), we derive asymptotic deterministic equivalents (ADE) of random variables therein. When N and K are large, the random variables in (6) will converge to their ADEs, so the expectations can be approximated by their corresponding ADEs. For example, if a random variable $x \rightarrow x_0$ and x_0 is deterministic, then x_0 is the ADE of x . So we can easily prove that $\mathbb{E}\{x\} \rightarrow x_0$ and $\text{var}\{x\} \rightarrow 0$.

1) ADE of $g_{k,k}$ and $g_{k,i}$ for $i \neq k$:

The following proposition gives the ADE of $g_{k,k}$ and $g_{k,i}$ for $i \neq k$.

Proposition 1: When $N \rightarrow \infty$, $g_{k,k} \rightarrow g_{k,k}^\circ$ and $|g_{k,i}|^2 \rightarrow g_{k,i}^{\circ 2}$, which are given by

$$\begin{aligned} g_{k,k}^\circ &= \frac{\sqrt{\lambda_{\text{RZF}}}\tilde{r}_k\tilde{t}_k^*\mu_2^\circ}{1 + |\tilde{t}_k|^2\mu_1^\circ}, \\ g_{k,i}^{\circ 2} &= \lambda_{\text{RZF}}|\tilde{r}_k|^2|\tilde{t}_i|^2 \\ &\times \left(\mu_3^\circ - \frac{|\tilde{t}_k|^2 2\text{Re}(\mu_2^\circ\mu_5^\circ)}{1 + |\tilde{t}_k|^2\mu_1^\circ} + \frac{|\tilde{t}_k|^4\mu_4^\circ|\mu_2^\circ|^2}{(1 + |\tilde{t}_k|^2\mu_1^\circ)^2} \right), i \neq k, \end{aligned} \quad (12)$$

in which

$$\begin{aligned} \mu_1^\circ &= \frac{1}{N}\text{Tr}(\mathbf{R}\Phi_k\mathbf{R}^H\mathbf{S}_k(\rho)), \\ \mu_2^\circ &= \frac{1}{N}\text{Tr}(\mathbf{T}\Phi_k\mathbf{R}^H\mathbf{S}_k(\rho)), \\ \mu_3^\circ &= \frac{1}{N} \frac{\frac{1}{N}\text{Tr}\{\mathbf{T}\Phi_k\mathbf{T}^H\mathbf{S}'_{ki}(\rho)\}}{\left(1 + \frac{1}{N}|\tilde{t}_i|^2\text{Tr}\{\mathbf{R}\Phi_i\mathbf{R}^H\mathbf{S}_{ki}(\rho)\}\right)^2}, \\ \mu_4^\circ &= \frac{1}{N} \frac{\frac{1}{N}\text{Tr}\{\mathbf{R}\Phi_k\mathbf{R}^H\mathbf{S}'_{ki}(\rho)\}}{\left(1 + \frac{1}{N}|\tilde{t}_i|^2\text{Tr}\{\mathbf{R}\Phi_i\mathbf{R}^H\mathbf{S}_{ki}(\rho)\}\right)^2}, \\ \mu_5^\circ &= \frac{1}{N} \frac{\frac{1}{N}\text{Tr}\{\mathbf{R}\Phi_k\mathbf{T}^H\mathbf{S}'_{ki}(\rho)\}}{\left(1 + \frac{1}{N}|\tilde{t}_i|^2\text{Tr}\{\mathbf{R}\Phi_i\mathbf{R}^H\mathbf{S}_{ki}(\rho)\}\right)^2}, \end{aligned} \quad (13)$$

where $\mathbf{S}_k(\rho)$ is given by, respectively

$$\mathbf{S}_k(\rho) = \left[\frac{1}{N} \sum_{i \in \Psi_k} \left\{ \frac{|\tilde{t}_i|^2 \mathbf{R}\Phi_i\mathbf{R}^H}{1 + e_i(\rho)} \right\} + \rho\mathbf{I}_N \right]^{-1}, \quad (14)$$

in which $\Psi_k = \{1, \dots, K\} \setminus \{k\}$ and $e_i(\rho)$ is the unique set of solutions to $K - 1$ equations given by (10) in the top of the next page; $\mathbf{S}_{ki}(\rho)$ is similar to $\mathbf{S}_k(\rho)$ with $\Psi_{ki} = \{1, \dots, K\} \setminus \{k, i\}$, and $\mathbf{S}'_{ki}(\rho)$ is calculated according to (A.1) with $\mathbf{B} = \mathbf{R}\Phi_i\mathbf{R}^H$.

Proof: See Appendix B. ■

Note that μ_3° , μ_4° and μ_5° are all $\mathcal{O}(N^{-1})$, and so is $|g_{k,i}|^2$. However, the interference term, i.e., the summation of $p_i|g_{k,i}|^2$ may not be $\mathcal{O}(N^{-1})$ when K is of the same order as N , especially in the high SNR region when ρ is small and the interference term dominates over the receiving noise. In the low SNR region, however, this interference term is much smaller than the noise and thus can be omitted. We just leave this term here in order to derive the achievable rates of MF from RZF.

2) Derivation of λ_{RZF} :

λ_{RZF} is a function of the combined channel matrix $\mathbf{H} = \mathbf{R}\mathbf{V}\tilde{\mathbf{T}}$. The wireless channel matrix \mathbf{V} therein has independent columns, while the entries in each column are correlated. We will derive the ADE of λ_{RZF} .

Proposition 2: When $N \rightarrow \infty$, $\lambda_{\text{RZF}} \rightarrow \lambda_{\text{RZF}}^\circ$, which is given by

$$\lambda_{\text{RZF}}^\circ = P_T \left[\sum_{k=1}^K \frac{\frac{1}{N}p_k|\tilde{t}_k|^2\text{Tr}\{\mathbf{R}\Phi_k\mathbf{R}^H\mathbf{S}'_k(\rho)\}}{\left(1 + \frac{1}{N}|\tilde{t}_k|^2\text{Tr}\{\mathbf{R}\Phi_k\mathbf{R}^H\mathbf{S}_k(\rho)\}\right)^2} \right]^{-1}, \quad (15)$$

$$e_i(\rho) = \frac{1}{N} \text{Tr} \left\{ |\tilde{t}_i|^2 \mathbf{R} \Phi_i \mathbf{R}^H \left[\frac{1}{N} \sum_{j \in \Psi_1} \left\{ \frac{|\tilde{t}_j|^2 \mathbf{R} \Phi_j \mathbf{R}^H}{1 + e_j(\rho)} \right\} + \rho \mathbf{I}_N \right]^{-1} \right\}, \quad i \in \Psi_1. \quad (10)$$

$$\gamma_{\text{MF},k} \rightarrow \frac{p_k |\tilde{r}_k \tilde{t}_k|^2 \left| \frac{1}{N} \text{Tr} \{ \mathbf{T} \Phi_k \mathbf{R}^H \} \right|^2}{\frac{\sigma_n^2}{P_T} \sum_{i=1}^K p_i |\tilde{t}_i|^2 \frac{1}{N} \text{Tr} \{ \mathbf{R} \Phi_i \mathbf{R}^H \} + \frac{1}{N} \sum_{i=1, i \neq k}^K p_i |\tilde{r}_k \tilde{t}_i|^2 \frac{1}{N} \text{Tr} \{ \mathbf{T} \Phi_k \mathbf{T}^H \mathbf{R} \Phi_i \mathbf{R}^H \}}. \quad (11)$$

in which $\mathbf{S}'_k(\rho)$ is calculated according to (A.1) with $\mathbf{B} = \mathbf{I}_N$ and $\mathbf{S}_k(\rho)$ given by (14).

Proof: See Appendix C. ■

Until this point, we have obtained all the ingredients required to derive the achievable rates of RZF. Combining the results of (12) and (15), we give the following theorem.

Theorem 1: When $N \rightarrow \infty$, the set of achievable rates with RZF is given by

$$R_{\text{RZF},k} = \log_2(1 + \gamma_{\text{RZF},k}), \quad \forall k = 1, \dots, K, \quad (16)$$

where

$$\gamma_{\text{RZF},k} \rightarrow \frac{p_k |g_{k,k}^\circ|^2}{\sigma_n^2 + \sum_{i=1, i \neq k}^K p_i g_{k,i}^{\circ 2}}. \quad (17)$$

Remark: MF and ZF can be treated as special cases of RZF. By setting $\rho \rightarrow 0$ in (17), one can obtain the achievable rates of ZF. However, the expression is quite similar to that of RZF and can not be simplified. Therefore, we omit the derivation for ZF, and only give the analysis of MF in the next subsection.

3) Achievable Rates of MF:

When MF is used for DL transmission and assuming perfect CSI at BS, the precoding matrix is given by

$$\mathbf{W}_{\text{MF}} = \sqrt{\lambda_{\text{MF}}} \mathbf{H}^H, \quad (18)$$

in which

$$\lambda_{\text{MF}} = \frac{P_T}{\mathbb{E} \text{Tr} \{ \mathbf{P} \mathbf{H} \mathbf{H}^H \}}. \quad (19)$$

The achievable rates of MF is obtained by setting $\rho \rightarrow \infty$ in Theorem 1, and is summarized in Corollary 1.

Corollary 1: When $N \rightarrow \infty$, the set of achievable rates with MF is $R_{\text{MF},k} = \log_2(1 + \gamma_{\text{MF},k})$, $\forall k = 1, \dots, K$, where $\gamma_{\text{MF},k}$ is given by (11) in the top of this page.

