Multipair Massive MIMO Relaying Systems With One-Bit ADCs and DACs

Chuili Kong, Student Member, IEEE, Amine Mezghani, Member, IEEE, Caijun Zhong, Senior Member, IEEE, A. Lee Swindlehurst, Fellow, IEEE, and Zhaoyang Zhang, Member, IEEE

Abstract—This paper considers a multipair amplify-and-forward massive multipair multiple-input multiple-output (MIMO) relaying system with one-bit analog-to-digital converters and one-bit digital-to-analog converters at the relay. The channel state information is estimated via pilot training, and then utilized by the relay to perform simple maximum-ratio combining/ratio transmission processing. Leveraging on the Bussgang decomposition, an exact achievable rate is derived for the system with correlated quantization noise. Based on this, a closed-form asymptotic approximation for the achievable rate is presented, thereby enabling efficient evaluation of the impact of key parameters on the system performance. Furthermore, power scaling laws are characterized to study the potential energy efficiency associated with deploying massive one-bit antenna arrays at the relay. In addition, a power allocation strategy is designed to compensate for the rate degradation caused by the coarse quantization. Our results suggest that the quality of the channel estimates depends on the specific orthogonal pilot sequences that are used, contrary to unquantized systems where any set of orthogonal pilot sequences gives the same result. Moreover, the sum rate gap between the double-quantized relay system and an ideal nonquantized system is a moderate factor of $4/n^2$ in the low power regime.

Index Terms—Massive MIMO, relays, one-bit quantization, power allocation.

I. INTRODUCTION

MULTIPAIR multiple-input multiple-output (MIMO) relaying networks have recently attracted considerable attention since they can provide a cost-effective way of achieving performance gains in wireless systems via coverage extension and maintaining a uniform quality of service. In such a system, multiple sources simultaneously exchange information with multiple destinations via a shared multiple-antenna relay. Hence, multi-user interference is the primary system bottleneck. The deployment of massive antenna arrays at the relay has been proposed to address this issue due to their ability to suppress interference, provide large array and spatial multiplexing gains, and in turn to yield large improvements in spectral and energy efficiency.

There has recently been considerable research interest in multipair massive MIMO relaying systems. For example, [1] derived the ergodic rate of the system when maximum ratio combining/ratio transmission (MRC/MRT) beamforming is employed and showed that the energy efficiency gain scales with the number of relay antennas in Rayleigh fading channels. Then, [2] extended the analysis to the Ricean fading case and obtained similar power scaling behavior. For full-duplex systems, [3], [4] analytically compared the performance of MRC/MRT and zero-forcing reception/transmission and characterized the impact of the number of user pairs on the spectral efficiency.

All the aforementioned works are based on the assumption of perfect hardware. However, a large number of antennas at the relay implies a very large deployment cost and significant energy consumption if a separate RF chain is implemented for each antenna in order to maintain full beamforming flexibility. In particular, the fabrication cost, chip area and power consumption of the analog-to-digital converters (ADCs) and the digital-to-analog converters (DACs) grow roughly exponentially with the number of quantization bits [5], [6]. The cumulative cost and power required to implement a relay with a very large array can be prohibitive, and thus it is desirable to investigate the use of cheaper and more energy-efficient components, such as low-resolution (e.g., one bit) ADCs and DACs. Fortunately, it has been shown in [7]–[11] that large arrays exhibit a certain resilience to RF hardware impairments that could be caused by such low-cost components.
A. Related Work

Several recent contributions have investigated the impact of low-resolution ADCs on the massive MIMO uplink [12]–[22]. For example, [15] optimized the training pilot length to maximize the spectral efficiency, while [16] revealed that in terms of overall energy efficiency, the optimal level of quantization is 4–5 bits. In [17], the Bussgang decomposition [23] was used to reformulate the nonlinear quantization using a second-order statistically equivalent linear operator, and to derive a linear minimum mean-squared error (LMMSE) channel estimator for one-bit ADCs. In [18], a near-optimal low complexity bit allocation scheme was presented for millimeter wave channels exhibiting sparsity. The work of [19] examined the impact of one-bit ADCs on wideband channels with frequency-selective fading, while [24], [25] investigated the impact of the large-scale fading and showed the importance of power control in quantized systems. Other work has focused on balancing the spectral and energy efficiency, either through the combined use of hybrid architectures with a small number of RF chains and low resolution ADCs, or using mixed ADCs architectures with high and low resolution.

In contrast to the uplink case, there are relatively fewer contributions that consider the massive MIMO downlink with low-resolution DACs. In [26], it was shown that performance approaching the unquantized case can be achieved using DACs with only 3–4 bits of resolution. The nearly optimal quantized Wiener precoder in terms of minimizing the mean-square error (MSE) with low-resolution DACs was studied in [27], and the resulting solution was shown to outperform the conventional Wiener precoder with 4–6 bits of resolution at high signal-to-noise ratio (SNR). For the case of one-bit DACs, [7], [28] showed that even simple MRT precoding can achieve reasonable results. In [29], an LMMSE precoder that exploits analog and digital processing was proposed by taking the quantization non-linearities into account, and different precoding schemes were compared in terms of uncoded bit error rate.

B. Contributions

All these prior works are for single-hop systems rather than dual-hop connections via a relay. Recently, [30] considered a relay-based system that uses mixed-resolution ADCs at the base station. Unlike [30], we consider a multipair amplify-and-forward (AF) relaying system where the relay uses both one-bit ADCs and one-bit DACs. Since channel estimation is critical for the one-bit receivers, and important work such as [31], [32] for channel estimation are missing, we take the channel training into account in the paper. The one-bit ADCs cause errors in the channel estimation stage and subsequently in the reception of the uplink data; then, after a linear transformation, the one-bit ADCs produce distortion when the downlink signal is coarsely quantized. In this paper, we present a detailed performance investigation of the achievable rate of such doubly quantized systems. In particular, the main contributions are summarized as follows:

- We investigate a multipair AF relaying system that employs one-bit ADCs and DACs at the relay and uses MRC/MRT beamforming to process the signals. We take the correlation of the quantization noise into account, and present an exact achievable rate by using the arcsine law. Then, we use asymptotic arguments to provide an approximate closed-form expression for the achievable rate. Numerical results demonstrate that the approximate rate is accurate in typical massive MIMO scenarios, even with only a moderate number of users.
- We show that the channel estimation accuracy of the quantized system depends on the specific orthogonal pilot matrix that is used, which is in contrast to unquantized systems where any orthogonal pilot sequence yields the same result. We consider the specific case of identity and Hadamard pilot matrices, and we show that the identity training scheme provides better channel estimation performance for users with weaker than average channels, while the Hadamard training sequence is better for users with stronger channels.
- We compare the achievable rate of different ADC and DAC configurations, and show that a system with one-bit DACs and perfect ADCs outperforms a system with one-bit ADCs and perfect DACs. Focusing on the low transmit power regime, we show that the sum rate of the relay system with one-bit ADCs and DACs is $4/\pi^2$ times that achievable with perfect ADCs and DACs. Also, it is shown that the transmit power of each source or the relay can be reduced inversely proportional to the number of relay antennas, while maintaining a given quality-of-service.
- We formulate a power allocation problem to allocate power to each source and the relay, subject to a sum power budget. Locally optimum solutions are obtained by solving a sequence of geometric programming (GP) problems. Our numerical results suggest that the power allocation strategy can efficiently compensate for the rate degradation caused by the coarse quantization.

C. Paper Outline and Notations

The remainder of the paper is organized as follows. Section II introduces the multipair AF relaying system model under consideration. Section III presents an approximate closed-form expression for the sum rate, and compares the rate achieved with different ADC and DAC configurations. Section IV formulates a power allocation problem to compensate for the rate loss caused by the coarse quantization. Numerical results are provided in Section V. Finally, Section VI summarizes the key findings.

**Notation:** We use bold upper case letters to denote matrices, bold lower case letters to denote vectors and lower case letters to denote scalars. The notation $(\cdot)^H$, $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^{-1}$ respectively represent the conjugate transpose operator, the conjugate operator, the transpose operator, and the inverse of a matrix. The Euclidian norm is denoted by $\| \cdot \|$, the absolute value by $| \cdot |$, and $[A]_{m,n}$ represents the $(m,n)$-th entry of $A$. Also, $x \sim \mathcal{CN}(0, \Sigma)$ denote a circularly symmetric complex Gaussian random vector with zero mean and covariance matrix $\Sigma$, while $I_k$ is the identity matrix of size $k$. The symbol $\otimes$ is the Kronecker product, vec $(A)$ represents a column vector containing the stacked columns of matrix $A$, diag $(B)$ denotes a diagonal matrix formed by the diagonal elements of matrix $B$, $\mathbb{R}^C$ and $\mathbb{C}$ stand for the real and imaginary part of $C$, respectively. Finally, the statistical expectation operator is represented by $\mathbb{E}(\cdot)$, the variance operator is $\text{Var}(\cdot)$, and the trace is denoted by $\text{tr}(\cdot)$. 
the large-scale path-loss, which is assumed to be constant over many coherence intervals and known a priori.

