

Blind Channel Estimation for Downlink Massive MIMO Systems with Imperfect Channel Reciprocity

Ribhu Chopra, Chandra R. Murthy, Himal A. Suraweera, and Erik G. Larsson

Abstract—We consider the performance of time-division duplex (TDD) massive multiple-input multiple-output (MIMO) with imperfect calibration of the transmit and receive radio frequency chains. By deriving the achievable signal-to-interference-plus-noise ratio (SINR) and the per-user bit error rate (BER) for constant modulus constellations, we establish that with linear precoding, reciprocity imperfections can result in substantial reduction of the array gain. To mitigate this loss, we propose an algorithm for blind estimation of the effective channel gain at each user. We show that given sufficiently many downlink symbols, this blind channel estimation algorithm restores the array gain. In addition, the proposed blind gain estimation algorithm can improve performance even under perfect reciprocity compared to standard hardening-based receivers. Following this, we derive the BERs for non-constant modulus constellations, viz. M -PAM and M -QAM. We corroborate all our derived results using numerical simulations.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) wireless technology with many antennas, known as large MIMO [1] for point-to-point communications, and as massive MIMO [2] for multiuser communications, are a key component of emerging new wireless standards [2]–[5]. Effectively, with multiple antennas, different data streams see quasi-orthogonal channels. Channel hardening also results, meaning that the effective channel gain seen by each user is substantially equal to a deterministic constant, which facilitates decoding at the user equipment (UE) without instantaneous channel state information (CSI). All these advantages of massive MIMO are contingent upon the availability of accurate CSI at the BS [6]. CSI errors at the BS due to noise, pilot contamination [7] and channel aging [2], [8]–[11] generally decrease performance.

The use of a large number of antennas at the BS makes the process of training in a massive MIMO system in the downlink prohibitively complex [6], [12]. Therefore, it is common to assume time division duplexed (TDD) operation with uplink training, resulting in reduced CSI acquisition times. However, while operating in TDD, the system is assumed to work under the assumption of perfect channel reciprocity [13]. However, in practice, the transmit and receive radio frequency (RF) chains at both the BS and the UEs are not perfectly calibrated [14], violating the channel reciprocity assumption. In this paper, we analytically study the detrimental effects of imperfect

reciprocity calibration on the achievable downlink rates in a massive MIMO system. We also present a blind channel estimation algorithm to mitigate these effects, and analyze its performance.

A. Related Work

The literature on reciprocity imperfections in massive MIMO can be broadly classified into two categories: techniques for reciprocity calibration [14]–[22], and performance analysis in the presence of reciprocity calibration errors [20], [23]–[25]. While these algorithms provide reasonably accurate calibration among the uplink and downlink RF chains, residual calibration errors still exist that can limit the system performance. Therefore, it is necessary to account for and analyze the performance of a massive MIMO system in the presence of reciprocity calibration errors.

The effects of residual reciprocity calibration errors have been modeled as both additive [24] and multiplicative [20], [23] errors in the channel coefficients available at the BS. The multiplicative error model is preferable due to its ability to capture RF mismatches. The effects of multiplicative reciprocity imperfections on the achievable signal to interference-plus-noise ratios (SINRs) of massive MIMO systems under linear precoding schemes have been derived in [23] and [25]. Specifically, in [25], the reciprocity errors are modeled as uniformly distributed random variables, and in [23], they are modeled as truncated Gaussian random variables. The latter model has some empirical evidence from the results available in [26]. The authors in [23] show that reciprocity imperfections lead to a saturation of the achievable SINRs for a system employing linear precoding. However, the current state-of-the-art on modeling the effects of reciprocity imperfections in massive MIMO implicitly assumes the effective downlink channel coefficients are perfectly known at the UE, which is normally not the case unless additional resources are expended on downlink pilots. Furthermore, the above-mentioned papers do not consider the effect of reciprocity imperfections on the BER for finite constellations.

B. Contributions

In this paper, we consider a typical massive MIMO system, where the CSI is acquired at the BS via uplink training, and the estimated CSI is used to precode data for downlink transmission. The UEs use the expected effective downlink channel gain for data detection, i.e., there is no downlink training [2]. We analyze the effect of reciprocity calibration errors in terms of the achievable SINR and the BER, for

R. Chopra is with the Dept. of EEE, IIT Guwahati, India. (ribhufec@iitg.ac.in). C. R. Murthy is with the Dept. of ECE, IISc Bangalore, India. (cmurthy@iisc.ac.in). H. A. Suraweera is with the Dept. of EEE, University of Peradeniya, Peradeniya 20400, Sri Lanka (himal@ee.pdn.ac.lk). E. G. Larsson is with the Dept. of Electrical Engineering (ISY), Linköping University, 58183 Linköping, Sweden (erik.g.larsson@liu.se).

both matched filter (MF) precoding (also known as maximal ratio transmission [27]) and regularized zero-forcing (RZF) precoding. Our approach differs from previous studies on the performance analysis of massive MIMO systems with reciprocity calibration imperfections [23], [28], where the exact knowledge of the downlink channel is assumed to be available at the UEs. We show that, in the absence of downlink CSI, reciprocity imperfections can result in severe loss in the achievable SINR. These losses can be mitigated by acquiring CSI at the UEs. To this end, we observe that the effective downlink channel of a precoded communication system is close to a positive real number. In the context of wireless sensor networks, it has been shown that, for such systems, the received data symbols can be used to perform blind channel estimation [29], [30]. Based on this, we present an algorithm for blind channel estimation at the UEs of a massive MIMO system and analyze its performance. The advantage of this approach is that there is no downlink pilot overhead, similar to the conventional massive MIMO systems. We summarize our main contributions as follows:

- 1) We model reciprocity calibration errors as multiplicative noise and derive expressions for the achievable SINR and the BER of a downlink massive MIMO system with binary phase shift keying (BPSK) modulation and linear precoding. The analysis requires the derivation of the deterministic equivalent of the signal and interference terms in the presence of random multiplicative noise, which is not available in the literature. The analysis is particularly involved under RZF precoding because of the joint dependence of the precoding vectors on the channels of all the users. (See Sections III and IV).
- 2) Motivated by the significant performance loss in the absence of CSI at the UEs, and based on the properties of the effective downlink channel under calibration imperfections, we develop and analyze a new blind algorithm for effective channel gain estimation at the UEs. (See Section V).
- 3) Using the properties of the effective estimated channel at the UE, we derive the achievable rate and BER for higher order constellations under estimated CSI at the UEs. (See Section V-C).
- 4) Via detailed simulations, we study the impact of system parameters such as the data and pilot signal-to-noise ratios (SNR), the number of BS antennas, and the maximum tolerable levels of reciprocity imperfections on the achievable rate and BER performance in a massive MIMO system. (See Section VI).

The key takeaway from this study is that in the absence of downlink CSI at the UE, reciprocity calibration imperfections can lead to a significant performance degradation in massive MIMO systems. However, this loss can be mitigated to a large extent by the use of appropriate blind channel estimation algorithms at the UEs.

Notation: Boldface lowercase and uppercase letters represent vectors and matrices, respectively. The k th column of the matrix \mathbf{A} is denoted by \mathbf{a}_k . $(\cdot)^H$ represents the conjugate transpose of a vector or a matrix. $\|\cdot\|_2$ and $\|\cdot\|_F$ respectively

represent the ℓ_2 -norm of a vector and the Frobenius norm of a matrix. $E[\cdot]$ and $\text{var}(\cdot)$ represent the mean and variance of a random variable. $Q(\cdot)$ represents the Gaussian Q function.

II. SYSTEM AND CHANNEL MODEL

A. System Model

We consider a single cell massive MIMO system with an N antenna (indexed as $i \in \{1, 2, \dots, N\}$) BS, and K single antenna UEs (indexed as $k \in \{1, \dots, K\}$), with $K \leq N$. We assume a TDD system with training in the uplink and no downlink training [2]. The data transmission frame, consisting of a total of T channel uses, is divided into three sub-frames, viz. uplink training, uplink data transmission, and downlink data transmission. We assume the channel to be block fading, and the frame duration T to be smaller than the channel coherence time. During the uplink training phase, the UEs transmit orthogonal pilots for estimating the uplink channels at the BS. Following this, during the uplink data transmission phase, the UEs transmit their individual data streams to the BS over the respective channels, and the BS decodes this using linear detectors obtained using the available channel estimates. During the downlink phase, the BS uses the available channel estimates to construct linear precoders and transmits K independent streams of data to the K UEs. In this paper, we focus on the performance achievable in the downlink sub-frame of the protocol.

The fading component of the uplink channel from the k th UE to the i th BS antenna is denoted by $h_{u,ik} = u_{t,k} b_{r,i} h_{ik}$, where $u_{t,k}$ is the complex valued gain associated with the transmit RF chain of the k th user, $b_{r,i}$ is the complex valued gain associated with the receive RF chain of the i th BS antenna, and h_{ik} is the fast fading component of the channel between the k th UE and the i th BS antenna. The composite uplink channel $h_{u,ik}$ is modeled as a zero mean circularly symmetric complex Gaussian (ZMCSG) random variable (rv). Similarly, the downlink channel between the i th BS antenna and the k th UE is given as, $h_{d,ik} = u_{r,k} b_{t,i} h_{ik}$, where $u_{r,k}$ is the complex valued gain associated with the receive RF chain of the k th user, and $b_{t,i}$ is the complex valued gain associated with the transmit RF chain of the i th BS antenna.

B. Reciprocity Calibration Error Model

We assume that both the BS and the UEs perform over-the-air calibration of their transmit and receive RF chains using one of the techniques in [17]. Note that, since these gains can be assumed to remain static over a long time duration, the control overhead involved is negligible. After the RF chain calibration, the BS assumes the uplink and downlink channels to be reciprocal, and consequently equal. That is, the BS assumes that the estimated downlink channel between the i th BS antenna and the k th UE, $\hat{h}_{d,ik}$ is equal to the available uplink channel estimate $\hat{h}_{u,ik}$.

However, the estimates of the RF chain gains at both the BS and the UEs are imperfect, which results in erroneous

estimates of the channel coefficients beyond the uplink channel estimation error. In order to model the additional error due to the imperfect calibration, we consider the quantities

$$\delta_{u,k} e^{j\phi_{u,k}} \triangleq \frac{u_{r,k}}{u_{t,k}} \text{ and } \delta_{b,i} e^{j\phi_{b,i}} \triangleq \frac{b_{t,i}}{b_{r,i}}. \quad (1)$$

Here, $\delta_{u,k}$, $\delta_{b,i}$ are the magnitude calibration errors, and $\phi_{u,k}$, $\phi_{b,i}$ are the phase calibration errors at the k th UE and i th BS antenna, respectively. We model both the magnitude and phase calibration errors to be independent and identically distributed (i.i.d.) rvs with finite moments. Further, we assume that the probability density function (pdf) of the phase calibration error is symmetric about zero. This is similar to [23] and [25], which consider particular distributions such as the truncated Gaussian or uniform distribution.

Now, letting $\mathbf{b}_r = [b_{r,1}, \dots, b_{r,N}]^T$, $\mathbf{u}_r = [u_{r,1}, \dots, u_{r,K}]^T$, $\mathbf{b}_t = [b_{t,1}, \dots, b_{t,N}]^T$, $\mathbf{u}_t = [u_{t,1}, \dots, u_{t,K}]^T$, the uplink channel matrix can be written as $\mathbf{H}_u = \text{diag}(\mathbf{b}_r) \mathbf{H} \text{diag}(\mathbf{u}_t) \in \mathbb{C}^{N \times K}$, with h_{ik} being the (ik) th entry of \mathbf{H} . Similarly, the transpose of the downlink channel matrix can be expressed as

$$\mathbf{H}_d = \text{diag}(\mathbf{b}_t) (\text{diag}(\mathbf{b}_r))^{-1} \mathbf{H}_u (\text{diag}(\mathbf{u}_t))^{-1} \text{diag}(\mathbf{u}_r). \quad (2)$$

In order to set the stage, consider the case where the uplink channel is available at the BS, i.e., $\hat{\mathbf{H}}_u = \mathbf{H}_u$. Then, defining $\mathbf{D}_b = \text{diag}([\delta_{b,1} e^{j\phi_{b,1}}, \dots, \delta_{b,N} e^{j\phi_{b,N}}]^T)$, $\mathbf{D}_u = \text{diag}([\delta_{u,1} e^{j\phi_{u,1}}, \dots, \delta_{u,K} e^{j\phi_{u,K}}]^T)$, it is easy to see that $\hat{\mathbf{H}}_d$ is related to \mathbf{H}_d as

$$\mathbf{H}_d = \mathbf{D}_b \hat{\mathbf{H}}_d \mathbf{D}_u. \quad (3)$$

This allows us to isolate the multiplicative effect of reciprocity calibrations errors from the additive errors due to channel estimation errors and pilot contamination.