Proof: This corollary is proved by setting $\rho \rightarrow \infty$ in Theorem 1. Since $\lim_{\rho \rightarrow \infty} \mathbf{S}_k(\rho) = \frac{1}{\rho} \mathbf{I}_N$, $\lim_{\rho \rightarrow \infty} \mathbf{S}'_k(\rho) = \frac{1}{\rho^2} \mathbf{I}_N$, $\lim_{\rho \rightarrow \infty} \mathbf{S}_{ki}(\rho) = \frac{1}{\rho} \mathbf{I}_N$, and $\lim_{\rho \rightarrow \infty} \mathbf{S}'_{ki}(\rho) = \frac{1}{\rho^2} \mathbf{R} \Phi_i \mathbf{R}^H$, we have

$$\begin{aligned} \lambda_{\text{MF}} &= \lim_{\rho \rightarrow \infty} \lambda_{\text{RZF}} = \frac{P_T}{\rho^2} \left[\sum_{k=1}^K p_k |\tilde{t}_k|^2 \frac{1}{N} \text{Tr} \{ \mathbf{R} \Phi_k \mathbf{R}^H \} \right]^{-1}, \\ \lim_{\rho \rightarrow \infty} g_{k,k}^\circ &= \lambda_{\text{MF}} |\tilde{r}_k \tilde{t}_k|^2 \left| \frac{1}{N} \text{Tr} \{ \mathbf{T} \Phi_k \mathbf{R}^H \} \right|^2, \\ \lim_{\rho \rightarrow \infty} g_{k,i}^{\circ 2} &= \lambda_{\text{MF}} |\tilde{r}_k \tilde{t}_i|^2 \frac{1}{\rho^2} \frac{1}{N^2} \text{Tr} \{ \mathbf{T} \Phi_k \mathbf{T}^H \mathbf{R} \Phi_i \mathbf{R}^H \}. \end{aligned} \quad (20)$$

Substituting (20) into (17) gives (11). ■

Although the achievable rates of RZF and MF can be calculated according to (17) and (11), it is not easy to get some insights on how HM affects the DL. Therefore, we will consider a simple scenario and show that there are performance bounds for MF and RZF in the high SNR region.

4) Performance Bound in the High SNR Region:

Assuming the HM parameters at BS as random variables while those at UEs remain as known constants and are well compensated, i.e., $\tilde{t}_i = \tilde{r}_i = 1, \forall i = 1, \dots, K$. In this way, we concentrate on how the number of antennas and the HM parameters at BS affect the DL. The following corollary shows there are performance bounds for both MF and RZF in the high SNR region.

Corollary 2: Assume that the HM parameters at BS r_i 's are i.i.d., t_i 's are i.i.d., r_i 's and t_i 's are independent. Furthermore, we assume $\Phi_k = \mathbf{I}_N, \forall i = 1, \dots, K$ and equal power allocation is used, i.e., $\mathbf{P} = \mathbf{I}_K$. Then for a communication system modeled as (3), when $P_T \rightarrow \infty$, as $N, K \rightarrow \infty$ and $\frac{K}{N} = \beta$, the received SINR of user k under MF $\gamma_{\text{MF},k}$ and RZF $\gamma_{\text{RZF},k}, \forall k = 1, \dots, K$ is given by, respectively

$$\gamma_{\text{MF},k} \rightarrow \frac{|\mathbb{E} \{ r_i^* t_i \}|^2}{\beta \mathbb{E} \{ |t_i|^2 \} \mathbb{E} \{ |r_i|^2 \}}, \quad (21)$$

$$\gamma_{\text{RZF},k} \rightarrow \frac{c_5^2}{\beta^3 c_6 c_1^2 - 2c_1 \beta^2 \text{Re} \{ c_5 c_8 \} + \beta c_7 |c_5|^2}, \quad (22)$$

where c_1, c_5, c_6, c_7 and c_8 are constants related to β and the statistics of the HM parameters. They are defined in (D.4) and (D.9) in Appendix D.

Proof: See Appendix D. ■

It can be seen from (21) and (22) that in the high SNR region, performance upper bounds exist for both MF and RZF. The performance loss due to HM are related to β and the statistics of the circuit gains at BS. In general, the performance loss of RZF is larger than that of MF. Because there is also a performance upper bound for MF (which is given by $\frac{1}{\beta}$ [10]) as a result of inter-user interference when no HM exists. From (21), we can see that the data rates loss due to HM is limited when the second order statistics of the HM parameters are small. In contrast, RZF converges to ZF when SNR is high, and the achievable rates without HM will grow almost linearly with respect to SNR. When HM exists, however, the achievable rates of RZF is upper bounded, as shown in Corollary 2.

IV. STUDY OF HM CALIBRATION SCHEMES

In this section, we carry out a study of HM calibration in the scenario outlined in Section II. With HM, the channels in

the UL and DL cannot be reciprocal. So the DL performance will be degraded due to the inaccurate precoding matrix calculated based on UL channel estimation. As analyzed in the previous section, HM significantly affects the DL performance. Therefore a channel calibration process is usually preferred. In this section, we first take ZF as an example to show the effects of HM. Then we compare two sorts of HM calibration schemes. The choice of ZF is based on the fact that ZF has a simpler expression than RZF and shows more insights on the HM's influence than MF, which will become clear in the following analysis.

When ZF is used for downlink transmission, and assuming perfect CSI at the BS, the precoding matrix is given by

$$\mathbf{W}_{\text{ZF}} = \sqrt{\lambda_{\text{ZF}}} (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H, \quad (23)$$

where

$$\begin{aligned} \lambda_{\text{ZF}} &= \frac{P_{\text{T}}}{\mathbb{E} \text{Tr} \left\{ (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{P} \mathbf{H} (\mathbf{H}^H \mathbf{H})^{-1} \right\}} \\ &= \lim_{\rho \rightarrow 0} \lambda_{\text{RZF}}^{\circ}. \end{aligned} \quad (24)$$

Substituting (23) into (3) results in

$$\mathbf{y} = \sqrt{\lambda_{\text{ZF}}} \mathbf{x} \sqrt{\mathbf{P}} [\mathbf{H}^H \mathbf{H}]^{-1} \mathbf{H}^H \tilde{\mathbf{H}} + \mathbf{n}. \quad (25)$$

Equation (25) can be expanded as

$$\begin{aligned} \mathbf{y} &= \sqrt{\lambda_{\text{ZF}}} \mathbf{x} \sqrt{\mathbf{P}} [\mathbf{H}^H \mathbf{H}]^{-1} \mathbf{H}^H \tilde{\mathbf{H}} + \mathbf{n} \\ &= \sqrt{\lambda_{\text{ZF}}} \mathbf{x} \sqrt{\mathbf{P}} [\tilde{\mathbf{T}}^H \mathbf{V}^H \mathbf{R}^H \mathbf{R} \mathbf{V} \tilde{\mathbf{T}}]^{-1} \tilde{\mathbf{T}}^H \mathbf{V}^H \mathbf{R}^H \mathbf{T} \mathbf{V} \tilde{\mathbf{R}} + \mathbf{n} \\ &= \sqrt{\lambda_{\text{ZF}}} \mathbf{x} \sqrt{\mathbf{P}} \tilde{\mathbf{T}}^{-1} [\mathbf{V}^H \mathbf{R}^H \mathbf{R} \mathbf{V}]^{-1} \mathbf{V}^H \mathbf{R}^H \mathbf{T} \mathbf{V} \tilde{\mathbf{R}} + \mathbf{n}. \end{aligned} \quad (26)$$

As can be seen from (26), if $\tilde{\mathbf{T}}^{-1} [\mathbf{V}^H \mathbf{R}^H \mathbf{R} \mathbf{V}]^{-1} \mathbf{V}^H \mathbf{R}^H \mathbf{T} \mathbf{V} \tilde{\mathbf{R}}$ is not diagonal, inter-user interference cannot be completely eliminated by ZF. This is caused by differences between \mathbf{R} and \mathbf{T} and the resulting biased beam directions in the DL. Hence a calibration scheme can be introduced to adapt ZF to HM.

A. Introduction to Partial and Full Calibration Schemes

In this subsection, we give an introduction of partial calibration and full calibration, and describe the two different ways to implement such calibration schemes.

1) Partial Calibration:

Partial calibration is used to eliminate HM's impact on DL beam directions without considering the HM parameters at the users. It adjusts the relative circuit gains of all the antennas at BS to a reference antenna. Approaches of such partial calibration can be found in [18], [20], [21].

When the HM parameters of the BS are obtained, there are basically two ways to carry out the HM calibration as mentioned briefly in [22]. One is to calibrate the DL channel after precoding, namely 'Post-precoding Calibration' (Post-Cal); the other one is to calibrate the estimated UL channel and use it for precoding, namely 'Pre-precoding Calibration' (Pre-Cal). For partial calibration, we refer to the two methods as Partial Post-Cal (P-Post-Cal) and Partial Pre-Cal (P-Pre-Cal), respectively.

For P-Post-Cal, let us suppose linear calibration is used as

$$\mathbf{W}_{\text{P-Post}} = \sqrt{\lambda_{\text{P-Post}}} [\mathbf{H}^H \mathbf{H}]^{-1} \mathbf{H}^H \mathbf{A}, \quad (27)$$

where the calibration matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$ is invertible. For P-Pre-Cal methods, let us suppose $\mathbf{A}^{-1} \mathbf{H}$ is used for precoding, and

$$\mathbf{W}_{\text{P-Pre}} = \sqrt{\lambda_{\text{P-Pre}}} [\mathbf{H}^H (\mathbf{A}^{-1})^H \mathbf{A}^{-1} \mathbf{H}]^{-1} \mathbf{H}^H (\mathbf{A}^{-1})^H. \quad (28)$$

The reason for introducing \mathbf{A}^{-1} in (28) is to unify the expression of sufficient conditions for nulling inter-user interference in (27) and (28).

One sufficient condition for zero downlink inter-user interference is given by $\mathbf{A} = \mathbf{R} \mathbf{T}^{-1}$ [21], [22]. Substituting $\mathbf{A} = \mathbf{R} \mathbf{T}^{-1}$ and (27) into (3) gives

$$\mathbf{y} = \sqrt{\lambda_{\text{P-Post}}} \mathbf{x} \sqrt{\mathbf{P}} \tilde{\mathbf{T}}^{-1} \tilde{\mathbf{R}} + \mathbf{n}, \quad (29)$$

The same holds for the P-Pre-Cal case. By choosing such \mathbf{A} , both P-Post-Cal and P-Pre-Cal methods are able to eliminate inter-user interference in the DL.

