

Full-Duplex Massive MIMO Relaying Systems with Low-Resolution ADCs

Chuili Kong, Caijun Zhong, Shi Jin, Sheng Yang, Hai Lin, Zhaoyang Zhang

► **To cite this version:**

Chuili Kong, Caijun Zhong, Shi Jin, Sheng Yang, Hai Lin, et al.. Full-Duplex Massive MIMO Relaying Systems with Low-Resolution ADCs. IEEE Transactions on Wireless Communications, Institute of Electrical and Electronics Engineers, 2017, 16 (8), pp.5033 - 5047. 10.1109/TWC.2017.2705041 . hal-01557926

HAL Id: hal-01557926

<https://hal-centralesupelec.archives-ouvertes.fr/hal-01557926>

Submitted on 28 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Full-Duplex Massive MIMO Relaying Systems with Low-Resolution ADCs

Chuili Kong, *Student Member, IEEE*, Caijun Zhong, *Senior Member, IEEE*, Shi Jin, *Member, IEEE*, Sheng Yang, *Member, IEEE*, Hai Lin, *Senior Member, IEEE*, and Zhaoyang Zhang, *Member, IEEE*

Abstract—This paper considers a multipair amplify-and-forward massive MIMO relaying system with low-resolution ADCs at both the relay and destinations. The channel state information (CSI) at the relay is obtained via pilot training, which is then utilized to perform simple maximum-ratio combining/maximum-ratio transmission processing by the relay. Also, it is assumed that the destinations use statistical CSI to decode the transmitted signals. Exact and approximated closed-form expressions for the achievable sum rate are presented, which enable the efficient evaluation of the impact of key system parameters on the system performance. In addition, optimal relay power allocation scheme is studied, and power scaling law is characterized. It is found that, with only low-resolution ADCs at the relay, increasing the number of relay antennas is an effective method to compensate for the rate loss caused by coarse quantization. However, it becomes ineffective to handle the detrimental effect of low-resolution ADCs at the destination. Moreover, it is shown that deploying massive relay antenna arrays can still bring significant power savings, i.e., the transmit power of each source can be cut down proportional to $1/M$ to maintain a constant rate, where M is the number of relay antennas.

Index terms— Amplify-and-forward, full-duplex, low-resolution ADC, massive MIMO, relaying.

I. INTRODUCTION

Full-duplex relaying has recently attracted considerable attention due to its substantial spectral efficiency gain over the conventional half-duplex relaying systems [1]. However, to realize the benefit, how to mitigate the loopback interference caused by the signal leakage from the relay output to its input, is one of the major issues to be tackled. Thus far, various techniques have been proposed to address this important issue, from separate transmit and receiver architecture to joint analog/digital filtering [2]. In parallel, the massive multiple-input multiple-output (MIMO) technique is able to significantly boost the spectral efficiency and effectively suppress interference, hence has also received great interests recently [3]–[6]. Therefore, it becomes a natural choice to combine these two promising technologies. The potential applications

of such a massive antenna full-duplex relay include millimeter wave communications, device-to-device communications and machine-to-machine systems.

A. Related Works and Motivation

The performance of one-way and two-way multipair full-duplex relaying systems has been respectively studied in [7], [8] and [9], where it was demonstrated that deploying a large antenna array at the relay helps eliminate both the inter-pair and loop interference, thereby substantially boosting the achievable sum rate. However, such performance gain comes at the price of increased hardware cost and power consumption due to the extra required high-resolution analog-to-digital converters (ADCs), which may be undesirable in practical system deployment. The reasons are three-fold: 1) The power consumption of ADCs scales exponentially with the resolution and linearly with the sampling rate. For instance, a typical flash ADC with b -bit and the sampling frequency f_s operates $f_s 2^b$ conversion steps per second [10]. 2) The fabrication cost of ADCs depends on the resolution. Thus, if high-resolution ADCs are adopted, the total financial cost will be a heavy burden for massive MIMO systems since each antenna requires a pair of ADCs to separately quantize the real and imaginary parts of signals. 3) The chip area of ADC increases exponentially with the resolution, which makes it difficult to put large number of collocated antenna arrays together. To resolve the above issues, a promising way is to use low-cost and power-efficient low-resolution or even one-bit ADCs unit to build radio frequency (RF) chains.

Low-resolution ADCs not only cause rate degradation, but also change some concluding remarks that have been made for unquantized systems. For example, the quality of the channel estimates depends on the set of orthogonal pilot sequences used, which is contrary to unquantized systems where any set of orthogonal pilot sequences gives the same result [11]. Moreover, compared to unquantized MIMO systems, the optimal length of training sequence is approximately 10 times more [12], [13]; while to achieve the same performance as that in a full channel state information (CSI) case, the length of pilot sequence increases to approximately 50 times the number of users [14], which is extremely long. Such a pilot overhead cannot be sustained and hence new channel estimation methods are proposed. For instance, [15] adopts a joint channel and data estimation approach to aid channel estimation and reduce pilot overhead. And other techniques such as expectation-maximization [11], [16], generalized approximate message

C. Kong, C. Zhong and Z. Zhang are with the Institute of Information and Communication Engineering, Zhejiang University, China. C. Zhong is also affiliated with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China (email: kcl_dut@163.com; caijun-zhong@zju.edu.cn; ning_ming@zju.edu.cn).

S. Jin is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (email: jinshi@seu.edu.cn).

S. Yang is with the Laboratoire des Signaux et Systèmes, CentraleSupélec, Gif-sur-Yvette Cedex 91192, France (email: sheng.yang@supelec.fr).

H. Lin is with the Department of Electrical and Information Systems, Osaka Prefecture University, Osaka 599-8531, Japan (email: lin@eis.osakafu-u.ac.jp).

passing [17], and maximum likelihood algorithms [18] have been proposed in the literature. Furthermore, the capacity achieving signals for single-input single-output system are no longer Gaussian distributions, and instead become discrete [19], [20]. In addition, the detectors might need to be modified by taking the quantization effects into account.

Recently, there has been a surge of research interests to understand the impact of coarse quantization effects on massive MIMO systems. The works [12], [21]–[26] have demonstrated that the massive antenna array has robustness against coarse quantization and capable of compensating for the performance loss caused by low-resolution ADCs. In particular, the work [12] studied the optimal training pilot length to maximize the spectral efficiency, while the work [24] revealed that the optimal number of quantization bits is 4 or 5 bits in terms of energy efficiency. For one-bit quantization systems, the power efficiency laws and energy-spectral efficiency tradeoff are characterized in [25] and [26], respectively. Moreover, various precoding methods such as spatio-temporal processing [27], minimum BER precoding [28], and hybrid beamforming [29] are studied. In addition, a mixed-ADC architecture is proposed to balance the spectral efficiency loss and power consumption [30]–[32]. Despite the spectral efficiency enhancement, adopting mixed-ADCs increases the hardware complexity since an ADC switch is required. Furthermore, the optimal input symbol distribution and codebook for limited feedback have been designed [33], [34]. However, all the aforementioned works consider a single-hop system with low-resolution ADCs being implemented at either the BS or the destination users. Only very recently, [35] introduces mixed-ADCs into the relay network, but it only considers the quantization at the BS. Therefore, the impact of low-resolution ADCs in a dual-hop system remains unknown.

B. Our Work and Contributions

Motivated by this, in this paper, we investigate the performance of full-duplex massive MIMO relaying systems with low-resolution ADCs at both the relay and the destinations. Specifically, we consider a multipair full-duplex relaying system using the amplify-and-forward (AF) protocol with simple maximum-ratio combining/maximum-ratio transmission (MRC/MRT) processing at the relay. First, CSI at the relay is acquired via pilot training, and the effect of low-resolution ADCs on the accuracy of CSI is characterized. Then, exact and approximated closed-form expressions for the sum rate are derived which enable efficient evaluation of the system's achievable sum rate. Moreover, based on the simple approximated sum rate expression, optimal relay power allocation strategy is characterized, and the power scaling law is studied. The findings of the paper suggest that, with only low-resolution ADCs at the relay, increasing the number of relay antennas is a promising method to compensate for the rate loss caused by the coarse quantization. However, compared to the infinite resolution ADC case, the required number of relay antennas doubles with one-bit ADCs. In addition, we show that the use of low-resolution ADCs at the destination is a major performance limiting factor, and

it is preferable to deploy the low-resolution ADCs at the relay and use high-resolution ADCs at the destination. Finally, it is revealed that, even with low-resolution ADCs, deploying massive relay antenna arrays can still bring significant power savings, i.e., the transmit power of each source can be cut down proportional to $1/M$ to maintain a constant rate, where M is the number of relay antennas.

C. Organization and Notations

The remainder of the paper is organized as follows: Section II introduces the multipair full-duplex relaying system model under consideration. Section III presents an exact closed-form expression for the sum rate. Section IV provides an accurate approximation for the sum rate, and gives a detailed evaluation of the impact of low-resolution ADCs on the system performance. 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. Moreover, $(\cdot)^H$, $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^{-1}$ represent the conjugate transpose operator, the conjugate operator, the transpose operator, and the matrix inverse, respectively. Also, $\|\cdot\|$ is the Euclidian norm, $|\cdot|$ is the absolute value, and $[\mathbf{A}]_{mn}$ gives the (m, n) -th entry of \mathbf{A} . In addition, $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes a circularly symmetric complex Gaussian random vector \mathbf{x} with zero mean and variance matrix $\mathbf{\Sigma}$, while \mathbf{I}_k is the identity matrix of size k . Finally, the statistical expectation operator is represented by $\mathbb{E}\{\cdot\}$, the variance operator is $\text{Var}(\cdot)$, and the trace is denoted by $\text{tr}(\cdot)$.

II. SYSTEM MODEL

We consider a multipair AF relaying system shown in Fig. 1, where K single-antenna user pairs, denoted as S_k and D_k , $k = 1, \dots, K$, aim to exchange information with each other with the assistance of a shared multi-antenna full-duplex relay. To reduce the implementation cost, it is assumed that low-resolution ADCs are used at both the relay and destinations D_k ¹, $k = 1, \dots, K$ [33], [34]. In addition, we assume that direct links between S_k and D_k do not exist due to large obstacles or severe shadowing [36].

During the i -th time slot, the sources S_k ($k = 1, \dots, K$), transmit the signals $\sqrt{p_S}x_{S,k}[i]$ satisfying $\mathbb{E}\{|x_{S,k}[i]|^2\} = 1$ to the relay, while the relay broadcasts the signal $\mathbf{x}_R[i]$ satisfying $\mathbb{E}\{\mathbf{x}_R[i]\mathbf{x}_R^H[i]\} = \frac{p_R}{M}\mathbf{I}_M$ simultaneously to all K destinations D_k , ($k = 1, \dots, K$). Hence, the received signals at the relay and the K destinations can be respectively expressed as

$$\mathbf{y}_R[i] = \sqrt{p_S}\mathbf{G}_{SR}\mathbf{x}_S[i] + \bar{\mathbf{G}}_{RR}\mathbf{x}_R[i] + \mathbf{n}_R[i], \quad (1)$$

$$\mathbf{y}_D[i] = \mathbf{G}_{RD}^T\mathbf{x}_R[i] + \mathbf{n}_D[i], \quad (2)$$

where $\mathbf{x}_S[i] \triangleq [x_{S,1}[i], x_{S,2}[i], \dots, x_{S,K}[i]]$. $\mathbf{G}_{SR} \in \mathbb{C}^{M \times K}$ and $\mathbf{G}_{RD} \in \mathbb{C}^{M \times K}$ denote the channels from the K sources

¹Note that the considered model is very generic, and the single antenna destination node is not constrained to be the mobile phone, it could be certain low-cost device such as sensor, where using low-resolution ADC is desirable.