C. Channel Calibration and Signaling Scheme

Now, we bring in the effect of uplink channel estimation error. During the uplink training phase, the BS uses the pilots transmitted by the UEs to obtain the minimum mean squared error (MMSE) estimate of the uplink channel matrix. Letting $\hat{h}_{u,ik}$ be the MMSE estimate of $h_{u,ik}$, it can be shown that

$$h_{u,ik} = a_k \hat{h}_{u,ik} + \bar{a}_k \tilde{h}_{u,ik}, \quad (4)$$

where $\tilde{h}_{u,ik} \sim \mathcal{CN}(0, 1)$ is the channel estimation error such that $E[\hat{h}_{u,ik} \tilde{h}_{u,ik}^*] = 0$, $a_k = \sqrt{\frac{\beta_k \mathcal{E}_{p,k}}{\beta_k \mathcal{E}_{p,k} + N_0}}$, with $\mathcal{E}_{p,k}$ being the pilot power transmitted by the k th user, β_k being the macroscopic fading coefficient for the k th user, N_0 being the variance of the AWGN, and $\bar{x} \triangleq \sqrt{1 - |x|^2}$ for any variable x . Defining $\mathbf{a} \triangleq [a_1, \dots, a_K]^T$ and $\mathbf{A} = \text{diag}(\mathbf{a})$, the overall uplink channel matrix can be written as

$$\mathbf{H}_u = \hat{\mathbf{H}}_u \mathbf{A} + \tilde{\mathbf{H}}_u \bar{\mathbf{A}}. \quad (5)$$

Combining (2), (3) and (5), we can write

$$\mathbf{H}_d = \mathbf{D}_b \hat{\mathbf{H}}_d \mathbf{D}_u \mathbf{A} + \tilde{\mathbf{H}}_d \bar{\mathbf{A}}, \quad (6)$$

with $\tilde{h}_{d,ik} \sim \mathcal{CN}(0, 1)$ being the i.i.d entries of $\tilde{\mathbf{H}}_d$. Thus, the downlink channel estimate differs from the true channel due

to both the additive noise during the training phase as well as the multiplicative error due to calibration imperfections.

The BS uses $\hat{\mathbf{H}}_d$ to design its combining matrix for uplink data reception and the precoding matrix for downlink data transmission. Now, if the signal to be transmitted to the k th user at the n th instant is $s_k[n]$, and $\mathcal{E}_{d,s,k}$ is its allocated energy, then, defining $\mathbf{s}[n] \triangleq [s_1[n], \dots, s_K[n]]^T$, $\mathcal{E} \triangleq [\mathcal{E}_{d,s,1}, \dots, \mathcal{E}_{d,s,K}]^T$, the signal transmitted by the BS at the n th instant is

$$\mathbf{x}[n] = \lambda \mathbf{V} \text{diag}(\sqrt{\mathcal{E}}) \mathbf{s}[n]. \quad (7)$$

Here, $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_K] \in \mathbb{C}^{N \times K}$ is the precoding matrix at the BS and the power normalization factor λ is such that

$$\lambda^2 E[\|\mathbf{V} \text{diag}(\sqrt{\mathcal{E}}) \mathbf{s}[n]\|_2^2] = \mathcal{E}_t, \quad (8)$$

where \mathcal{E}_t is the total downlink energy at the BS.

The signal received at the k th user is therefore given as,

$$\begin{aligned} y_k[n] &= \sqrt{\beta_k} \mathbf{h}_{d,k}^T \mathbf{x}[n] + \sqrt{N_0} w_k[n] \\ &= \lambda \sum_{l=1}^K \sqrt{\beta_k \mathcal{E}_{d,s,l}} \mathbf{h}_{d,k}^T \mathbf{v}_l s_l[n] + \sqrt{N_0} w_k[n], \end{aligned} \quad (9)$$

with $\mathbf{h}_{d,k}$ denoting the k th column of \mathbf{H}_d and $w_k[n] \sim \mathcal{CN}(0, 1)$ being the AWGN at the k th user. We see that the effective scalar channel coefficient for the k th user's symbol, $s_k[n]$, is $g_{kk} = \sqrt{\beta_k \mathcal{E}_{d,s,k}} \mathbf{h}_{d,k}^T \mathbf{v}_k$. In the absence of any forward link training, the UEs use $\hat{g}_{kk} = \sqrt{\beta_k \mathcal{E}_{d,s,k}} E[\mathbf{h}_{d,k}^T \mathbf{v}_k]$ as the effective scalar channel for downlink data. As shown in the following sections, this leads to significant performance degradation in case of inaccurate channel calibration.

III. PERFORMANCE UNDER MFP

In this section, we discuss the effects of inaccurate reciprocity calibration on the performance of a massive MIMO system employing MFP. Due to the large system dimensions, the effective channel and interference coefficients can be accurately approximated using their deterministic equivalents (DEs). Each UE can then use the DE of the downlink channel coefficient to decode its intended data stream. This technique, known as the use and forget analysis [2], works well for massive MIMO systems with no calibration errors. However, in the absence of the knowledge of calibration errors at the UEs, the effective channel coefficient will be different from the DE calculated at the UEs under the assumption of perfect reciprocity calibration, leading to a loss in performance. Here, we quantify this performance loss by using the achievable SINR and BER as metrics.

Under MFP, the precoding vector for the l th stream is $\mathbf{v}_l = \hat{\mathbf{h}}_{d,l}^*$. Consequently, with $E[\|\mathbf{s}[n]\|_2^2] = 1$,

$$\lambda^2 = \frac{\mathcal{E}_t}{E[\text{Tr}\{\hat{\mathbf{H}}_d^* \text{diag}(\mathcal{E}) \hat{\mathbf{H}}_d^T\}]} = \frac{\mathcal{E}_t}{\sum_{l=1}^K \mathcal{E}_{d,s,l}}. \quad (10)$$

Also, the received signal at the k th user can be written as

$$y_k[n] = N \hat{g}_{kk} s_k[n] + N \tilde{g}_{kk} s_k[n]$$

$$+ N \sum_{\substack{l=1 \\ l \neq k}}^K g_{kl} s_l[n] + \sqrt{N_0} w_k[n], \quad (11)$$

where $\hat{g}_{kk} = \frac{\lambda}{N} a_k d_{u,k} \sqrt{\beta_k \mathcal{E}_{d,s,k}} \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \hat{\mathbf{h}}_{d,k}^*$, $\tilde{g}_{kk} = \frac{\lambda}{N} \bar{a}_k \sqrt{\beta_k \mathcal{E}_{d,s,k}} \tilde{\mathbf{h}}_{d,k}^T \hat{\mathbf{h}}_{d,k}^*$, and $g_{kl} = \frac{\lambda}{N} \sqrt{\beta_k \mathcal{E}_{d,s,l}} \hat{\mathbf{h}}_{d,k}^T \hat{\mathbf{h}}_{d,l}^*$ are the effective channel coefficients for the desired signal, the self interference of the k th user, and the interference channel between the streams of the k th and l th users, respectively. Here, $d_{u,k}$ denotes the k th diagonal element of \mathbf{D}_u .

Since there is no downlink training, the UE uses the expected value of the channel assuming perfect calibration as the effective channel coefficient, $\dot{g}_{kk} = \frac{\lambda}{N} a_k \sqrt{\beta_k \mathcal{E}_{d,s,k}} E[\hat{\mathbf{h}}_{d,k}^T \hat{\mathbf{h}}_{d,k}^*]$, for data decoding. It is easy to show that this reduces to

$$\dot{g}_{kk} = \lambda a_k \sqrt{\beta_k \mathcal{E}_{d,s,k}}. \quad (12)$$

The received signal is therefore expressed as

$$y_k[n] = N \dot{g}_{kk} s_k[n] + N (\hat{g}_{kk} - \dot{g}_{kk}) s_k[n] + N \tilde{g}_{kk} s_k[n] + N \sum_{\substack{l=1 \\ l \neq k}}^K g_{kl} s_l[n] + \sqrt{N_0} w_k[n], \quad (13)$$

with the first term corresponding to the desired signal, and all the other terms to noise and interference. Consequently, under the assumption of the data streams to different users being independent of each other and the channel coefficients, the instantaneous SINR for the k th user can be written as,

$$\gamma_k = \frac{N^2 |\dot{g}_{kk}|^2}{N^2 (|\hat{g}_{kk} - \dot{g}_{kk}|^2 + |\tilde{g}_{kk}|^2 + \sum_{\substack{l=1 \\ l \neq k}}^K |g_{kl}|^2) + N_0}. \quad (14)$$

When the BS is equipped with a large number of antennas, all the effective channel coefficients converge to their DEs [8, Lemma 1]. Also, since the calibration errors are assumed to be identically distributed across the RF chains and the phase calibration errors are symmetric, these DEs can be written as

$$\hat{g}_{kk} - \lambda d_{u,k} a_k \sqrt{\beta_k \mathcal{E}_{d,s,k}} E[\delta_b \cos(\phi_b)] \xrightarrow{\text{a.s.}} 0, \quad (15)$$

$$|\tilde{g}_{kk}|^2 - \frac{\lambda^2}{N} \bar{a}_k^2 \beta_k \mathcal{E}_{d,s,k} \xrightarrow{\text{a.s.}} 0, \quad (16)$$

$$|g_{kl}|^2 - \frac{\lambda^2}{N} \beta_k \mathcal{E}_{d,s,l} \xrightarrow{\text{a.s.}} 0, \quad k \neq l. \quad (17)$$

Under perfect reciprocity calibration, i.e., $\delta_u = \delta_b = 1$, $\phi_u = \phi_b = 0$, we see that the DE of \hat{g}_{kk} in (15) is the same as the channel estimate used at the receiver, \dot{g}_{kk} , given in (12). Thus, the receiver's estimate of the channel is asymptotically unbiased. However, in the presence of calibration errors, \hat{g}_{kk} is a biased estimate of \dot{g}_{kk} .

Substituting these into (14), and defining $\mathcal{F}_{d,s,k} = \mathcal{E}_{d,s,k} \frac{\mathcal{E}_t}{\sum_{l=1}^K \mathcal{E}_{d,s,l}}$, we obtain

$$\gamma_k = \frac{N a_k^2 \beta_k \mathcal{F}_{d,s,k}}{\bar{a}_k^2 \beta_k \mathcal{F}_{d,s,k} + \beta_k \sum_{\substack{l=1 \\ l \neq k}}^K \mathcal{F}_{d,s,l} + N_0} \quad (18)$$

for perfect reciprocity calibrations, and (19) in the presence of calibration errors, and conditioned on a particular realization of

the calibration error $\delta_{u,k}$ at the UE. Note that, the expectation operator here represents the spatial average of the calibration error across all the BS antennas, for a particular realization of the calibration errors. On comparing (18) with (19), we observe that the reciprocity imperfections result in an additional interference term in the denominator that scales linearly with the number of antennas. The achievable downlink rate for the k th user in either case can be written as

$$R_k = \frac{T-K}{T} \log_2(1 + \gamma_k). \quad (20)$$

Now, in the presence of imperfect reciprocity calibration, with an increasing number of BS antennas, the achievable SINR with MFP saturates to

$$\lim_{N \rightarrow \infty} \gamma_k = \frac{1}{|1 - d_{u,k} E[\delta_b \cos(\phi_b)]|^2}, \quad (21)$$

thereby leading to the suppression of the array gain of massive MIMO and a saturation of the achievable rate.

Using (18) and (19), we can calculate the BER for any given signal constellation. However, for constant modulus constellations such as M-PSK, one can do slightly better by noting that maximum likelihood (ML) decoding only requires the phase of the received signal for data detection; the magnitude estimate is not necessary. For example, considering BPSK, the probability of error for the k th user with and without reciprocity imperfections are given by

$$P_{e,k} = Q \left(\sqrt{\frac{N a_k^2 \beta_k \mathcal{F}_{d,s,k} E^2[\delta_b \cos(\phi_b)]}{\bar{a}_k^2 \beta_k \mathcal{F}_{d,s,k} + \beta_k \sum_{\substack{l=1 \\ l \neq k}}^K \mathcal{F}_{d,s,l} + N_0}} \delta_{u,k} \cos(\phi_{u,k}) \right), \quad (22)$$

and

$$P_{e,k} = Q \left(\sqrt{\frac{N a_k^2 \beta_k \mathcal{F}_{d,s,k}}{\bar{a}_k^2 \beta_k \mathcal{F}_{d,s,k} + \beta_k \sum_{\substack{l=1 \\ l \neq k}}^K \mathcal{F}_{d,s,l} + N_0}} \right), \quad (23)$$

respectively. In this case, we see that calibration imperfections only affect the numerator of the BER expression, thus preserving the array gain advantage of massive MIMO.