2) Full Calibration:

It can be seen from (29) that the HM parameters at the users will change the average received signal power. Although the entries in $\tilde{\mathbf{T}}^{-1} \tilde{\mathbf{R}}$ can be estimated at each user as part of the DL channel as pointed out in [24], it affects the stability of the DL performance. The received signal power of each user is unpredictable at the BS without knowledge of each user's HM parameters, and the resulting achievable rates will vary in a wide range. Another consideration would be that the the partial calibration algorithm in [21] is not suitable for RZF and some non-linear precoding schemes which needs accurate information of the gains of each user's sub-channels. Therefore, full calibration becomes necessary, which exchanges pilot signals between the BS and the users through the air interface and estimates the HM parameters of both the BS and the users.

Many different versions of full calibration can be found in previous works, e.g., [14] and [25], and detailed discussion of these schemes are beyond the scope of this paper. Here for convenience, a demonstration of such calibration is described as follows.

For the Post-Cal methods, an additional calibration matrix $\mathbf{B} = \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{T}}$ is introduced to (29). The uplink channel from the j -th user to a reference antenna at BS is represented as $h_{1j} = r_1 v_{1j} \tilde{t}_j$, and its downlink counterpart is $h'_{1j} = t_1 v_{1j} \tilde{r}_j$. h_{1j} can be obtained at BS using uplink pilots, while h'_{1j} must be reported to BS by user j . Then we have

$$\frac{h_{1j}}{h'_{1j}} = \frac{r_1 v_{1j} \tilde{t}_j}{t_1 v_{1j} \tilde{r}_j} = \frac{r_1 \tilde{t}_j}{t_1 \tilde{r}_j} \Rightarrow \frac{\tilde{t}_j}{\tilde{r}_j} = \frac{h_{1j}}{h'_{1j}} \frac{t_1}{r_1} \triangleq \frac{h_{1j}}{h'_{1j}} \mu. \quad (30)$$

Thus the calibration matrix \mathbf{B} is constructed as $\mathbf{B} = \text{diag} \{ \mu h_{11}/h'_{11}, \dots, \mu h_{1K}/h'_{1K} \}$.

Once \mathbf{A} and \mathbf{B} are acquired, the Post-Cal scheme is carried out as

$$\mathbf{W}_{\text{F-Post}} = \sqrt{\lambda_{\text{F-Post}}} \mathbf{B} [\mathbf{H}^H \mathbf{H}]^{-1} \mathbf{H}^H \mathbf{A}, \quad (31)$$

and then the received signal vector at K users is given by

$$\begin{aligned} \mathbf{y} &= \sqrt{\lambda_{\text{F-Post}}} \mathbf{x} \mathbf{B} \tilde{\mathbf{T}}^{-1} \tilde{\mathbf{R}} + \mathbf{n} \\ &= \sqrt{\lambda_{\text{F-Post}}} \mathbf{x} + \mathbf{n}. \end{aligned}$$

which indicates no inter-user interference due to HM exists in the system. Therefore, by using calibration matrices \mathbf{A} and \mathbf{B} , all the impact of HM in the DL is compensated for ZF. We refer to the calibration scheme based on (31) as Full Post Calibration (F-Post-Cal).

For the Pre-Cal scheme, with the same \mathbf{A} and \mathbf{B} as F-Post-cal, we have

$$\begin{aligned} \mathbf{A}^{-1}\mathbf{H}\mathbf{B}^{-1} &= \mathbf{T}\mathbf{R}^{-1}\mathbf{V}\tilde{\mathbf{T}}^{-1}\tilde{\mathbf{R}} \\ &= \mathbf{T}\mathbf{V}\tilde{\mathbf{R}} = \tilde{\mathbf{H}}. \end{aligned}$$

Thus the accurate values of DL channel coefficients are obtained by multiplying \mathbf{A}^{-1} and \mathbf{B}^{-1} to the UL channel matrix $\tilde{\mathbf{H}}$. Any precoding scheme can be carried out based on the calibrated UL channel matrix $\mathbf{A}^{-1}\mathbf{H}\mathbf{B}^{-1}$. We call this method Full Pre-Cal (F-Pre-Cal). After calibration, the equivalent ZF precoding matrix for F-Pre-Cal becomes

$$\mathbf{W}_{\text{F-Pre}} = \sqrt{\lambda_{\text{F-Pre}}}[\tilde{\mathbf{H}}^{\text{H}}\tilde{\mathbf{H}}]^{-1}\tilde{\mathbf{H}}^{\text{H}}.$$

Remarks: 1) When comparing partial and full calibration approaches, there is a tradeoff between complexity and performance. It is apparent that full calibration demands assistance from the users and thus will cost more system resources. 2) In general full calibration has superior performance and the F-Pre-cal scheme is able to suit any precoding scheme.

B. Comparison between Post-Cal and Pre-Cal Schemes

After partial calibration, the equivalent ZF Precoding matrices for P-Post-Cal and P-Pre-Cal become, respectively

$$\begin{aligned} \mathbf{W}_{\text{P-Post}} &= \sqrt{\lambda_{\text{P-Post}}}[\mathbf{H}^{\text{H}}\mathbf{H}]^{-1}\mathbf{H}^{\text{H}}\mathbf{A} \\ &= \sqrt{\lambda_{\text{P-Post}}}\tilde{\mathbf{T}}^{-1}[\mathbf{V}^{\text{H}}\mathbf{R}^{\text{H}}\mathbf{R}\mathbf{V}]^{-1}\mathbf{V}^{\text{H}}\mathbf{R}^{\text{H}}\mathbf{R}\mathbf{T}^{-1}, \\ \mathbf{W}_{\text{P-Pre}} &= \sqrt{\lambda_{\text{P-Pre}}}[\mathbf{H}^{\text{H}}(\mathbf{A}^{-1})^{\text{H}}\mathbf{A}^{-1}\mathbf{H}]^{-1}\mathbf{H}^{\text{H}}(\mathbf{A}^{-1})^{\text{H}} \\ &= \sqrt{\lambda_{\text{P-Pre}}}\tilde{\mathbf{T}}^{-1}[\mathbf{V}^{\text{H}}\mathbf{T}^{\text{H}}\mathbf{T}\mathbf{V}]^{-1}\mathbf{V}^{\text{H}}\mathbf{T}^{\text{H}}. \end{aligned}$$

From the above equations, it can be seen that the only difference between $\mathbf{W}_{\text{P-Post}}$ and $\mathbf{W}_{\text{P-Pre}}$ is the additional term $\mathbf{A} = \mathbf{R}\mathbf{T}^{-1}$ in $\mathbf{W}_{\text{P-Post}}$ if \mathbf{R} and \mathbf{T} have the same distributions. However, the performance is not the same because of $\lambda_{\text{P-Post}}$ and $\lambda_{\text{P-Pre}}$. Actually when a transmit power constraint exists, the performance of P-Pre-Cal methods is better than the P-Post-Cal schemes, which is shown in the following proposition.

Proposition 3: Provided that the magnitude of the HM parameters of both the BS and the users are independently and identically distributed, then we have $\lambda_{\text{P-Pre}} \geq \lambda_{\text{P-Post}}$.

Proof: See Appendix E. ■

For full calibration, the equivalent ZF precoding matrices for F-Post-Cal and F-Pre-Cal become, respectively

$$\begin{aligned} \mathbf{W}_{\text{F-Post}} &= \sqrt{\lambda_{\text{F-Post}}}\mathbf{B}[\mathbf{H}^{\text{H}}\mathbf{H}]^{-1}\mathbf{H}^{\text{H}}\mathbf{A}, \\ \mathbf{W}_{\text{F-Pre}} &= \sqrt{\lambda_{\text{F-Pre}}}[\tilde{\mathbf{H}}^{\text{H}}\tilde{\mathbf{H}}]^{-1}\tilde{\mathbf{H}}^{\text{H}}. \end{aligned}$$

Note that the difference between the F-Pre-Cal and F-Post-Cal methods for ZF precoding lies in the equivalent channel used for data transmission. F-Post-Cal calibrates the DL channel according to the UL channel. The equivalent channel used for data transmission is actually the UL one. However, the F-Pre-Cal use the UL channel to predict the real DL channel and

TABLE I
PARAMETERS USED IN THE SIMULATIONS

noise power spectrum density	-174 dBm/Hz
bandwidth	20 MHz
τ_k	$U(0, 1)$
path loss exponent α	3.5
$d_k(\text{m})$	$U(30, 80)$

therefore the equivalent channel is the DL channel. Similar to the partial calibration schemes, F-Pre-Cal outperforms F-Post-Cal as stated in the following Proposition.

Proposition 4: Provided that the magnitude of the HM parameters of both the BS and the users are independently and identically distributed, we have $\lambda_{\text{F-Pre}} \geq \lambda_{\text{F-Post}}$.

Proof: The proof is similar to that of Proposition 3 and thus omitted. ■

The additional term \mathbf{A} introduced in $\mathbf{W}_{\text{P-Post}}$ generates a smaller power factor and thus degrades the overall performance. The main reason is that when the diagonal entries of \mathbf{A} has a mean value greater than 1, it causes a relatively larger power consumption. When a power constraint exists, the power factor used to normalize the transmit power becomes smaller, and the average received SNR decreases consequently. For the same reason that the diagonal entries of \mathbf{B} and \mathbf{A} have a mean value greater than 1, we have $\lambda_{\text{F-Pre}} > \lambda_{\text{F-Post}}$.

Remark: The results of Proposition 3 and Proposition 4 can be extended to RZF. The proofs and discussions follow a similar procedure as ZF and are thus omitted here. For MF, however, Pre-Cal and Post-Cal methods have the same power normalization factor, and therefore have the same performance.