A. Channel Estimation

We assume training pilots are used to estimate the channel matrices \( G_{SR} \) and \( G_{RD} \), as in other massive MIMO AF relaying systems [34]. Therefore, during each coherence interval of length \( \tau_p \) (in symbols), all sources simultaneously transmit their mutually orthogonal pilot sequences \( \Phi_S \in \mathbb{C}^{q \times K} \) satisfying \( \Phi_S^H \Phi_S = \tau_p I_K \) to the relay while the destinations remain silent (\( K \leq \tau_p < \tau_p/2 \)). Afterwards, all destinations simultaneously transmit their mutually orthogonal pilot sequences \( \Phi_D \in \mathbb{C}^{q \times K} \) satisfying \( \Phi_D^H \Phi_D = \tau_p I_K \) to the relay while the sources remain silent.

Since the channels \( G_{SR} \) and \( G_{RD} \) are estimated in the same fashion, we focus only on the first link \( G_{SR} \). The received training signal at the relay is given by

\[
Y_p = \sqrt{p}G_{SR}\Phi_S^H + N_p, \tag{3}
\]

where \( p_R \) represents the transmit power of each pilot symbol, and \( N_p \) denotes the noise at the relay, which has i.i.d. \( \mathcal{CN}(0, 1) \) elements. After vectorizing the matrix \( Y_p \), we obtain

\[
y_p = \text{vec}(Y_p) = \Phi_S g_{SR} + \mathbf{n}_p, \tag{4}
\]

where \( \Phi_S = \Phi_S \otimes \sqrt{p} I_M \), \( g_{SR} = \text{vec}(G_{SR}) \), and \( \mathbf{n}_p = \text{vec}(N_p) \).

1) One-Bit ADCs: After the one-bit ADCs, the quantized signal can be expressed as

\[
r_p = Q(y_p), \tag{5}
\]

where \( Q(\cdot) \) denotes the one-bit quantization operation, which separately processes the real and imaginary parts of the signal. Therefore, the output set of the one-bit ADCs is \( \frac{1}{\sqrt{2}} \{ \pm 1 \pm j \} \).

Using the Bussgang decomposition [23], [35], \( \tau_p \) can be represented by a linear signal component and an uncorrelated quantization noise \( q_p \):

\[
r_p = \mathbf{A}_p y_p + q_p, \tag{6}
\]

where \( \mathbf{A}_p \) is the linear operator obtained by minimizing the power of the quantization noise \( \text{E}\{||q_p||^2\} \), and is expressed as [8], [17]

\[
A_p = \sqrt{\frac{\tau_p}{\pi}} \text{diag}(R_{yy_yy})^{-1/2}, \tag{7}
\]

where

\[
R_{yy_yy} = \Phi_S \mathbf{D}_{SR} \Phi_S^H + I_M \tau_p, \tag{8}
\]

with \( \mathbf{D}_{SR} = (\mathbf{D}_{SR} \otimes I_M) \) and \( \mathbf{D}_{SR} \) being a diagonal matrix whose elements are \( \mathbf{D}_{SR_{kk}^{\text{rd}}} = \beta_{SR_{kk}^{\text{rd}}} \) for \( k = 1, \ldots, K \).

The covariance matrix of the quantization noise \( q_p \) is given by [8], [17]

\[
\mathbf{R}_{qq_{qq}} = \frac{2}{\pi} \left( \arcsin(J) + j \arcsin(K) \right) - \frac{2}{\pi} (J + jK), \tag{9}
\]

where

\[
J = \text{diag} \left( R_{yy_yy} \right)^{-1/2} \otimes \text{diag} \left( R_{yy_yy} \right)^{-1/2}, \tag{10}
\]

\[
K = \text{diag} \left( R_{yy_yy} \right)^{-1/2} \otimes \text{diag} \left( R_{yy_yy} \right)^{-1/2}. \tag{11}
\]
2) **LMMSE Estimator**: Based on the observation $r_p$ and the training pilots $\Phi_S$, we use the LMMSE technique to estimate $G_{SR}$. Hence, the estimated channel $\hat{g}_{SR}$ is given by

$$\hat{g}_{SR} = R_{\hat{g}_{SR}r_p} R_{r_p}^{-1} r_p,$$  \hfill (12)

As a result, the covariance matrix of the estimated channel $\hat{g}_{SR}$ is expressed as

$$R_{\hat{g}_{SR} \hat{g}_{SR}^H} = R_{\hat{g}_{SR}r_p} R_{r_p}^{-1} R_{r_p}^{-1} R_{r_p} R_{\hat{g}_{SR}r_p}.$$

$$= \hat{D}_{SR} \hat{\Phi}_S^H \left( \hat{\Phi}_S \hat{D}_{SR} \hat{\Phi}_S^H + A_p A_p^H + R_{q_i q_p} \right)^{-1} \hat{\Phi}_S \hat{D}_{SR}. \hfill \hfill (13)$$

where $\hat{\Phi}_S = A_p \hat{\Phi}_S$.

**Remark 1**: From (13), we can see that $R_{\hat{g}_{SR} \hat{g}_{SR}^H}$ is a non-trivial function of $\hat{\Phi}_S$, which indicates that the quality of the channel estimates depends on the specific realization of the pilot sequence, which is contrary to unquantized systems where any set of orthogonal pilot sequences gives the same result.

**Remark 2**: Although our conclusion in Remark 1 is obtained based on the LMMSE estimator, it also holds for the maximum likelihood estimator [36].

In the following, we study the performance of two specific pilot sequences to show how the pilot matrix affects the channel estimation. Here, we choose $\tau_p = K$, which is the minimum possible value to get unambiguous channel estimates and also simpler to analyze than the case $\tau_p > K$.

a) **Identity Matrix**: In this case, $\Phi_S = \sqrt{K} I_K$, and hence we have

$$R_{r_p r_p} = K_p \hat{D}_{SR} + I_{MK}. \hfill \hfill (14)$$

Consequently,

$$A_p = \sqrt{\frac{2}{\pi}} \left( K_p \hat{D}_{SR} + I_{MK} \right)^{-1/2} = \hat{A}_p \otimes I_M, \hfill \hfill (15)$$

$$R_{q_i q_p} = \left( 1 - \frac{2}{\pi} \right) I_{MK}, \hfill \hfill (16)$$

where $\hat{A}_p$ is a diagonal matrix with $[\hat{A}_p]_{kk} = \alpha_{p,k} = \sqrt{\frac{1}{\pi} \sum_{i=1}^{K} \beta_{SR,k} + 1}$. Substituting (15) and (16) into (13), we obtain

$$R_{\hat{g}_{SR} \hat{g}_{SR}^H} = Q_{SR}^{(1)} \otimes I_M, \hfill \hfill (17)$$

where $Q_{SR}^{(1)}$ is a diagonal matrix with elements

$$[Q_{SR}^{(1)}]_{kk} = \sigma_{SR,k}^2 = \frac{2}{\pi} \frac{K_p \beta_{SR,k}^2}{K_p \beta_{SR,k}^2 + 1}. \hfill \hfill (18)$$

b) **Hadamard Matrix**: In this case, every element of $\Phi_S$ is $+1$ or $-1$, and hence we have

$$A_p = \sqrt{\frac{2}{\pi}} \frac{1}{K_p \sum_{k=1}^{K} \beta_{SR,k} + 1} I_{MK} \hfill \hfill (19)$$

$$R_{q_i q_p} = \left( 1 - \frac{2}{\pi} \right) I_{MK}. \hfill \hfill (20)$$

Substituting (19) and (20) into (13), we obtain

$$R_{\hat{g}_{SR} \hat{g}_{SR}^H} = Q_{SR}^{(2)} \otimes I_M, \hfill \hfill (21)$$

where $Q_{SR}^{(2)}$ is a diagonal matrix with entries

$$[Q_{SR}^{(2)}]_{kk} = \kappa_{SR,k}^2 = \frac{K_p \alpha_p^2 \beta_{SR,k}^2}{K_p \alpha_p^2 \beta_{SR,k}^2 + \alpha_p^2 + 1 - \frac{2}{\pi}}, \hfill \hfill (22)$$

where

$$\alpha_p = \sqrt{\frac{2}{\pi}} \frac{1}{K_p \sum_{k=1}^{K} \beta_{SR,k} + 1}. \hfill \hfill (23)$$

For both cases, the channels from the sources to the relay $g_{SR,k}$ can be decomposed as

$$g_{SR,k} = \hat{g}_{SR,k} + e_{SR,k}, \hfill \hfill (24)$$

where $e_{SR,k}$ is the estimation error vector which includes the effect of the coarse quantization. The elements of $\hat{g}_{SR,k}$ and $e_{SR,k}$ are respectively distributed as $CN(0, \sigma_{SR,k}^2)$ and $CN(0, \kappa_{SR,k}^2)$ when $G_{SR}$ is an identity matrix, while they are distributed as $CN(0, \sigma_{SR,k}^2)$ and $CN(0, \kappa_{SR,k}^2)$ when $G_{SR}$ is a Hadamard matrix, where $\sigma_{SR,k}^2 = \beta_{SR,k}^2 - \sigma_{RD,k}^2$ and $\kappa_{SR,k}^2 = \beta_{SR,k}^2 - \kappa_{RD,k}^2$.