We next examine the performance of an RZF beamforming system under calibration imperfections.

IV. PERFORMANCE ANALYSIS WITH RZF PRECODING

As before, we assume that the UE uses the expected effective channel under accurate reciprocity calibration for data decoding. Using DE analysis, we show that in addition to a bias in the effective channel coefficient of the desired signal, calibration imperfections also impair the interference canceling ability of RZF precoding. Hence, calibration imperfections affect RZF precoding more than MFP.

1) *Power Control*: The RZF precoding matrix, $\mathbf{V} = \mathbf{Q}^{-1} \hat{\mathbf{H}}_d^*$, where $\mathbf{Q} = (\hat{\mathbf{H}}_d^* \text{diag}(\mathcal{E}) \hat{\mathbf{H}}_d^T + \epsilon \mathbf{I}_N)$, and ϵ is a regularization parameter [9]. The precoding vector for the l th data stream is $\mathbf{v}_l = (\hat{\mathbf{H}}_d^* \text{diag}(\mathcal{E}) \hat{\mathbf{H}}_d^T + \epsilon \mathbf{I}_N)^{-1} \hat{\mathbf{h}}_{d,l}^*$.

$$\gamma_k = \frac{Na_k^2 \beta_k \mathcal{F}_{d,s,k}}{Na_k^2 \beta_k \mathcal{F}_{d,s,k} |1 - \delta_{u,k} E[\delta_b \cos(\phi_b)]|^2 + \bar{a}_k^2 \beta_k \mathcal{F}_{d,s,k} + \beta_k \sum_{\substack{l=1 \\ l \neq k}}^K \mathcal{F}_{d,s,l} + N_0} \quad (19)$$

Invoking the energy constraint from (8), we obtain

$$\begin{aligned} \lambda^2 &= \frac{\mathcal{E}_t}{E[\text{Tr}\{\mathbf{Q}^{-1} \hat{\mathbf{H}}_d^* \text{diag}(\mathcal{E}) \hat{\mathbf{H}}_d^T \mathbf{Q}^{-1}\}]} \\ &= \frac{\mathcal{E}_t}{E[\text{Tr}\{\mathbf{Q}^{-1}\} - \epsilon \text{Tr}\{\mathbf{Q}^{-2}\}]} \end{aligned} \quad (24)$$

Using [9, Theorem 4], it can be shown that the DE of λ^2 reduces to

$$\lambda^2 - \frac{\mathcal{E}_t}{\text{Tr}\{\mathbf{T}(\epsilon)\} - \text{Tr}\{\epsilon \mathbf{T}^2(\epsilon)\}} \xrightarrow{\text{a.s.}} 0. \quad (25)$$

with

$$\mathbf{T}(\epsilon) = \left(\sum_{k=1}^K \frac{\mathcal{E}_{d,s,k}}{1 + e_k(\epsilon)} + \epsilon \right)^{-1} \mathbf{I}_N, \quad (26)$$

where $e_k(\epsilon) = \lim_{N \rightarrow \infty} e_k^n(\epsilon)$ can be calculated iteratively as

$$e_k^n(\epsilon) = N \mathcal{E}_{d,s,k} \left(\sum_{l=1}^K \frac{\mathcal{E}_{d,s,l}}{1 + e_l^{n-1}(\epsilon)} + \epsilon \right)^{-1}, \quad (27)$$

with the initialization $e_k^0(\epsilon) = \frac{1}{\epsilon}$.

2) *Received Signal Model*: The signal received at the k th user can be expressed as

$$\begin{aligned} y_k[n] &= N \hat{g}_{kk} s_k[n] + N \sum_{\substack{l=1 \\ l \neq k}}^K \hat{g}_{kl} s_l[n] \\ &\quad + N \sum_{l=1}^K \tilde{g}_{kl} s_l[n] + \sqrt{N_0} w_k[n], \end{aligned} \quad (28)$$

where $\hat{g}_{kk} = \frac{\lambda}{N} a_k d_{u,k} \sqrt{\beta_k \mathcal{E}_{d,s,k}} \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \mathbf{Q}^{-1} \hat{\mathbf{h}}_{d,k}^*$ denotes the effective channel for the desired signal, $\hat{g}_{kl} = \frac{\lambda}{N} a_l d_{u,k} \sqrt{\beta_k \mathcal{E}_{d,s,l}} \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \mathbf{Q}^{-1} \hat{\mathbf{h}}_{d,l}^*$ denotes the effective inter-user interference channel between the l th data stream and the k th user, and $\tilde{g}_{kl} = \frac{\lambda}{N} \bar{a}_l \sqrt{\beta_k \mathcal{E}_{d,s,k}} \hat{\mathbf{h}}_{d,k}^T \mathbf{Q}^{-1} \hat{\mathbf{h}}_{d,l}^*$ denotes the interference channel due to CSI acquisition errors.

3) *SINR Expressions*: Based on the above, we show in Appendix A that the achievable SINRs with and without reciprocity calibration errors can be expressed as

$$\begin{aligned} \gamma_k &= \frac{N^2 a_k^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{k,\mathbf{I}}^2}{(1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})^2}}{N^2 |\hat{g}_{kk} - \dot{g}_{kk}|^2 + N^2 \sum_{\substack{l=1 \\ l \neq k}}^K |\hat{g}_{kl}|^2 + N^2 \sum_{l=1}^K |\tilde{g}_{kl}|^2 + N_0}, \end{aligned} \quad (29)$$

and

$$\gamma_k = \frac{N^2 a_k^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{k,\mathbf{I}}^2}{(1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})^2}}{N^2 \sum_{\substack{l=1 \\ l \neq k}}^K |\hat{g}_{kl}|^2 + N^2 \sum_{l=1}^K |\tilde{g}_{kl}|^2 + N_0}, \quad (30)$$

respectively. In (29) and (30), $\mu_{k,\mathbf{B}} \triangleq \hat{\mathbf{h}}_{d,k}^T \mathbf{B} \mathbf{Q}^{-1} \hat{\mathbf{h}}_{d,k}^*$ for a matrix \mathbf{B} , and,

$$\begin{aligned} |\hat{g}_{kk} - \dot{g}_{kk}|^2 &= a_k^2 \frac{\lambda^2}{N^2} \beta_k \mathcal{E}_{d,s,k} \frac{1}{(1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})^2} \\ &\quad \times (\mu_{k,\mathbf{I}}^2 + \delta_{u,k}^2 \mu_{k,\mathbf{D}_b}^2 - 2 \Re\{d_{u,k} \mu_{k,\mathbf{D}_b} \mu_{k,\mathbf{I}}\}). \end{aligned} \quad (31)$$

It is shown in Appendix A that the DEs of μ_{k,\mathbf{D}_b} and $\mu_{k,\mathbf{I}}$ reduce to

$$\mu_{k,\mathbf{D}_b} - E[\delta_b \cos \phi_b] \text{tr}\{\dot{\mathbf{T}}_k(\epsilon)\} \xrightarrow{\text{a.s.}} 0, \quad (32)$$

$$\mu_{k,\mathbf{I}} - \text{tr}\{\dot{\mathbf{T}}_k(\epsilon)\} \xrightarrow{\text{a.s.}} 0, \quad (33)$$

and that of $|\tilde{g}_{kl}|^2$ is,

$$|\tilde{g}_{kl}|^2 = \bar{a}_l^2 \frac{\lambda^2}{N^2} \beta_k \mathcal{E}_{d,s,k} \frac{|\nu_{kl}|^2}{(1 + \mathcal{E}_{d,s,k} \mu_{l,\mathbf{I}})^2}, \quad (34)$$

such that,

$$|\nu_{kl}|^2 - \text{tr}\{\dot{\mathbf{T}}_l^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0, \quad (35)$$

$$\dot{\mathbf{T}}_k(\epsilon) = \left(\sum_{\substack{l=1 \\ l \neq k}}^K \frac{\mathcal{E}_{d,s,l}}{1 + \dot{e}_{k,l}(\epsilon)} + \epsilon \right)^{-1} \mathbf{I}_N, \quad (36)$$

and $\dot{e}_{k,l}(\epsilon)$ can be evaluated by iterating the following equation over n , starting at $n = 1$, with the initialization $\dot{e}_{k,l}^0(\epsilon) = \frac{1}{\epsilon}$:

$$\dot{e}_{k,l}^n(\epsilon) = N \mathcal{E}_{d,s,l} \left(\sum_{\substack{m=1 \\ m \neq k}}^K \frac{\mathcal{E}_{d,s,m}}{1 + \dot{e}_{k,m}^{n-1}(\epsilon)} + \epsilon \right)^{-1}. \quad (37)$$

Similarly,

$$\bar{g}_{kk} = a_k \frac{\lambda}{N} \sqrt{\beta_k \mathcal{E}_{d,s,k}} \frac{\mu_{k,\mathbf{I}}}{1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}}}, \quad (38)$$

and

$$\begin{aligned} |\hat{g}_{kl}|^2 &= \delta_{u,k}^2 a_l^2 \frac{\lambda^2}{N^2} (\beta_k \mathcal{E}_{d,s,l}) \left(\frac{1}{1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right)^2 \\ &\quad \times \left(\dot{\mu}_{kl,\mathbf{D}_b}^2 + \mathcal{E}_{d,s,k}^2 \frac{\dot{\mu}_{kk,\mathbf{D}_b}^2 \dot{\mu}_{kl,\mathbf{I}}^2}{(1 + \mathcal{E}_{d,s,k} \dot{\mu}_{kk,\mathbf{I}})^2} \right. \\ &\quad \left. - 2 \mathcal{E}_{d,s,k} \frac{\Re\{\dot{\mu}_{kk,\mathbf{D}_b} \dot{\mu}_{kl,\mathbf{I}} \dot{\mu}_{kl,\mathbf{D}_b}^*\}}{(1 + \mathcal{E}_{d,s,k} \dot{\mu}_{kk,\mathbf{I}})} \right), \end{aligned} \quad (39)$$

in the presence of reciprocity calibration errors, and,

$$\begin{aligned} |\hat{g}_{kl}|^2 &= a_l^2 \frac{\lambda^2}{N^2} (\beta_k \mathcal{E}_{d,s,l}) \left(\frac{1}{1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right)^2 \\ &\quad \times \left(\dot{\mu}_{kl,\mathbf{I}}^2 + \frac{\mathcal{E}_{d,s,k}^2 \dot{\mu}_{kk,\mathbf{I}}^2 \dot{\mu}_{kl,\mathbf{I}}^2}{(1 + \mathcal{E}_{d,s,k} \dot{\mu}_{kk,\mathbf{I}})^2} - \frac{2 \mathcal{E}_{d,s,k} \dot{\mu}_{kk,\mathbf{I}} |\dot{\mu}_{kl,\mathbf{I}}|^2}{(1 + \mathcal{E}_{d,s,k} \dot{\mu}_{kk,\mathbf{I}})} \right), \end{aligned} \quad (40)$$

with perfect reciprocity calibration, wherein, $\dot{\mu}_{kk,\mathbf{D}_b} =$

$E[\delta_b \cos \phi_b] \text{tr}\{\ddot{\mathbf{T}}_{kl}(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $\dot{\mu}_{kk,\mathbf{I}} - \text{tr}\{\ddot{\mathbf{T}}_{kl}(\epsilon)\} \xrightarrow{\text{a.s.}} 0$,
 $\dot{\mu}_{kl,\mathbf{I}} = \hat{\mathbf{h}}_{d,k}^T \mathbf{Q}_{kl}^{-1} \hat{\mathbf{h}}_{d,l}^* \xrightarrow{\text{a.s.}} 0$, $\dot{\mu}_{kl,\mathbf{I}}^2 - \text{Tr}\{\ddot{\mathbf{T}}_{kl}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$,
 $|\dot{\mu}_{kl,\mathbf{D}_b}|^2 - E^2[\delta_b \cos \phi_b] \text{Tr}\{\ddot{\mathbf{T}}_{kl}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, and $\dot{\mu}_{kl,\mathbf{D}_b} \dot{\mu}_{kl,\mathbf{I}}^* - E[\delta_b \cos \phi_b] \text{Tr}\{\ddot{\mathbf{T}}_{kl}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, with

$$\ddot{\mathbf{T}}_{lk}(\epsilon) = \left(\sum_{\substack{m=1 \\ m \neq k,l}}^K \frac{\mathcal{E}_{d,s,m}}{1 + \ddot{e}_{kl,m}(\epsilon)} + \epsilon \right)^{-1} \mathbf{I}_N, \quad (41)$$

and $\ddot{e}_{k,l}(\epsilon)$ can be evaluated by iterating the following equation over n , starting at $n = 1$, with the initialization $\ddot{e}_{k,l}^0(\epsilon) = \frac{1}{\epsilon}$:

$$\ddot{e}_{kl,m}^n(\epsilon) = N \mathcal{E}_{d,s,m} \left(\sum_{\substack{p=1 \\ p \neq k,l}}^K \frac{\mathcal{E}_{d,s,p}}{1 + \ddot{e}_{kl,p}^{n-1}(\epsilon)} + \epsilon \right)^{-1}. \quad (42)$$