V. NUMERICAL RESULTS

Similar to [22], the simulations are carried out by assuming that the magnitudes of t_i , r_i , \tilde{t}_i , \tilde{r}_i are identically and uniformly distributed in the range of $(1 - 0.5\delta, 1 + 0.5\delta)$, i.e., $U(1 - 0.5\delta, 1 + 0.5\delta)$. For simplicity, we assume that the phase of the HM parameters are well compensated in the simulations. The larger δ is, the more severer HM is. The correlation matrix of each user's channel, i.e., Φ_k , $k = 1, \dots, K$, is modeled as

$$\Phi_{k,i,j} = \begin{cases} \text{PL}_k \tau_k^{j-i} & i \leq j \\ \text{PL}_k (\tau_k^{j-i})^* & i > j \end{cases}$$

where $|\tau_k| \leq 1$, and the path loss of the k -th user, $\text{PL}_k(\text{dB}) = -38.46 - \alpha 10 \log_{10}(d_k)$, in which α is the path loss exponent and d_k is the distance between the BS and user k . The parameters used in the simulations are summarized in Table I. All the simulation results are averaged over randomly generated HM parameters and channel realizations.

A. Performance Loss due to HM

In this subsection, the performance loss due to HM is evaluated through simulations. The transmit power is equally allocated among all the users. The loss caused by HM in terms of achievable rates is illustrated in Fig. 3 for both MF and RZF, and the bit error rate (BER) of QPSK is shown in Fig. 4, when $N = 100, 10$, $\delta = 0.4$. In all the figures, 'Perfect'

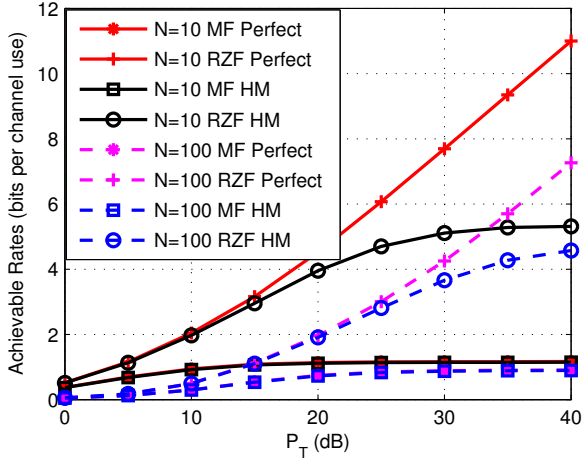


Fig. 3. Achievable rates per user with and without HM when $\beta = 0.4$, $\delta = 0.4$ for HM.

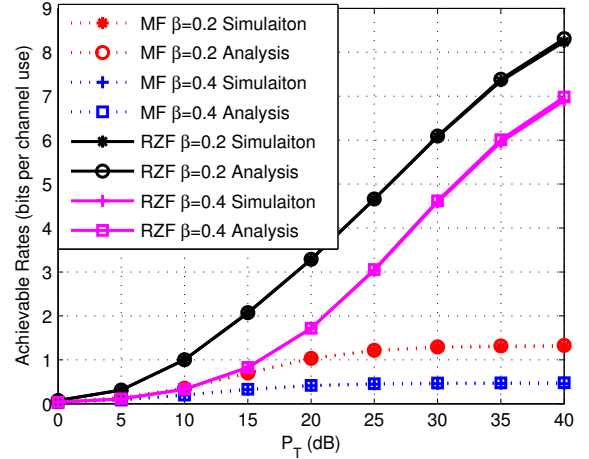


Fig. 5. Achievable rates of MF, RZF when $N = 100$, $\delta = 0.2$.

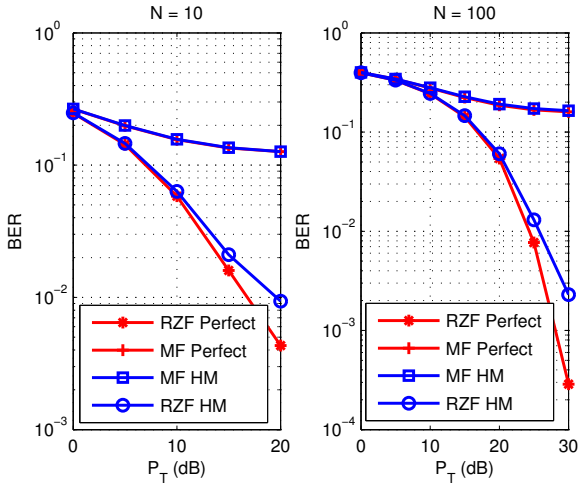


Fig. 4. Bit Error Rates of QPSK with and without HM when $\beta = 0.4$, $\delta = 0.4$ for HM.

means perfect UL/DL hardware with no mismatch and 'HM' stands for scenarios with HM. It is worth mentioning that the simulation results of $N = 10$ are better than those of $N = 100$, because the channel gain is normalized to $\frac{1}{N}$ (please see description of v_k in Section II), and 10 times more users are served when $N = 100$ with the same $\beta = \frac{K}{N}$.

As can be seen in Fig. 3 and Fig. 4, the performance of RZF is much better than MF in terms of achievable rates and BER. A significant performance loss is observed for RZF in the high SNR region, while MF performs almost the same. In terms of achievable rates, when HM exists there are upper bounds for both RZF and MF in the high SNR region, as is predicted in Corollary 2. For BER results of QPSK modulation, MF undergoes an error floor around 10^{-1} in the high SNR region, and RZF experiences a performance loss of 5 dB due to HM.

From the simulation results, we can conclude that the performance of MF with and without HM is comparable, while there is significant performance loss in the high SNR region for RZF when HM exists. This is due to the fact that the performance of MF is limited by the severe inter-user interference, so it is less sensitive to HM as compared

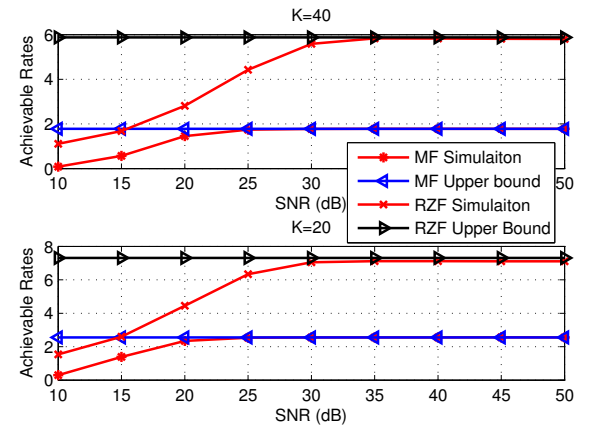


Fig. 6. Upper Bounds of MF, RZF when $N = 100$, $\delta = 0.4$.

with RZF. In order to improve the performance of MF, a smaller value of β is required. While for RZF, taking HM into consideration can lead to remarkably better downlink performance.

B. Achievable rates of RZF and MF with HM

Fig. 5 illustrates the analysis and Monte-Carlo simulation results on achievable rates of RZF and MF, where $\delta = 0.2$, $N = 100$, $\beta = 0.2, 0.4$. The analytical achievable rates with RZF and MF are given by Theorem 1 and Corollary 1, respectively. In order to show the accuracy of the analysis under different power allocation schemes, the transmit power for each user is randomly allocated in the simulation, i.e., $p_k \sim U(0, 2)$. It can be seen from Fig. 5 that as the transmit power P_T increases, the achievable rates of both RZF and MF increase, and the higher achievable rates per user can be obtained for smaller value of β . The analysis results of MF and RZF match well with the simulation results.

In Fig. 6, we consider a special case when $\Phi_k = \mathbf{I}_N$ and $\tilde{\mathbf{T}} = \tilde{\mathbf{R}} = \mathbf{I}_K$. The SNR therein is defined as $\frac{P_T}{\sigma^2}$. The upper bounds for MF and RZF which are given in Corollary 2 are very tight, especially when K is larger. When HM exists, the achievable rates of RZF do not grow linearly with respect to

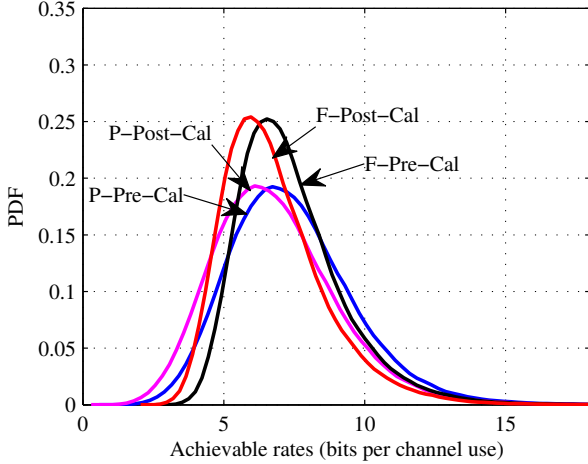


Fig. 7. Empirical PDF of achievable rates with ZF when $N = 100$, $K = 40$, $\delta = 1.0$ and $P_T = 40$ dB.

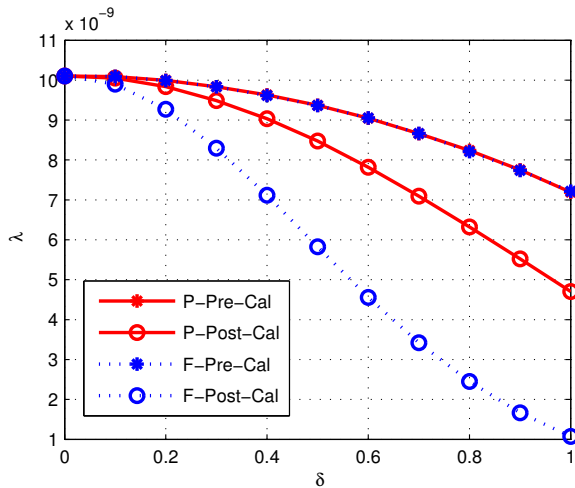


Fig. 8. Power normalization factors of different calibration schemes when $N = 100$, $K = 40$, and $P_T = 40$ dB.

SNR. When $N = 100$, as SNR approaches 30 dB, the data rates saturate. So it is not beneficial to increase the transmit power any more.

C. Calibration for HM

In order to compare the performance of the four calibration schemes discussed in Section IV, i.e., P-Pre-Cal, P-Post-Cal, F-Pre-Cal and F-Post-Cal, we show simulation results in terms of the empirical PDF of the achievable rates per user in severe HM ($\delta = 1.0$), which is useful to show the overall calibration performance.