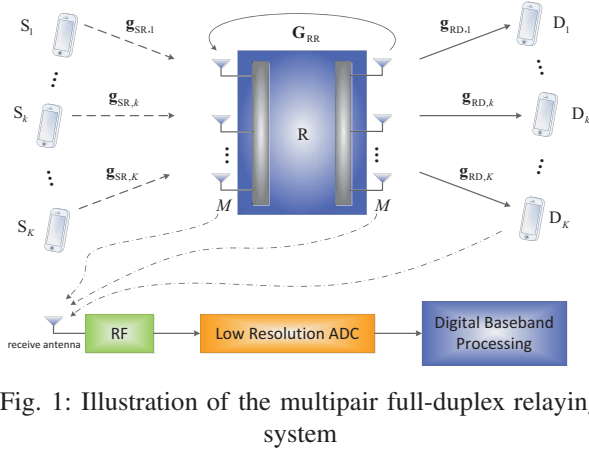


Fig. 1: Illustration of the multipair full-duplex relaying system

to the relay and from the relay to the K destinations, respectively, which account for both small-scale and large-scale fading effects. More specifically, the k -th columns of \mathbf{G}_{SR} and \mathbf{G}_{RD} are given by $\mathbf{g}_{\text{SR},k} \sim \mathcal{CN}(\mathbf{0}, \beta_{\text{SR},k} \mathbf{I}_M)$ and $\mathbf{g}_{\text{RD},k} \sim \mathcal{CN}(\mathbf{0}, \beta_{\text{RD},k} \mathbf{I}_M)$, where $\beta_{\text{SR},k}$ and $\beta_{\text{RD},k}$ model the large-scale path-loss effect, which are assumed to remain constant over many coherence intervals and known a priori. Also, $\bar{\mathbf{G}}_{\text{RR}} \in \mathbb{C}^{M \times M}$ represents the loop interference channel at the full-duplex relay. In addition, $\mathbf{n}_{\text{R}}[i]$ and $\mathbf{n}_{\text{D}}[i]$ denote the additive white Gaussian noise (AWGN) at the relay and the K destinations, respectively. The elements of $\mathbf{n}_{\text{R}}[i]$ and $\mathbf{n}_{\text{D}}[i]$ are assumed to be independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$.

To model the receivers with low-resolution ADCs, we focus on the non-uniform quantizer, and adopt the additive quantization noise model (AQN) as in [21], [22], [29], [37] for tractable analysis, shown in Fig. 2. As such, the outputs of the ADCs corresponding to input $\mathbf{y}_{\text{R}}[i]$ at the relay and $y_{\text{D},k}[i]$ ($y_{\text{D},k}[i]$ is the k -th element of $\mathbf{y}_{\text{D},k}[i]$) at D_k are denoted as

$$\tilde{\mathbf{y}}_{\text{R}}[i] = \alpha \mathbf{y}_{\text{R}}[i] + \tilde{\mathbf{n}}_{\text{R}}[i], \quad (3)$$

$$\tilde{y}_{\text{D},k}[i] = \theta y_{\text{D},k}[i] + \tilde{n}_{\text{D},k}[i], \quad (4)$$

respectively, where parameters α and θ are determined by the number of quantization bits b of ADCs, which indicate the resolution of ADC. For instance, $\alpha = \theta = 1$ implies perfect ADCs. For $b \leq 5$, the typical values of α and θ are listed in Table I, while for $b > 5$, they can be approximated by α (or θ) = $1 - \frac{\pi\sqrt{3}}{2} 2^{-2b}$ [38]. Also, $\tilde{\mathbf{n}}_{\text{R}}[i]$ and $\tilde{n}_{\text{D},k}[i]$ represent the additive Gaussian quantization noise at the relay and D_k , respectively. For a fixed channel realization \mathbf{G}_{SR} , \mathbf{G}_{RD} , and $\bar{\mathbf{G}}_{\text{RR}}$, the covariance of $\tilde{\mathbf{n}}_{\text{R}}[i]$ and $\tilde{n}_{\text{D},k}[i]$ are respectively given by

$$\mathbf{R}_{\tilde{\mathbf{n}}_{\text{R}}[i]} = \alpha(1 - \alpha) \text{diag}(\mathbf{E}\{\mathbf{y}_{\text{R}}[i]\mathbf{y}_{\text{R}}[i]^H\}), \quad (5)$$

$$R_{\tilde{n}_{\text{D},k}[i]} = \theta(1 - \theta) \mathbf{E}\{|y_{\text{D},k}[i]|^2\}. \quad (6)$$

Note that this AQNM model is realistic enough since the quantization noise variance not only depends on the number of quantization bit but also scales up with the received power. Also, [29] has proved that this AQNM model is accurate at the low signal-to-noise-ratio regime in which our considered system is most likely to operate.

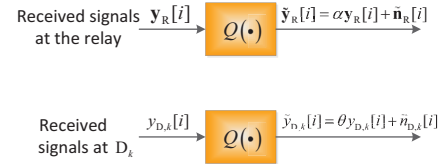


Fig. 2: Outputs of receivers with AQNM.

TABLE I: α and θ for different ADC quantization bits b .

b	1	2	3	4	5
α (or θ)	0.6366	0.8825	0.96546	0.990503	0.997501

A. Channel Estimation

We utilize pilots to estimate the channel matrices \mathbf{G}_{SR} and \mathbf{G}_{RD} , which is a technique widely used in the literature [39], [40]. Therefore, during each coherence interval of length τ_c (in symbols), the channel training occupies $2\tau_p$ symbols (the minimum length of τ_p equals to the number of users K). First, the destinations remain silent and all sources transmit simultaneously their mutually orthogonal pilot sequences $\Phi_{\text{S}} \in \mathbb{C}^{K \times \tau_p}$ to the relay. After that, all destinations transmit their mutually orthogonal pilot sequences $\Phi_{\text{D}} \in \mathbb{C}^{K \times \tau_p}$ to the relay whilst the destinations remain silent². For analytical tractability, we adopt the same pilot sequences as in [11]. Thus, the received signals at the relay's receive and transmit antenna arrays are [7]

$$\mathbf{Y}_{\text{rp}} = \sqrt{\tau_p p_p} \mathbf{G}_{\text{SR}} \Phi_{\text{S}} + \mathbf{N}_{\text{rp}}, \quad (7)$$

$$\mathbf{Y}_{\text{tp}} = \sqrt{\tau_p p_p} \mathbf{G}_{\text{RD}} \Phi_{\text{D}} + \mathbf{N}_{\text{tp}}, \quad (8)$$

respectively, where p_p is the transmit power of each pilot symbol; \mathbf{N}_{rp} and \mathbf{N}_{tp} denote the noise at the receive and transmit antenna arrays of the relay, respectively, with i.i.d. $\mathcal{CN}(0, 1)$ elements.

With low-resolution ADC receivers, the resulting quantized signals at the relay's receive and transmit antenna arrays respectively read as

$$\tilde{\mathbf{Y}}_{\text{rp}} = \alpha \mathbf{Y}_{\text{rp}} + \mathbf{N}_{\text{rq}}, \quad (9)$$

$$\tilde{\mathbf{Y}}_{\text{tp}} = \alpha \mathbf{Y}_{\text{tp}} + \mathbf{N}_{\text{tq}}, \quad (10)$$

where \mathbf{N}_{rq} and \mathbf{N}_{tq} represent the quantization noise, whose covariance matrices are respectively given by

$$\mathbf{R}_{\mathbf{N}_{\text{rq}}} = \mathbf{E}\{\mathbf{N}_{\text{rq}}\mathbf{N}_{\text{rq}}^H\} = \alpha(1 - \alpha) \text{diag}(\mathbf{E}\{\mathbf{Y}_{\text{rp}}\mathbf{Y}_{\text{rp}}^H\}), \quad (11)$$

$$\mathbf{R}_{\mathbf{N}_{\text{tq}}} = \mathbf{E}\{\mathbf{N}_{\text{tq}}\mathbf{N}_{\text{tq}}^H\} = \alpha(1 - \alpha) \text{diag}(\mathbf{E}\{\mathbf{Y}_{\text{tp}}\mathbf{Y}_{\text{tp}}^H\}). \quad (12)$$

Assuming that the relay employs the minimum mean-square-error (MMSE) estimator, the channel matrices \mathbf{G}_{SR} and

²Note that the channel training scheme where all sources and destinations transmit their pilots simultaneously can obtain the same channel estimation accuracy as that in our proposed training approach, but with increased computational complexity.

\mathbf{G}_{RD} can be decomposed as

$$\mathbf{G}_{\text{SR}} = \hat{\mathbf{G}}_{\text{SR}} + \mathbf{E}_{\text{SR}}, \quad (13)$$

$$\mathbf{G}_{\text{RD}} = \hat{\mathbf{G}}_{\text{RD}} + \mathbf{E}_{\text{RD}}, \quad (14)$$

where \mathbf{E}_{SR} and \mathbf{E}_{RD} are the estimation error matrices of \mathbf{G}_{SR} and \mathbf{G}_{RD} . Due to the orthogonality property of MMSE estimators and the fact that $\hat{\mathbf{G}}_{\text{SR}}$, \mathbf{E}_{SR} , $\hat{\mathbf{G}}_{\text{RD}}$, and \mathbf{E}_{RD} are complex Gaussian distributed, these matrices are independent of each other. By rewriting (13) and (14) in vector form, we have

$$\mathbf{g}_{\text{SR},k} = \hat{\mathbf{g}}_{\text{SR},k} + \mathbf{e}_{\text{SR},k}, \quad (15)$$

$$\mathbf{g}_{\text{RD},k} = \hat{\mathbf{g}}_{\text{RD},k} + \mathbf{e}_{\text{RD},k}, \quad (16)$$

where $\hat{\mathbf{g}}_{\text{SR},k}$, $\mathbf{e}_{\text{SR},k}$, $\hat{\mathbf{g}}_{\text{RD},k}$, and $\mathbf{e}_{\text{RD},k}$ are the k -th columns of $\hat{\mathbf{G}}_{\text{SR}}$, \mathbf{E}_{SR} , $\hat{\mathbf{G}}_{\text{RD}}$, and \mathbf{E}_{RD} , respectively, which are mutually independent.

The distributions of these vectors are given in the following lemma that is essential to the ensuing analysis.

Lemma 1: The elements of $\hat{\mathbf{g}}_{\text{SR},k}$, $\mathbf{e}_{\text{SR},k}$, $\hat{\mathbf{g}}_{\text{RD},k}$, and $\mathbf{e}_{\text{RD},k}$ are independent Gaussian random variables with zero mean, variance $\sigma_{\text{SR},k}^2 = \frac{\alpha\tau_p p_p \beta_{\text{SR},k}^2}{1+\tau_p p_p \beta_{\text{SR},k}^2}$, $\tilde{\sigma}_{\text{SR},k}^2 = \frac{\beta_{\text{SR},k} + (1-\alpha)\tau_p p_p \beta_{\text{SR},k}^2}{1+\tau_p p_p \beta_{\text{SR},k}^2}$, $\sigma_{\text{RD},k}^2 = \frac{\alpha\tau_p p_p \beta_{\text{RD},k}^2}{1+\tau_p p_p \beta_{\text{RD},k}^2}$, and $\tilde{\sigma}_{\text{RD},k}^2 = \frac{\beta_{\text{RD},k} + (1-\alpha)\tau_p p_p \beta_{\text{RD},k}^2}{1+\tau_p p_p \beta_{\text{RD},k}^2}$, respectively.

Proof: See Appendix A. \square

B. Data Transmission

During the i -th time slot, the relay amplifies the previously received symbol and broadcasts it to the K destinations. Thus, we have

$$\mathbf{x}_{\text{R}}[i] = \gamma \mathbf{F} \tilde{\mathbf{y}}_{\text{R}}[i - d], \quad (17)$$

where $\mathbf{F} \in \mathbb{C}^{M \times M}$ is the linear processing matrix to be specified shortly, γ is an amplification constant factor which is chosen to satisfy the power constraint at the relay, and $d \geq 1$ denotes the processing delay.

1) *Loop Interference:* Since the relay operates in the full-duplex mode, it suffers from loop interference. However, since the relay is aware of its transmitted signal, some form of self-interference mitigation method can be applied, such that the remaining self-interference is sufficiently weak and can be treated as additional noise $\hat{\mathbf{x}}_{\text{R}}[i]$ [9], [41] with the same power constraint as $\mathbf{x}_{\text{R}}[i]$, i.e., $\mathbb{E}\{\hat{\mathbf{x}}_{\text{R}}[i]\hat{\mathbf{x}}_{\text{R}}^H[i]\} = \frac{p_{\text{R}}}{M}\mathbf{I}_M$. Therefore, (1) can be re-expressed as

$$\mathbf{y}_{\text{R}}[i] = \sqrt{p_{\text{S}}}\mathbf{G}_{\text{SR}}\mathbf{x}_{\text{S}}[i] + \mathbf{G}_{\text{RR}}\hat{\mathbf{x}}_{\text{R}}[i] + \mathbf{n}_{\text{R}}[i], \quad (18)$$

where \mathbf{G}_{RR} models the residual loopback interference channel due to imperfect cancellation [42]. As in [7], [9], we assume that the entries of \mathbf{G}_{RR} are i.i.d. $\mathcal{CN}(0, \sigma_{\text{LI}}^2)$ with σ_{LI}^2 denoting the level of loop interference.