In what follows we define $\hat{G}_{SR} = [\hat{g}_{SR,1}, \ldots, \hat{g}_{SR,K}]$ and $E_{SR} = [e_{SR,1}, \ldots, e_{SR,K}]$.

Similarly, the channels from the relay to the destinations $g_{RD,k}$ can be decomposed as

$$g_{RD,k} = \hat{g}_{RD,k} + e_{RD,k}, \hfill \hfill (25)$$

where $\hat{g}_{RD,k}$ is the estimated channel and $e_{RD,k}$ is the estimation error vector which includes the effect of the coarse quantization. The elements of $\hat{g}_{RD,k}$ and $e_{RD,k}$ are respectively distributed as $CN(0, \sigma_{RD,k}^2)$ and $CN(0, \kappa_{RD,k}^2)$ when $G_{RD}$ is an identity matrix, while they are $CN(0, \sigma_{RD,k}^2)$ and $CN(0, \kappa_{RD,k}^2)$ when $G_{RD}$ is a Hadamard matrix, where

$$\sigma_{RD,k}^2 = \frac{2}{\pi} \frac{K_p \beta_{RD,k}^2}{K_p \beta_{RD,k}^2 + 1} \hfill \hfill (26)$$

$$\kappa_{RD,k}^2 = \frac{K_p \alpha_p^2 \beta_{RD,k}^2}{K_p \alpha_p^2 \beta_{RD,k}^2 + \alpha_p^2 + 1 - \frac{2}{\pi}}, \hfill \hfill (27)$$

with

$$\alpha_p = \sqrt{\frac{2}{\pi}} \frac{1}{K_p \sum_{k=1}^{K} \beta_{RD,k} + 1}. \hfill \hfill (28)$$

and $\sigma_{RD,k}^2 = \beta_{RD,k}^2 - \sigma_{RD,k}^2$ and $\kappa_{RD,k}^2 = \beta_{RD,k}^2 - \kappa_{RD,k}^2$. We also define $G_{RD} = [\hat{g}_{RD,1}, \ldots, \hat{g}_{RD,K}]$ and $E_{RD} = [e_{RD,1}, \ldots, e_{RD,K}]$.

For the channel from the $k$-th source to the relay, the MSE is given by

$$MSE_{SR,k} = \mathbb{E} \left\{ \left| \hat{g}_{SR,k} - g_{SR,k} \right|^2 \right\}. \hfill \hfill (29)$$

Based on the above results, we have $MSE_{SR,k} = \sigma_{SR,k}^2$ for the identity matrix and $MSE_{SR,k} = \kappa_{SR,k}^2$ for the Hadamard matrix. The following proposition compares the MSE of the two approaches.

**Proposition 1**: For estimating the channel $g_{SR,k}$, the identity matrix is preferable to the Hadamard matrix for user $k$ if $\beta_{SR,k} < \frac{1}{K} \sum_{i=1}^{K} \beta_{SR,i}$, and vice versa.
Proof: The proof is trivial since \( \tilde{\sigma}_{SR,k}^2 < \tilde{\kappa}_{SR,k}^2 \) if \( \beta_{SR,k} < \frac{1}{\pi} \sum_{i=1}^{K} \beta_{SR,i} \).

Proposition 1 reveals that the accuracy of the individual channel estimates depends on the particular choice of the orthogonal training scheme, contrary to the ideal case without quantization. This finding agrees with the results in [17], [36]. More precisely, the scaled identity matrix is beneficial for any user with higher path loss than the average. This is because a weak user benefits from being the only one transmitting at a given time, without the presence of stronger users that dominate the behavior of the ADC. In the case of Hadamard matrix, all users are transmitting simultaneously, resulting in an average quantization noise level for all users jointly, which is advantageous for users with stronger channels.

The question of optimizing the pilot sequence for a given performance metric is an interesting one, but is beyond the scope of the paper. For simplicity, we will assume the identity matrix approach in which each user’s channel is estimated one at a time.

B. Data Transmission

1) Quantization with One-Bit ADCs: With one-bit ADCs at the receiver, the resulting quantized signals can be expressed as

\[
\tilde{y}_R = \tilde{Q}(y_R) = A_x y_R + q_d,
\]

where \( A_x \) is the linear operator, which is chosen to make the quantization noise \( q_d \) uncorrelated with \( y_R \). By adopting the same technique in the previous subsection, we have

\[
A_x = \sqrt{\frac{2}{\pi}} \text{diag} (R_{y_a,y_a})^{-1/2}
\]

\[
R_{q_d,q_d} = \frac{2}{\pi} (\arcsin(X) + j \arcsin(Y)) - \frac{2}{\pi} (X + j Y),
\]

where

\[
X = \text{diag} (R_{y_a,y_a})^{-1/2} \Re (R_{x_a,x_a} y_a) \text{diag} (R_{x_a,x_a})^{-1/2}
\]

\[
Y = \text{diag} (R_{y_a,y_a})^{-1/2} \Im (R_{x_a,x_a} y_a) \text{diag} (R_{x_a,x_a})^{-1/2}
\]

\[
R_{y_a,y_a} = G_{SR} P_x G_{SR}^H + I_M.
\]

2) Digital Linear Processing: We assume that the relay adopts an AF protocol to process the quantized signals by one-bit ADCs \( \tilde{y}_R \), yielding

\[
x_R = W \tilde{y}_R,
\]

where \( W = G_{SR}^H P_x G_{SR}^H \) for MRC/MRT beamforming.\(^2\)

3) Quantization with One-Bit DACs: Assuming one-bit DACs at the transmitter, the resulting quantized signals to be sent by the relay’s transmit antennas can be expressed as:

\[
\tilde{x}_R = \tilde{Q}(x_R) = A_d x_R + q_d,
\]

where \( A_d \) is the linear operator, and \( q_d \) is the quantization noise at the relay’s transmit antennas, which is uncorrelated with \( x_R \).

Due to the one-bit DACs, we have \( \mathbb{E} \left\{ |\tilde{x}_R|^2 \right\} = M \). Therefore, the normalization factor \( \gamma \) (c.f. (2)) can be expressed as

\[
\gamma = \sqrt{\frac{PR}{M}}
\]

Following in the same fashion as with the ADCs derivations, we obtain [7]

\[
A_d = \sqrt{\frac{2}{\pi}} \text{diag} (R_{x_a,x_a})^{-1/2}
\]

\[
R_{q_d,q_d} = \frac{2}{\pi} (\arcsin(U) + j \arcsin(V)) - \frac{2}{\pi} (U + j V),
\]

where

\[
U = \text{diag} (R_{x_a,x_a})^{-1/2} \Re (R_{x_a,x_a} y_a) \text{diag} (R_{x_a,x_a})^{-1/2}
\]

\[
V = \text{diag} (R_{x_a,x_a})^{-1/2} \Im (R_{x_a,x_a} y_a) \text{diag} (R_{x_a,x_a})^{-1/2}
\]

\[
R_{y_a,y_a} = W R \hat{y}_R W^H
\]

\[
R_{\hat{y}_R} = A_d R_{y_a,y_a} A_d^H + R_{q_d,q_d}.
\]

III. ACHIEVABLE RATE ANALYSIS

In this section, we investigate the achievable rate of the considered system. In particular, we first provide an expression for the exact achievable rate, which is applicable to arbitrary system configurations. Then we use asymptotic arguments to derive an approximate rate to provide some key insights.