As illustrated in Fig. 7, the PDF curves of the full calibration schemes (F-Post-Cal and F-Pre-Cal) is shaper than those of the partial calibration approaches (F-Post-Cal and F-Pre-Cal). Therefore, the full calibration provides a much more stable performance. Moreover, it can be seen that there is performance loss with post-precoding calibration schemes (F-Post-Cal and P-Post-Cal) compared with the pre-precoding counterparts. As described in Proposition 3 and Proposition 4, the loss comes from the additional terms \mathbf{A} and \mathbf{B} introduced in the precoding matrix, which decreases the value

of the power normalization factor λ_{ZF} , and the received SNR degrades consequently.

Fig. 8 illustrates how the value of δ affects the power normalization factor of different calibration schemes. As δ becomes larger, the power normalization factors of all the four schemes decrease. As can be seen from the plots, Pre-Cal schemes outperform Post-Cal schemes, especially for large δ . Although P-Pre-Cal has almost the same power normalization factor as F-Pre-Cal, the achievable rates performance of F-Pre-Cal is more stable, as is shown in Fig. 7.

VI. CONCLUSION

In this paper, the problem of HM in large-scale antenna systems has been studied. By handling expectations and asymptotic deterministic equivalents of a series of random variables, we have derived analytical expressions that provide sets of achievable rates for both MF and RZF. In addition, the upper bounds on achievable rates of MF and RZF been investigated. The results show that when N and K are large, the downlink SINR of each user is upper bounded for a given β , and is related to the statistics of random HM parameters. Moreover, we have compared different HM calibration schemes and shown that Pre-Cal outperforms Post-Cal schemes. The analytical results are verified by simulations.

APPENDIX A USEFUL LEMMAS

Lemma 2 (Matrix Inversion Lemma): Let \mathbf{A} be Hermitian invertible. Then for any vector $\mathbf{x}, \mathbf{y} \in \mathbb{C}^N$ and any scalar $\rho \in \mathbb{C}$, such that $\mathbf{A} + \rho\mathbf{x}\mathbf{y}^H$ is invertible, and

$$\begin{aligned} (\mathbf{A} + \rho\mathbf{x}\mathbf{y}^H)^{-1} &= \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{x}\mathbf{y}^H\mathbf{A}^{-1}}{1 + \mathbf{y}^H\mathbf{A}^{-1}\mathbf{x}} \\ &= \mathbf{A}^{-1} \left(\mathbf{I} - \frac{\mathbf{x}\mathbf{y}^H\mathbf{A}^{-1}}{1 + \mathbf{y}^H\mathbf{A}^{-1}\mathbf{x}} \right) \\ &= \left(\mathbf{I} - \frac{\mathbf{A}^{-1}\mathbf{x}\mathbf{y}^H}{1 + \mathbf{y}^H\mathbf{A}^{-1}\mathbf{x}} \right) \mathbf{A}^{-1}, \end{aligned}$$

and

$$\begin{aligned} \mathbf{y}^H(\mathbf{A} + \rho\mathbf{x}\mathbf{y}^H)^{-1} &= \frac{\mathbf{y}^H\mathbf{A}^{-1}}{1 + \rho\mathbf{y}^H\mathbf{A}^{-1}\mathbf{x}}, \\ (\mathbf{A} + \rho\mathbf{x}\mathbf{y}^H)^{-1}\mathbf{x} &= \frac{\mathbf{A}^{-1}\mathbf{x}}{1 + \rho\mathbf{y}^H\mathbf{A}^{-1}\mathbf{x}}. \end{aligned}$$

Proof: The proof of the first part can be found in [31], and the second part is obtained accordingly. ■

Lemma 3 (Lemma 4 in [28]): Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ and $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \frac{1}{N}\Phi_{\mathbf{x}})$, $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \frac{1}{N}\Phi_{\mathbf{y}})$. Assume that \mathbf{A} has uniformly bounded spectral norm (with respect to N) and that \mathbf{x} and \mathbf{y} are mutually independent and independent of \mathbf{A} . Then

for all $p \geq 1$,

- a) $\mathbb{E} \left\{ \left| \mathbf{x}^H \mathbf{A} \mathbf{x} - \frac{1}{N} \text{Tr}(\mathbf{A} \Phi_x) \right|^p \right\} = \mathcal{O} \left(N^{-\frac{p}{2}} \right),$
- b) $\mathbf{x}^H \mathbf{A} \mathbf{x} - \frac{1}{N} \text{Tr}(\mathbf{A} \Phi_x) \rightarrow 0,$
- c) $\mathbf{x}^H \mathbf{A} \mathbf{y} \rightarrow 0,$
- d) $\mathbb{E} \left\{ \left| (\mathbf{x}^H \mathbf{A} \mathbf{x})^2 - \left(\frac{1}{N} \text{Tr}(\mathbf{A} \Phi_x) \right)^2 \right| \right\} \rightarrow 0.$

Lemma 4 ([28], see also [27]): Let $\mathbf{V} \in \mathbb{C}^{N \times K}$, $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_K]$ with independent columns and $\mathbf{v}_i \sim \mathcal{CN}(0, \frac{1}{N} \Phi_i)$. Assume that $\mathbf{A} \in \mathbb{C}^{N \times N}$ and matrices Φ_k , $k = 1, \dots, K$ have uniformly bounded spectral norms (with respect to N). Then when $N \rightarrow \infty$, $K \rightarrow \infty$ with fixed ratio, $\forall \rho > 0$

$$\frac{1}{N} \text{Tr} \left\{ \mathbf{A} [\mathbf{V} \mathbf{V}^H + \rho \mathbf{I}_N]^{-1} \right\} \rightarrow \frac{1}{N} \text{Tr} \{ \mathbf{A} \mathcal{S}(\rho) \},$$

where

$$\mathcal{S}(\rho) = \left[\frac{1}{N} \sum_{i=1}^K \left\{ \frac{\Phi_i}{1 + e_i(\rho)} \right\} + \rho \mathbf{I}_N \right]^{-1},$$

in which $\forall i = 1, \dots, K$,

$$\begin{aligned} e_i(\rho) &= \frac{1}{N} \text{Tr} \left\{ \Phi_i \left[\frac{1}{N} \sum_{j=1}^K \left\{ \frac{\Phi_j}{1 + e_j(\rho)} \right\} + \rho \mathbf{I}_N \right]^{-1} \right\}, \\ &= \frac{1}{N} \text{Tr} \left\{ \Phi_i \left[\frac{1}{N} \sum_{j=1}^K \left\{ \frac{\Phi_j}{1 + e_j(\rho)} \right\} + \rho \mathbf{I}_N \right]^{-1} \right\}, \end{aligned}$$

Moreover, let $\mathbf{B} \in \mathbb{C}^{N \times N}$ be hermitian nonnegative definite with uniformly bounded spectral norm (with respect to N). We have

$$\begin{aligned} &\frac{1}{N} \text{Tr} \left\{ \mathbf{A} [\mathbf{V} \mathbf{V}^H + \rho \mathbf{I}_N]^{-1} \mathbf{B} [\mathbf{V} \mathbf{V}^H + \rho \mathbf{I}_N]^{-1} \right\} \\ &\rightarrow \frac{1}{N} \text{Tr} \{ \mathbf{A} \mathcal{S}'(\rho) \}, \end{aligned}$$

where

$$\mathcal{S}'(\rho) = \mathcal{S}(\rho) \mathbf{B} \mathcal{S}(\rho) + \mathcal{S}(\rho) \frac{1}{N} \sum_{k=1}^K \frac{\Phi_k e'_k(\rho)}{[1 + e_k(\rho)]^2} \mathcal{S}(\rho), \quad (\text{A.1})$$

and $e'(\rho) = [e'_1(\rho), \dots, e'_K(\rho)]^T$ is calculated as

$$e'(\rho) = [\mathbf{I}_K - \mathbf{J}(\rho)]^{-1} \mathbf{u}(\rho), \quad (\text{A.2})$$

where $\mathbf{J}(\rho) \in \mathbb{C}^{K \times K}$ and $\mathbf{u}(\rho) \in \mathbb{C}^K$ are defined as

$$\begin{aligned} [\mathbf{J}(\rho)]_{kl} &= \frac{\frac{1}{N} \text{Tr}[\Phi_k \mathcal{S}(\rho) \Phi_l \mathcal{S}(\rho)]}{N[1 + e_l(\rho)]^2}, \quad 1 \leq k, l \leq K, \\ [\mathbf{u}(\rho)]_k &= \frac{1}{N} \text{Tr}[\Phi_k \mathcal{S}(\rho) \mathbf{B} \mathcal{S}(\rho)], \quad 1 \leq k \leq K. \end{aligned} \quad (\text{A.3})$$

APPENDIX B PROOF OF PROPOSITION 1

Let us start from the ADE of $g_{k,k}$.
Expand \mathbf{Q} as

$$\mathbf{Q} = \left[\mathbf{R} \mathbf{V}_{-k} \tilde{\mathbf{T}}_{-k} \tilde{\mathbf{T}}_{-k}^H \mathbf{V}_{-k}^H \mathbf{R}^H + |\tilde{t}_k|^2 \mathbf{R} \mathbf{v}_k \mathbf{v}_k^H \mathbf{R}^H + \rho \mathbf{I}_N \right]^{-1}, \quad (\text{B.1})$$

in which

$$\begin{aligned} \mathbf{V}_{-k} &= [\mathbf{v}_1, \dots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \dots, \mathbf{v}_K], \\ \tilde{\mathbf{T}}_{-k} &= \text{diag} \{ \tilde{t}_1, \dots, \tilde{t}_{k-1}, \tilde{t}_{k+1}, \dots, \tilde{t}_K \}. \end{aligned}$$

Let $\mathbf{Q}_{-k} = [\mathbf{R} \mathbf{V}_{-k} \tilde{\mathbf{T}}_{-k} \tilde{\mathbf{T}}_{-k}^H \mathbf{V}_{-k}^H \mathbf{R}^H + \rho \mathbf{I}]^{-1}$ and using Lemma 2 in Appendix A, we get

$$g_{k,k} = \frac{\sqrt{\lambda_{\text{RZF}}} \tilde{t}_k \tilde{t}_k^* \mu_2}{1 + |\tilde{t}_k|^2 \mu_1}. \quad (\text{B.2})$$

where

$$\begin{aligned} \mu_1 &= \mathbf{v}_k^H \mathbf{R}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_k, \\ \mu_2 &= \mathbf{v}_k^H \mathbf{R}^H \mathbf{Q}_{-k} \mathbf{T} \mathbf{v}_k. \end{aligned} \quad (\text{B.3})$$