4) *BER Expressions*: Similar to MF precoding, we can use (29) and (30) to obtain the BER with RZF precoding. Also, constant modulus constellations can be used to avoid self interference due to the mismatch in the magnitudes of \dot{g}_{kk} and g_{kk} . With calibration errors, the BER with BPSK modulation and RZF precoding becomes

$$P_{e,k} = Q \left(\sqrt{\frac{2N^2 a_k^2 \delta_{u,k}^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{k,\mathbf{D}_b}^2}{(1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})^2} \cos(\phi_{u,k})}{N^2 \sum_{\substack{l=1 \\ l \neq k}}^K |\hat{g}_{kl}|^2 + N^2 \sum_{l=1}^K |\tilde{g}_{kl}|^2 + N_0}} \right), \quad (43)$$

with $Q(\cdot)$ being the Gaussian Q -function [31]. In the absence of calibration errors, the BER is given by

$$P_{e,k} = Q \left(\sqrt{\frac{2N^2 a_k^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{k,\mathbf{I}}^2}{(1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})^2}}{N^2 \sum_{\substack{l=1 \\ l \neq k}}^K |\hat{g}_{kl}|^2 + N^2 \sum_{l=1}^K |\tilde{g}_{kl}|^2 + N_0}} \right). \quad (44)$$

5) *Discussion*: Comparing (29) with (30), and (43) with (44), we observe that the performance deterioration due to the mismatch in the actual channel coefficient (i.e., the desired signal component) between the BS and the UE has an identical effect on systems employing both MFP and RZF precoding. Also, comparing (39) with (40), the effective channel coefficient for the interfering stream due to the l th users data at the k th user is different with and without calibration errors. That is, with no calibration errors, and if $\mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}} \gg 1$, then $|\hat{g}_{kl}|^2 \approx 0$, which is the expected behavior for RZF precoding. However, under imperfect calibration, $|\hat{g}_{kl}|^2$ takes a nonzero value, leading to increased inter-stream interference. Hence, reciprocity calibration errors affect RZF precoding more severely than MFP.

Broadly, we have seen that calibration errors can cause a significant degradation in the performance of massive MIMO systems. Accurate CSI can be made available at the UEs via one of the following mechanisms:

- 1) Blind downlink channel estimation at the UEs.
- 2) Limited downlink training from the BS on the precoded channels to the different users.

- 3) Full downlink training from the BS, ignoring channel reciprocity.

Out of the above techniques, the first technique fits well into the canonical massive MIMO frame structure, and requires the least training overhead. The second technique requires partial downlink training, and thus involves a larger overhead. This technique also requires a change in the canonical TDD massive MIMO frame structure, and therefore needs to be examined separately. The third technique requires full downlink training, and is essentially the same as frequency division duplexed (FDD) massive MIMO [12]. Therefore, we consider blind channel estimation at the UEs for mitigating reciprocity calibration errors. The key idea is as follows. For precoded systems, the effective channel coefficients have a phase angle close to zero with high probability. This allows us to exploit techniques similar to those developed in [29], [30], [32] for blind channel estimation at the UEs.

V. BLIND CHANNEL ESTIMATION

We now consider blind channel estimation using the data symbols transmitted by the BS. Assuming the downlink transmission to consist of B symbols, and defining $\mathbf{s}_k = [s_k[1], \dots, s_k[B]]^T$, $\mathbf{y}_k = [y_k[1], \dots, y_k[B]]^T$, with $k \in \{1, \dots, K\}$, we can write \mathbf{y}_k as

$$\mathbf{y}_k = N(\hat{g}_{kk} + \tilde{g}_{kk})\mathbf{s}_k + N \sum_{\substack{l=1 \\ l \neq k}}^K g_{kl} \mathbf{s}_l + \sqrt{N_0} \mathbf{w}_k, \quad (45)$$

for both MFP and RZF transmission by the BS. Since the symbols come from zero mean unit variance constellations, and are independent across different users, we can approximate $\frac{1}{B} \mathbf{s}_k^H \mathbf{s}_l \approx \delta[k - l]$. Similarly, since $\mathbf{w}_k[n] \sim \mathcal{CN}(0, 1)$, and is independent of the symbols sent to any UE, we can approximate $\frac{1}{B} \mathbf{w}_k^H \mathbf{s}_l \approx 0$, $\forall k$.

Defining $\mathbf{g}_{kk} \triangleq \hat{g}_{kk} + \tilde{g}_{kk}$ as the effective coefficient for the k th data stream, it is easy to show that

$$\frac{1}{B} \|\mathbf{y}_k\|_2^2 \approx N^2 |\mathbf{g}_{kk}|^2 + N^2 \sum_{\substack{l=1 \\ l \neq k}}^K |g_{kl}|^2 + N_0. \quad (46)$$

Approximating $|g_{kl}|^2 = \beta_k \mathcal{F}_{d,s,l}$ for MFP, and

$$|g_{kl}|^2 = a_k^2 \frac{\lambda^2}{N} \beta_k \mathcal{E}_{d,s,k} \frac{\text{tr}\{\mathbf{T}_l^2(\epsilon)\}}{(1 + \mathcal{E}_{d,s,k} \mu_{l,\mathbf{I}})^2},$$

for RZF, we can estimate \mathbf{g}_{kk} as,

$$\hat{\mathbf{g}}_{kk} = \frac{1}{N} \sqrt{\left(\frac{1}{B} \|\mathbf{y}\|_2^2 - \beta_k \sum_{\substack{l=1 \\ l \neq k}}^K \mathcal{F}_{d,s,l} - N_0 \right)}, \quad (47)$$

and

$$\hat{\mathbf{g}}_{kk} = \frac{1}{N} \sqrt{\left(\frac{1}{B} \|\mathbf{y}\|_2^2 - a_k^2 \frac{\lambda^2}{N} \beta_k \mathcal{E}_{d,s,k} \frac{\text{tr}\{\mathbf{T}_l^2(\epsilon)\}}{(1 + \mathcal{E}_{d,s,k} \mu_{l,\mathbf{I}})^2} - N_0 \right)}, \quad (48)$$

for MFP and RZF precoding, respectively. Note that we choose the real valued positive square root of the power term as the channel estimate because the effective channel is close to a positive real number with high probability. This technique is similar to the gain estimation technique discussed in [33]. However, this coarse estimate of the effective channel gain does not account for the phase offset due to the calibration imperfections of the UEs RF chains. Still, it can be used by the k th user to estimate the transmitted sequence as

$$\hat{s}_k[n] = \arg \min_{s \in \mathcal{S}} |y_k[n] - N\hat{\mathbf{g}}_{kk}s|^2, \quad (49)$$

with \mathcal{S} being the symbol constellation being employed.

Letting, $\hat{\mathbf{s}}_k = [\hat{s}_k[1] \dots \hat{s}_k[B]]$, we can equivalently write,

$$\mathbf{y}_k = N\mathbf{g}_{kk}\hat{\mathbf{s}}_k + N\mathbf{g}_{kk}\tilde{\mathbf{s}}_k + N\sum_{\substack{l=1 \\ l \neq k}}^K g_{kl}\mathbf{s}_l + \sqrt{N_0}\mathbf{w}_k, \quad (50)$$

with $\tilde{\mathbf{s}}_k = \mathbf{s}_k - \hat{\mathbf{s}}_k$ being the symbol error vector for the k th user, such the $E[\tilde{\mathbf{s}}_k^H \tilde{\mathbf{s}}_k] = 0$, and $\frac{1}{B}E[\tilde{\mathbf{s}}_k^H \tilde{\mathbf{s}}_k] \leq 4P_{e,k}$ [29], with $P_{e,k}$ being the symbol error probability (SEP) over the k th stream. With $\hat{\mathbf{s}}_k$ known at the UE, all terms except the first can be considered as noise and interference, and an updated estimate of \mathbf{g}_{kk} can be obtained by solving

$$\hat{\mathbf{g}}_{kk} = \arg \min \|\mathbf{y}_k - \mathbf{g}_{kk}\hat{\mathbf{s}}_k\|_2^2, \quad (51)$$

which can be reduced to

$$\hat{\mathbf{g}}_{kk} = \frac{1}{N} \frac{1}{\|\hat{\mathbf{s}}_k\|_2^2} \hat{\mathbf{s}}_k^H \mathbf{y}_k \approx \frac{1}{BN} \hat{\mathbf{s}}_k^H \mathbf{y}_k. \quad (52)$$

The new estimate of \mathbf{g}_{kk} can again be used to solve for a more accurate estimate of \mathbf{s}_k . Hence, (49) and (52) can be used iteratively to improve the quality of the estimates of \mathbf{g}_{kk} . This algorithm alternately minimizes the cost function

$$[\hat{\mathbf{s}}_k, \hat{\mathbf{g}}_{kk}] = \arg \min_{\mathbf{s} \in \mathcal{S}^B, \mathbf{g}} \|\mathbf{y}_k - N\mathbf{g}\mathbf{s}\|_2^2, \quad (53)$$

which is lower bounded by zero. Since the two sub problems, viz. channel estimation and symbol detection, are solved optimally, the objective function is guaranteed to decrease in each iteration. Therefore, the cost function is guaranteed to converge to a local optimum, a characteristic shared by all iterative alternating minimization algorithms.

A. Mean Squared Channel Estimation Error

Assuming that the estimated data symbol vector at the k th user is $\hat{\mathbf{s}}_k$ and the SEP at the k th UE is $P_{e,k}$, we can express \mathbf{y}_k as

$$\mathbf{y}_k = N\mathbf{g}_{kk}\hat{\mathbf{s}}_k + \zeta_k \quad (54)$$

with the entries $\zeta_k[n]$ of ζ_k being iid, and approximated as ZMCSCG noise having a variance $4N^2P_{e,k}|\mathbf{g}_{kk}|^2 + N^2\sum_{\substack{l=1 \\ l \neq k}}^K |g_{kl}|^2 + N_0$, and thus

$$\hat{\mathbf{g}}_{kk} = \mathbf{g}_{kk} + \frac{1}{N\|\hat{\mathbf{s}}_k\|_2^2} \hat{\mathbf{s}}_k^H \zeta_k, \quad (55)$$

with the second term corresponding to the estimation error. The mean squared channel estimation error is given by

$$E \left[\left| \frac{1}{N\|\hat{\mathbf{s}}_k\|_2^2} \hat{\mathbf{s}}_k^H \zeta_k \right|^2 \right] = \frac{1}{B} \left(4P_{e,k}|\mathbf{g}_{kk}|^2 + \sum_{\substack{l=1 \\ l \neq k}}^K |g_{kl}|^2 + \frac{N_0}{N^2} \right). \quad (56)$$

The MSE can be easily computed using the DEs of all the terms in (56) derived earlier, for both MFP and RZF precoding.

B. Achievable SINR with Blind Channel Estimation

In order to analyze the SINR and BER performance of blind channel estimation, we consider that the number of downlink symbols B is large (it is shown via simulations (see Fig. 6) in Section VI that $B \approx 100$ is sufficient), so that the iterative estimation scheme described above returns an accurate estimate of the effective channel \mathbf{g}_{kk} . Hence, the SINR k th stream under MFP can be expressed as

$$\gamma_k = \frac{(NE[\delta_b \cos(\phi_b)]\delta_{u,k}^2 a_k^2 + \bar{a}_k^2)\beta_k \mathcal{F}_{d,s,k}}{\beta_k \sum_{\substack{l=1 \\ l \neq k}}^K \mathcal{F}_{d,s,l} + N_0}. \quad (57)$$

Similarly, under RZF precoding, this becomes

$$\gamma_{k,rzf} = \frac{N^2 \left(a_k^2 \delta_{u,k}^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{k,D_b}^2}{(1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})^2} + |\tilde{g}_{kl}|^2 \right)}{N^2 \sum_{\substack{l=1 \\ l \neq k}}^K |\hat{g}_{kl}|^2 + N_0}, \quad (58)$$

Comparing (18), and (30) with (57) and (58), blind channel estimation by the UEs can also lead to the suppression of self-interference at the UEs as well as an increase in the effective gain for the desired signal. However, the blind channel estimator works under the assumption that the phase of the effective channel coefficient \mathbf{g}_{kk} is smaller than the rotational symmetry of the constellation being used. When this is violated, catastrophic detection errors occur during iterative blind channel estimation/detection phase. We discuss this phenomenon in the next section, where we present a more accurate analysis of the BER with blind channel estimation at the UEs. We also derive the BER for higher order constellations.