A. Exact Achievable Rate Analysis

We consider the realistic case where the \( K \) destinations do not have access to the instantaneous CSI, which is a typical assumption in the massive MIMO literature since the dissemination of instantaneous CSI leads to excessively high computational and signaling costs for very large antenna arrays. Hence, \( D_k \) uses only statistical CSI to decode the desired signal. Combining (1), (2), (30), (34), and (35) yields the received signal at the \( k \)-th destination

\[
y_{D,k} = \gamma \sqrt{\frac{P_S}{P_d,k}} \mathbb{E} \left\{ g_{RD,k}^T A_d W A_d g_{SR,k} \right\} x_{S,k} + \tilde{n}_{D,k},
\]

where

\[
\tilde{n}_{D,k} = \gamma \sqrt{\frac{P_S}{P_d,k}} \mathbb{E} \left\{ g_{RD,k}^T A_d W A_d g_{SR,k} \right\} x_{S,k}
\]

\[
+ \gamma \sum_{i \neq k} \sqrt{\frac{P_S}{P_d,i}} \mathbb{E} \left\{ g_{RD,i}^T A_d W A_d g_{SR,i} \right\} x_{S,i} + \gamma \sum_{i \neq k} \mathbb{E} \left\{ g_{RD,i}^T A_d W A_d n_{D,k} \right\} \text{interpair interference}
\]

\[
+ \gamma \mathbb{E} \left\{ g_{RD,k}^T A_d W q_d \right\} + \gamma \mathbb{E} \left\{ g_{RD,k}^T q_d \right\} + \mathbb{E} \left\{ n_{D,k} \right\} \text{quantization noise of ADCs}
\]

\[
\text{quantization noise of DACs}
\]

where \( n_{D,k} \) is the \( k \)-th element of the noise vector \( n_D \). Noticing that the “desired signal” and the “effective noise” in (39) are uncorrelated, and on the fact that the worst-case uncorrelated additive noise is independent Gaussian, we obtain the following...
achievable rate for the $k$-th destination:

$$R_k = \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + \frac{A_k}{B_k + C_k + D_k + E_k + F_k + \frac{1}{\gamma}} \right),$$  \hspace{0.5cm} (40)

where

$$A_k = p_{s,k} |E \{ g_{RD,k}^T A_d W_A E_{SR,k} \} |^2$$  \hspace{0.5cm} (41)

$$B_k = p_{s,k} \text{Var} \{ g_{RD,k}^T A_d W_A E_{SR,k} \}$$  \hspace{0.5cm} (42)

$$C_k = \sum_{i \neq k} p_s, i |E \{ |g_{RD,k}^T A_d W_A E_{SR,i} |^2 \}$$  \hspace{0.5cm} (43)

$$D_k = E \{ \| E_{RD,k} A_d W_A \|^2 \}$$  \hspace{0.5cm} (44)

$$E_k = E \{ |g_{RD,k}^T A_d W_A E_{SR,i} A_d^H g_{RD,k} |^2 \}$$  \hspace{0.5cm} (45)

$$F_k = E \{ |g_{RD,k}^T A_d E_{SR,i} |^2 \}.$$  \hspace{0.5cm} (46)

### B. Asymptotic Simplifications

As we can see, the matrices $R_{q,q_i}, A_d$, and $R_{q_i,q_i}$ all involve arctan functions, which does not give much insight into how the rate changes with various parameters. To facilitate the analysis, we focus on the asymptotic regime for a large number of users ($K \gg 1$), in which (33) can be approximated by

$$R_{x,y} \approx \text{diag} (R_{x,y}) \approx \left( 1 + \sum_{k=1}^{K} p_{s,k} \beta_{SR,k} \right) I_M.$$  \hspace{0.5cm} (47)

Substituting (47) into (31) and (32), we have

$$A_k \approx \sqrt{\frac{\gamma}{\pi}} \sqrt{\frac{1}{1 + \sum_{k=1}^{K} p_{s,k} \beta_{SR,k}}} I_M = \alpha_d I_M.$$  \hspace{0.5cm} (48)

$$R_{q,q_i} \approx \left( 1 - \frac{2}{\pi} \right) I_M.$$  \hspace{0.5cm} (49)

Similarly, asymptotically we have

$$R_{x,y} \approx \text{diag} (R_{x,y}) \approx \eta_d I_M,$$  \hspace{0.5cm} (50)

where

$$\eta_d = M \left( \alpha_a^2 + 1 - \frac{2}{\pi} \right) \sum_{k=1}^{K} \sigma_{SR,k}^2 \sigma_{RD,k}^2$$

$$+ M \alpha_a^2 \sum_{k=1}^{K} \sigma_{SR,k}^2 \sigma_{RD,k}^2 M \sigma_{SR,k}^2 \sigma_{RD,k}^2 + \sum_{k=1}^{K} p_{s,i} \beta_{SR,i}.$$  \hspace{0.5cm} (51)

Note that the justification for calculating the approximate $R_{x,y}$ can be found in the first part of Appendix A.

As a result, the matrices $A_d$ and $R_{q_i,q_i}$ can be approximated by

$$A_d \approx \sqrt{\frac{2}{\pi \eta_d}} I_M = \alpha_d I_M$$  \hspace{0.5cm} (52)

$$R_{q_i,q_i} \approx \left( 1 - \frac{2}{\pi} \right) I_M.$$  \hspace{0.5cm} (53)

### C. Approximate Rate Analysis

In this section, we derive a simpler closed-form approximation for the achievable rate. Substituting (48), (49), (52), and (53) into (40), the exact achievable rate $R_k$ can be approximated by

$$\tilde{R}_k = \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + \frac{\tilde{A}_k}{\tilde{B}_k + \tilde{C}_k + \tilde{D}_k + \tilde{E}_k + \tilde{F}_k + \tilde{G}_k} \right),$$  \hspace{0.5cm} (54)

where

$$\tilde{A}_k = p_{s,k} |E \{ g_{RD,k}^T W_{SR,k} \} |^2,$$  \hspace{0.5cm} (55)

$$\tilde{B}_k = p_{s,k} \text{Var} \{ g_{RD,k}^T W_{SR,k} \},$$  \hspace{0.5cm} (56)

$$\tilde{C}_k = \sum_{i \neq k} p_s, i |E \{ |g_{RD,k}^T W_{SR,i} A_d^H g_{RD,k} |^2 \},$$  \hspace{0.5cm} (57)

$$\tilde{D}_k = E \{ \| E_{RD,k} W \|^2 \},$$  \hspace{0.5cm} (58)

$$\tilde{E}_k = \left( 1 - \frac{2}{\pi} \right) \frac{1}{\alpha_a^2} E \{ \| g_{RD,k} W \|^2 \},$$  \hspace{0.5cm} (59)

$$\tilde{F}_k = \left( 1 - \frac{2}{\pi} \right) \frac{1}{\alpha_a^2} E \{ \| g_{RD,k} \|^2 \},$$  \hspace{0.5cm} (60)

$$\tilde{G}_k = \frac{1}{\gamma^2 \alpha_a^2}.$$  \hspace{0.5cm} (61)

With this expression, we can compute $\tilde{R}_k$ by using random matrix theory and present a closed-form approximate rate for the $k$-th destination, as formalized in the following theorem.

**Theorem 1**: With one-bit ADCs and DACs and MRC/MRT beamforming at the relay, the approximate achievable rate of the $k$-th destination is given by (54), where

$$\tilde{A}_k = p_{s,k} M^4 \sigma_{SR,k}^2 \sigma_{RD,k}^2,$$  \hspace{0.5cm} (62)

$$\tilde{B}_k = p_{s,k} M^2 (M \sigma_{SR,k}^2 \sigma_{RD,k}^2 + \beta_{SR,k} t_k),$$  \hspace{0.5cm} (63)

$$\tilde{C}_k = M^2 \sum_{i \neq k} p_s, i (M \sigma_{SR,i}^2 \sigma_{RD,i} \beta_{RD,i} + \beta_{SR,i} t_k),$$  \hspace{0.5cm} (64)

$$\tilde{D}_k = M^2 t_k,$$  \hspace{0.5cm} (65)

$$\tilde{E}_k = \frac{\pi}{2} - 1 \left( 1 + \sum_{k=1}^{K} p_{s,k} \beta_{SR,k} \right) M^2 t_k,$$  \hspace{0.5cm} (66)

$$\tilde{F}_k = \beta_{RD,k} \left( \frac{\pi}{2} - 1 \right) M^3 \sum_{k=1}^{K} p_{s,k} \sigma_{SR,k}^2 \sigma_{RD,k}^2 + \beta_{RD,k} \frac{M^2 \pi^2}{2} \left( \frac{\pi}{2} - 1 \right) \left( 1 + \sum_{k=1}^{K} p_{s,k} \beta_{SR,k} \right) \sum_{k=1}^{K} \sigma_{SR,k}^2 \sigma_{RD,k}^2,$$  \hspace{0.5cm} (67)

$$\tilde{G}_k = \frac{M^3 \pi^2}{2 \pi} \sum_{k=1}^{K} p_{s,k} \sigma_{SR,k}^2 \sigma_{RD,k}^2 + \frac{M^2 \pi^2}{4 \pi} \left( 1 + \sum_{k=1}^{K} p_{s,k} \beta_{SR,k} \right) \sum_{k=1}^{K} \sigma_{SR,k}^2 \sigma_{RD,k}^2.$$  \hspace{0.5cm} (68)
with \( t_k = M \sigma_{\text{RD},k}^4 \sigma_{\text{SR},k}^2 + \beta_{\text{RD},k} \sum_{n=1}^{K} \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \).