2) *Linear Processing:* The relay treats the channel estimates as the true channels for subsequent linear processing. With MRC/MRT processing, \mathbf{F} is given by

$$\mathbf{F} = \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H. \quad (19)$$

Recall that γ is chosen to meet the power constraint at the relay, after some simple algebraic manipulations, γ can be obtained as (20), shown on the top of the next page.

III. SYSTEM DESIGN GUIDELINES

This section investigates the achievable rate with low-resolution ADCs both at the relay and destinations. In particular, an exact closed-form expression is derived for the system's achievable rate. Furthermore, we provide an answer to the important question of where to deploy the low-resolution ADCs to attain best performance.

Substituting (18), (17), (2), and (3) into (4), we have

$$\begin{aligned} \tilde{\mathbf{y}}_{\text{D},k}[i] &= \alpha\theta\gamma\sqrt{p_{\text{S}}}\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},k} x_{\text{S},k}[i - d] \\ &+ \alpha\theta\gamma\sqrt{p_{\text{S}}}\sum_{j \neq k} \mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},j} x_{\text{S},j}[i - d] \\ &+ \alpha\theta\gamma\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{G}_{\text{RR}} \hat{\mathbf{x}}_{\text{R}}[i - d] \\ &+ \alpha\theta\gamma\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{n}_{\text{R}}[i - d] \\ &+ \theta\gamma\mathbf{g}_{\text{RD},k}^T \mathbf{F} \tilde{\mathbf{n}}_{\text{R}}[i - d] + \theta n_{\text{D},k}[i] + \tilde{n}_{\text{D},k}[i]. \end{aligned} \quad (21)$$

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 is extremely costly for very large antenna array³. Hence, D_k uses only statistical CSI to decode the desired signal [7], [43]. Thus, we have

$$\tilde{\mathbf{y}}_{\text{D},k}[i] = \underbrace{\alpha\theta\gamma\sqrt{p_{\text{S}}}\mathbf{E}\{\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},k}\}}_{\text{desired signal}} x_{\text{S},k}[i - d] + \underbrace{n_k^{\text{eff}}[i]}_{\text{effective noise}}, \quad (22)$$

where $n_k^{\text{eff}}[i]$ is the effective noise given by

$$\begin{aligned} n_k^{\text{eff}}[i] &= \\ &\underbrace{\alpha\theta\gamma\sqrt{p_{\text{S}}}\left(\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},k} - \mathbb{E}\{\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},k}\}\right)}_{\text{estimation error}} x_{\text{S},k}[i - d] \\ &+ \underbrace{\alpha\theta\gamma\sqrt{p_{\text{S}}}\sum_{j \neq k} \mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},j} x_{\text{S},j}[i - d]}_{\text{interpair interference}} \\ &+ \underbrace{\alpha\theta\gamma\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{G}_{\text{RR}} \hat{\mathbf{x}}_{\text{R}}[i - d]}_{\text{loop interference}} \\ &+ \underbrace{\alpha\theta\gamma\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{n}_{\text{R}}[i - d]}_{\text{noise at the relay}} + \underbrace{\theta\gamma\mathbf{g}_{\text{RD},k}^T \mathbf{F} \tilde{\mathbf{n}}_{\text{R}}[i - d]}_{\text{quantization noise at the relay}} \\ &+ \underbrace{\theta n_{\text{D},k}[i]}_{\text{noise at } \text{D}_k} + \underbrace{\tilde{n}_{\text{D},k}[i]}_{\text{quantization noise at } \text{D}_k}. \end{aligned} \quad (23)$$

Noticing that the “desired signal” and the “effective noise” in (22) are uncorrelated, and capitalizing on the fact that the worst-case uncorrelated additive noise is independent Gaussian noise, we obtain the following achievable rate of k -th user.

³Note that there are two typical ways to obtain the instantaneous CSI, i.e., downlink training and feedback. However, both methods incur huge overheads, hence is highly undesirable.

$$\gamma = \sqrt{\frac{p_R}{\alpha^2 (p_S E \{ \|\mathbf{F}\mathbf{G}_{\text{SR}}\|^2 \} + \frac{p_R}{M} E \{ \|\mathbf{F}\mathbf{G}_{\text{RR}}\|^2 \} + E \{ \|\mathbf{F}\|^2 \}) + E \{ \|\mathbf{F}\tilde{\mathbf{U}}_R\|^2 \}}}. \quad (20)$$

Theorem 1: With low-resolution ADCs at the relay and D_k , the achievable rate of k -th user is given by

$$R_k = \frac{\tau_c - 2\tau_p}{\tau_c} \log_2(1 + \text{SINR}_k), \quad (24)$$

where

$$\text{SINR}_k = \frac{A_k}{B_k + C_k + D_k + E_k + F_k + G_k + H_k}, \quad (25)$$

with

$$A_k = p_S M^4 \sigma_{\text{SR},k}^4 \sigma_{\text{RD},k}^4, \quad (26)$$

$$B_k = p_S M^3 \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^2 (\beta_{\text{SR},k} \sigma_{\text{RD},k}^2 + \beta_{\text{RD},k} \sigma_{\text{SR},k}^2) \quad (27)$$

$$+ p_S M^2 \beta_{\text{SR},k} \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2,$$

$$C_k = p_S M^3 \sum_{j \neq k} (\sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \beta_{\text{SR},j} + \beta_{\text{RD},k} \sigma_{\text{SR},j}^4 \sigma_{\text{RD},j}^2) \quad (28)$$

$$+ p_S M^2 \sum_{j \neq k} \beta_{\text{SR},j} \beta_{\text{RD},k} \sum_{n \neq k,j} \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2$$

$$+ p_S M^2 \sum_{j \neq k} \beta_{\text{SR},j} \beta_{\text{RD},k} \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^2$$

$$+ p_S M^2 \sum_{j \neq k} \beta_{\text{SR},j} \beta_{\text{RD},k} \sigma_{\text{SR},j}^2 \sigma_{\text{RD},j}^2,$$

$$D_k = M^2 \sigma_{\text{LI}}^2 p_R \left(M \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 + \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \right),$$

$$E_k = M^2 \left(M \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 + \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \right), \quad (29)$$

$$F_k = \frac{1-\alpha}{\alpha} M^3 p_S \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \left(\sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \quad (30)$$

$$+ \frac{1-\alpha}{\alpha} M^3 (p_R \sigma_{\text{LI}}^2 + 1) \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4$$

$$+ \frac{1-\alpha}{\alpha} M^2 (p_R \sigma_{\text{LI}}^2 + 1) \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2$$

$$+ \frac{1-\alpha}{\alpha} M^2 p_S \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \left(\sigma_{\text{SR},n}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right),$$

$$G_k = \frac{1}{\alpha^2 \gamma^2}, \quad (31)$$

$$\begin{aligned} H_k &= \frac{1-\theta}{\theta} p_S M^3 \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \left(M \sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \quad (32) \\ &+ \frac{1-\theta}{\theta} p_S M^2 \sigma_{\text{SR},k}^2 \left(M \sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \beta_{\text{RD},k} \sum_{i=1}^K \sigma_{\text{RD},i}^2 \\ &+ \frac{1-\theta}{\alpha \theta} M^3 (p_R \sigma_{\text{LI}}^2 + 1) \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \\ &+ \frac{1-\theta}{\alpha \theta} M^2 (p_R \sigma_{\text{LI}}^2 + 1) \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \\ &+ \frac{(1-\alpha)(1-\theta)}{\alpha \theta} M^3 p_S \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \left(\sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \\ &+ \frac{(1-\alpha)(1-\theta)}{\alpha \theta} p_S M^2 \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^4 \sigma_{\text{RD},n}^2 \\ &+ \frac{(1-\alpha)(1-\theta)}{\alpha \theta} p_S M^2 \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \sum_{i=1}^K \beta_{\text{SR},i} \\ &+ \frac{1-\theta}{\alpha^2 \theta \gamma^2}, \end{aligned}$$

where γ is given by (III) (shown on the top of the next page).

Proof: See Appendix B \square

Theorem 1 presents an exact closed-form expression for the achievable rate of D_k , which is valid for arbitrary configuration of relay antenna number and user pairs, which enables efficient evaluation of the achievable rate. In addition, it also reveals the impact of key system parameters on the achievable rate. For instance, It can be observed that R_k is an increasing function with respect to M , suggesting the benefits of deploying large antenna array at the relay. Furthermore, R_k decreases with K , which is also intuitive since a larger number of user pairs results in more severe inter-pair interference. Moreover, R_k reduces if α and/or θ become small, indicating that using low-resolution ADCs at the relay and/or the destinations always degrades the achievable rate.

IV. SUM RATE APPROXIMATION

In the previous section, an exact expression has been derived for the achievable rate of k -th user. However, the expression is rather involved, and is not amenable for further manipulations. Motivated by this, we now present a relatively simple large-scale approximation for sufficiently large M . Based on which, the optimal relay power allocation scheme is studied and the asymptotic power scaling law is characterized.

Proposition 1: With low-resolution ADCs at the relay and D_k , for sufficiently large M , R_k can be accurately approximated by \tilde{R}_k , which is given by

$$\tilde{R}_k = \frac{\tau_c - 2\tau_p}{\tau_c} \log_2 \left(1 + \widetilde{\text{SINR}} \right), \quad (34)$$

$$\gamma = \frac{1}{M} \sqrt{\frac{p_R}{p_S \sum_{i=1}^K \sigma_{SR,i}^4 \sigma_{RD,i}^2 (M\alpha^2 + \alpha(1-\alpha)) + \alpha \sum_{i=1}^K \sigma_{SR,i}^2 \sigma_{RD,i}^2 \left(p_S \sum_{j=1}^K \beta_{SR,j} + p_R \sigma_{LI}^2 + 1 \right)}}}. \quad (33)$$

where

$$\widetilde{\text{SINR}} = \frac{\tilde{A}_k}{\tilde{B}_k + \tilde{C}_k + \tilde{D}_k + \tilde{E}_k + \tilde{F}_k + \tilde{G}_k + \tilde{H}_k} \quad (35)$$

with

$$\tilde{A}_k = p_S M \sigma_{SR,k}^2 \sigma_{RD,k}^2, \quad (36)$$

$$\tilde{B}_k = p_S (\beta_{SR,k} \sigma_{RD,k}^2 + \beta_{RD,k} \sigma_{SR,k}^2), \quad (37)$$

$$\tilde{C}_k = p_S \sum_{j \neq k} \left(\sigma_{RD,k}^2 \beta_{SR,j} + \frac{\beta_{RD,k} \sigma_{SR,j}^4 \sigma_{RD,j}^2}{\sigma_{SR,k}^2 \sigma_{RD,k}^2} \right), \quad (38)$$

$$\tilde{D}_k = p_R \sigma_{LI}^2 \sigma_{RD,k}^2, \quad (39)$$

$$\tilde{E}_k = \sigma_{RD,k}^2, \quad (40)$$

$$\tilde{F}_k = \frac{1-\alpha}{\alpha} \sigma_{RD,k}^2 \left(p_S \left(\sigma_{SR,k}^2 + \sum_{i=1}^K \beta_{SR,i} \right) + p_R \sigma_{LI}^2 + 1 \right), \quad (41)$$

$$\tilde{G}_k = \frac{p_S}{p_R \sigma_{SR,k}^2 \sigma_{RD,k}^2} \sum_{i=1}^K \sigma_{SR,i}^4 \sigma_{RD,i}^2, \quad (42)$$

$$\tilde{H}_k = \frac{1-\theta}{\theta} p_S \left(M \sigma_{SR,k}^2 \sigma_{RD,k}^2 + \frac{\sigma_{SR,k}^2 \beta_{RD,k}}{\sigma_{RD,k}^2} \sum_{i=1}^K \sigma_{RD,i}^2 \right) \quad (43)$$

$$\begin{aligned} & + \frac{1-\theta}{\theta} p_S \sigma_{RD,k}^2 \left(\sum_{i=1}^K \beta_{SR,i} + \frac{p_R \sigma_{LI}^2 + 1}{\alpha p_S} \right) \\ & + \frac{(1-\alpha)(1-\theta)}{\alpha \theta} \sigma_{RD,k}^2 p_S \left(\sigma_{SR,k}^2 + \sum_{i=1}^K \beta_{SR,i} \right) \\ & + \frac{(1-\theta) p_S \sum_{i=1}^K \sigma_{SR,i}^4 \sigma_{RD,i}^2}{\theta p_R \sigma_{SR,k}^2 \sigma_{RD,k}^2}. \end{aligned}$$

Proof: By ignoring the insignificant terms in the large M regime in Equations from (26) to (33), the desired result can be obtained after some simple algebraic manipulations. \square

Despite being obtained under the assumption of large M , the above approximation turns out to be rather accurate even for moderate number of antennas, i.e., $M = 64$, as will be shown in Section V. In addition, we observe that the quantization noise at the relay (corresponding to the term \tilde{F}_k) is a function with respect to α whereas the quantization noise at the destination (corresponding to the term \tilde{H}_k) depends on both α and θ . This is expected since the quantization noise scales with the power of input signals and the quantization level of ADCs. For the quantization noise at the relay, the input signals are only quantized once by the low-resolution ADCs at the relay; while for the quantization noise at the destinations, the input signals are double-quantized by the

low-resolution ADCs at the relay and destinations. Finally, we see that the quantization noise at the destination is the most significant term which has the same order as the desired signal (corresponding to the term \tilde{A}_k). Unlike the quantization noise at the relay which can be mitigated by exploiting the large antenna array, the quantization noise at the destination can not be effectively suppressed, hence, is the major performance limiting factor as shown in the following corollary.