According to Lemma 3,

$$\begin{aligned} \mu_1 &\rightarrow \frac{1}{N} \text{Tr} (\mathbf{R}^H \mathbf{Q}_{-k} \mathbf{R} \Phi_k), \\ \mu_2 &\rightarrow \frac{1}{N} \text{Tr} (\mathbf{R}^H \mathbf{Q}_{-k} \mathbf{T} \Phi_k). \end{aligned} \quad (\text{B.4})$$

Because \mathbf{Q}_{-k} is still random, using Lemma 4 gives,

$$\begin{aligned} \mu_1 &\rightarrow \frac{1}{N} \text{Tr} (\mathbf{R} \Phi_k \mathbf{R}^H \mathcal{S}_k(\rho)) \triangleq \mu_1^\circ, \\ \mu_2 &\rightarrow \frac{1}{N} \text{Tr} (\mathbf{T} \Phi_k \mathbf{R}^H \mathcal{S}_k(\rho)) \triangleq \mu_2^\circ, \end{aligned} \quad (\text{B.5})$$

where $\mathcal{S}(\rho)$ is given by (14).

Substituting (B.5) and (B.4) into (B.2), we have

$$g_{k,k} \rightarrow \frac{\sqrt{\lambda_{\text{RZF}}} \tilde{t}_k \tilde{t}_k^* \mu_2^\circ}{1 + |\tilde{t}_k|^2 \frac{1}{N} \mu_1^\circ} \triangleq g_{k,k}^\circ. \quad (\text{B.6})$$

Now let us handle the ADE of $|g_{k,i}|^2$. Since $g_{k,i} = \sqrt{\lambda_{\text{RZF}}} \tilde{t}_i^* \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q} \mathbf{T} \mathbf{v}_k \tilde{t}_k$, we have

$$\sum_{i \neq k} |g_{k,i}|^2 = \lambda_{\text{RZF}} |\tilde{t}_k \tilde{t}_i|^2 \mathbf{v}_k^H \mathbf{T}^H \mathbf{Q} \mathbf{R} \mathbf{v}_i \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q} \mathbf{T} \mathbf{v}_k. \quad (\text{B.7})$$

Applying Lemma 2 to \mathbf{Q} yields

$$\mathbf{Q} = \mathbf{Q}_{-k} - \frac{|\tilde{t}_k|^2 \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_k \mathbf{v}_k^H \mathbf{R}^H \mathbf{Q}_{-k}}{1 + |\tilde{t}_k|^2 \mathbf{v}_k^H \mathbf{R}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_k}. \quad (\text{B.8})$$

Substituting (B.8) into (B.7) gives

$$\begin{aligned} |g_{k,i}|^2 &= \lambda_{\text{RZF}} |\tilde{t}_k|^2 |\tilde{t}_i|^2 \\ &\times \left(\mu_3 - \frac{|\tilde{t}_k|^2 2 \text{Re}(\mu_2 \mu_5)}{1 + |\tilde{t}_k|^2 \mu_1} + \frac{|\tilde{t}_k|^4 |\mu_4| |\mu_2|^2}{(1 + |\tilde{t}_k|^2 \mu_1)^2} \right), \end{aligned} \quad (\text{B.9})$$

where

$$\begin{aligned} \mu_3 &= \mathbf{v}_k^H \mathbf{T}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_i \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q}_{-k} \mathbf{T} \mathbf{v}_k, \\ \mu_4 &= \mathbf{v}_k^H \mathbf{R}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_i \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_k, \\ \mu_5 &= \mathbf{v}_k^H \mathbf{T}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_i \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_k, \end{aligned} \quad (\text{B.10})$$

Now the remaining work is to derive the ADEs of μ_3 , μ_4 and μ_5 . From (B.10), we have

$$\begin{aligned} \mu_3 &= \mathbf{v}_k^H \mathbf{T}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_i \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q}_{-k} \mathbf{T} \mathbf{v}_k \\ &\stackrel{(a)}{\rightarrow} \frac{1}{N} \text{Tr} \left\{ \mathbf{T} \Phi_k \mathbf{T}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_i \mathbf{v}_i^H \mathbf{V}_{-k}^H \mathbf{R}^H \mathbf{Q}_{-k} \right\} \\ &= \frac{1}{N} \text{Tr} \left\{ \mathbf{T} \Phi_k \mathbf{T}^H \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_i \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q}_{-k} \right\} \\ &\stackrel{(b)}{=} \frac{1}{N} \frac{\text{Tr} \left\{ \mathbf{T} \Phi_k \mathbf{T}^H \mathbf{Q}_{-ki} \mathbf{R} \mathbf{v}_i \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q}_{-ki} \right\}}{\left(1 + |\tilde{t}_i|^2 \mathbf{v}_i^H \mathbf{R}^H \mathbf{Q}_{-ki} \mathbf{R} \mathbf{v}_i\right)^2}, \end{aligned} \quad (\text{B.11})$$

in which (a) follows Lemma 3 and (b) is achieved using Lemma 2. Using Lemma 3 and Lemma 4, we get

$$\begin{aligned} \mu_3 &\rightarrow \frac{1}{N} \frac{\frac{1}{N} \text{Tr} \left\{ \mathbf{T} \Phi_k \mathbf{T}^H \mathbf{Q}_{-ki} \mathbf{R} \Phi_i \mathbf{R}^H \mathbf{Q}_{-ki} \right\}}{\left(1 + \frac{1}{N} |\tilde{t}_i|^2 \text{Tr} \left\{ \mathbf{R} \Phi_i \mathbf{R}^H \mathbf{S}_{ki}(\rho) \right\}\right)^2} \\ &\rightarrow \frac{1}{N} \frac{\frac{1}{N} \text{Tr} \left\{ \mathbf{T} \Phi_k \mathbf{T}^H \mathbf{S}'_{ki}(\rho) \right\}}{\left(1 + \frac{1}{N} |\tilde{t}_i|^2 \text{Tr} \left\{ \mathbf{R} \Phi_i \mathbf{R}^H \mathbf{S}_{ki}(\rho) \right\}\right)^2} \\ &\triangleq \mu_3^\circ. \end{aligned} \quad (\text{B.12})$$

where \mathbf{Q}_{-ki} is similar to \mathbf{Q}_{-k} by removing \mathbf{v}_i and \tilde{t}_i from \mathbf{V}_{-k} and $\tilde{\mathbf{T}}_{-k}$, respectively; $\mathbf{S}_{ki}(\rho)$ is similar to $\mathbf{S}_k(\rho)$ with $\Psi_{ki} = \{1, \dots, K\} \setminus \{k, i\}$, $\mathbf{S}'_{ki}(\rho)$ is calculated according to (A.1) with $\mathbf{B} = \mathbf{R} \Phi_i \mathbf{R}^H$.

Similarly, we have

$$\begin{aligned} \mu_4 &\rightarrow \frac{1}{N} \frac{\frac{1}{N} \text{Tr} \left\{ \mathbf{R} \Phi_k \mathbf{R}^H \mathbf{S}'_{ki}(\rho) \right\}}{\left(1 + \frac{1}{N} |\tilde{t}_i|^2 \text{Tr} \left\{ \mathbf{R} \Phi_i \mathbf{R}^H \mathbf{S}_{ki}(\rho) \right\}\right)^2} \triangleq \mu_4^\circ, \\ \mu_5 &\rightarrow \frac{1}{N} \frac{\frac{1}{N} \text{Tr} \left\{ \mathbf{R} \Phi_k \mathbf{T}^H \mathbf{S}'_{ki}(\rho) \right\}}{\left(1 + \frac{1}{N} |\tilde{t}_i|^2 \text{Tr} \left\{ \mathbf{R} \Phi_i \mathbf{R}^H \mathbf{S}_{ki}(\rho) \right\}\right)^2} \triangleq \mu_5^\circ. \end{aligned} \quad (\text{B.13})$$

Substituting (B.5), (B.12) and (B.13) into (B.9) yields (12).

APPENDIX C PROOF OF PROPOSITION 2

In order to derive λ_{RZF} in (7), we need to handle the ADE of $\text{Tr} \left\{ \mathbf{P} [\mathbf{H}^H \mathbf{H} + \rho \mathbf{I}_K]^{-1} \mathbf{H}^H \mathbf{H} [\mathbf{H}^H \mathbf{H} + \rho \mathbf{I}_K]^{-1} \right\}$.