**Proof:** See Appendix A.

From Theorem 1, we can more readily see the impact of key parameters on the achievable rate. For instance, \( \hat{R}_k \) decreases with the number of user pairs \( K \) according to (63)–(68). This is expected since a higher number of users increases the amount of inter-user interference. In addition, \( \hat{R}_k \) is an increasing function of \( M \) since the signal-to-interference plus noise ratio (SINR) of (54) can be expressed as \( a M^2 + b M \), where \( a \) and \( b \) are independent of \( M \), which reveals that increasing the number of relay’s antennas always boosts the system performance. As \( p_{S,k} \) approaches infinity, \( \hat{R}_k \) converges to a constant that is independent of \( p_{S,k} \). In this case, the system becomes interference-limited.

To quantify the impact of the double quantization on system performance, in the following corollaries we compare the achievable rate with different ADC and DAC configurations.

**Corollary 1:** With perfect ADCs and one-bit DACs, and using MRC/MRT beamforming, the achievable rate of the \( k \)-th destination can be approximated as

\[
R^\text{pA}_k = \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + \frac{\hat{A}_k}{\hat{B}_k + \hat{C}_k + \hat{D}_k + \left( \frac{\pi}{2} - 1 \right) M \beta_{\text{RD},k} \hat{\alpha}_d + \frac{\pi}{2} \hat{G}_k} \right),
\]

(69)

where

\[
\hat{\alpha}_d = M \sum_{k=1}^{K} \hat{\sigma}_{\text{SR},k}^2 \hat{\sigma}_{\text{RD},k}^2 + M \sum_{k=1}^{K} \hat{\sigma}_{\text{SR},k}^2 \hat{\sigma}_{\text{RD},k}^2 \left( M p_{S,k} \hat{\sigma}_{\text{SR},k}^2 + \sum_{i=1}^{K} p_{S,i} \beta_{\text{SR},i} \right),
\]

(70)

with \( \hat{\sigma}_{\text{SR},k}^2 = \frac{K \beta_{\text{SR},k} p_{S,k}}{K \beta_{\text{SR},k} p_{S,k} + 1} \) and \( \hat{\sigma}_{\text{RD},k}^2 = \frac{K \beta_{\text{RD},k} p_{p} - 1}{K \beta_{\text{RD},k} p_{p} + 1} \); \( \hat{A}_k, \hat{B}_k, \hat{C}_k, \hat{D}_k \) can be obtained by replacing \( \hat{\sigma}_{\text{SR},k}^2 \) and \( \hat{\sigma}_{\text{RD},k}^2 \) with \( \sigma_{\text{SR},k}^2 \) and \( \sigma_{\text{RD},k}^2 \) in \( \hat{A}_k, \hat{B}_k, \hat{C}_k, \hat{D}_k \), respectively.

**Corollary 2:** With perfect DACs and one-bit ADCs, and using MRC/MRT beamforming, the achievable rate of the \( k \)-th destination can be approximated as

\[
R^\text{pD}_k = \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + \frac{\hat{A}_k}{\hat{B}_k + \hat{C}_k + \hat{D}_k + \left( \frac{\pi}{2} - 1 \right) \hat{G}_k} \right).
\]

(71)

**Corollary 3:** With perfect ADCs and Dacs, and using MRC/MRT beamforming, the achievable rate of the \( k \)-th destination can be approximated as

\[
R^p_k = \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + \frac{\hat{A}_k}{\hat{B}_k + \hat{C}_k + \hat{D}_k + \hat{G}_k} \right).
\]

(72)

Corollaries 1–3 together with Theorem 1 provide four cases with different ADC/DAC configurations at the relay: 1) Case I: perfect ADCs and DACs; 2) Case II: perfect ADCs and one-bit DACs; 3) Case III: one-bit ADCs and perfect DACs; 4) Case IV: one-bit ADCs and Dacs. The relative performance of these four configurations is described below.

**Proposition 2:** As the number of relay antennas becomes very large, we have

\[
R^\text{pA}_k > R^\text{pD}_k > R^\text{RD}_k > \hat{R}_k.
\]

(73)

**Proof:** See Appendix B.

Proposition 2 indicates that the rate of the system with perfect ADCs and one-bit DACs is higher than that of one-bit ADCs and perfect DACs system. This is because one-bit ADCs cause both channel estimation errors and rate degradation, while one-bit DACs only lead to a rate reduction. For what follows, we define the three rate ratios

\[
[\delta_1, \delta_2, \delta_3] = \left[ \frac{R^\text{pA}_k}{\hat{R}_k}, \frac{R^\text{pD}_k}{\hat{R}_k}, \frac{R^\text{RD}_k}{\hat{R}_k} \right].
\]

(74)

We will compare these ratios for low SNR situations where massive MIMO systems are likely to operate. Here, we consider two cases: a) the transmit power of each source scales as \( p_S = E_S/M \) (where we define \( p_S = p_{S,k} \) for \( k = 1, \ldots, K \)) with fixed \( E_S \), while \( p_p \) and \( p_R \) are fixed. This case focuses on the potential power savings of the sources; b) the transmit power of the relay scales as \( p_R = E_R/M \) with fixed \( E_R \), while \( p_S \) and \( p_R \) are fixed. This case focuses on the potential power savings of the relay.

**Proposition 3:** With \( p_S = E_S/M \), and \( E_S, p_p, p_R \) fixed, we have

\[
\frac{\hat{R}_k}{\hat{R}_k} \rightarrow \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + \frac{2}{\pi} E_S \alpha_{\text{SR},k}^2 \right),
\]

(75)

\[
\frac{R^\text{p}_k}{\hat{R}_k} \rightarrow \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + E_S \alpha_{\text{SR},k}^2 \right),
\]

(76)

as \( M \rightarrow \infty \). In addition, if \( E_S \rightarrow 0 \), the rate ratios are given by

\[
[\delta_1, \delta_2, \delta_3] = [1, 4/\pi^2, 4/\pi^2].
\]

(77)

**Proposition 4:** With \( p_R = E_R/M \), and \( E_R, p_p, p_S \) fixed, we have

\[
\frac{\hat{R}_k}{\hat{R}_k} \rightarrow \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + \frac{2 E_R \alpha_{\text{SR},k}^2 \sigma_{\text{RD},k}^2}{\pi \sum_{k=1}^{K} \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^2} \right),
\]

(78)

\[
\frac{R^\text{p}_k}{\hat{R}_k} \rightarrow \frac{\tau_c - 2K}{2\tau_c} \log_2 \left( 1 + E_R \frac{\sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^2}{\sum_{k=1}^{K} \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^2} \right),
\]

(79)

as \( M \rightarrow \infty \). In addition, if \( E_R \rightarrow 0 \), the rate ratios are given by

\[
[\delta_1, \delta_2, \delta_3] = [2/\pi, 2/\pi, 4/\pi^2].
\]

(80)

From Propositions 3 and 4, we can see that the transmit power of each source or the relay can be reduced proportional to \( 1/M \) to maintain a constant rate, which means the system with one-bit ADCs and Dacs has the same power scaling laws as the perfect hardware case, which is an encouraging result. In addition, for both Propositions 3 and 4, \( \delta_1 = 4/\pi^2 \), revealing that for the double-quantized system, the rate ratio is \( 4/\pi^2 \) times less than the perfect ADC/DAC case, for low transmit power at the sources or low transmit power at the relay. This result is the
same as that in the system which is only quantized once [39]. Interestingly, focusing on the values of $\delta_1$ and $\delta_2$, we observe that the process to achieve the final scaling of $4/\pi^2$ is quite different. For low $p_S$ case, the value $4/\pi^2$ only results from $\delta_2 = 4/\pi^2$, implying that the rate loss is only caused by the one-bit ADCs. In contrast, for the low $p_R$ case, the value $4/\pi^2$ is generated by $\delta_1 = 2/\pi$ and $\delta_2 = 2/\pi$, indicating that the rate degradation comes from both one-bit ADCs and one-bit DACs.

IV. POWER ALLOCATION

In one-bit quantized systems, having users with a wide range of $\beta_k$ values causes significant rate degradation [24], [25]. To address this issue, we formulate a power allocation problem maximizing the sum rate of the system for a given total power budget $P_T$, i.e., $\sum_{k=1}^{K} p_{S,k} + p_{R} \leq P_T$.