Corollary 1: As $M \rightarrow \infty$, the rate of k -th user with low-resolution ADCs at both the relay and destinations converges to

$$\tilde{R}_k \rightarrow \frac{\tau_c - 2\tau_p}{\tau_c} \log_2 \left(1 + \frac{\theta}{1-\theta} \right). \quad (44)$$

Proof: Starting from Proposition 1, by keeping only the most significant terms in (34), the desired result can be obtained after some simple algebraic manipulations. \square

Corollary 1 indicates that, in the asymptotic large antenna array regime, the rate of k -th user converges to a finite limit. In addition, the limit is independent of the resolution level of relay ADCs, and is only determined by the resolution level of destination ADCs, which indicates that using large antenna array at the relay cannot compensate for the rate loss caused by low-resolution ADCs at the destinations.

A. Power scaling law

In this subsection, we investigate the potential for power saving in the data transmission phase due to the deployment of very large antenna array at the relay, when the channel estimation accuracy is fixed. Specifically, we assume that p_p is fixed, while $p_S = E_S/M$ and $p_R = E_R/M$, where E_S and E_R are constants that do not scale with M . Then, as $M \rightarrow \infty$, the rate of k -th user is provided by the following corollary.

Corollary 2: For fixed p_p , E_S and E_R , if $p_S = E_S/M$ and $p_R = E_R/M$, as $M \rightarrow \infty$, we have

$$\tilde{R}_k \rightarrow \frac{\tau_c - 2\tau_p}{\tau_c} \log_2 \left(1 + \frac{\theta}{1-\theta + \frac{1}{\alpha E_S \sigma_{SR,k}^2} + \frac{\sum_{i=1}^K \sigma_{SR,i}^4 \sigma_{RD,i}^2}{E_R \sigma_{SR,k}^4 \sigma_{RD,k}^4}} \right). \quad (45)$$

Corollary 2 reveals a rather remarkable result, that despite the low-resolution ADCs at both the relay and destinations, the transmit power of each source and the relay can still be cut down proportionally to $1/M$ while maintaining a desired constant rate, which is the same as the case with ideal infinite-resolution ADCs.

B. Power allocation

From Proposition 1, it is not difficult to observe that \tilde{R}_k improves when the transmit power of sources p_S increases.

As such, each user should transmit at the maximum power. However, due to the existence of loopback interference, the relationship between \tilde{R}_k and p_R is more complicated. As can be readily shown that $\tilde{R}_k \rightarrow 0$ when $p_R \rightarrow 0$ or $p_R \rightarrow \infty$. Hence, using the maximum relay power does not necessarily yield the best performance. Therefore, we now optimize the relay transmit power with the objective of maximizing the achievable sum rate of the system. Specifically, the optimization problem can be formulated as

$$\mathcal{P}_1 : \underset{p_R}{\text{maximize}} \quad \sum_{k=1}^K \tilde{R}_k \quad (46)$$

$$\text{subject to} \quad p_R \geq 0. \quad (47)$$

Due to the involved sum rate expression, analytical characterization of the optimal p_R^* is intractable. However, the optimal solution can be efficiently obtained by one-dimensional search such as bisection method. To gain further insights, we consider the special homogeneous case where all links have the same large-scale fading, e.g., $\beta_{\text{SR},k} = \beta_{\text{RD},k} = 1$. And we have the following important corollary:

Corollary 3: In a homogeneous communication setting, i.e., $\beta_{\text{SR},k} = \beta_{\text{RD},k} = 1$, the optimal relay transmit power is given by

$$p_R^* = \sqrt{\frac{\alpha p_S K}{\sigma_{\text{LI}}^2}}. \quad (48)$$

Proof: When $\beta_{\text{SR},k} = \beta_{\text{RD},k} = 1$, the problem \mathcal{P}_1 is equivalent to minimizing the function $f(p_R) = b + cp_R + \frac{d}{p_R}$, where $b = \sigma^2 (2Kp_S + \frac{1-\alpha}{\alpha\theta} p_S (\sigma^2 + K) + \frac{1-\alpha\theta}{\alpha\theta} + \frac{1-\theta}{\theta} p_S (2K + M\sigma^2))$, $c = \frac{\sigma^2 \sigma_{\text{LI}}^2}{\alpha\theta}$, and $d = \frac{psK\sigma^2}{1+\tau_p p_p}$ with $\sigma^2 = \frac{\alpha\tau_p p_p}{1+\tau_p p_p}$. Since $f''(p_R) = 2dx^{-3} > 0$, $f(p_R)$ is a convex function. Thus, the optimal transmit power of the relay p_R^* is obtained by solving $f'(p_R) = 0$. \square

Corollary 3 shows that the optimal transmit relay power is a function of the resolution level of ADCs at the relay, the transmit power of users, the number of user pairs, and the residual loop interference power, but is independent of the channel estimation accuracy, resolution level of ADCs at the destination, and the relay antenna number. First of all, we see that less power should be used when the residual loopback interference is high. This is reasonable since increasing the relay power would result in higher residual loopback interference. In contrast, if K becomes large, we should increase p_R to serve the additional users. Now, we turn our attention to the impact of the resolution of ADCs at the relay. It is observed that the optimal relay power increases as the resolution of relay ADCs improves. This is also intuitive, since higher-resolution ADCs result in less quantization errors, hence, the benefit of using large transmit power becomes more significant. For instance, considering the special cases with perfect ADCs and one-bit ADCs, i.e., $\alpha = 1$ and $\alpha = 0.6366$, the difference between the optimal p_R^* of the two cases can be computed as $\Delta p_R^* = (1 - \sqrt{0.0366}) \sqrt{\frac{psK}{\sigma_{\text{LI}}^2}} = 0.2 \sqrt{\frac{psK}{\sigma_{\text{LI}}^2}}$, which implies 20% less power for the case with one-bit ADCs.

C. Deploying low-resolution ADCs at the relay or the destinations?

At this point, it is also worth noting that there exists tradeoff between α and θ , i.e., the rate \tilde{R}_k may remain unchanged by jointly adjusting α and θ , indicating that it is possible to increase the ADC resolution at the relay to compensate for the performance loss due to low resolution ADCs at the destination or vice versa.

Now, let us consider two extreme cases: 1) $\alpha = \alpha_1 \neq 1, \theta = 1$, namely, low-resolution ADCs at the relay and infinite resolution ADCs at the destination. 2) $\alpha = 1, \theta = \theta_2 \neq 1$, namely, infinite resolution ADCs at the relay and low-resolution ADCs at the destination, where α_1 and θ_2 respectively denote the quantization level of the two cases.

If $\alpha = \alpha_1 \neq 1, \theta = 1$, \tilde{R}_k reduces to

$$\tilde{R}_k^{\text{R}} = \frac{\tau_c - 2\tau_p}{\tau_c} \log_2 \left(1 + \frac{\tilde{A}_k}{\tilde{B}_k + \tilde{C}_k + \tilde{D}_k + \tilde{E}_k + \tilde{F}_k + \tilde{G}_k} \right), \quad (49)$$

which can be written in the following form

$$\tilde{R}_k^{\text{R}} = \frac{\tau_c - 2\tau_p}{\tau_c} \log_2 (1 + aM), \quad (50)$$

where a is a constant determined by other system parameters such as α but independent of M . This result suggests that deploying large antenna array at the relay is an effective method to compensate for the rate degradation caused by low-resolution ADCs at the relay.

If $\alpha = 1, \theta = \theta_2 \neq 1$, \tilde{R}_k reduces to

$$\tilde{R}_k^{\text{D}} = \frac{\tau_c - 2\tau_p}{\tau_c} \log_2 \left(1 + \frac{\tilde{A}_k}{\tilde{B}_k + \tilde{C}_k + \tilde{D}_k + \tilde{E}_k + \tilde{G}_k + \tilde{H}_k} \right), \quad (51)$$

where

$$\begin{aligned} \tilde{H}_k &= \frac{1-\theta}{\theta} p_S \left(M\sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^2 + \frac{\sigma_{\text{SR},k}^2 \beta_{\text{RD},k}}{\sigma_{\text{RD},k}^2} \sum_{i=1}^K \sigma_{\text{RD},i}^2 \right) \\ &+ \frac{1-\theta}{\theta} p_S \sigma_{\text{RD},k}^2 \left(\sum_{i=1}^K \beta_{\text{SR},i} + \frac{p_R \sigma_{\text{LI}}^2 + 1}{p_S} \right). \end{aligned} \quad (52)$$

In contrast to the previous case, both the denominator and numerator scales with M , indicating that increasing the number of relay antennas is ineffective if the performance bottleneck is due to low-resolution ADCs at the destination. The above observations reveal the asymmetric impact of low-resolution ADCs at the relay and destination, and also shed key design insights on how to allocate the ADCs to achieve optimal performance. Specifically, we have the following important result:

Corollary 4: If $\alpha_1 = \theta_2$, the achievable rate of case 1 is larger than that of case 2, i.e., $\tilde{R}_k^{\text{R}} \geq \tilde{R}_k^{\text{D}}$.

Proof: Noticing that $\tilde{F}_k < \tilde{H}_k$ as $\alpha_1 = \theta_2 \neq 1$ and $\tilde{F}_k = \tilde{H}_k = 0$ when $\alpha_1 = \theta_2 = 1$, the desired result can be easily obtained. \square

Corollary 4 suggests that, in terms of maximizing the achievable rate, it is preferable to deploy the low-resolution ADCs at the relay and infinite resolution ADCs at the destination.

D. Comparison between full-duplex and half-duplex modes

We now compare the performance of the full- and half-duplex modes. For the half-duplex mode, the sources and the relay separately occupy $\frac{\tau_c - 2\tau_p}{2\tau_c}$ time to transmit signals. For fair comparison, the transmit powers of each source and the relay in half-duplex mode should be twice of the powers in full-duplex mode, to ensure that the total energy spent in a coherence interval for both modes are the same. Therefore, the achievable rate of k -th user in half-duplex mode is given by

$$\tilde{R}_k^H = \frac{\tau_c - 2\tau_p}{2\tau_c} \log_2 \left(1 + \frac{2\tilde{A}_k}{2\tilde{B}_k + 2\tilde{C}_k + \tilde{E}_k + \tilde{F}_k + \tilde{G}_k + \tilde{H}_k} \right), \quad (53)$$

where

$$\tilde{F}_k = \frac{1 - \alpha}{\alpha} \sigma_{RD,k}^2 \left(2p_S \left(\sigma_{SR,k}^2 + \sum_{i=1}^K \beta_{SR,i} \right) + 1 \right), \quad (54)$$

$$\tilde{H}_k = \frac{1 - \theta}{\theta} 2p_S \left(M \sigma_{SR,k}^2 \sigma_{RD,k}^2 + \frac{\sigma_{SR,k}^2 \beta_{RD,k}}{\sigma_{RD,k}^2} \sum_{i=1}^K \sigma_{RD,i}^2 \right) \quad (55)$$

$$+ \frac{1 - \theta}{\theta} 2p_S \sigma_{RD,k}^2 \left(\sum_{i=1}^K \beta_{SR,i} + \frac{1}{2\alpha p_S} \right)$$

$$+ \frac{(1 - \alpha)(1 - \theta)}{\alpha \theta} \sigma_{RD,k}^2 2p_S \left(\sigma_{SR,k}^2 + \sum_{i=1}^K \beta_{SR,i} \right)$$

$$+ \frac{(1 - \theta) p_S \sum_{i=1}^K \sigma_{SR,i}^4 \sigma_{RD,i}^2}{\theta p_R \sigma_{SR,k}^2 \sigma_{RD,k}^2}.$$

It is difficult to tell which mode is better, since the achievable system performance depends on various parameters such as the transmit powers, channel gains, the number of relay antennas, and the loop interference level. In particular, the loop interference level and the number of relay antennas play critical roles. If all the other parameters are fixed and only the loop interference level changes, the full-duplex mode outperforms the half-duplex mode when $\sigma_{LI}^2 \leq \sigma_{LI,0}^2$, where $\sigma_{LI,0}^2$ is the root of $\sum_{k=1}^K \tilde{R}_k = \sum_{k=1}^K \tilde{R}_k^H$. Similarly, if only the number of relay antennas changes, the full-duplex mode outperforms the half-duplex mode when $M \geq M_0$, where M_0 is the root of $\sum_{k=1}^K \tilde{R}_k = \sum_{k=1}^K \tilde{R}_k^H$.