C. BER Analysis with Blind Channel Estimation

In this section, we analyze the probability of error performance of massive MIMO systems with imperfect reciprocity calibration, for higher order constellations. Here, we drop the time index of different signals for the sake of simplicity.

Let the symbol transmitted to the k th user be the i th symbol s_i from the constellation \mathcal{S} . The symbol received at the k th user, y_k , is given by

$$y_k = N\mathbf{g}_{kk}s_i + \zeta_k. \quad (59)$$

The union bound on the SEP can be written as [31, Chapter 4]

$$P_{e,k} \leq \sum_{s_i \in \mathcal{S}} P(s_i) \sum_{s_j \in \mathcal{S}, j \neq i} P(s_i \rightarrow s_j), \quad (60)$$

where $P(s_i \rightarrow s_j)$ is the pairwise error probability (PEP) of detecting the transmitted symbol s_i as s_j . Assuming that the

blind channel estimate at the receiver is accurate, and for MFP, the PEP between the i th and j th symbols of a constellation \mathcal{S} can be written as

$$P(s_i \rightarrow s_j) = Q \left(\frac{N |\mathbf{g}_{kk}(s_i - s_j)|}{\sqrt{N^2 \sum_{l=1, l \neq k}^K |g_{kl}|^2 + N_0}} \right). \quad (61)$$

Similarly, for RZF, $P(s_i \rightarrow s_j)$ becomes

$$P(s_i \rightarrow s_j) = Q \left(\frac{N |\mathbf{g}_{kk}(s_i - s_j)|}{\sqrt{N^2 \sum_{l=1, l \neq k}^K (|\hat{g}_{kl}|^2 + |\tilde{g}_{kl}|^2) + N_0}} \right). \quad (62)$$

In order to accurately estimate \mathbf{g}_{kk} blindly from the data received as per (59), we need the effective channel coefficient \mathbf{g}_{kk} to be real and positive. However, as shown in the previous sections, channel estimation errors at the BS coupled with reciprocity imperfections can lead to the phase of \mathbf{g}_{kk} being large. If the phase of \mathbf{g}_{kk} exceeds the rotational symmetry of the constellation \mathcal{S} , catastrophic detection errors occur at the UE. This phenomenon, known as *channel corruption*, is common to all blind channel estimation based detection schemes operating on rotationally symmetrical constellations [29]. When channel corruption occurs, all the received symbols to be decoded incorrectly with high probability, and hence the SEP can be upper bounded by unity. In view of this, we can express the overall SEP at the k th UE as

$$P_{e,k} \approx P_{cc,k} + (1 - P_{cc,k})P_{e,CSI,k} \quad (63)$$

where $P_{cc,k}$ is the probability of channel corruption over the k th UE's channel, and $P_{e,CSI,k}$, the SEP for the k th UE under perfect CSI at the receiver. Since a large number of data symbols are used for blind channel estimation, it is safe to assume the availability of perfect CSI at the UEs when channel corruption does not occur.

1) *Derivation of $P_{cc,k}$* : If the symbol constellation \mathcal{S} has a rotational symmetry φ , such that, for M-PSK $\varphi = 2\pi/M$, the event of channel corruption occurs whenever $\angle \mathbf{g}_{kk} > \frac{\varphi}{2}$. Hence, the probability of channel corruption over the k th UE's channel becomes

$$P_{cc,k} = \Pr \left\{ \angle \mathbf{g}_{kk} > \frac{\varphi}{2} \right\}. \quad (64)$$

It is shown in Appendix B that, conditioned on $\phi_{u,k}$,

$$\begin{aligned} P_{cc,k} = & Q \left(\frac{|\hat{g}_{kk}| \cos(\phi_{u,k})}{\sqrt{\frac{|\tilde{g}_{kk}|^2}{2}}} \right) \\ & + Q \left(\frac{|\hat{g}_{kk}| \left(\tan\left(\frac{\varphi}{2}\right) \cos(\phi_{u,k}) - \sin(\phi_{u,k}) \right)}{\sqrt{\frac{|\tilde{g}_{kk}|^2}{2} (1 + \tan^2 \frac{\varphi}{2})}} \right) \\ & \times \left(1 - Q \left(\frac{|\hat{g}_{kk}| \cos(\phi_{u,k})}{\sqrt{\frac{|\tilde{g}_{kk}|^2}{2}}} \right) \right) \mathbb{1}_{\{\varphi < \pi\}}. \quad (65) \end{aligned}$$

2) *Derivation of $P_{e,CSI,k}$ for M-PSK Signaling*: Using (57) and (58), it is easy to show that for an M-PSK ($M \geq 4$) con-

stellation, $P_{e,CSI,k}$ can be approximated as (66), and (67) [31], respectively. Combining the above results, we can obtain the overall probability of error for PSK signaling using (63).

3) *Derivation of $P_{e,CSI,k}$ for QAM Signaling*: The probability of error for the M -PAM constellation with perfect CSI at the UE, and a channel coefficient \mathbf{g}_{kk} , and MFP being used at the BS is given as (68) [31]. Similarly, the BER with perfect CSI at the UE, and RZF precoding can be written as (69). Since the QAM constellation can be viewed as a superposition of two orthogonal PAM constellations, we have

$$P_{e,CSI,k}^{\text{QAM}} = 2P_{e,CSI,k}^{\text{PAM}} - (P_{e,CSI,k}^{\text{PAM}})^2. \quad (70)$$

Similar to the previous subsection, the overall probabilities of error for PAM and QAM become,

$$P_{e,k}^{\text{PAM}} = P_{cc,k}^{\text{PAM}} + (1 - P_{cc,k})P_{e,CSI,k}^{\text{PAM}} \quad (71)$$

$$P_{e,k}^{\text{QAM}} = P_{cc,k}^{\text{QAM}} + (1 - P_{cc,k})P_{e,CSI,k}^{\text{QAM}} \quad (72)$$

where $P_{cc,k}^{\text{PAM}}$ and $P_{cc,k}^{\text{QAM}}$ can be calculated using the facts that the rotational symmetries of PAM and QAM are $\varphi = \pi$, and $\varphi = \frac{\pi}{2}$, respectively.

VI. SIMULATION RESULTS

In this section, we present Monte Carlo simulation results to validate the derived theory and obtain further insights into the effect of calibration errors on the performance of massive MIMO systems. We consider a single cell narrowband massive MIMO system with an $N = 512$ antenna BS. The system is assumed to operate at a carrier frequency of 2 GHz, and a bandwidth of 1 MHz. For all the experiments, we assume a path loss inversion based uplink pilot power control, such that the pilot SNR at the BS is 10 dB for all the users.

In Fig. 1, we plot the normalized downlink channel gain for a user against the quantum of phase error for different phase error distributions, and a gain error uniformly distributed in the range $[0.98, 1.02]$. For generating the uniformly distributed phase error, we assume it to be distributed in the range $[-\phi_0, \phi_0]$. In case of truncated Gaussian distributed phase error, we truncate samples from a zero mean unit variance Gaussian distribution at $\pm\phi_0$. For wrapped Gaussian and wrapped Laplacian distributions respectively, we generate samples from zero mean Gaussian and Laplacian distributions having variances equal to ϕ_0^2 . Our main observation is that in all the cases, the simulated result closely follows the theoretically derived result, and the effect of phase error, while detrimental to the achievable gain, is largely independent of the distribution of phase error, if they have the same variance. Therefore, in the subsequent experiments, we assume the phase error to be uniformly distributed.

In Fig. 2, we plot the achievable BER of a system employing MFP and BPSK modulation under perfect reciprocity calibration for different numbers of UEs, to illustrate the limiting effect of inter-user interference in MFP based systems. The figure also illustrates the accuracy of the deterministic equivalent expressions: the markers show the simulated performance for a *single* randomly chosen instantiation of the channel, while the solid curves show the BER achieved according to the DE expressions in (23).

$$P_{e,\text{CSI},k}^{\text{PSK}} \approx E_{\delta_{u,k}} \left[Q \left(\sqrt{\frac{(NE[\delta_b \cos(\phi_b)]\delta_{u,k}^2 a_k^2 + \bar{a}_k^2)\beta_k \mathcal{F}_{d,s,k}}{\beta_k \sum_{l \neq k}^K \mathcal{F}_{d,s,l} + N_0}} \sin\left(\frac{\pi}{M}\right)} \right) \right] \quad (66)$$

$$P_{e,\text{CSI},k}^{\text{PSK}} \approx E_{\delta_{u,k}} \left[Q \left(\sqrt{\frac{2N^2 \left(a_k^2 \delta_{u,k}^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{\mathbf{D},k}^2}{(1+\mathcal{E}_{d,s,k} \mu_{\mathbf{I},k})^2} + |\tilde{g}_{kl}|^2 \right)}{N^2 \sum_{l \neq k}^K (|\hat{g}_{kl}|^2 + |\tilde{g}_{kl}|^2) + N_0}} \sin\left(\frac{\pi}{M}\right)} \right) \right] \quad (67)$$

$$P_{e,\text{CSI},k}^{\text{PAM}} = \frac{2(M-1)}{M} E_{\delta_{u,k}} \left[Q \left(\sqrt{\frac{6}{M^2-1} \frac{(NE[\delta_b \cos(\phi_b)]\delta_{u,k}^2 a_k^2 + \bar{a}_k^2)\beta_k \mathcal{F}_{d,s,k}}{\beta_k \sum_{l \neq k}^K \mathcal{F}_{d,s,l} + N_0}} \right) \right] \quad (68)$$

$$P_{e,\text{CSI},k}^{\text{PAM}} = \frac{2(M-1)}{M} E_{\delta_{u,k}} \left[Q \left(\sqrt{\frac{6}{M^2-1} \frac{2N^2 \left(a_k^2 \delta_{u,k}^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{\mathbf{D},k}^2}{(1+\mathcal{E}_{d,s,k} \mu_{\mathbf{I},k})^2} + |\tilde{g}_{kl}|^2 \right)}{N^2 \sum_{l \neq k}^K (|\hat{g}_{kl}|^2 + |\tilde{g}_{kl}|^2) + N_0}} \right) \right] \quad (69)$$

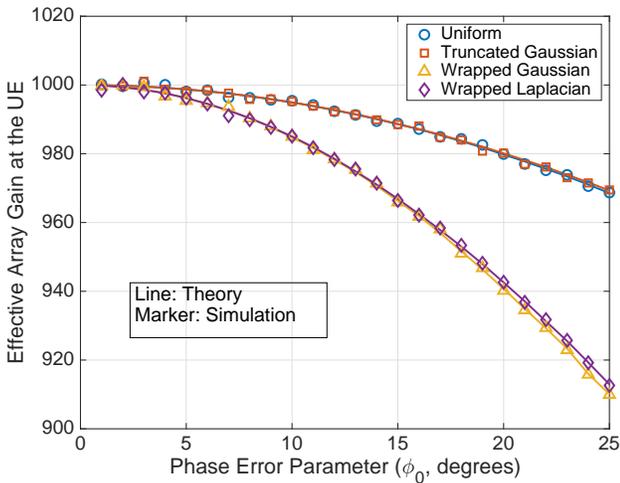


Fig. 1: Effective array gain at the UE for MFP transmission versus the phase error (degrees) for different distributions of the phase error.

In Fig. 3, we plot the achievable rate of a single user system against the number of BS antennas, at a transmit SNR of 10 dB. Here, we focus on the single user system to better isolate the effects of reciprocity calibration errors. We observe that, for a large number of BS antennas, phase calibration errors result in a saturation in the achievable rate, thereby negating the array gain advantage of massive MIMO systems. Also, the dashed lines in Fig. 3 represent the achievable rates under the assumption of the availability of calibration error information at the UEs, as in [23]. It is observed that this assumption overestimates the achievable rate. In fact, in the single user case, calibration error only slightly decreases the numerator term of the SINR expression, leading to an almost negligible loss in rate due to calibration errors. This underscores the importance of accounting for the effect of

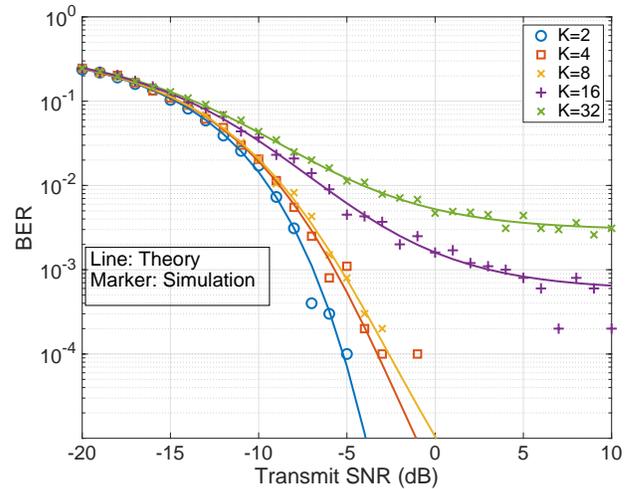


Fig. 2: BER vs. transmit SNR for different numbers of users for BPSK under MFP and 512 BS antennas.

calibration errors in the decoding process at the receiver.