From (7), we have

$$\begin{aligned} &\text{Tr} \left\{ \mathbf{P} [\mathbf{H}^H \mathbf{H} + \rho \mathbf{I}_K]^{-1} \mathbf{H}^H \mathbf{H} [\mathbf{H}^H \mathbf{H} + \rho \mathbf{I}_K]^{-1} \right\} \\ &= \text{Tr} \left\{ \mathbf{Q} \mathbf{H} \mathbf{P} \mathbf{H}^H \mathbf{Q} \right\} \\ &= \sum_{k=1}^K \text{Tr} \left\{ p_k \mathbf{Q} \mathbf{h}_k \mathbf{h}_k^H \mathbf{Q} \right\} \\ &\stackrel{(a)}{=} \sum_{k=1}^K \frac{p_k |\tilde{t}_k|^2 \text{Tr} \left\{ \mathbf{Q}_{-k} \mathbf{R} \mathbf{v}_k \mathbf{v}_k^H \mathbf{R}^H \mathbf{Q}_{-k} \right\}}{\left(1 + |\tilde{t}_k|^2 \mathbf{v}_k^H \mathbf{R}^H \mathbf{Q}_k \mathbf{R} \mathbf{v}_k\right)^2}, \end{aligned}$$

where (a) follows Lemma 2. Applying Lemma 3 and Lemma

4 to the above equation gives

$$\begin{aligned} &\text{Tr} \left\{ \mathbf{P} [\mathbf{H}^H \mathbf{H} + \rho \mathbf{I}_K]^{-1} \mathbf{H}^H \mathbf{H} [\mathbf{H}^H \mathbf{H} + \rho \mathbf{I}_K]^{-1} \right\} \\ &\rightarrow \sum_{k=1}^K \frac{\frac{1}{N} p_k |\tilde{t}_k|^2 \text{Tr} \left\{ \mathbf{R} \Phi_k \mathbf{R}^H \mathbf{Q}_{-k}^2 \right\}}{\left(1 + \frac{1}{N} |\tilde{t}_k|^2 \text{Tr} \left\{ \mathbf{R} \Phi_k \mathbf{R}^H \mathbf{Q}_{-k} \right\}\right)^2} \\ &\rightarrow \sum_{k=1}^K \frac{\frac{1}{N} p_k |\tilde{t}_k|^2 \text{Tr} \left\{ \mathbf{R} \Phi_k \mathbf{R}^H \mathbf{S}'_k(\rho) \right\}}{\left(1 + \frac{1}{N} |\tilde{t}_k|^2 \text{Tr} \left\{ \mathbf{R} \Phi_k \mathbf{R}^H \mathbf{S}_k(\rho) \right\}\right)^2}, \end{aligned} \quad (\text{C.1})$$

in which $\mathbf{S}'_k(\rho)$ is calculated according to (A.1) with $\mathbf{B} = \mathbf{I}_N$ and $\mathbf{S}_k(\rho)$ given by (14).

APPENDIX D PROOF OF COROLLARY 2

The proof for MF is by direct application of Law of Large Numbers (LLN) to (11). As N and $K \rightarrow \infty$

$$\begin{aligned} \frac{1}{N} \text{Tr} \left\{ \mathbf{R}^H \mathbf{T} \right\} &\rightarrow \mathbb{E} \left\{ r_i^* t_i \right\}, \\ \frac{1}{N} \text{Tr} \left\{ \mathbf{R} \mathbf{R}^H \right\} &\rightarrow \mathbb{E} \left\{ |r_i|^2 \right\}, \\ \frac{1}{N} \text{Tr} \left\{ \mathbf{R}^H \mathbf{T} \mathbf{T}^H \mathbf{R} \right\} &\rightarrow \mathbb{E} \left\{ |r_i|^2 \right\} \mathbb{E} \left\{ |t_i|^2 \right\}, \\ \frac{1}{K-1} \sum_{i=1, i \neq k}^K |\tilde{t}_i^* \tilde{r}_k|^2 &\rightarrow 1. \end{aligned} \quad (\text{D.1})$$

Substituting (D.1) into (11) results in (21).

The proof for RZF is more complicated. We first focus on the expressions of $\mathbf{S}_k(\rho)$ and $\mathbf{S}'_k(\rho)$. Note that both $\mathbf{S}_k(\rho)$ and $\mathbf{S}'_k(\rho)$ are diagonal under the assumption made in Corollary 2.

From (10), we get $e_1(\rho) = e_2(\rho) = \dots = e_K(\rho) \triangleq e_0(\rho)$. According to LLN, when $N \rightarrow \infty$

$$e_0(\rho) = \frac{1}{N} \text{Tr} \left\{ \mathbf{R} \mathbf{R}^H \mathbf{S}_k(\rho) \right\} \rightarrow \mathbb{E} \left\{ |r_i|^2 s_{k,i}(\rho) \right\}, \quad (\text{D.2})$$

where $s_{k,i}(\rho)$ is the i -th diagonal element of $\mathbf{S}_k(\rho)$ in (14), and is given by

$$s_{k,i}(\rho) \rightarrow \frac{1}{\beta |r_i|^2 [1 + e_0(\rho)]^{-1} + \rho}. \quad (\text{D.3})$$

Substituting (D.3) into (D.2) and by simple algebraic manipulation, when $\rho \rightarrow 0$ we have

$$\mathbb{E} \left\{ \frac{|r_i|^2}{|r_i|^2 + c_1} \right\} = \beta, \quad (\text{D.4})$$

in which $c_1 = e_0(\rho)\rho/\beta$, is a constant related to β and the statistics of r_i . Once c_1 is obtained by solving (D.4), $e_0(\rho)$ is achieved directly as $e_0(\rho) = c_1\beta/\rho$. Consequently, we get

$$\begin{aligned} s_{k,i}(\rho) &\rightarrow \frac{1}{\beta |r_i|^2 [1 + e_0(\rho)]^{-1} + \rho} \\ &\rightarrow \frac{\rho^{-1}}{|r_i|^2 / c_1 + 1}. \end{aligned} \quad (\text{D.5})$$

The derivation for $\mathbf{S}'_k(\rho)$ is similar. According to LLN

$$\begin{aligned} \frac{1}{N} \text{Tr} \left\{ \mathbf{S}(\rho) \mathbf{S}(\rho) \right\} &\rightarrow \rho^{-2} \mathbb{E} \left\{ \frac{1}{(|r_i|^2 / c_1 + 1)^2} \right\} \triangleq \rho^{-2} c_2, \\ \frac{1}{N} \text{Tr} \left\{ \mathbf{S}(\rho) \mathbf{B} \mathbf{S}(\rho) \right\} &\rightarrow \rho^{-2} \mathbb{E} \left\{ \frac{|r_i|^2}{(|r_i|^2 / c_1 + 1)^2} \right\} \triangleq \rho^{-2} c_3, \end{aligned}$$

and thus we have

$$\begin{aligned} [\mathbf{J}(\rho)]_{kl} &\rightarrow \frac{c_2}{c_1^2 \beta K}, \quad 1 \leq k, l \leq K, \\ [\mathbf{u}(\rho)]_k &\rightarrow \rho^{-2} c_3, \quad 1 \leq k \leq K. \end{aligned} \quad (\text{D.6})$$

Substituting (D.6) into (A.2) yields

$$e'_i(\rho) \rightarrow \rho^{-2} \frac{c_3}{1 - c_2(c_1^2 \beta)^{-1}} \triangleq \rho^{-2} c_4. \quad (\text{D.7})$$

and consequently

$$\begin{aligned} s'_i(\rho) &= s_i^2(\rho) |r_i|^2 + s_i^2(\rho) \frac{1}{N} \sum_{k=1}^K \frac{\rho^{-2} c_4}{[1 + c_1 \beta / \rho]^2} \\ &\rightarrow \rho^{-2} \frac{|r_i|^2 + c_4 c_1^{-2} \beta^{-1}}{(|r_i|^2 / c_1 + 1)^2}. \end{aligned} \quad (\text{D.8})$$

Substituting (D.8) and (D.5) into (13) gives

$$\begin{aligned} \mu_1^\circ &\rightarrow e_0(\rho) = c_1 \beta \rho^{-1}, \\ \mu_2^\circ &\rightarrow \rho^{-1} \mathbb{E} \left\{ \frac{t_i r_i^*}{|r_i|^2 / c_1 + 1} \right\} \triangleq \rho^{-1} c_5, \\ \mu_3^\circ &\rightarrow \frac{1}{N} (c_1 \beta \rho^{-1})^{-2} \mathbb{E} \left\{ \rho^{-2} |t_i|^2 \frac{|r_i|^2 + c_4 c_1^{-2} \beta^{-1}}{(|r_i|^2 / c_1 + 1)^2} \right\} \triangleq \frac{c_6}{N}, \\ \mu_4^\circ &\rightarrow \frac{1}{N} (c_1 \beta \rho^{-1})^{-2} \mathbb{E} \left\{ \rho^{-2} |r_i|^2 \frac{|r_i|^2 + c_4 c_1^{-2} \beta^{-1}}{(|r_i|^2 / c_1 + 1)^2} \right\} \triangleq \frac{c_7}{N}, \\ \mu_5^\circ &\rightarrow \frac{1}{N} (c_1 \beta \rho^{-1})^{-2} \mathbb{E} \left\{ \rho^{-2} r_i t_i^* \frac{|r_i|^2 + c_4 c_1^{-2} \beta^{-1}}{(|r_i|^2 / c_1 + 1)^2} \right\} \triangleq \frac{c_8}{N}, \end{aligned} \quad (\text{D.9})$$

when $N \rightarrow \infty$ and $\rho \rightarrow 0$. By simple algebraic manipulation, we get

$$\begin{aligned} \lambda_{\text{RZF}} &\rightarrow \frac{P_{\text{T}}}{\beta c_7}, \quad g_{k,k}^\circ \rightarrow \sqrt{\lambda_{\text{RZF}}} \frac{c_5}{c_1 \beta}, \\ g_{k,i}^{\circ 2} &\rightarrow \frac{1}{N} \lambda_{\text{RZF}} \left[c_6 - \frac{2 \text{Re}\{c_5 c_8\}}{c_1 \beta} + \frac{c_7 |c_5|^2}{c_1^2 \beta^2} \right]. \end{aligned} \quad (\text{D.10})$$

and thus

$$\gamma_{\text{RZF},k} \rightarrow \frac{c_5^2}{\beta^3 c_6 c_1^2 - 2 c_1 \beta^2 \text{Re}\{c_5 c_8\} + \beta c_7 |c_5|^2}. \quad (\text{D.11})$$

APPENDIX E PROOF OF PROPOSITION 3

Before the proof of this proposition, a useful lemma is introduced first.

Lemma 5 (Lemma 11, [32]): For any semi-positive definite matrices \mathbf{A} and \mathbf{E} with dimension $N \times N$, the following inequality holds

$$\text{Tr}(\mathbf{A}\mathbf{E}) \geq \sum_{i=1}^N \lambda_i(\mathbf{A}) \lambda_{N-i+1}(\mathbf{E}),$$

where $\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq \dots \geq \lambda_N(\mathbf{A})$ are the ordered eigenvalues of \mathbf{A} .