V. NUMERICAL RESULTS

In this section, we present numerical results to validate the previous theoretical analysis. For all illustrative examples, the length of the coherence interval is $\tau_c = 196$ (symbols), chosen according to the LTE standard. The length of pilot sequence is $\tau_p = K$. Also, we set the large-scale fading coefficient $\beta_{SR,k} = \beta_{RD,k} = 1$ for simplicity.

A. Validation of analytical results

Fig. 3 illustrates the impact of relay antenna number on the sum rate of the K destinations. Note that the curves associated with ‘‘Numerical results’’ are generated by Monte-Carlo simulations according to (24) by averaging over 10^4 independent channel realizations, the ‘‘Exact results’’ curves are plotted according to Theorem 1, and the ‘‘Approximations’’ curves are obtained based on Proposition 1. As can be readily observed, the ‘‘Exact results’’ and ‘‘Numerical results’’ curves overlap, which validates our exact analysis. Also, the gap between ‘‘Approximations’’ and ‘‘Numerical results’’ curves is sufficiently small, especially when the number of relay antenna is large. In addition, we observe the sum rate saturates in the high signal-to-noise ratio (SNR) regime. Intuitively, the system becomes interference-limited at high SNR.

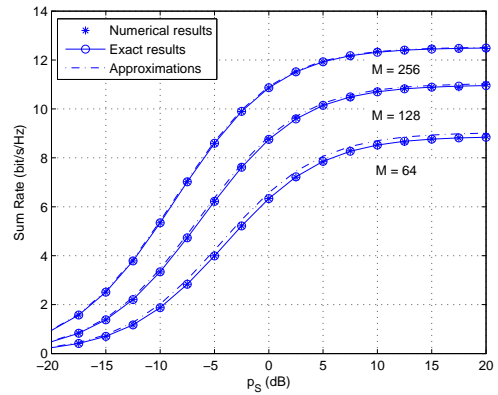


Fig. 3: Sum rate versus p_S for $K = 5$, $\alpha = \theta = 0.8825$ (two-bit ADC), $p_p = 10$ dB, $p_R = 10$ dB, and $\sigma_{LI}^2 = 0$ dB.

B. Where to deploy the low-resolution ADCs?

Fig. 4 demonstrates the asymmetric effect of low-resolution ADCs at the relay and destinations on the sum rate of the K destinations. Recall that $\alpha \neq 1, \theta = 1$ represents employing low-resolution ADCs at the relay only, while $\alpha = 1, \theta \neq 1$ refers to using low-resolution ADCs at the destinations only. We observe a substantial performance gap between the case with $\alpha \neq 1, \theta = 1$ and the case with $\alpha = 1, \theta \neq 1$ when the number of quantization bits b is small, as predicted by Corollary 4. However, when the number of quantization bits b increases, the two curves converge to the same rate. This is because both α and θ approach to 1 as b is large enough ($b \geq 6$ bits in this example), as such the system behaves as that with perfect ADCs.

C. Can we use more antennas to compensate for the coarse quantization?

Fig. 5 plots the sum rate of the K destinations versus the number of relay antennas for different number of quantization bits at the destinations. Note that the ‘‘Sum rate’’ and ‘‘Sum rate limit’’ curves are generated by (24) and (44), respectively. As expected, higher number of quantization bits of ADCs at the destinations results in better sum rate. Also, the sum

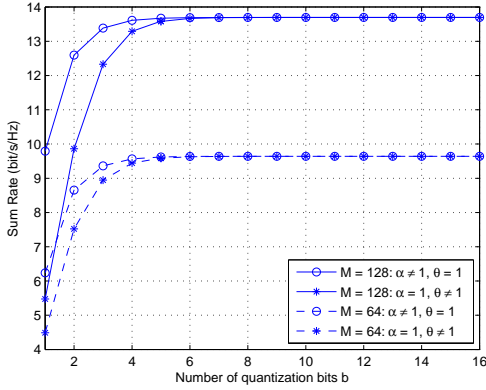


Fig. 4: Sum rate versus b with full-resolution ADCs at the relay or the destinations ($\alpha = 1$ or $\theta = 1$) for $K = 5$, $p_S = 0$ dB, $p_P = 0$ dB, $p_R = 0$ dB, and $\sigma_{LI}^2 = -10$ dB.

rate is an increasing function with respect to the number of relay antennas, and converges to a finite limit determined by b as M becomes large as predicated by Corollary 1, which indicates that using more relay antennas is not an effective approach to compensate for the rate loss due to low resolution ADCs at the destinations. However, the claim would be quite different in the case of low-resolution ADCs at the relay only. As illustrated in Fig. 6, increasing the number of relay antennas can effectively compensate for the rate degradation due to coarse quantization. Nevertheless, the required number of antennas is closely related to the resolution level of ADCs. For instance, compared to the perfect ADCs case, the one-bit system requires approximately twice ($305/158 = 1.93$) antennas to achieve a sum rate of 15 bit/s/Hz (marked by a solid black line), while the 3-bit ADCs system merely needs an additional $9/158 = 5.7\%$ more antennas.

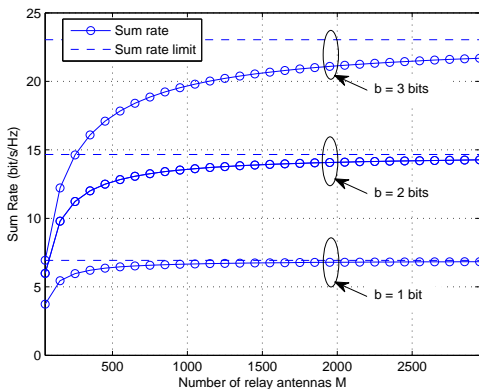


Fig. 5: Sum rate versus M for $K = 5$, $\alpha = 0.8825$, $p_S = 0$ dB, $p_P = 0$ dB, $p_R = 0$ dB, and $\sigma_{LI}^2 = -10$ dB.

D. How does coarse quantization affect the relay transmit power?

Fig. 7 illustrates the impact of relay transmit power on the achievable sum rate of the K destinations with different loopback interference level. As can be readily observed,

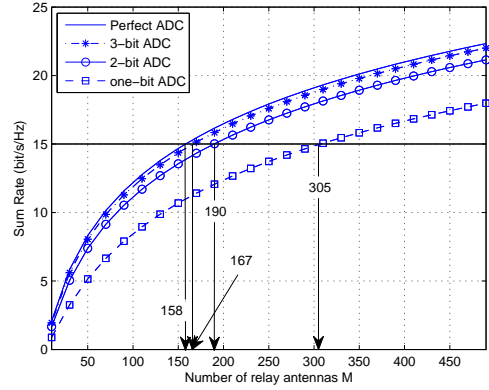


Fig. 6: Sum rate versus M for $K = 5$, $\theta = 1$, $p_S = 0$ dB, $p_P = 0$ dB, $p_R = 0$ dB, and $\sigma_{LI}^2 = -10$ dB.

there exists an optimal relay transmit power maximizing the sum rate. Also, with one-bit ADCs at the relay, the optimal relay transmit power decreases if the loop interference level increases ($p_R^* = 7.51$ dB for $\sigma_{LI}^2 = -20$ dB, and $p_R^* = 2.51$ dB for $\sigma_{LI}^2 = -10$ dB), which aligns with Corollary 3. In addition, we can see that, regardless of the loop interference level, the relay with one-bit ADCs should transmit approximately 20% less power than the perfect ADC case.

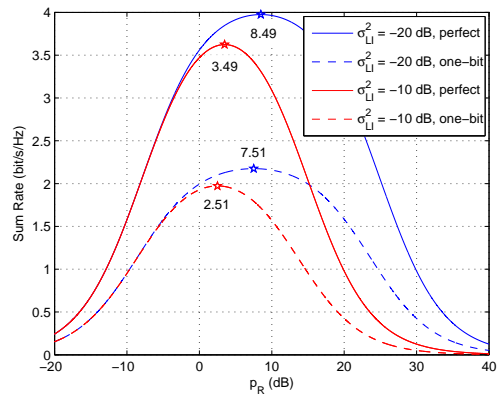


Fig. 7: Sum rate versus p_R for $K = 5$, $M = 64$, $\theta = 0.8825$, $p_S = -10$ dB, and $p_P = -10$ dB.

E. Power scaling law

Fig. 8 plots the required transmit power of each source to maintain a given sum rate of the K destinations 5 bits/s/Hz. Note that the curves associated with “One-bit ADC”, “2-bit ADC” and “Perfect ADC” are obtained by setting $\alpha = \theta = 0.6366$, $\alpha = \theta = 0.8825$, and $\alpha = \theta = 1$, respectively. We can see that, when the number of relay antennas increases, the required transmit powers are significantly reduced. Also, lower-resolution ADC costs more power to achieve the target sum rate. For instance, compared to the perfect ADC case, the one-bit and two-bit ADCs require approximately 10 dB and 2.5 dB more power, respectively. In addition, when σ_{LI}^2 becomes large, we need more transmit power. For instance,

with one-bit ADC and $M = 200$, the required transmit power of each source increases from -6.25 dB for $\sigma_{LI}^2 = -20$ dB to -1.25 dB for $\sigma_{LI}^2 = 0$ dB.

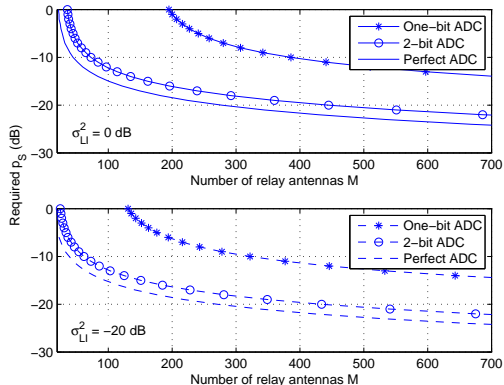


Fig. 8: Required p_s versus the number of relay antennas for $K = 5$, $p_p = 0$ dB, and $p_R = K p_s$.

F. Full-duplex vs. half-duplex modes

Fig. 9 shows the sum rate versus the loop interference levels when the relay and destinations employ 2-bit ADCs. As expected, full-duplex outperforms half-duplex relaying at low σ_{LI}^2 . However, when σ_{LI}^2 is high, loop interference dominates and hence the full-duplex mode is inferior. In such case, we can deploy more antennas to mitigate the effect of loop interference. For instance, by increasing the number of relay antennas from $M = 100$ to $M = 200$, the operating region where the full-duplex relaying works better enlarges, i.e., from $\sigma_{LI,0}^2 = 13.5$ dB to $\sigma_{LI,0}^2 = 15.5$ dB. This fact is further emphasized in Fig. 10. By focusing on the point $M = 185$ in Fig. 10, we observe that the half-duplex is superior with perfect ADCs while is inferior with 2-bit ADCs compared to the full-duplex mode, which indicates that the half-duplex is more sensitive to the low-resolution ADCs. The reason is that the quantization noise scales with the power of input signals, which becomes large since the transmit powers are doubled in the half-duplex mode.

VI. CONCLUSION

We analyzed the achievable sum rate of a multipair full-duplex massive antenna relaying system assuming that both the relay and the destinations use low-resolution ADCs. Exact and approximated closed-form expressions for the achievable sum rate were derived, based on which, the impact of key system parameters was characterized. The findings suggested that, deploying massive relay antenna arrays is an effective approach to compensate for the rate loss due to low-resolution ADCs at the relay, yet becomes ineffective to deal with the rate degradation due to the low-resolution ADCs at the destinations, which indicates that it is important to use higher resolution ADCs at the destination. In addition, it was revealed that, despite the use of low-resolution ADCs, employing massive antenna array at the relay enables significant power savings, i.e., the transmit power of each source can be scaled down proportional to $1/M$, to maintain a constant rate.