In Fig. 4, we plot the BER for a 64 user system with varying degrees of phase calibration errors. We observe that while RZF precoding leads to cancellation of inter-user interference, reciprocity calibration errors result in imperfect inter-stream interference cancellation. Consequently, the BER saturates at high SNRs, as predicted by the results derived in (44).

In Fig. 5, we plot the rates achievable by an 8 user RZF system under calibration imperfections for different numbers of BS antennas. Similar to the MFP case, an increased phase error results in saturation in the achievable rates, limiting the system performance. Also similar to the MFP case, the dashed lines represent the achievable rates with the knowledge of reciprocity calibration imperfections at the UEs. The assumption of the knowledge of calibration imperfections at the UEs

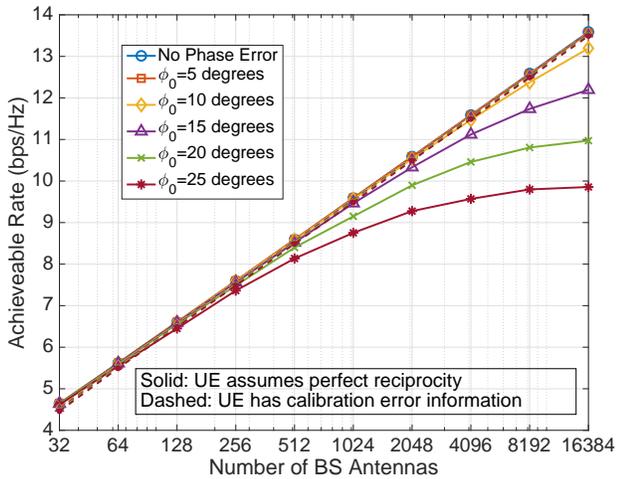


Fig. 3: Achievable rates for a single user under MFP for different degrees of calibration error.

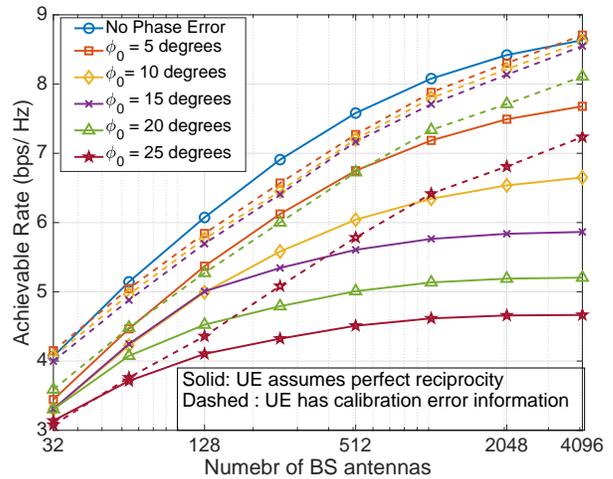


Fig. 5: Achievable rates for a 8 users under RZF for different degrees of calibration error.

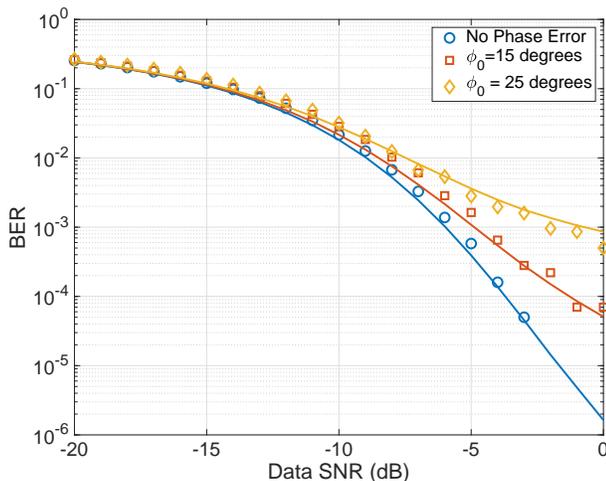


Fig. 4: BER vs. SNR for 64 users under BPSK and RZF precoding for different amounts of phase calibration errors.

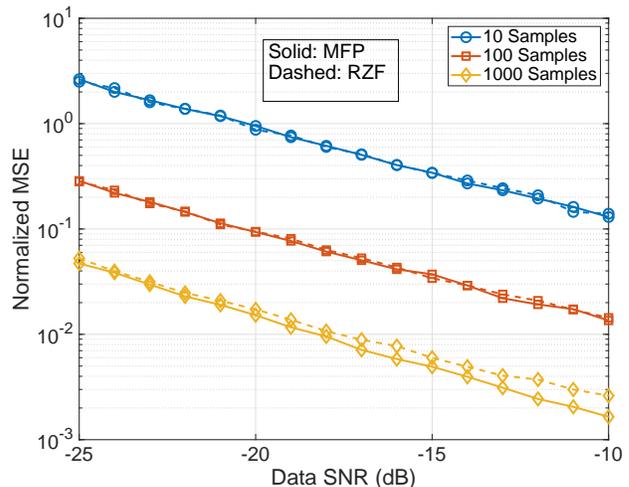


Fig. 6: Normalized MSE of the blind channel estimation algorithm for different cases.

not only leads to an overestimation of the achievable rates, it also fails to capture the saturation behavior as the number of antennas is increased.

In Fig. 6, we plot the normalized mean squared error (NMSE), that is, the ratio of the mean squared estimation error of the channel coefficient \mathbf{g}_{kk} to the mean squared value of \mathbf{g}_{kk} , as a function of the data SNR, for different numbers of data symbols received at the UE. We see that a ten-fold increase in the number of data symbols results in an almost 10 dB reduction in the normalized MSE, as predicted by (56). Further, the performance of RZF and MFP are similar, with the performance of RZF degrading slightly at high SNRs and when a large number of data samples are available. RZF is more sensitive to calibration errors compared to MFP, especially in interference limited scenarios. Hence, RZF suffers from a slightly higher symbol error rate, which in turn leads to an increased NMSE compared to MFP.

In Fig. 7, we plot the achievable BERs for a two user system employing 4-PAM, with the blind channel estimation scheme presented in Sec. V. Blind channel estimation is observed to reverse the effects of calibration errors for both MFP and RZF precoding, and resulting in a performance comparable to that of a perfectly-calibrated massive MIMO system. It is also observed that RZF attempts to cancel the inter-user interference at the transmitter side, and is therefore more sensitive to calibration errors than MFP: a 25 degree calibration error results in a performance loss of about 1 dB for MFP, but a 1.7 dB loss for RZF. Blind channel estimation improves the performance by obtaining better channel estimates at the UE side, but this does not help with reducing the inter-user interference in RZF introduced due to calibration errors. Therefore, the performance improvement offered by blind channel estimation, about 1 dB, is roughly the same for both RZF and MFP. The net outcome is that the

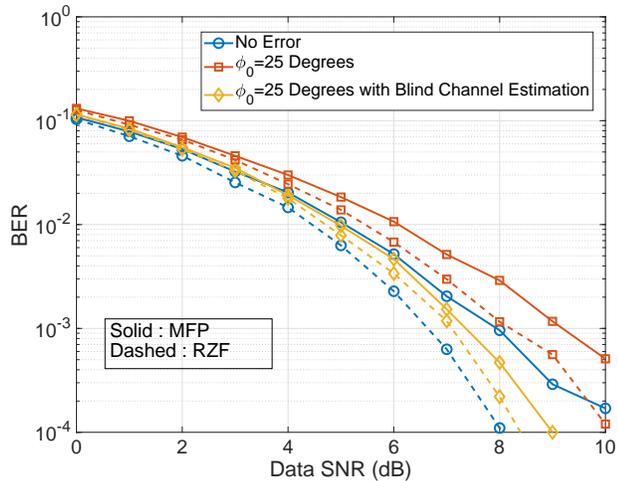


Fig. 7: BER vs. SNR for 2 users under 4 PAM with and without blind channel estimation.

performance of RZF with blind channel estimation is slightly worse than that of RZF without calibration errors and no channel estimation at the UEs, while the performance of MFP with blind channel estimation is slightly better than that of MFP without calibration errors and no channel estimation at the UEs. Hence, the detrimental effect of calibration errors and the benefit of blind channel estimation depends on a variety of factors such as the pilot signal power, the number of users, the number of base station antennas, the modulation order, the precoding scheme used, etc.

In Fig. 8, we plot the achievable rate against the number of BS antennas in a 4 user massive MIMO system with and without blind channel estimation at the UEs. For both MFP and RZF, blind channel estimation results in a slightly improved performance as compared to the no error case, an effect that can be attributed to the observations made in (57) and (58). Also, in the absence of any calibration imperfections, RZF precoding performs much better than MFP. However, both the precoding schemes have near identical performance in the presence of calibration imperfections, thus validating the fact that reciprocity imperfections affect RZF precoded systems more than MFP systems.

In Fig 9, we plot the achievable BERs under BPSK for an RZF system, with a phase error coefficient $\phi_0 = 25$ degrees and blind channel estimation. We observe that blind channel estimation results in almost identical performance of the system for different numbers of users. However, the performance worsens by about 0.5 dB at a BER of 10^{-6} , as the number of UEs increases from $K = 2$ to $K = 16$. This is due to the uncanceled interference caused by the reciprocity imperfections at the BS.

VII. CONCLUSIONS

We considered the downlink performance of massive MIMO with reciprocity calibration errors. We derived the achievable rates for both MFP and RZF precoding and showed that reciprocity calibration errors can result in significant losses in

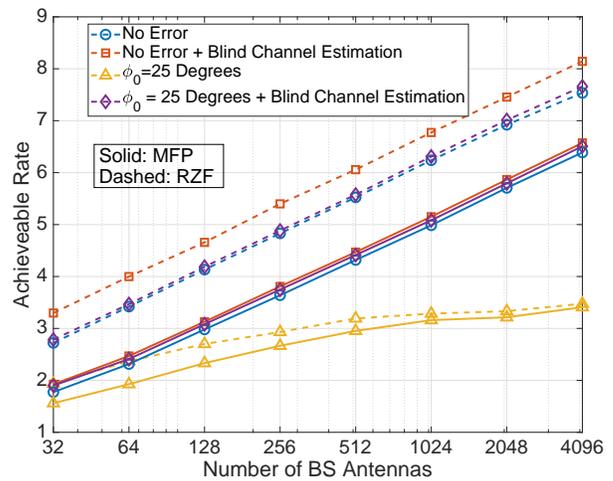


Fig. 8: Achievable rates for a 4 user system with and without blind channel estimation.

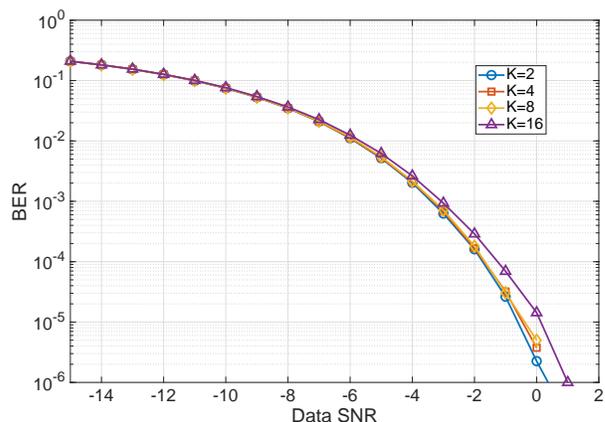


Fig. 9: BER with RZF precoding for different numbers of users under blind channel estimation, with $\phi_0 = 25$ degrees.

the achievable rates, and may substantially degrade the array gain offered by massive MIMO. Then, using the fact that the effective downlink channel gain is close to a positive real number with high probability, we presented a blind algorithm for the estimation of this gain. We showed that, for sufficiently large block lengths, it is possible to obtain an accurate estimate of the effective downlink channel, and that the use of this approach can restore the array gain. Also, our discussion in this paper is limited to a frequency flat single carrier system. Extending the calibration error model and the analysis to a multi-carrier, frequency selective fading channel is an interesting direction for future work.