Now we are able to prove the proposition. Rewrite the two P-Precoding matrices for P-Pre-Cal and P-Post-Cal as follows

$$\begin{aligned} \mathbf{W}_{\text{P-Post}} &= \sqrt{\lambda_{\text{P-Post}}} \tilde{\mathbf{T}}^{-1} [\mathbf{V}^H \mathbf{R}^H \mathbf{R} \mathbf{V}]^{-1} \mathbf{V}^H \mathbf{R}^H \mathbf{R} \mathbf{T}^{-1}, \\ \mathbf{W}_{\text{P-Pre}} &= \sqrt{\lambda_{\text{P-Pre}}} \tilde{\mathbf{T}}^{-1} [\mathbf{V}^H \mathbf{T}^H \mathbf{T} \mathbf{V}]^{-1} \mathbf{V}^H \mathbf{T}^H. \end{aligned} \quad (\text{E.1})$$

Denote $\mathbf{G}_{\text{P-Post}} = \tilde{\mathbf{T}}^{-1} [\mathbf{V}^H \mathbf{R}^H \mathbf{R} \mathbf{V}]^{-1} \mathbf{V}^H \mathbf{R}^H$ and $\mathbf{G}_{\text{P-Pre}} = \tilde{\mathbf{T}}^{-1} [\mathbf{V}^H \mathbf{T}^H \mathbf{T} \mathbf{V}]^{-1} \mathbf{V}^H \mathbf{T}^H$. Therefore according to (24), we have

$$\begin{aligned} \lambda_{\text{P-Post}} &= \frac{P}{\mathbb{E} \text{Tr}\{\mathbf{A}^H \mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}} \mathbf{A}\}}, \\ \lambda_{\text{P-Pre}} &= \frac{P}{\mathbb{E} \text{Tr}\{\mathbf{G}_{\text{P-Pre}}^H \mathbf{P} \mathbf{G}_{\text{P-Pre}}\}}. \end{aligned} \quad (\text{E.2})$$

Note that if we replace \mathbf{T} with \mathbf{R} in $\mathbf{G}_{\text{P-Pre}}$, $\mathbf{G}_{\text{P-Pre}}$ will be exactly the same as $\mathbf{G}_{\text{P-Post}}$. Because \mathbf{T} and \mathbf{R} have the same distribution, we have $\mathbb{E} \text{Tr}\{\mathbf{G}_{\text{P-Pre}}^H \mathbf{P} \mathbf{G}_{\text{P-Pre}}\} = \mathbb{E} \text{Tr}\{\mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}}\}$. Therefore, comparing $\lambda_{\text{P-Post}}$ and $\lambda_{\text{P-Pre}}$ is equivalent to comparing $\mathbb{E} \text{Tr}\{\mathbf{A}^H \mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}} \mathbf{A}\}$ and $\mathbb{E} \text{Tr}\{\mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}}\}$.

Using Lemma 5, we achieve

$$\begin{aligned} &\mathbb{E} \text{Tr}\{\mathbf{A}^H \mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}} \mathbf{A}\} \\ &\geq \mathbb{E} \left\{ \sum_{i=1}^N a_{N-i+1}^2 \lambda_i(\mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}}) \right\}, \end{aligned}$$

where $a_j = \frac{|r_j|}{|t_j|}$ is the magnitude of the j -th diagonal entry of \mathbf{A} . Note that when $N \rightarrow \infty$, $K \rightarrow \infty$, the empirical distribution of the eigenvalues of $\mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}}$ will converge to a deterministic distribution. Therefore

$$\begin{aligned} &\mathbb{E} \text{Tr}\{\mathbf{A}^H \mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}} \mathbf{A}\} \\ &\geq \sum_{i=1}^N \mathbb{E}\{a_{N-i+1}^2\} \mathbb{E}\{\lambda_i(\mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}})\} \\ &= \mathbb{E}\{a_j^2\} \mathbb{E} \text{Tr}\{\mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}}\}. \end{aligned} \quad (\text{E.3})$$

According to Jensen's inequality, we have $\mathbb{E}\{\frac{1}{|r_j|}\} \geq \frac{1}{\mathbb{E}\{|r_j|\}}$. Thus $\mathbb{E}\{a_j^2\} \geq 1$. Substituting this into (E.3) finally gives

$$\begin{aligned} &\mathbb{E} \text{Tr}\{\mathbf{A}^H \mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}} \mathbf{A}\} \\ &\geq \mathbb{E} \text{Tr}\{\mathbf{G}_{\text{P-Post}}^H \mathbf{P} \mathbf{G}_{\text{P-Post}}\}. \end{aligned}$$

and therefore $\lambda_{\text{P-Pre}} \geq \lambda_{\text{P-Post}}$.

REFERENCES

- [1] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker, and P. A. Hoeher, "Multiple-antenna techniques for wireless communications - a comprehensive literature survey," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 2, pp. 87–105, 2009.
- [2] J. Evans and D. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2059–2078, 2000.
- [3] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A New Approach for Mutual Information Analysis of Large Dimensional Multi-Antenna Channels," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 3987–4004, Sep. 2008.
- [4] V. Raghavan, A. M. Sayeed, and V. V. Veeravalli, "Semiunitary Precoding for Spatially Correlated MIMO Channels," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1284–1298, Mar. 2011.
- [5] R. Muharar and J. Evans, "Downlink Beamforming with Transmit-Side Channel Correlation: A Large System Analysis," in *Proc. 2011 IEEE International Conference on Communications (ICC)*. Ieee, Jun. 2011, pp. 1–5.
- [6] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [7] F. Rusek, D. Persson, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

- [8] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [9] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Linear Pre-Coding Performance in Measured Very-Large MIMO Channels," in *Proc. 2011 IEEE Vehicular Technology Conference (VTC Fall)*. IEEE, Sep. 2011, pp. 1–5.
- [10] W. Zhang, B. Du, C. Pan, and M. Chen, "Downlink SINR Distribution in Multiuser Large Scale Antenna Systems with Conjugate Beamforming," in *Proc. IEEE GlobeCom '13*, 2013.
- [11] R. C. de Lamare, "Massive MIMO Systems: Signal Processing Challenges and Future Trends," *URSI Radio Science Bulletin*, 2013.
- [12] J. Choi, D. Love, and P. Bidigare, "Downlink Training Techniques for FDD Massive MIMO Systems: Open-Loop and Closed-Loop Training with Memory," *IEEE Journal of Selected Topics in Signal Processing*, vol. PP, no. 99, 2014.
- [13] G. Smith, "A Direct Derivation of a Single-Antenna Reciprocity Relation for the Time Domain," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 6, pp. 1568–1577, Jun. 2004.
- [14] M. Guillaud, D. T. Slock, and R. Knopp, "A practical method for wireless channel reciprocity exploitation through relative calibration," *ISSPA*, pp. 403–406, 2005.
- [15] Z.-J. Liu, D.-C. Sun, J.-L. Wang, and K.-C. Yi, "Impact and compensation of I/Q imbalance on channel reciprocity of time-division-duplexing multiple-input multiple-output systems," *IET Communications*, vol. 7, no. 7, pp. 663–672, May 2013.
- [16] C. Zhang, Z. Xiao, B. Gao, L. Su, and D. Jin, "Three-Stage Treatment of TX/RX IQ Imbalance and Channel with CFO for SC-FDE Systems," *IEEE Communications Letters*, vol. 18, no. 2, pp. 297–300, February 2014.
- [17] S. Han, C. Yang, G. Wang, D. Zhu, and M. Lei, "Coordinated Multi-Point Transmission Strategies for TDD Systems with Non-Ideal Channel Reciprocity," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4256–4270, Oct. 2013.
- [18] A. Bourdoux, B. Come, and N. Khaled, "Non-reciprocal transceivers in OFDM/SDMA systems: Impact and mitigation," in *Proc. IEEE Radio and Wireless Conference, 2003*, 2003, pp. 183–186.
- [19] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative Channel Reciprocity Calibration in MIMO/TDD Systems," *Future Network and Mobile Summit*, pp. 1–10, 2010.
- [20] J. Shi, Q. Luo, and M. You, "An efficient method for enhancing TDD over the air reciprocity calibration," in *Proc. 2011 IEEE Wireless Communications and Networking Conference*. Ieee, Mar. 2011, pp. 339–344.
- [21] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical Many-Antenna Base Stations," in *Proc. Mobicom '12*. ACM Press, 2012, p. 53.
- [22] R. Rogalin, O. Y. Bursalioglu, H. C. Papadopoulos, G. Caire, and A. F. Molisch, "Hardware-impairment compensation for enabling distributed large-scale MIMO," *2013 Information Theory and Applications Workshop (ITA)*, pp. 1–10, Feb. 2013.
- [23] M. Guillaud and F. Kaltenberger, "Towards practical channel reciprocity exploitation: Relative calibration in the presence of frequency offset," in *Proc. 2013 IEEE Wireless Communications and Networking Conference (WCNC)*. Ieee, Apr. 2013, pp. 2525–2530.
- [24] R. Rogalin, O. Y. Bursalioglu, H. Papadopoulos, G. Caire, A. F. Molisch, A. Michaloliakos, V. Balan, and K. Psounis, "Scalable Synchronization and Reciprocity Calibration for Distributed Multiuser MIMO," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 1815–1831, Apr. 2014.
- [25] L. Su, C. Yang, G. Wang, and M. Lei, "Retrieving Channel Reciprocity for Coordinated Multi-Point Transmission with Joint Processing," *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1541–1553, May 2014.
- [26] F. Huang, Y. Wang, J. Geng, and D. Yang, "Antenna mismatch and calibration problem in coordinated multi-point transmission system," *IET Communications*, vol. 6, no. 3, p. 289, 2012.
- [27] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels Under Limited Feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.
- [28] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [29] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot Contamination and Precoding in Multi-Cell TDD Systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [30] C. Peel, B. Hochwald, and A. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication-Part I: Channel Inversion and Regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, Jan. 2005.
- [31] X. Zhang, *Matrix Analysis and Applications*. Beijing: Tsinghua University Press, 2004.
- [32] D. Palomar, J. Cioffi, and M. Lagunas, "Joint TX-RX Beamforming Design for Multicarrier MIMO Channels: A Unified Framework for Convex Optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.