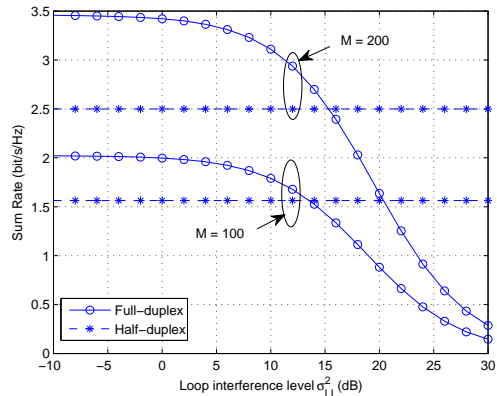


Fig. 9: Sum rate versus σ_{LI}^2 for $K = 5$, $\alpha = \theta = 0.8825$, $p_p = -10$ dB, $p_s = -10$ dB, and $p_R = -10$ dB.

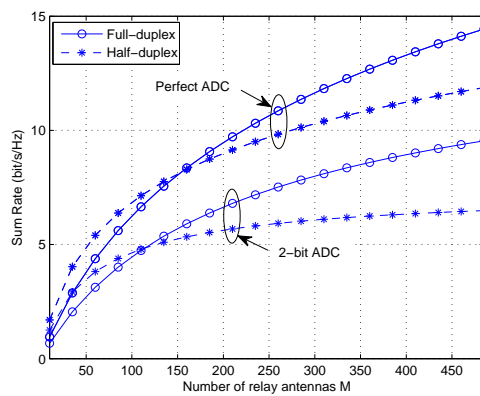


Fig. 10: Sum rate versus M for $K = 5$, $\sigma_{LI}^2 = 16$ dB, $p_p = 0$ dB, $p_s = 0$ dB, and $p_R = 0$ dB.

APPENDIX A PROOF OF THEOREM 1

We focus on the derivation of $\hat{\mathbf{G}}_{SR}$ only since $\hat{\mathbf{G}}_{RD}$ can be deduced in the same fashion.

To start with, we substitute (7) into (9), and obtain

$$\tilde{\mathbf{Y}}_{rp}^H = \alpha \sqrt{\tau_p p_p} \Phi_S^H \mathbf{G}_{SR}^H + \alpha \mathbf{N}_{rp}^H + \mathbf{N}_{rq}^H. \quad (56)$$

Then, following the standard MMSE channel estimation technique, $\hat{\mathbf{G}}_{SR}^H$ is given by

$$\hat{\mathbf{G}}_{SR}^H = \mathbf{E} \left\{ \mathbf{G}_{SR}^H \tilde{\mathbf{Y}}_{rp} \right\} \left(\mathbf{E} \left\{ \tilde{\mathbf{Y}}_{rp}^H \tilde{\mathbf{Y}}_{rp} \right\} \right)^{-1} \tilde{\mathbf{Y}}_{rp}^H. \quad (57)$$

We first calculate

$$\begin{aligned} \mathbf{E} \left\{ \mathbf{G}_{SR}^H \tilde{\mathbf{Y}}_{rp} \right\} &= \mathbf{E} \left\{ \alpha \sqrt{\tau_p p_p} \mathbf{G}_{SR}^H \mathbf{G}_{SR} \Phi_S \right\} \\ &= M \alpha \sqrt{\tau_p p_p} \mathbf{D}_{SR} \Phi_S. \end{aligned} \quad (58)$$

and

$$\begin{aligned}
& \mathbb{E} \left\{ \tilde{\mathbf{Y}}_{\text{rp}}^H \tilde{\mathbf{Y}}_{\text{rp}} \right\} \\
&= \mathbb{E} \left\{ \alpha^2 \tau_p p_p \Phi_S^H \mathbf{G}_{\text{SR}}^H \mathbf{G}_{\text{SR}} \Phi_S \right\} + \alpha^2 \mathbf{I}_{\tau_p} + \mathbb{E} \left\{ \mathbf{N}_{\text{rq}}^H \mathbf{N}_{\text{rq}} \right\} \\
&= M \alpha^2 \tau_p p_p \Phi_S^H \mathbf{D}_{\text{SR}} \Phi_S + M \alpha^2 \mathbf{I}_{\tau_p} \\
&+ \alpha (1 - \alpha) \text{diag} \left(\mathbb{E} \left\{ \mathbf{Y}_{\text{rp}}^H \mathbf{Y}_{\text{rp}} \right\} \right) \\
&= M \alpha^2 \tau_p p_p \Phi_S^H \mathbf{D}_{\text{SR}} \Phi_S + M \alpha^2 \mathbf{I}_{\tau_p} \\
&+ M \alpha (1 - \alpha) \text{diag} \left(\tau_p p_p \Phi_S^H \mathbf{D}_{\text{SR}} \Phi_S + \mathbf{I}_{\tau_p} \right).
\end{aligned} \tag{59}$$

Here, we choose $\tau_p = K$, and assume that Φ_S and Φ_D are identity matrices, i.e., $\Phi_S = \Phi_D = \mathbf{I}_K$, for trackable analysis as in [11]. As a result, we have

$$\text{diag} \left(\Phi_S^H \mathbf{D}_{\text{SR}} \Phi_S \right) = \mathbf{D}_{\text{SR}}, \tag{60}$$

thus,

$$\mathbb{E} \left\{ \tilde{\mathbf{Y}}_{\text{rp}}^H \tilde{\mathbf{Y}}_{\text{rp}} \right\} = M \alpha \left(\tau_p p_p \mathbf{D}_{\text{SR}} + \mathbf{I}_{\tau_p} \right). \tag{61}$$

Using the property of $(\mathbf{A}^H \mathbf{A} + a \mathbf{I}_n)^{-1} \mathbf{A}^H = \mathbf{A}^H (\mathbf{A} \mathbf{A}^H + a \mathbf{I}_m)^{-1}$ ($\mathbf{A} \in \mathbb{C}^{m \times n}$), and then substituting (58) and (61) into (57), we have

$$\hat{\mathbf{G}}_{\text{SR}}^H = \alpha \left(\mathbf{I}_K + \frac{\mathbf{D}_{\text{SR}}^{-1}}{\tau_p p_p} \right)^{-1} \left(\mathbf{G}_{\text{SR}} + \frac{\mathbf{N}_{\text{rp}}}{\sqrt{\tau_p p_p}} + \frac{\mathbf{N}_{\text{rq}}}{\alpha \sqrt{\tau_p p_p}} \right). \tag{62}$$

As a result, we arrive the desired results after some simple mathematical derivations.

APPENDIX B PROOF OF THEOREM 1

Capitalizing on the results of [7], the achievable rate of k -th user is given by (24), where $A_k = p_S \mathbb{E} \left\{ \mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},k} \right\}^2$ is the desired signal power, $B_k = p_S \text{Var} \left(\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},k} \right)$ is the estimation error, $C_k = p_S \sum_{j \neq k} \mathbb{E} \left\{ \left| \mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{SR},j} \right|^2 \right\}$ is the interpair interference, $D_k = \frac{p_R}{M} \mathbb{E} \left\{ \left\| \mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{g}_{\text{RR}} \right\|^2 \right\}$, is the loop interference, $E_k = \mathbb{E} \left\{ \left\| \mathbf{g}_{\text{RD},k}^T \mathbf{F} \right\|^2 \right\}$ is the noise at the relay, $F_k = \frac{1}{\alpha^2} \mathbb{E} \left\{ \left| \mathbf{g}_{\text{RD},k}^T \mathbf{F} \tilde{\mathbf{n}}_{\text{r}}[i-d] \right|^2 \right\}$ is the quantization noise at the relay, $G_k = \frac{1}{\alpha^2 \gamma^2}$ is the noise at the k -th destination, $H_k = \frac{1}{\alpha^2 \theta^2 \gamma^2} \mathbb{E} \left\{ \left| \tilde{n}_{\text{D},k}[i] \right|^2 \right\}$ is the quantization noise at the k -th destination. Besides these terms, we also need to calculate the normalization factor γ in (20). In the following, we compute them one by one.

1) Compute γ :

$$\begin{aligned}
& \text{(a) } \mathbb{E} \left\{ \left\| \mathbf{F} \mathbf{G}_{\text{SR}} \right\|^2 \right\} = \mathbb{E} \left\{ \text{tr} \left(\mathbf{G}_{\text{SR}}^H \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \mathbf{G}_{\text{SR}} \right) \right\} \\
&= \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{SR}}^H \mathbf{G}_{\text{SR}} \mathbf{G}_{\text{SR}}^H \hat{\mathbf{G}}_{\text{SR}} \right\} \mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^T \hat{\mathbf{G}}_{\text{RD}}^* \right\} \right) = \\
&M^3 \sum_{i=1}^K \sigma_{\text{SR},i}^4 \sigma_{\text{RD},i}^2 + M^2 \sum_{i=1}^K \sigma_{\text{SR},i}^2 \sigma_{\text{RD},i}^2 \sum_{j=1}^K \beta_{\text{SR},j}. \\
& \text{(b) } \mathbb{E} \left\{ \left\| \mathbf{F} \mathbf{G}_{\text{RR}} \right\|^2 \right\} = \text{tr} \left(\mathbb{E} \left\{ \mathbf{G}_{\text{RR}} \mathbf{G}_{\text{RR}}^H \right\} \mathbb{E} \left\{ \mathbf{F}^H \mathbf{F} \right\} \right) = \\
&\text{tr} \left(M \sigma_{\text{LI}}^2 \mathbb{E} \left\{ \mathbf{F}^H \mathbf{F} \right\} \right) = M^3 \sigma_{\text{LI}}^2 \sum_{i=1}^K \sigma_{\text{SR},i}^2 \sigma_{\text{RD},i}^2.
\end{aligned}$$

$$\begin{aligned}
& \text{(c) } \mathbb{E} \left\{ \left\| \mathbf{F} \right\|^2 \right\} = \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{SR}}^H \hat{\mathbf{G}}_{\text{SR}} \right\} \mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^T \hat{\mathbf{G}}_{\text{RD}}^* \right\} \right) = \\
&M^2 \sum_{i=1}^K \sigma_{\text{SR},i}^2 \sigma_{\text{RD},i}^2 \\
& \text{(d) } \mathbb{E} \left\{ \left\| \mathbf{F} \tilde{\mathbf{n}}_{\text{r}}[i-d] \right\|^2 \right\} = \alpha (1 - \alpha) \\
&\text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(p_S \mathbf{G}_{\text{SR}} \mathbf{G}_{\text{SR}}^H + \frac{p_R \mathbf{G}_{\text{RR}} \mathbf{G}_{\text{RR}}^H}{M} + \mathbf{I}_M \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right).
\end{aligned}$$

Firstly, we compute

$$\begin{aligned}
& \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\mathbf{G}_{\text{SR}} \mathbf{G}_{\text{SR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right) \\
&= \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{SR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right) \\
&+ \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\mathbf{E}_{\text{SR}} \mathbf{E}_{\text{SR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right).
\end{aligned} \tag{63}$$

By utilizing the fact that the channel matrices $\hat{\mathbf{G}}_{\text{SR}}$, \mathbf{E}_{SR} , and $\hat{\mathbf{G}}_{\text{RD}}$ are independent of each other, $\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{SR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \right\} = M \mathbf{A}$ (where \mathbf{A} is a $K \times K$ diagonal matrix with $[\mathbf{A}]_{kk} = \sigma_{\text{SR},k}^4 + \sigma_{\text{SR},k}^2 \sum_{i=1}^K \sigma_{\text{SR},i}^2$), and $\mathbb{E} \left\{ \text{diag} \left(\mathbf{E}_{\text{SR}} \mathbf{E}_{\text{SR}}^H \right) \right\} = \sum_{i=1}^K \tilde{\sigma}_{\text{SR},i}^2 \mathbf{I}_M$, we have

$$\begin{aligned}
& \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\mathbf{G}_{\text{SR}} \mathbf{G}_{\text{SR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right) \\
&= \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{SR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \right\} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right) \\
&+ \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{SR}}^H \left\{ \text{diag} \left(\mathbf{E}_{\text{SR}} \mathbf{E}_{\text{SR}}^H \right) \right\} \hat{\mathbf{G}}_{\text{SR}} \right\} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right) \\
&= M^2 \left(\sum_{j=1}^K \sigma_{\text{SR},j}^4 \sigma_{\text{RD},j}^2 + \sum_{i=1}^K \sigma_{\text{SR},i}^2 \sum_{j=1}^K \sigma_{\text{SR},j}^2 \sigma_{\text{RD},j}^2 \right) \\
&+ M^2 \sum_{i=1}^K \tilde{\sigma}_{\text{SR},i}^2 \sum_{j=1}^K \sigma_{\text{SR},j}^2 \sigma_{\text{RD},j}^2 \\
&= M^2 \left(\sum_{j=1}^K \sigma_{\text{SR},j}^4 \sigma_{\text{RD},j}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \sum_{j=1}^K \sigma_{\text{SR},j}^2 \sigma_{\text{RD},j}^2 \right).
\end{aligned} \tag{64}$$