APPENDIX A

DERIVATION OF THE SINR WITH RZF PRECODING

We first define the matrices \mathbf{Q}_k and \mathbf{Q}_{kl} as,

$$\mathbf{Q}_k \triangleq \sum_{\substack{l=1 \\ l \neq k}}^K \mathcal{E}_{d,s,l} \hat{\mathbf{h}}_{d,l}^* \hat{\mathbf{h}}_{d,l}^T + \epsilon \mathbf{I}_N, \quad (73)$$

and

$$\mathbf{Q}_{kl} \triangleq \sum_{\substack{m=1 \\ m \neq k,l}}^K \mathcal{E}_{d,s,m} \hat{\mathbf{h}}_{d,m}^* \hat{\mathbf{h}}_{d,m}^T + \epsilon \mathbf{I}_N. \quad (74)$$

Now, the effective channel coefficient for the desired signal, \hat{g}_{kk} , is given as $\hat{g}_{kk} = \frac{\lambda}{N} a_k d_{u,k} \sqrt{\beta_k \mathcal{E}_{d,s,k}} \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \mathbf{Q}^{-1} \hat{\mathbf{h}}_{d,k}^*$, and we know that [34],

$$\begin{aligned} \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \mathbf{Q}^{-1} \hat{\mathbf{h}}_{d,k}^* &= \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b (\mathbf{Q}_k + \mathcal{E}_{d,s,k} \hat{\mathbf{h}}_{d,k}^* \hat{\mathbf{h}}_{d,k}^T)^{-1} \hat{\mathbf{h}}_{d,k}^* \\ &= \frac{\hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \mathbf{Q}_k^{-1} \hat{\mathbf{h}}_{d,k}^*}{1 + \mathcal{E}_{d,s,k} \hat{\mathbf{h}}_{d,k}^T \mathbf{Q}_k^{-1} \hat{\mathbf{h}}_{d,k}^*}. \end{aligned} \quad (75)$$

Defining $\mu_{k,\mathbf{B}} \triangleq \hat{\mathbf{h}}_{d,k}^T \mathbf{B} \mathbf{Q}_k^{-1} \hat{\mathbf{h}}_{d,k}^*$, we can write \hat{g}_{kk} as

$$\hat{g}_{kk} = a_k d_{u,k} \lambda \sqrt{\beta_k \mathcal{E}_{d,s,k}} \frac{\mu_{k,\mathbf{D}_b}}{N(1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})}. \quad (76)$$

Similarly,

$$\tilde{\mathbf{h}}_{d,k}^T \mathbf{Q}^{-1} \hat{\mathbf{h}}_{d,l}^* = \tilde{\mathbf{h}}_{d,k}^T \left(\frac{\mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,l}^*}{1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right), \quad (77)$$

and letting $\nu_{kl} \triangleq \tilde{\mathbf{h}}_{d,k}^T \mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,l}^*$, we get,

$$\tilde{g}_{kl} = \bar{a}_l \lambda \sqrt{\beta_k \mathcal{E}_{d,s,k}} \frac{\nu_{kl}}{N(1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}})}. \quad (78)$$

Also,

$$\begin{aligned} \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \mathbf{Q}^{-1} \hat{\mathbf{h}}_{d,l}^* &= \left(\frac{1}{1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right) \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,l}^* \\ &= \left(\frac{1}{1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right) \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b (\mathbf{Q}_{kl} + \mathcal{E}_{d,s,k} \hat{\mathbf{h}}_{d,k}^* \hat{\mathbf{h}}_{d,k}^T)^{-1} \hat{\mathbf{h}}_{d,l}^* \\ &= \left(\frac{1}{1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right) \hat{\mathbf{h}}_{d,k}^T \mathbf{D}_b \\ &\quad \times \left(\mathbf{Q}_{kl}^{-1} - \mathcal{E}_{d,s,k} \frac{\mathbf{Q}_{kl}^{-1} \hat{\mathbf{h}}_{d,k}^* \hat{\mathbf{h}}_{d,k}^T \mathbf{Q}_{kl}^{-1}}{1 + \mathcal{E}_{d,s,k} \hat{\mathbf{h}}_{d,k}^T \mathbf{Q}_{kl}^{-1} \hat{\mathbf{h}}_{d,k}^*} \right) \hat{\mathbf{h}}_{d,l}^* \\ &= \left(\frac{1}{1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right) \left(\dot{\mu}_{kl,\mathbf{D}_b} - \mathcal{E}_{d,s,k} \frac{\dot{\mu}_{kk,\mathbf{D}_b} \dot{\mu}_{kl,\mathbf{I}}}{1 + \mathcal{E}_{d,s,k} \dot{\mu}_{kk,\mathbf{I}}} \right), \end{aligned} \quad (79)$$

where $\dot{\mu}_{kl,\mathbf{B}} \triangleq \hat{\mathbf{h}}_{d,k}^T \mathbf{B} \mathbf{Q}_{kl}^{-1} \hat{\mathbf{h}}_{d,l}^*$.

Consequently,

$$\begin{aligned} \hat{g}_{kl} &= \frac{1}{N} \lambda a_l d_{u,k} \sqrt{\beta_k \mathcal{E}_{d,s,l}} \left(\frac{1}{1 + \mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right) \\ &\quad \times \left(\dot{\mu}_{kl,\mathbf{D}_b} - \mathcal{E}_{d,s,k} \frac{\dot{\mu}_{kk,\mathbf{D}_b} \dot{\mu}_{kl,\mathbf{I}}}{1 + \mathcal{E}_{d,s,k} \dot{\mu}_{kk,\mathbf{I}}} \right). \end{aligned} \quad (80)$$

Moreover, $\mu_{k,\mathbf{I}} \triangleq \hat{\mathbf{h}}_{d,k}^T \mathbf{Q}_k^{-1} \hat{\mathbf{h}}_{d,k}^*$ and from [9], [35], we get the following: $\mu_{k,\mathbf{I}} - \text{tr}\{\dot{\mathbf{T}}_k(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $\mu_{k,\mathbf{D}_b} - E[\delta_b \cos \phi_b] \text{tr}\{\dot{\mathbf{T}}_k(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $\nu_{kl} = \tilde{\mathbf{h}}_{d,k}^T \mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,l}^* \xrightarrow{\text{a.s.}} 0$, $|\nu_{kl}|^2 = |\tilde{\mathbf{h}}_{d,k}^T \mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,l}^*|^2 = \tilde{\mathbf{h}}_{d,k}^T \mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,l}^* \hat{\mathbf{h}}_{d,l}^T \mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,k}^*$, with $\tilde{\mathbf{h}}_{d,k}^T \mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,l}^* \hat{\mathbf{h}}_{d,l}^T \mathbf{Q}_l^{-1} \hat{\mathbf{h}}_{d,k}^* - \text{tr}\{\dot{\mathbf{T}}_l^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $\dot{\mu}_{kl,\mathbf{I}} = \hat{\mathbf{h}}_{d,k}^T \mathbf{Q}_{kl}^{-1} \hat{\mathbf{h}}_{d,l}^* \xrightarrow{\text{a.s.}} 0$, and $|\hat{\mathbf{h}}_{d,k}^T \mathbf{Q}_{kl}^{-1} \hat{\mathbf{h}}_{d,l}^*|^2 - \text{Tr}\{\dot{\mathbf{T}}_{lk}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$.

Also, under the assumption of independence of the RF chain gains and calibration errors, $|\dot{\mu}_{kl,\mathbf{D}_b}|^2 - E^2[\delta_b \cos \phi_b] \text{Tr}\{\dot{\mathbf{T}}_{kl}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, and $\dot{\mu}_{kl,\mathbf{D}_b} \dot{\mu}_{kl,\mathbf{I}} - E[\delta_b \cos \phi_b] \text{Tr}\{\dot{\mathbf{T}}_{kl}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$.

If the UEs assume perfect reciprocity, then the approxima-

tion for the effective channel available at the k th user becomes

$$\bar{g}_{kk} = a_k \frac{\lambda}{N} \sqrt{\beta_k \mathcal{E}_{d,s,k}} \frac{\mu_{k,\mathbf{I}}}{1 + \mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}}}. \quad (81)$$

The received signal $y_k[n]$ can therefore be written as

$$\begin{aligned} y_k[n] &= N \dot{g}_{kk} s_k[n] + N(\hat{g}_{kk} - \dot{g}_{kk}) s_l[n] \\ &\quad + N \sum_{\substack{l=1 \\ l \neq k}}^K \hat{g}_{kl} s_l[n] + N \sum_{l=1}^K \tilde{g}_{kl} s_l[n] + \sqrt{N} w_k[n], \end{aligned} \quad (82)$$

with all the terms except the first contributing to additive noise and interference, and being uncorrelated with the desired signal. Therefore, the SINRs for the signals received at the UEs with and without reciprocity calibration errors can be reduced to (29) and (30).

APPENDIX B DERIVATION OF $P_{cc,k}$

Considering the fact that the maximum rotational symmetry of a constellation is π , a channel corruption will result whenever \mathbf{g}_{kk} lies in the left half of the complex plane. Based on this, we can split (64) as

$$\begin{aligned} P_{cc,k} &= \Pr \left\{ |\angle \mathbf{g}_{kk}| > \frac{\pi}{2} \right\} \\ &\quad + \left(\Pr \left\{ |\angle \mathbf{g}_{kk}| > \frac{\varphi}{2} \mid |\angle \mathbf{g}_{kk}| < \frac{\pi}{2} \right\} \right) \\ &\quad \left(1 - \Pr \left\{ |\angle \mathbf{g}_{kk}| > \frac{\pi}{2} \right\} \right) \mathbb{1}_{\{\varphi < \pi\}}. \end{aligned} \quad (83)$$

Here, $\mathbb{1}_{\{A\}}$ is the indicator function that evaluates to 1 when the event A is true, and evaluates to 0 when A is false.

Next, as $\angle \mathbf{g}_{kk} = \tan^{-1} \left(\frac{\Im\{\mathbf{g}_{kk}\}}{\Re\{\mathbf{g}_{kk}\}} \right)$, the above equation can be rewritten as

$$\begin{aligned} P_{cc,k} &= \Pr \{ \Re\{\mathbf{g}_{kk}\} < 0 \} \\ &\quad + 2 \left(\Pr \left\{ \Im\{\mathbf{g}_{kk}\} > \tan \left(\frac{\varphi}{2} \right) \Re\{\mathbf{g}_{kk}\} \right\} \right) \\ &\quad \left(1 - \Pr \{ \Re\{\mathbf{g}_{kk}\} < 0 \} \right) \mathbb{1}_{\{\varphi < \pi\}}. \end{aligned} \quad (84)$$

Now, $\mathbf{g}_{kk} = \hat{g}_{kk} + \tilde{g}_{kk}$, with \hat{g}_{kk} converging to a constant, expressed as, $\hat{g}_{kk} = |\hat{g}_{kk}| e^{j\phi_{u,k}}$, and \tilde{g}_{kk} being a zero mean circularly symmetric r.v. with a variance approximated by the DE of $|\tilde{g}_{kk}|^2$. Also, since \tilde{g}_{kk} is the sum of a large number of i.i.d. terms, it can be approximated as a ZMCSCG r.v.. Consequently, both $\Re\{\mathbf{g}_{kk}\}$ and $\Im\{\mathbf{g}_{kk}\}$ are real valued Gaussian r.v.s, and hence,

$$\begin{aligned} &\Pr \left\{ \Im\{\mathbf{g}_{kk}\} > \tan \left(\frac{\varphi}{2} \right) \Re\{\mathbf{g}_{kk}\} \right\} \\ &= \Pr \left\{ \Im\{\tilde{g}_{kk}\} - \tan \left(\frac{\varphi}{2} \right) \Re\{\tilde{g}_{kk}\} \right. \\ &\quad \left. > |\hat{g}_{kk}| \left(\tan \left(\frac{\varphi}{2} \right) \cos(\phi_{u,k}) - \sin(\phi_{u,k}) \right) \right\}. \\ &= E_{\phi_u} \left[Q \left(\frac{|\hat{g}_{kk}| \left(\tan \left(\frac{\varphi}{2} \right) \cos(\phi_{u,k}) - \sin(\phi_{u,k}) \right)}{\sqrt{\frac{|\tilde{g}_{kk}|^2}{2} (1 + \tan^2 \frac{\varphi}{2})}} \right) \right]. \end{aligned} \quad (85)$$

Similarly,

$$\Pr\{\Re\{\mathbf{g}_{kk}\} < 0\} = E_{\phi_u} \left[Q \left(\frac{|\hat{g}_{kk}| \cos(\phi_{u,k})}{\sqrt{\frac{|\hat{g}_{kk}|^2}{2}}} \right) \right]. \quad (86)$$

These expressions can be substituted in (64) to obtain (65).

APPENDIX C

SINR EXPRESSIONS UNDER CORRELATED CHANNELS

We assume that the uplink channel \mathbf{H}_u is correlated, and is given as, $\mathbf{H}_u = \mathbf{R}^{1/2} \mathbf{F}$, with the columns of $\mathbf{F} \text{in} \mathbb{C}^{N \times K}$ distributed as ZMCSG random variables, and $\mathbf{R} \in \mathbb{C}^{N \times N}$ being the channel correlation matrix. Then, by following a similar procedure as in the earlier sections, we can write the achievable SINR with MFP at the BS and reciprocity calibration errors, and without blind channel estimation at the UEs as (87) at the top of the next page.