A. Problem Formulation

Defining $p_S = [p_{S,1}, \ldots, p_{S,K}]^T$, the problem is expressed as

$$\mathcal{P}_1 : \text{maximize} \quad \frac{\tau - 2K}{2\tau_c} \sum_{k=1}^{K} \log_2 (1 + \gamma_k)$$

subject to

$$\gamma_k = \frac{p_{S,k}}{\xi_k}, \quad k = 1, \ldots, K$$

$$\sum_{i=1}^{K} p_{S,i} + p_r \leq P_T$$

$$p_S \geq 0, \quad p_r \geq 0,$$

where

$$\xi_k = \sum_{i=1}^{K} p_{S,i} a_{k,i} + p_r \left( \sum_{i=1}^{K} b_{k,i} p_{S,i} + c_k \right) + d_k$$

$$c_k = \frac{\pi^2}{4M^2 \sigma^4_{SR,k} \sigma^4_{RD,k}} \sum_{n=1}^{2} \sigma^2_{SR,n} \sigma^2_{RD,n}$$

$$d_k = \frac{\pi^2}{2M \sigma^2_{SR,k}} + \frac{\pi^2}{4M^2 \sigma^4_{SR,k} \sigma^4_{RD,k}} \sum_{n=1}^{K} \sigma^2_{SR,n} \sigma^2_{RD,n},$$

and $a_{k,i}$ and $b_{k,i}$ are respectively given by (88) and (89), shown at the bottom of this page.

Since $\log (\cdot)$ is an increasing function, problem $\mathcal{P}_1$ can be reformulated as

$$\mathcal{P}_2 : \text{minimize} \quad \prod_{k=1}^{K} (1 + \gamma_k)^{-1}$$

subject to

$$\gamma_k \leq \frac{p_{S,k}}{\xi_k}, \quad k = 1, \ldots, K$$

$$\sum_{i=1}^{K} p_{S,i} + p_r \leq P_T$$

$$p_S \geq 0, \quad p_r \geq 0,$$

which can be identified as a complementary geometric program (CGP) [40]. Note that the equality constraints (82) of $\mathcal{P}_1$ have been replaced with inequality constraints (91). Since the objective function of $\mathcal{P}_2$ decreases with $\gamma_k$, we can guarantee that the inequality constraints (91) must be active at any optimal solution of $\mathcal{P}_2$, which means that problem $\mathcal{P}_2$ is equivalent to $\mathcal{P}_1$.

B. Successive Approximation Algorithm

CGP problems are in general nonconvex. Fortunately, we can first approximate the CGP by solving a sequence of GP problems. Then, each GP can be solved very efficiently with standard convex optimization tools such as CVX. The key idea is to use a monomial function $\omega_k \gamma_k^{\mu_k}$ to approximate $1 + \gamma_k$ near an arbitrary point $\gamma_k > 0$. To make the approximation accurate, we need to ensure that

$$\begin{cases}
1 + \gamma_k = \omega_k \gamma_k^{\mu_k} \\
\mu_k \omega_k \gamma_k^{\mu_k-1} = 1.
\end{cases}$$

These results will hold if the parameters $\omega_k$ and $\mu_k$ are chosen as $\omega_k = \gamma_k^{\theta-1} (1 + \gamma_k)$ and $\mu_k = \frac{1}{(\theta-1) \gamma_k}$. At each iteration, the GP is obtained by replacing the posynomial objective function with its best local monomial approximation near the solution obtained at the previous iteration. The following algorithm shows the steps to solving $\mathcal{P}_2$.

We have neglected $\omega_k$ in the objective function of $\mathcal{P}_2$ since they are constants at each iteration and do not affect the problem solution. Also, some trust region constraints are added, i.e., $\theta^{-1} \gamma_k \leq \gamma_k \leq \theta \gamma_k$, which limits how much the variables are allowed to differ from the current guess $\gamma_k$. The parameter $\theta > 1$ controls the desired accuracy. More precisely, when $\theta$ is close to 1, it provides good accuracy for the monomial approximation.
Algorithm 1: Successive Approximation Algorithm for $\mathcal{P}_2$.

1) Initialization. Define a tolerance $\epsilon$ and parameter $\theta$. Set $j = 1$, and set the initial value of $\hat{\gamma}_k$ according to the SINR in Theorem 1 with $p_{s,k} = \frac{1}{2}$ and $p_r = \frac{P}{2}$.

2) Iteration $j$. Compute $\mu_k = \frac{\gamma_k}{1 + \gamma_k}$. Then, solve the following GP problem $\mathcal{P}_3$:

$$\mathcal{P}_3: \text{minimize } \prod_{k=1}^{K} \gamma_k^{\frac{\mu_k}{\theta}} \quad (95)$$

subject to

$$\theta^{-1} \hat{\gamma}_k \leq \gamma_k \leq \theta \hat{\gamma}_k, \quad k = 1, \ldots, K \quad (96)$$

$$\gamma_k p_{s,k} \xi_k \leq 1, \quad k = 1, \ldots, K \quad (97)$$

$$\sum_{i=1}^{K} p_{s,i} + p_r \leq P_l \quad (98)$$

$$p_s \geq 0, p_r \geq 0. \quad (99)$$

Denote the optimal solutions by $\gamma^{(j)}_k$, for $k = 1, \ldots, K$.

3) Stopping criterion. If $\max_k |\gamma^{(j)}_k - \hat{\gamma}_k| < \epsilon$, stop; otherwise, go to step 4).

4) Update initial values. Set $\tilde{\gamma}_k = \gamma^{(j)}_k$, and $j = j + 1$. Go to step 2).

![Algorithm 1](image)

In this section, we present numerical results to validate previous analytical results and demonstrate the benefits of the power allocation algorithm.

V. NUMERICAL RESULTS

In this section, we validate the theoretical derivations. For simplicity, we set the large-scale fading coefficients as $\beta_{SR} = \beta_{RD} = 1$ and adopt an equal power allocation strategy, i.e., $p_{s,k} = p_s$.

![Fig. 3](image)

As discussed in [41], $\theta = 1.1$ offers a good tradeoff between accuracy and convergence speed.

![Fig. 2](image)

but with slower convergence speed, and vice versa if $\theta$ is large. As discussed in [41], $\theta = 1.1$ offers a good tradeoff between accuracy and convergence speed.

B. Validation of Analytical Results

In this section, we validate the theoretical derivations. For simplicity, we set the large-scale fading coefficients as $\beta_{SR} = \beta_{RD} = 1$ and adopt an equal power allocation strategy, i.e., $p_{s,k} = p_s$.

Fig. 3 shows the sum rate versus the number of user pairs $K$. The curves associated with “Exact numerical results” and “Approximate numerical results” are respectively generated by Monte-Carlo simulations according to (40) and (54) by averaging over $10^3$ independent channel realizations, and the “Theoretical results” curves are obtained based on Theorem 1. As can be seen, there exists a gap between “Exact numerical results” (where the matrices $\mathbf{R}_{a,q_a}$ and $\mathbf{R}_{q_a,q_d}$ are diagonal, which means that the quantization noise is correlated) and “Approximate numerical results” (where the matrices $\mathbf{R}_{a,q_a}$ and $\mathbf{R}_{q_a,q_d}$ are approximated by identity matrices) when the number of user pairs is small, while the gap narrows and finally disappears as $K$ becomes large. The reason is that the correlation effect is stronger with smaller $K$ and weaker with larger $K$. In this example, our approximate model is very accurate when the number of user pairs is greater than 15. In addition, we observe that the “Approximate numerical results” curve overlaps with that for the “Theoretical results”, which verifies our analytical derivations in Theorem 1.

Fig. 4 shows the sum rate versus the number of relay antennas.

![Fig. 3](image)

From Fig. 4(a), we can see that when $K = 10$, the gap between the exact and approximate numerical results increases with the number of relay antennas. This suggests that for large antenna arrays, the correlation of the quantization noise becomes important and cannot be neglected. However, we are interested in the typical massive MIMO setup where the ratio between the number of relay antennas and the users is on the order of about $M/K = 10$, and thus we plot Fig. 4(b). In this figure, we can see the gap slightly narrows (from 0.2791 bit/s/Hz at $M = 80$ to 0.2505 bit/s/Hz at $M = 200$) as the number of relay antennas increases, which indicates that our approximate model is accurate for massive MIMO scenarios.

![Fig. 4](image)

Fig. 5(a) shows the transmit power $p_s$ of each source required to maintain a given sum rate of 5 bit/s/Hz. We can see that when the number of relay antennas increases, the required $p_s$ sig-
Fig. 4. Sum rate versus the number of relay antennas \( M \) for \( p_S = 10 \) dB, \( p_R = 10 \) dB, and \( p_p = 10 \) dB.