Then, following the same way for calculating $\text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\mathbf{G}_{\text{SR}} \mathbf{G}_{\text{SR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right)$, we obtain

$$\begin{aligned}
& \text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\mathbf{G}_{\text{RR}} \mathbf{G}_{\text{RR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right) \\
&= M^3 \sigma_{\text{LI}}^2 \sum_{j=1}^K \sigma_{\text{SR},j}^2 \sigma_{\text{RD},j}^2,
\end{aligned} \tag{65}$$

$$\text{tr} \left(\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \right\} \right) = M^2 \sum_{j=1}^K \sigma_{\text{SR},j}^2 \sigma_{\text{RD},j}^2. \tag{66}$$

Substituting (64), (65), and (66) into (d), we have

$$\begin{aligned}
& \mathbb{E} \left\{ \left\| \mathbf{F} \tilde{\mathbf{n}}_{\text{r}}[i-d] \right\|^2 \right\} \\
&= \alpha (1 - \alpha) M^2 p_S \left(\sum_{j=1}^K \sigma_{\text{SR},j}^4 \sigma_{\text{RD},j}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \sum_{j=1}^K \sigma_{\text{SR},j}^2 \sigma_{\text{RD},j}^2 \right) \\
&+ \alpha (1 - \alpha) M^2 \sum_{j=1}^K \sigma_{\text{SR},j}^2 \sigma_{\text{RD},j}^2 (p_R \sigma_{\text{LI}}^2 + 1).
\end{aligned} \tag{67}$$

To this end, combining the results (a), (b), (c), and (d), we obtain (69), shown on the top of the next page.

$$\gamma = \frac{1}{M} \sqrt{\frac{p_R}{p_S \sum_{i=1}^K \sigma_{SR,i}^4 \sigma_{RD,i}^2 (M\alpha^2 + \alpha(1-\alpha)) + \alpha \sum_{i=1}^K \sigma_{SR,i}^2 \sigma_{RD,i}^2 \left(p_S \sum_{j=1}^K \beta_{SR,j} + p_R \sigma_{LI}^2 + 1 \right)}}}. \quad (69)$$

2) Calculate A_k :

$$\begin{aligned} & \mathbb{E} \{ \mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,k} \} \\ &= \mathbb{E} \{ \hat{\mathbf{g}}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \} \mathbb{E} \{ \hat{\mathbf{g}}_{SR,k}^H \hat{\mathbf{g}}_{SR,k} \} \\ &= M^2 \sigma_{SR,k}^2 \sigma_{RD,k}^2. \end{aligned} \quad (70)$$

Consequently, we have

$$A_k = \alpha^2 \theta^2 \gamma^2 p_S M^4 \sigma_{SR,k}^4 \sigma_{RD,k}^4. \quad (71)$$

3) Calculate B_k :

$$\begin{aligned} & \mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,k}|^2 \} \\ &= \mathbb{E} \left\{ \sum_{n=1}^K \sum_{l=1}^K \mathbf{g}_{RD,k}^T \hat{\mathbf{g}}_{RD,n}^* \hat{\mathbf{g}}_{SR,n}^H \mathbf{g}_{SR,k} \mathbf{g}_{SR,k}^H \hat{\mathbf{g}}_{SR,l} \hat{\mathbf{g}}_{RD,l}^T \hat{\mathbf{g}}_{RD,k}^* \right\}, \end{aligned} \quad (72)$$

which can be decomposed into three different cases:

a) for $n \neq l \neq k$, we have $\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,k}|^2 \} = 0$.

b) for $n = l \neq k$, we have

$$\begin{aligned} & \mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,k}|^2 \} \\ &= \mathbb{E} \left\{ \sum_{n=1}^K \mathbf{g}_{RD,k}^T \hat{\mathbf{g}}_{RD,n}^* \hat{\mathbf{g}}_{SR,n}^H \mathbf{g}_{SR,k} \mathbf{g}_{SR,k}^H \hat{\mathbf{g}}_{SR,n} \hat{\mathbf{g}}_{RD,n}^T \hat{\mathbf{g}}_{RD,k}^* \right\} \\ &= M^2 \beta_{SR,k} \beta_{RD,k} \sum_{n \neq k} \sigma_{SR,n}^2 \sigma_{RD,n}^2. \end{aligned} \quad (73)$$

c) for $n = l = k$, we have

$$\begin{aligned} & \mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,k}|^2 \} \\ &= \mathbb{E} \{ \mathbf{g}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \hat{\mathbf{g}}_{SR,k}^H \mathbf{g}_{SR,k} \mathbf{g}_{SR,k}^H \hat{\mathbf{g}}_{SR,k} \hat{\mathbf{g}}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \} \\ &= \mathbb{E} \{ \hat{\mathbf{g}}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \hat{\mathbf{g}}_{SR,k}^H \hat{\mathbf{g}}_{SR,k} \hat{\mathbf{g}}_{SR,k}^H \hat{\mathbf{g}}_{SR,k} \hat{\mathbf{g}}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \} \\ &+ \mathbb{E} \{ \hat{\mathbf{g}}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \hat{\mathbf{g}}_{SR,k}^H \mathbf{e}_{SR,k} \mathbf{e}_{SR,k}^H \hat{\mathbf{g}}_{SR,k} \hat{\mathbf{g}}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \} \\ &+ \mathbb{E} \{ \mathbf{e}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \hat{\mathbf{g}}_{SR,k}^H \hat{\mathbf{g}}_{SR,k} \hat{\mathbf{g}}_{SR,k}^H \hat{\mathbf{g}}_{SR,k} \hat{\mathbf{g}}_{RD,k}^T \mathbf{e}_{RD,k}^* \} \\ &+ \mathbb{E} \{ \mathbf{e}_{RD,k}^T \hat{\mathbf{g}}_{RD,k}^* \hat{\mathbf{g}}_{SR,k}^H \mathbf{e}_{SR,k} \mathbf{e}_{SR,k}^H \hat{\mathbf{g}}_{SR,k} \hat{\mathbf{g}}_{RD,k}^T \mathbf{e}_{RD,k}^* \} \\ &= M^2 (M+1)^2 \sigma_{SR,k}^4 \sigma_{RD,k}^4 + M^2 (M+1) \sigma_{SR,k}^2 \tilde{\sigma}_{SR,k}^2 \sigma_{RD,k}^4 \\ &+ M^2 (M+1) \sigma_{SR,k}^4 \sigma_{RD,k}^2 \tilde{\sigma}_{RD,k}^2 + M^2 \sigma_{SR,k}^2 \tilde{\sigma}_{SR,k}^2 \sigma_{RD,k}^2 \tilde{\sigma}_{RD,k}^2. \end{aligned} \quad (75)$$

Finally, combining a), b), and c), we obtain

$$\begin{aligned} B_k &= \alpha^2 \theta^2 \gamma^2 p_S \left(\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,k}|^2 \} - |\mathbb{E} \{ \mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,k} \}|^2 \right) \\ &= \alpha^2 \theta^2 \gamma^2 p_S M^3 \sigma_{SR,k}^2 \sigma_{RD,k}^2 (\beta_{SR,k} \sigma_{RD,k}^2 + \beta_{RD,k} \sigma_{SR,k}^2) \\ &+ \alpha^2 \theta^2 \gamma^2 p_S M^2 \beta_{SR,k} \beta_{RD,k} \sum_{n=1}^K \sigma_{SR,n}^2 \sigma_{RD,n}^2. \end{aligned} \quad (77)$$

4) Calculate C_k :

We first rewrite $\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,j}|^2 \}$ for $j \neq k$ as

$$\begin{aligned} & \mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,j}|^2 \} \\ &= \mathbb{E} \left\{ \sum_{n=1}^K \sum_{l=1}^K \mathbf{g}_{RD,k}^T \hat{\mathbf{g}}_{RD,n}^* \hat{\mathbf{g}}_{SR,n}^H \mathbf{g}_{SR,j} \mathbf{g}_{SR,j}^H \hat{\mathbf{g}}_{SR,l} \hat{\mathbf{g}}_{RD,l}^T \hat{\mathbf{g}}_{RD,k}^* \right\}. \end{aligned} \quad (78)$$

Next, (78) can be split into six different cases:

a) for $n \neq l \neq k$, we have $\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,j}|^2 \} = 0$.

b) for $n = l \neq k$, we have $\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,j}|^2 \} = M^2 \beta_{SR,j} \beta_{RD,k} \sum_{n \neq k,j} \sigma_{SR,n}^2 \sigma_{RD,n}^2$.

c) for $n = l = k$, we have $\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,j}|^2 \} = M^2 \beta_{SR,j} \sigma_{SR,k}^2 \sigma_{RD,k}^2 (M \sigma_{RD,k}^2 + \beta_{RD,k})$.

d) for $n = l = j$, we have $\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,j}|^2 \} = M^2 \beta_{RD,k} \sigma_{SR,j}^2 \sigma_{RD,j}^2 (M \sigma_{SR,j}^2 + \beta_{SR,j})$.

e) for $n = k, l = j$, we have $\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,j}|^2 \} = 0$.

f) for $n = j, l = k$, we have $\mathbb{E} \{ |\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{SR,j}|^2 \} = 0$.

Altogether, C_k is given by

$$C_k = \alpha^2 \theta^2 \gamma^2 p_S M^3 \sum_{j \neq k} (\sigma_{SR,k}^2 \sigma_{RD,k}^4 \beta_{SR,j} + \beta_{RD,k} \sigma_{SR,j}^4 \sigma_{RD,j}^2) \quad (79)$$

$$\begin{aligned} & + \alpha^2 \theta^2 \gamma^2 p_S M^2 \sum_{j \neq k} \beta_{SR,j} \beta_{RD,k} \sum_{n \neq k,j} \sigma_{SR,n}^2 \sigma_{RD,n}^2 \\ & + \alpha^2 \theta^2 \gamma^2 p_S M^2 \sum_{j \neq k} \beta_{SR,j} \beta_{RD,k} \sigma_{SR,k}^2 \sigma_{RD,k}^2 \\ & + \alpha^2 \theta^2 \gamma^2 p_S M^2 \sum_{j \neq k} \beta_{SR,j} \beta_{RD,k} \sigma_{SR,j}^2 \sigma_{RD,j}^2. \end{aligned}$$

5) Calculate D_k and E_k :

Following the same technique for deriving B_k , we can obtain

$$\begin{aligned} & \mathbb{E} \{ \|\mathbf{g}_{RD,k}^T \mathbf{F} \mathbf{g}_{RR}\|^2 \} \\ &= M^3 \sigma_{LI}^2 \left(M \sigma_{SR,k}^2 \sigma_{RD,k}^4 + \beta_{RD,k} \sum_{n=1}^K \sigma_{SR,n}^2 \sigma_{RD,n}^2 \right), \\ & \mathbb{E} \{ \|\mathbf{g}_{RD,k}^T \mathbf{F}\|^2 \} \\ &= M^2 \left(M \sigma_{SR,k}^2 \sigma_{RD,k}^4 + \beta_{RD,k} \sum_{n=1}^K \sigma_{SR,n}^2 \sigma_{RD,n}^2 \right). \end{aligned} \quad (81)$$

Thus, D_k and E_k are given by

$$D_k = \alpha^2 \theta^2 \gamma^2 M^3 \sigma_{\text{LI}}^2 p_{\text{R}} \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \quad (82)$$

$$+ \alpha^2 \theta^2 \gamma^2 M^2 \sigma_{\text{LI}}^2 p_{\text{R}} \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2,$$

$$E_k =$$

$$\alpha^2 \theta^2 \gamma^2 M^2 \left(M \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 + \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \right). \quad (83)$$

6) Calculate F_k :