Also, the achievable SINR with RZF at the BS and reciprocity calibration errors, without blind channel estimation at the UEs as (88) at the top of the next page, where $|\hat{g}_{kk} - \dot{g}_{kk}|^2 = a_k^2 \frac{\lambda^2}{N^2} \beta_k \mathcal{E}_{d,s,k} \frac{1}{(1+\mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})^2} (\mu_{k,\mathbf{I}}^2 + \delta_{u,k}^2 \mu_{k,\mathbf{D}_b}^2 - 2\Re\{d_{u,k} \mu_{k,\mathbf{D}_b} \mu_{k,\mathbf{I}}\})$. Also, $\mu_{k,\mathbf{D}_b} - E[\delta_b \cos \phi_b] \text{tr}\{\dot{\mathbf{T}}_k(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $\mu_{k,\mathbf{I}} - \text{tr}\{\dot{\mathbf{T}}_k(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $|\tilde{g}_{kl}|^2 = \bar{a}_l^2 \frac{\lambda^2}{N^2} \beta_k \mathcal{E}_{d,s,k} \frac{|\nu_{kl}|^2}{(1+\mathcal{E}_{d,s,k} \mu_{l,\mathbf{I}})^2}$, such that, $|\nu_{kl}|^2 - \text{tr}\{\dot{\mathbf{T}}_l^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$,

$$\dot{\mathbf{T}}_k(\epsilon) = \left(\sum_{l=1, l \neq k}^K \frac{\mathcal{E}_{d,s,l}}{1 + \dot{e}_{k,l}(\epsilon)} + \epsilon \right)^{-1} \mathbf{R}, \quad (89)$$

and $\dot{e}_{k,l}(\epsilon)$ can be evaluated by iterating the following equation over n , starting at $n = 1$, with the initialization $\dot{e}_{k,l}^0(\epsilon) = \frac{1}{\epsilon}$:

$$\dot{e}_{k,l}^n(\epsilon) = \text{Tr}\{\mathbf{R}\} \mathcal{E}_{d,s,l} \left(\sum_{\substack{m=1 \\ m \neq k}}^K \frac{\mathcal{E}_{d,s,m}}{1 + \dot{e}_{k,m}^{n-1}(\epsilon)} + \epsilon \right)^{-1}. \quad (90)$$

Similarly, $\bar{g}_{kk} = a_k \frac{\lambda}{N} \sqrt{\beta_k \mathcal{E}_{d,s,k} \frac{\mu_{k,\mathbf{I}}}{1+\mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}}}}$, and $|\hat{g}_{kl}|^2 = \delta_{u,k}^2 a_l^2 \frac{\lambda^2}{N^2} (\beta_k \mathcal{E}_{d,s,l}) \left(\frac{1}{1+\mathcal{E}_{d,s,l} \mu_{l,\mathbf{I}}} \right)^2 \left(\dot{\mu}_{kl,\mathbf{D}_b}^2 + \mathcal{E}_{d,s,k}^2 \frac{\dot{\mu}_{kk,\mathbf{D}_b} \dot{\mu}_{kl,\mathbf{I}}^2}{(1+\mathcal{E}_{d,s,k} \mu_{kk,\mathbf{I}})^2} - 2\mathcal{E}_{d,s,k} \frac{\Re\{\dot{\mu}_{kk,\mathbf{D}_b} \dot{\mu}_{kl,\mathbf{I}} \mu_{kk,\mathbf{I}}^*\}}{(1+\mathcal{E}_{d,s,k} \mu_{kk,\mathbf{I}})} \right)$, wherein, $\dot{\mu}_{kk,\mathbf{D}_b} - E[\delta_b \cos \phi_b] \text{tr}\{\dot{\mathbf{T}}_{kl}(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $\dot{\mu}_{kk,\mathbf{I}} - \text{tr}\{\dot{\mathbf{T}}_{kl}(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $\dot{\mu}_{kl,\mathbf{I}} = \hat{\mathbf{h}}_{d,k}^T \mathbf{Q}_{kl}^{-1} \hat{\mathbf{h}}_{d,l}^* \xrightarrow{\text{a.s.}} 0$, $\dot{\mu}_{kl,\mathbf{I}}^2 - \text{Tr}\{\dot{\mathbf{T}}_{kl}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, $|\dot{\mu}_{kl,\mathbf{D}_b}|^2 - E^2[\delta_b \cos \phi_b] \text{Tr}\{\dot{\mathbf{T}}_{kl}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, and $\dot{\mu}_{kl,\mathbf{D}_b} \dot{\mu}_{kl,\mathbf{I}}^* - E[\delta_b \cos \phi_b] \text{Tr}\{\dot{\mathbf{T}}_{kl}^2(\epsilon)\} \xrightarrow{\text{a.s.}} 0$, with

$$\dot{\mathbf{T}}_{lk}(\epsilon) = \left(\sum_{\substack{m=1 \\ m \neq k,l}}^K \frac{\mathcal{E}_{d,s,m}}{1 + \ddot{e}_{k,l,m}(\epsilon)} + \epsilon \right)^{-1} \mathbf{R}, \quad (91)$$

and $\ddot{e}_{k,l}(\epsilon)$ can be evaluated by iterating the following equation over n , starting at $n = 1$, with the initialization $\ddot{e}_{k,l}^0(\epsilon) = \frac{1}{\epsilon}$:

$$\ddot{e}_{kl,m}^n(\epsilon) = \text{Tr}^2\{\mathbf{R}\} \mathcal{E}_{d,s,m} \left(\sum_{\substack{p=1 \\ p \neq k,l}}^K \frac{\mathcal{E}_{d,s,p}}{1 + \ddot{e}_{kl,p}^{n-1}(\epsilon)} + \epsilon \right)^{-1}. \quad (92)$$

Similarly, the achievable SINRs with blind channel estimation under RZF and MFP can be respectively given as,

$$\gamma_k = \frac{(\text{Tr}\{\mathbf{R}\} E[\delta_b \cos(\phi_b)] \delta_{u,k}^2 a_k^2 + \bar{a}_k^2) \beta_k \mathcal{F}_{d,s,k}}{\beta_k \sum_{l=1, l \neq k}^K \mathcal{F}_{d,s,l} + N_0}, \quad (93)$$

$$\gamma_{k,\text{rzf}} = \frac{\left(a_k^2 \delta_{u,k}^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{k,\mathbf{D}_b}^2}{(1+\mathcal{E}_{d,s,k} \mu_{k,\mathbf{I}})^2} + |\tilde{g}_{kl}|^2 \right)}{\text{Tr}^2\{\mathbf{R}\} \sum_{l=1, l \neq k}^K |\tilde{g}_{kl}|^2 + N_0}. \quad (94)$$

REFERENCES

- [1] A. Chockalingam and B. S. Rajan, *Large MIMO Systems*. New York, NY, USA: Cambridge University Press, 2014.
- [2] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [3] 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.
- [4] T. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 847–860, Oct. 2014.
- [5] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [6] R. Chopra, C. Murthy, H. A. Suraweera, and E. Larsson, "Performance analysis of FDD massive MIMO systems under channel aging," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1094–1108, Feb. 2018.
- [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] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *J. Commun. and Networks*, vol. 15, no. 4, pp. 338–351, Aug. 2013.
- [9] A. K. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1428–1442, Feb. 2017.
- [10] C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou, and Z. Zhang, "Sum-rate and power scaling of massive MIMO systems with channel aging," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4879–4893, Dec. 2015.
- [11] A. K. Papazafeiropoulos and T. Ratnarajah, "Deterministic equivalent performance analysis of time-varying massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5795–5809, Oct. 2015.
- [12] A. Decurvinge, M. Guillaud, and D. T. M. Stock, "Channel covariance estimation in massive MIMO frequency division duplex systems," in *Proc. IEEE Global Commun. Conf. Workshops*, San Diego, CA, USA, Dec. 2015, pp. 1–6.
- [13] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.
- [14] X. Jiang, A. Decurvinge, K. Gopala, F. Kaltenberger, M. Guillaud, D. Stock, and L. Deneire, "A framework for over-the-air reciprocity calibration for TDD massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 5975–5990, Sep. 2018.
- [15] K. Nishimori, K. Cho, Y. Takatori, and T. Hori, "Automatic calibration method using transmitting signals of an adaptive array for TDD systems," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1636–1640, Nov. 2001.
- [16] A. Bourdoux, B. Come, and N. Khaled, "Non-reciprocal transceivers in OFDM/SDMA systems: impact and mitigation," in *Proc. Radio and Wireless Conf. (RAWCON)*, Boston, MA, USA, Aug. 2003, pp. 183–186.

$$\gamma_k = \frac{\text{Tr}\{\mathbf{R}\} a_k^2 \beta_k \mathcal{F}_{d,s,k}}{\text{Tr}\{\mathbf{R}\} a_k^2 \beta_k \mathcal{F}_{d,s,k} |1 - \delta_{u,k} E[\delta_b \cos(\phi_b)]|^2 + \bar{a}_k^2 \beta_k \mathcal{F}_{d,s,k} + \beta_k \sum_{l=1, l \neq k}^K \mathcal{F}_{d,s,l} + N_0} \quad (87)$$

$$\gamma_k = \frac{\text{Tr}^2\{\mathbf{R}\} a_k^2 \lambda^2 \beta_k \mathcal{E}_{d,s,k} \frac{\mu_{k,1}}{(1 + \mathcal{E}_{d,s,k} \mu_{k,1})^2}}{\text{Tr}^2\{\mathbf{R}\} |\hat{g}_{kk} - \dot{g}_{kk}|^2 + \text{Tr}^2\{\mathbf{R}\} \sum_{l=1, l \neq k}^K |\hat{g}_{kl}|^2 + \text{Tr}^2\{\mathbf{R}\} \sum_{l=1}^K |\tilde{g}_{kl}|^2 + N_0}, \quad (88)$$

- [17] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems," in *Proc. Future Network Mobile Summit*, Florence, Italy, Jun. 2010, pp. 1–10.
- [18] J. Shi, Q. Luo, and M. You, "An efficient method for enhancing TDD over the air reciprocity calibration," in *Proc. WCNC*, Cancun, Mexico, Mar. 2011, pp. 339–344.
- [19] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proc. 18th Annual Int. Conf. on Mobile Computing and Networking*, Istanbul, Turkey, 2012, pp. 53–64.
- [20] 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 Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1815–1831, Apr. 2014.
- [21] J. Vieira, F. Rusek, and F. Tufvesson, "Reciprocity calibration methods for massive MIMO based on antenna coupling," in *Proc. Globecom*, Austin, TX, USA, Dec. 2014, pp. 3708–3712.
- [22] J. Vieira, F. Rusek, O. Edfors, S. Malkowsky, L. Liu, and F. Tufvesson, "Reciprocity calibration for massive MIMO: Proposal, modeling, and validation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3042–3056, May 2017.
- [23] D. Mi, M. Dianati, L. Zhang, S. Muhaidat, and R. Tafazolli, "Massive MIMO performance with imperfect channel reciprocity and channel estimation error," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3734–3749, Sep. 2017.
- [24] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov 2014.
- [25] W. Zhang, H. Ren, C. Pan, M. Chen, R. C. de Lamare, B. Du, and J. Dai, "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.
- [26] D. Inserra and A. M. Tonello, "Characterization of hardware impairments in multiple antenna systems for DoA estimation," *J. Electrical and Computer Engg.*, vol. 2011, no. 1, pp. 1–10, Dec. 2011.
- [27] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1458–1461, Oct. 1999.
- [28] V. Adinarayana, Muralikrishna, and P. K. Rajesh Kumar, "Performance comparison of linear precoders for TDD large scale MIMO system with imperfect reciprocity and channel estimations," *Wireless Pers Commun*, pp. 1–13, Oct. 2019.
- [29] A. Manesh, C. Murthy, and R. Annavajjala, "Physical layer data fusion via distributed co-phasing with general signal constellations," *IEEE Trans. Signal Process.*, vol. 63, no. 17, pp. 4660–4672, Sep. 2015.
- [30] R. Chopra, C. R. Murthy, and R. Annavajjala, "Multistream distributed cophasing," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 1042–1057, Feb. 2017.
- [31] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2001.
- [32] R. Chopra, R. Annavajjala, and C. R. Murthy, "Distributed cophasing with autonomous constellation selection," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5798–5811, Nov. 2017.
- [33] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, May 2017.
- [34] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY: Cambridge University Press, 2012.
- [35] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*, 1st ed. Cambridge, UK: Cambridge University Press, 2011.