Fig. 5. Required \( p_S \) and rate ratio versus the number of relay antennas \( M \) for \( K = 5 \), \( p_p = 10 \) dB, and \( p_S = 10 \) dB.

Fig. 6. Required \( p_R \) and rate ratio versus the number of relay antennas \( M \) for \( K = 5 \), \( p_p = 10 \) dB, and \( p_S = 10 \) dB.

C. Power Allocation

Fig. 7 illustrates the impact of the optimal power allocation scheme on the sum rate when all users experience different large-scale fading. The large-scale fading coefficients are arbitrarily generated by \( \beta_{SR,k} = z_k (r_{SR,k}/r_0)^\kappa \) and \( \beta_{RD,k} = z_k (r_{RD,k}/r_0)^\kappa \), where \( z_k \) is a log-normal random variable with standard deviation 8 dB, \( r_{SR,k} \) and \( r_{RD,k} \) respectively represent the distances from the sources and destinations to the relay, \( \kappa = 3.8 \) is the path loss exponent, and \( r_0 \) denotes the guard interval which specifies the nearest distance between the sources and destinations.
users and the relay. The relay is located at the center of a cell with a radius of 1000 meters and $r_0 = 100$ meters. We choose $\beta_{SR} = [0.2688, 0.0368, 0.00025, 0.1398, 0.0047]$, and $\beta_{RD} = [0.0003, 0.00025, 0.0050, 0.0794, 0.0001]$. As a benchmark scheme for comparison, we also plot the sum rate with uniform power allocation, i.e., $p_S = \frac{P_T}{2K}$ and $p_R = \frac{P_T}{2}$. For uniform power allocation, we can see that the rate of Case I is the highest, Case IV is the lowest, while Case II outperforms Case III. These results are in agreement with Proposition 2. In addition, we observe that the optimal power allocation strategy significantly boosts the sum rate. For the case of $M = 200$, we see that the optimal power allocation strategy provides a 31% improvement for the perfect ADCs and DACs scenario, and a 43% increase for the one-bit ADCs and DACs scenario. This demonstrates the great importance of power allocation in quantized systems.

VI. CONCLUSION

We have analyzed the achievable rate of a multipair half-duplex massive antenna relaying system assuming that one-bit ADCs and DACs are deployed and MRC/MRT beamforming is used at the relay. An approximate closed-form expression for the achievable rate was derived, based on which the impact of key system parameters was characterized. It was shown that the sum rate with one-bit ADCs and DACs is $4/\pi^2$ times less than that achieved by an unquantized system in the low power regime. Despite the rate loss due to the use of one-bit ADCs and DACs, employing massive antenna arrays still enables significant power savings; i.e., the transmit power of each source or the relay can be reduced proportional to $1/M$ to maintain a constant rate, as in the unquantized case. Finally, we show that a good power allocation strategy can substantially compensate for the rate loss caused by the coarse quantization.

APPENDIX A
JUSTIFICATION OF THEOREM 1

The end-to-end SINR given in (54) consists of six expectation terms: 1) desired signal $A_k$; 2) estimation error $B_k$; 3) interpair interference $\hat{C}_k$; 4) noise at the relay $\hat{D}_k$; 5) quantization noise of ADCs $\hat{E}_k$; 6) quantization noise of DACs $\hat{F}_k$. Besides, we also need to calculate an approximation of $R_{XX,x}$. We compute them one by one as follows.

1) Approximate $R_{XX,x}$:

\[
R_{XX,x} = E \left\{ \hat{G}_{RD}^H \hat{G}_{SR} \hat{R}_{\text{y,y,y}} \hat{G}_{SR} \hat{G}_{RD}^T \right\} \\
\approx \alpha_2^2 E \left\{ \hat{G}_{RD}^H \hat{G}_{SR} \hat{G}_{SR} P_S \hat{G}_{SR} \hat{G}_{RD}^T \right\} \\
+ \left( \alpha_3^2 + 1 - \frac{2}{\pi} \right) E \left\{ \hat{G}_{RD}^H \hat{G}_{SR} \hat{G}_{SR} \hat{G}_{RD}^T \right\}.
\]

(100)

By using the fact that $E \left\{ ||\sigma_{SR,k}||^2 \right\} = M (M + 1) \beta_{SR,k}$, we have

\[
E \left\{ \hat{G}_{RD}^H \hat{G}_{SR} \hat{G}_{SR} P_S \hat{G}_{SR} \hat{G}_{RD}^T \right\} \\
= E \left\{ \hat{G}_{RD}^H \hat{G}_{SR} \hat{G}_{SR} \hat{G}_{RD}^T \right\} \\
+ E \left\{ \hat{G}_{RD}^H \hat{G}_{SR} \hat{G}_{SR} P_S \hat{G}_{SR} \hat{G}_{RD}^T \right\} \\
= M \sum_{k=1}^{K} \sigma_{SR,k}^2 \sigma_{RD,k}^2 \left( M p_S, k \sigma_{SR,k}^2 + \sum_{i=1}^{K} p_S, i \beta_{SR,i} \right) I_M
\]

(101)

Then, by substituting (101) and (102) into (100), we directly obtain

\[
R_{XX,x} \approx M \left( \alpha_2^2 + 1 - \frac{2}{\pi} \right) \frac{1}{M} \sum_{k=1}^{K} \sigma_{SR,k}^2 \sigma_{RD,k}^2 I_M
\]

(103)

3) $\hat{B}_k$:

\[
E \left\{ ||\sigma_{SR,k}^H \hat{G}_{RD}^H \hat{G}_{SR} \sigma_{SR,k} ||^2 \right\} =
\]

\[
E \left\{ \sum_{m=1}^{K} \sum_{n=1}^{K} \sigma_{RD,m}^T \sigma_{RD,m} \sigma_{SR,n}^H \sigma_{SR,n} \right\}.
\]

(106)

which can be decomposed into three different cases:
a) for \( m = n = k \),
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,k}^H |^2 \right\} \\
= E \left\{ |\hat{G}_{SR,k} |^2 |\hat{G}_{RD,k} |^2 \right\}
\]
\[
+ E \left\{ |\hat{G}_{RD,k} |^2 |\hat{G}_{SR,k}^H |^2 \right\}
\]
\[
+ E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{SR,k}^H |^2 \right\}
\]
\[
+ E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{SR,k} |^2 \right\}
\]
\[
= M^2 (M + 1)^2 \sigma_{SR,k}^2 \sigma_{RD,k}^2
\]
\[
+ M^2 (M + 1) \sigma_{SR,k}^4 \sigma_{RD,k}^2 \sigma_{RD,k}^4
\]
\[
+ M^2 (M + 1) \sigma_{RD,k}^2 \sigma_{SR,k}^4 \sigma_{RD,k}^2
\]
\[
+ M^2 \sigma_{SR,k}^2 \sigma_{RD,k}^2 \sigma_{RD,k}^2 \sigma_{RD,k}^2 .
\]
(107)

b) for \( m \neq n \neq k \),
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,k}^H |^2 \right\} = 0.
\]
(108)

c) for \( m \neq n \neq k \),
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,k}^H |^2 \right\} = 0.
\]
(109)

Combining a), b), and c), and by utilizing the fact of \( \sigma_{SR,k}^2 + \sigma_{SR,k}^2 = \beta_{SR,k} \) and \( \sigma_{RD,k}^2 + \sigma_{RD,k}^2 = \beta_{RD,k} \), we have
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,k}^H |^2 \right\}
\]
\[
= p_{SR,k} M^2 \sigma_{SR,k}^2 \sigma_{RD,k}^2
\]
\[
+ p_{SR,k} M^2 \sigma_{SR,k}^4 \sigma_{RD,k}^4
\]
\[
+ p_{SR,k} M^2 \sigma_{SR,k}^2 \sigma_{RD,k}^4 \beta_{SR,k}
\]
\[
+ p_{SR,k} M^2 \beta_{SR,k} \beta_{RD,k}
\]
\[
K \sum_{n=1}^K \sigma_{SR,n} \sigma_{RD,n}.
\]
(110)

Thus,
\[
\tilde{B}_k = p_{SR,k} M^2 \sigma_{SR,k}^4 \sigma_{RD,k}^4
\]
\[
+ p_{SR,k} M^2 \beta_{SR,k} \beta_{RD,k}
\]
\[
K \sum_{n=1}^K \sigma_{SR,n} \sigma_{RD,n}.
\]
(111)