By using the fact that the channel matrices \mathbf{G}_{RD} and $\hat{\mathbf{G}}_{\text{RD}}$ are independent of \mathbf{G}_{SR} , $\hat{\mathbf{G}}_{\text{SR}}$, and \mathbf{G}_{RR} , we have

$$\begin{aligned} & \mathbb{E} \{ \|\mathbf{g}_{\text{RD},k}^T \mathbf{F} \tilde{\mathbf{n}}_{\text{R}}[i-d]\|^2 \} \\ &= \alpha (1-\alpha) \mathbb{E} \left\{ \mathbf{g}_{\text{RD},k}^T \hat{\mathbf{G}}_{\text{RD}}^* \mathbb{E} \{ \mathbf{B} \} \hat{\mathbf{G}}_{\text{RD}}^T \mathbf{g}_{\text{RD},k}^* \right\}, \end{aligned} \quad (84)$$

where

$$\mathbb{E} \{ \mathbf{B} \} = \quad (85)$$

$$\mathbb{E} \left\{ \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(p_{\text{S}} \mathbf{G}_{\text{SR}} \mathbf{G}_{\text{SR}}^H + \frac{p_{\text{R}} \mathbf{G}_{\text{RR}} \mathbf{G}_{\text{RR}}^H}{M} + \mathbf{I}_M \right) \hat{\mathbf{G}}_{\text{SR}} \right\}.$$

Then, following the same fashion as for deducing $\mathbb{E} \{ \|\mathbf{F} \tilde{\mathbf{n}}_{\text{R}}[i-d]\|^2 \}$, we can easily obtain

$$\mathbb{E} \left\{ \mathbf{g}_{\text{RD},k}^T \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\mathbf{G}_{\text{SR}} \mathbf{G}_{\text{SR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \mathbf{g}_{\text{RD},k}^* \right\} \quad (86)$$

$$\begin{aligned} &= M^2 \beta_{\text{RD},k} \sum_{n \neq k} \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \left(\sigma_{\text{SR},n}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \\ &+ M^2 \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^2 \left(M \sigma_{\text{RD},k}^2 + \beta_{\text{RD},k} \right) \left(\sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right), \end{aligned}$$

$$\mathbb{E} \left\{ \mathbf{g}_{\text{RD},k}^T \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \text{diag} \left(\mathbf{G}_{\text{RR}} \mathbf{G}_{\text{RR}}^H \right) \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \mathbf{g}_{\text{RD},k}^* \right\} \quad (87)$$

$$= M^3 \sigma_{\text{LI}}^2 \left(M \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 + \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \right),$$

$$\mathbb{E} \left\{ \mathbf{g}_{\text{RD},k}^T \hat{\mathbf{G}}_{\text{RD}}^* \hat{\mathbf{G}}_{\text{SR}}^H \hat{\mathbf{G}}_{\text{SR}} \hat{\mathbf{G}}_{\text{RD}}^T \mathbf{g}_{\text{RD},k}^* \right\} \quad (88)$$

$$= M^3 \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 + M^2 \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2.$$

Substituting (86), (87), and (88) into (84), we have

$$F_k = \frac{1-\alpha}{\alpha} M^3 p_{\text{S}} \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \left(\sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \quad (89)$$

$$+ \frac{1-\alpha}{\alpha} M^3 (p_{\text{R}} \sigma_{\text{LI}}^2 + 1) \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4$$

$$+ \frac{1-\alpha}{\alpha} M^2 (p_{\text{R}} \sigma_{\text{LI}}^2 + 1) \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2$$

$$+ \frac{1-\alpha}{\alpha} M^2 p_{\text{S}} \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \left(\sigma_{\text{SR},n}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right).$$

7) Calculate H_k :

$$\mathbb{E} \{ \|\tilde{\mathbf{n}}_{\text{D},k}[i]\|^2 \} = \theta (1-\theta) \mathbb{E} \{ |y_{\text{D},k}[i]|^2 \} \quad (90)$$

$$= \alpha^2 \theta (1-\theta) \gamma^2 p_{\text{S}} \mathbb{E} \{ \|\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{G}_{\text{SR}}\|^2 \}$$

$$+ \alpha^2 \theta (1-\theta) \gamma^2 \frac{p_{\text{R}}}{M} \mathbb{E} \{ \|\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{G}_{\text{RR}}\|^2 \}$$

$$+ \alpha^2 \theta (1-\theta) \gamma^2 \mathbb{E} \{ \|\mathbf{g}_{\text{RD},k}^T \mathbf{F}\|^2 \}$$

$$+ \theta (1-\theta) \gamma^2 \mathbb{E} \{ \|\mathbf{g}_{\text{RD},k}^T \mathbf{F} \tilde{\mathbf{n}}_{\text{R}}[i-d]\|^2 \} + \theta (1-\theta).$$

We first calculate

$$\mathbb{E} \{ \|\mathbf{g}_{\text{RD},k}^T \mathbf{F} \mathbf{G}_{\text{SR}}\|^2 \} \quad (91)$$

$$= M^3 \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \left(M \sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right)$$

$$M^2 \sigma_{\text{SR},k}^2 \left(M \sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \beta_{\text{RD},k} \sum_{i=1}^K \sigma_{\text{RD},i}^2.$$

Then, by substituting (91), (80), (81), and (84) into (90), we arrive the following result:

$$H_k = \frac{1-\theta}{\theta} p_{\text{S}} M^3 \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \left(M \sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \quad (92)$$

$$+ \frac{1-\theta}{\theta} p_{\text{S}} M^2 \sigma_{\text{SR},k}^2 \left(M \sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right) \beta_{\text{RD},k} \sum_{i=1}^K \sigma_{\text{RD},i}^2$$

$$+ \frac{1-\theta}{\alpha \theta} M^3 (p_{\text{R}} \sigma_{\text{LI}}^2 + 1) \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4$$

$$+ \frac{1-\theta}{\alpha \theta} M^2 (p_{\text{R}} \sigma_{\text{LI}}^2 + 1) \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2$$

$$+ \frac{(1-\alpha)(1-\theta)}{\alpha \theta} M^3 p_{\text{S}} \sigma_{\text{SR},k}^2 \sigma_{\text{RD},k}^4 \left(\sigma_{\text{SR},k}^2 + \sum_{i=1}^K \beta_{\text{SR},i} \right)$$

$$+ \frac{(1-\alpha)(1-\theta)}{\alpha \theta} p_{\text{S}} M^2 \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2$$

$$+ \frac{(1-\alpha)(1-\theta)}{\alpha \theta} p_{\text{S}} M^2 \beta_{\text{RD},k} \sum_{n=1}^K \sigma_{\text{SR},n}^2 \sigma_{\text{RD},n}^2 \sum_{i=1}^K \beta_{\text{SR},i}$$

$$+ \frac{1-\theta}{\alpha^2 \theta \gamma^2}.$$

Finally, combining (69), (71), (77), (79), (82), (83), (89), (92), and (24), and after some simple algebraic manipulation, we complete the proof.

REFERENCES

- [1] Q. Shi, M. Hong, X. Gao, E. Song, Y. Cai, and W. Xu, "Joint source-relay design for full-duplex MIMO AF relay systems," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6118–6131, Dec. 2016.
- [2] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.
- [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] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," accepted to appear in *IEEE Trans. Veh. Technol.*, earlier access available at IEEEXplore.
- [5] H. Xie, B. Wang, F. Gao, and S. Jin, "A full-space spectrum-sharing strategy for massive MIMO cognitive radio," *IEEE J. Select. Areas Commun.*, vol. 34, no. 10, pp. 2537–2549, Oct. 2016.

- [6] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access*, vol. 4, pp. 7313–7321, Nov. 2016.
- [7] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Oct. 2014.
- [8] X. Xia, Y. Xu, K. Xu, D. Zhang, and W. Ma, "Full-duplex massive MIMO AF relaying with semiblind gain control," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5797–5804, July 2016.
- [9] Z. Zhang, Z. Chen, M. Shen, and B. Xia, "Spectral and energy efficiency of multipair two-way full-duplex relay systems with massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 848–863, Apr. 2016.
- [10] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.
- [11] M. T. Ivrlač and J. A. Nossek, "On MIMO channel estimation with single-bit signal-quantization," in *Proc. ITG Workshop Smart Antennas*, 2007.
- [12] L. Fan, D. Qiao, S. Jin, C.-K. Wen, and M. Matthaiou, "Optimal pilot length for uplink massive MIMO systems with low-resolution ADC," in *Proc. IEEE SAM*, July 2016.
- [13] Y. Li, C. Tao, L. Liu, A. Mezghani and A. L. Swindlehurst, "How much training is needed in one-bit massive MIMO systems at low SNR?," in *Proc. IEEE GLOBECOM*, Dec. 2016.
- [14] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC," [Online] Available: <https://arxiv.org/pdf/1404.7736.pdf>
- [15] C.-K. Wen, C.-J. Wang, S. Jin, K.-K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2541–2556, May 2016.
- [16] T. M. Lok and V.-W. Wei, "Channel estimation with quantized observations," in *Proc. IEEE ISIT*, Aug. 1998, pp. 333.
- [17] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath Jr., "Channel estimation in millimeter wave MIMO systems with one-bit quantization," in *Proc. IEEE Asilomar Conf. Signals Syst. Comput.*, Nov. 2014, pp. 957–961.
- [18] J. Choi, J. Mo, and R. W. Heath Jr., "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2005–2018, May 2016.
- [19] A. Mezghani and J. Nossek, "Analysis of rayleigh-fading channels with 1-bit quantized output," in *Proc. IEEE ISIT*, July 2008, pp. 260–264.
- [20] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009.
- [21] L. Fan, S. Jin, C.-K. Wen, and H. Zhang, "Uplink achievable rate for massive MIMO systems with low-resolution ADC," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2186–2189, Dec. 2015.
- [22] J. Zhang, S. Sun, and Z. Wang, "On the spectral efficiency of massive MIMO systems with low-resolution ADCs," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 842–845, May 2016.
- [23] D. Qiao, W. Tan, Y. Zhao, C.-K. Wen, and S. Jin, "Spectral efficiency for massive MIMO zero-forcing receiver with low-resolution ADC," in *Proc. IEEE WCSP*, Oct. 2016.
- [24] D. Verenzuela, E. Björnson, and M. Matthaiou, "Hardware design and optimal ADC resolution for uplink massive MIMO systems," in *Proc. IEEE SAM*, July 2016.
- [25] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," [Online] Available: <https://arxiv.org/pdf/1612.03271.pdf>
- [26] Y. Li, C. Tao, A. Mezghani, A. L. Swindlehurst, G. Seco-Granados, and L. Liu, "Optimal design of energy and spectral efficiency tradeoff in one-bit massive MIMO systems," [Online] Available: <https://arxiv.org/pdf/1609.07427.pdf>
- [27] A. Gokceoglu, E. Björnson, E. G. Larsson, and M. Valkama, "Spatio-temporal waveform design for multi-user massive MIMO downlink with 1-bit receivers," *IEEE J. Sel. Topics Signal Process.*, Nov. 2016.
- [28] H. Jedda, J. A. Nossek, and A. Mezghani, "Minimum BER precoding in 1-bit massive MIMO systems," in *Proc. IEEE SAM*, July 2016.
- [29] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath Jr., "Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs," [Online] Available: <http://arxiv.org/pdf/1605.00668.pdf>
- [30] N. Liang and W. Zhang, "Mixed-ADC massive MIMO uplink in frequency-selective channels," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4652–4666, Nov. 2016.
- [31] C. Mollen, J. Choi, E. G. Larsson, and R. W. Heath Jr., "One-bit ADCs in wideband massive MIMO systems with OFDM transmission," in *Proc. IEEE ICASSP*, Mar. 2016.
- [32] N. Liang and W. Zhang, "Mixed-ADC massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 983–997, Apr. 2016.
- [33] J. Mo and R. W. Heath Jr., "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, Oct. 2015.
- [34] J. Mo and R. W. Heath Jr., "Limited feedback in multiple-antenna systems with one-bit quantization," in *Proc. IEEE Asilomar Conf. Signals Syst. Comput.*, Nov. 2015.
- [35] J. Liu, J. Xu, W. Xu, S. Jin, and X. Dong, "Multiuser massive MIMO relaying with mixed-ADC receiver," *IEEE Signal Process. Lett.*, Dec. 2016.
- [36] M. Tao and R. Wang, "Linear precoding for multi-pair two-way MIMO relay systems with max-min fairness," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5361–5370, Oct. 2012.
- [37] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *Proc. IEEE ISIT*, July 2012.
- [38] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.
- [39] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay network," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.
- [40] F. Gao, R. Zhang, and Y.-C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.
- [41] L. J. Rodríguez, N. H. Tran, and T. L.-Ngoc, "Performance of full-duplex AF relaying in the presence of residual self-interference," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1752–1764, Sep. 2014.
- [42] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074–3085, Sep. 2011.
- [43] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.