4) \( \hat{C}_k \):
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,k}^H |^2 \right\}
\]
\[
= E \left\{ \sum_{m=1}^K \sum_{n=1}^K |\hat{G}_{RD,m}^{\ast} \hat{G}_{RD,m} \hat{G}_{SR,n} \hat{G}_{SR,n}^H |^2 \right\}
\]
(112)

which can be decomposed as six cases:

a) for \( m \neq n \neq k \),
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,k}^H |^2 \right\} = 0.
\]
(113)

b) for \( m \neq k \neq i \),
\[
E \left\{ \sum_{n \neq k} |\hat{G}_{SR,n} \hat{G}_{SR,n}^H |^2 \right\}
\]
\[
= M^2 \beta_{SR,k} \beta_{RD,k} \sum_{n \neq k} \sigma_{SR,n}^2 \sigma_{RD,n}^2 .
\]
(114)

c) for \( m = n = k \neq i \),
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,i}^H |^2 \right\}
\]
\[
= M^2 \sigma_{SR,k}^2 \sigma_{RD,k}^2 \beta_{SR,i} \left( (M+1) \sigma_{RD,k}^2 + \sigma_{RD,k}^2 \right)
\]
(115)

d) for \( m = n = i \neq k \),
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,i}^H |^2 \right\}
\]
\[
= M^2 \sigma_{SR,k}^2 \sigma_{RD,i}^2 \sigma_{RD,k}^2 \beta_{SR,i} \left( (M+1) \sigma_{SR,i}^2 + \sigma_{SR,i}^2 \right).
\]
(116)
e) for \( m = i, n = k \),
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,i}^H |^2 \right\} = 0.
\]
f) for \( m = k, n = i \),
\[
E \left\{ |\hat{G}_{RD,k}^{\ast} \hat{G}_{RD,k} \hat{G}_{SR,i}^H |^2 \right\} = 0.
\]

Combining a), b), c), d), e), and f), we have
\[
\hat{C}_k = M^2 \sum_{i \neq k} p_{SR,i} \beta_{SR,i} \beta_{RD,k} \sum_{n=1}^K \sigma_{SR,n} \sigma_{RD,n}
\]
\[
+ M^2 \sum_{i \neq k} p_{SR,i} \left( \sigma_{SR,k}^4 \sigma_{RD,k}^4 \beta_{SR,i} + \sigma_{SR,i}^4 \sigma_{RD,i} \beta_{RD,k} \right).
\]
(117)

5) \( \hat{D}_k \): Following the same approach as with the derivations of \( \tilde{B}_k \), we obtain
\[
\hat{D}_k = M^2 \sigma_{SR,k}^4 \sigma_{RD,k}^4 + M^2 \beta_{RD,k} \sum_{n=1}^K \sigma_{SR,n} \sigma_{RD,n} .
\]
(118)

6) \( \hat{E}_k \): By using the fact that \( \hat{E}_k = (1 - \frac{2}{\sigma^2}) \frac{1}{\sigma^2} \hat{D}_k \), we obtain the result for \( \hat{E}_k \).

7) \( \hat{F}_k \): By using the fact that \( \hat{F}_k = \frac{1-\frac{2}{\sigma^2}}{\sigma^2} E \left\{ |\hat{G}_{RD,k} |^2 \right\} = (1 - \frac{2}{\sigma^2}) \frac{M \beta_{RD,k} \alpha^2}{\sigma^2} .
\]

8) \( \hat{G}_k \): Combining (36), (48), and (52), we can find the value of \( \hat{G}_k \).

**APPENDIX B**

**PROOF OF PROPOSITION 2**

We can readily observe that \( R_k^{DA} > R_k^{AB} \) and \( R_k^{RD} > R_k^{DR} \). Thus, we only focus on comparing \( R_k^{DR} \) and \( R_k^{RD} \). Due to the fact that \( \sigma_{SR,k}^2 = \frac{\tau}{2} \sigma_{SR,k}^2 \) and \( \sigma_{RD,k}^2 = \frac{\tau}{2} \sigma_{RD,k}^2 \) (c.f. (18), (26), and Corollary 1), and by neglecting the low order terms as \( M \rightarrow \infty \), the ratio between the SINR of \( R_k^{DA} \) and that of \( R_k^{RD} \) can be 

expressed as

\[ f_1 = \frac{2}{\pi} \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{P_A}{R_k^A} - 1 \to f_2 \]

\[ f_2 = \frac{2}{\pi} \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{P_A}{R_k^A} - 1 \to f_1 \]

(119)

where

\[ f_1 = \frac{2}{\pi} p_{S,k} \sigma_{SR,k}^2 \tau_{RD,k} \beta_{RD,k} + \frac{2}{\pi} p_{S,k} \sigma_{RD,k}^2 \beta_{SR,k} + \frac{2}{\pi} \sum_{i \neq k} p_{S,i} \left( \sigma_{SR,i}^2 \tau_{RD,i} \beta_{RD,i} + \sigma_{RD,i}^2 \beta_{SR,i} \right) \]

\[ + \frac{2}{\pi} \sum_{i \neq k} p_{S,i} \left( \sigma_{SR,i}^2 \tau_{RD,i} \beta_{RD,i} + \sigma_{RD,i}^2 \beta_{SR,i} \right) \]

(120)

Since \( f_1 < f_2 \), we conclude that \( R_k^A > R_k^D \).

REFERENCES


Chuli Kong (S’15) received the B.S. degree in electronics and information engineering from the Dalian University of Technology, Dalian, China, in 2013. She is currently working toward the Ph.D. degree in information and communication engineering at the Zhejiang University, Hangzhou, China. From 2016 to 2017, she was a Visiting Scholar with the University of California, Irvine, CA, USA. Her research interests include massive MIMO systems and relaying communications. She was an Exemplary Reviewer for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS in 2017.

Amine Mezghani (S’08–M’16) received the Dipl.-Ing. degree in electrical engineering from the Technische Universität München, Munich, Germany, and the “Diplome d’Ingénieur” degree from the École Centrale de Paris, Paris, France, both in 2006. He received the Ph.D. degree in electrical engineering from the Technische Universität München, in 2015. In 2017, he was a Postdoctoral Fellow with the University of Texas, Austin, TX, USA. Prior to this, he was a Postdoctoral Scholar with the Department of Electrical Engineering and Computer Science, University of California, Irvine, CA, USA. His current research interests include millimeter-wave massive MIMO, hardware constrained communication theory, and signal processing under low-resolution analog-to-digital and digital-to-analog converters. He was the recipient of the Rohde & Schwarz Outstanding Dissertation Award in 2016.

Caijun Zhong (S’07–M’10–SM’14) received the B.S. degree in information engineering from the Xi’an Jiaotong University, Xi’an, China, in 2004. He received the M.S. degree in information security and the Ph.D. degree in telecommunications both from the University College London, London, U.K., in 2006 and 2010, respectively. From September 2009 to September 2011, he was a Research Fellow with the Institute for Electronics, Communications and Information Technologies, Queen’s University Belfast, Belfast, U.K. Since September 2011, he has been with the Zhejiang University, Hangzhou, China, where he is currently an Associate Professor. His research interests include massive MIMO systems, full-duplex communications, wireless power transfer, and physical layer security.

Dr. Zhong is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, EURASIP Journal on Wireless Communications and Networking, and Journal of Communications and Network. He is the recipient of the 2013 IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award. He and his co-authors have been awarded a Best Paper Award at the WCSP 2013. He was an Exemplary Reviewer for IEEE TRANSACTIONS ON COMMUNICATIONS in 2014.

A. Lee Swindlehurst (F’04) received the B.S. and M.S. degrees in electrical engineering from Brigham Young University (BYU), Provo, UT, USA, in 1985 and 1986, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1991. From 1990 to 2007, he was with the Department of Electrical and Computer Engineering, BYU, where he was the Department Chair from 2003 to 2006. During 1996–1997, he held a joint appointment as a Visiting Scholar with the Uppsala University, Uppsala, Sweden, and the Royal Institute of Technology, Stockholm, Sweden. During 2006–2007, he was a Vice President of Research for ArrayComm LLC in San Jose, California, CA, USA. Since 2007 he has been a Professor with the Department of Electrical Engineering and Computer Science, University of California, Irvine, CA, USA, where he was an Associate Dean for research and graduate studies with the Samueli School of Engineering, Irvine, CA, USA, from 2013 to 2016. From 2014 to 2017, he was a Hans Fischer Senior Fellow with the Institute for Advanced Studies, Technical University of Munich, Munich, Germany. He has authored or co-authored more than 300 publications in his research areas. His research interests include array signal processing for radar, wireless communications, and biomedical applications. He was the inaugural Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING. He was the recipient of 2000 IEEE W. R. G. Baker Prize Paper Award, the 2006 IEEE Communications Society Stephen O. Rice Prize in the Field of Communication Theory, the 2006 and 2010 IEEE Signal Processing Society’s Best Paper Awards, and the 2017 IEEE Signal Processing Society Donald G. Fink Overview Paper Award.