Massive MIMO relaying with hybrid processing

Fozooni, M., Matthaiou, M., Jin, S., & Alexandropoulos, G. C. (2016). Massive MIMO relaying with hybrid processing. In 2016 IEEE International Conference on Communications (ICC). Institute of Electrical and Electronics Engineers (IEEE). DOI: 10.1109/ICC.2016.7510972

Published in:
2016 IEEE International Conference on Communications (ICC)

Document Version:
Peer reviewed version

Massive MIMO Relaying with Hybrid Processing

Milad Fozooni*, Michail Matthaiou*, Shi Jin†, and George C. Alexandropoulos‡
*School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K.
†National Mobile Communications Research Laboratory, Southeast University, Nanjing, China
‡Mathematical and Algorithmic Sciences Lab, France Research Center, Huawei Technologies Co. Ltd., Paris, France
Emails: {mfozooni01, m.matthaiou}@qub.ac.uk, jinshi@seu.edu.cn, george.alexandropoulos@huawei.com

***Abstract*—Massive multiple-input multiple-output (MIMO) relaying is a promising technological paradigm which can offer high spectral efficiency and substantially improved coverage. Yet, these configurations face some formidable challenges in terms of digital signal processing (DSP) power consumption and circuitry complexity, since the number of radio frequency (RF) chains may scale with the number of antennas at the relay station. In this paper, we advocate that performing a portion of the power-intensive DSP in the analog domain, using simple phase shifters and with a reduced number of RF paths, can address these challenges. In particular, we consider a multipair amplify-and-forward (AF) relay system with maximum ratio combining/transmission (MRC/MRT) and we determine the asymptotic spectral efficiency for this hybrid analog/digital architecture. After that, we extend our analytical results to account for heavily quantized analog phase shifters and show that the performance loss with 2 quantization bits is only 10%.**

I. Introduction

Massive MIMO is a promising way to reap all advantages of a MIMO system, such as power and multiplexing gains in a larger scale [1]–[3]. It has also been extensively investigated over the past years, thanks to its ability to cancel out noise, inter-user interference and fast fading. Fortunately, all these advantages can be obtained with simple linear signal processing [3]–[5]. On the other hand, MIMO relay systems have been intensively studied since they can provide extended coverage and enhance the spectral efficiency, particularly at the edges of cells [6]. However, they typically require an extremely complex power allocation and precoder/decoder design [7]–[10]. Therefore, a relaying system with a massive number of antennas at the relay station has emerged as a viable candidate to address the aforementioned challenges.

Massive relaying is a fairly new research area which has been investigated from different viewpoints. In [1], a massive relay is considered to overcome the detrimental effects of loop interference in full-duplex operation. There are also some other research efforts which investigate the spectral efficiency of massive relaying and derive asymptotic scaling laws [4], [7], [11]. However, having one RF chain dedicated to each antenna imposes several challenges in terms of DSP power consumption and circuitry complexity such that this fully digital architecture may scale badly especially in the mm-wave regime [12]. Recently, this critical issue has been addressed by researchers in other fields [12]–[15] and many scholars hold the view that the best suitable solution is a hybrid structure consisting of a digital baseband processor and an analog RF beamformer/combiner. A considerable amount of literature assumes a hybrid analog/digital transceivers for different communications applications [12]–[17], but not in the context of relaying. More recently, [18] assumed a half-duplex relay system where each node is equipped with a hybrid beamformer, but hybrid processing is performed on the nodes not the relay, while no spectral efficiency characterization is being presented either. Motivated by the above discussion, this paper investigates, for the first time ever, the performance of a multipair massive relaying where part of the DSP on the relay station is performed in the analog domain, using simple analog phase shifters. In particular, we analytically determine the asymptotic end-to-end spectral efficiency by considering MRC/MRT processing, where the number of antennas grows up without bound. Then, we elaborate on three power saving strategies and deduce their asymptotic power scaling laws. These laws reveal important physical insights and tradeoffs between the transmit power of user nodes and relay. Finally, we consider the case of quantized phase shifters and work out the performance degradation for small number of quantization bits. Our numerical results indicate that (a) hybrid processing can offer a very satisfactory performance with a substantially lower power consumption and number of RF chains; and (b) 2 bits of quantization cause a minor performance degradation (approximately 10%).

Notation: Upper and lower case bold-face letters denote matrices and vectors, respectively. Also, the symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $\mathrm{Tr}(\cdot)$, $\|\cdot\|$, and $\|\cdot\|_F$ indicate the transpose, conjugate, conjugate transpose, trace operator, Euclidean norm, and Frobenius norm, respectively. In addition, the symbol $[\cdot]_{m,n}$ returns the (m,n)-th element of a matrix. Also, we define the phase and absolute value of a complex number z with $\angle z$ and $|z|$, respectively. Furthermore, $\mathbb{E}[\cdot]$ is the expectation operation, and $\mathbf{I}_N$ refers to the $N \times N$ identity matrix.

II. System Model

Consider a system model as shown in Fig. 1, where a group of K sources, S_k with $k = 1, 2, \ldots, K$, communicate with their own destinations, D_k, via a single one-way relay, R. All sources and destinations are equipped with a single antenna while the relay is equipped with N antennas on each side. Furthermore, the direct link among the K pairs does not exist due to large path loss and heavy shadowing; to keep our analysis simple, full channel state information (CSI) is available and we ignore hardware imperfections [19]. Users send their data streams through a narrowband

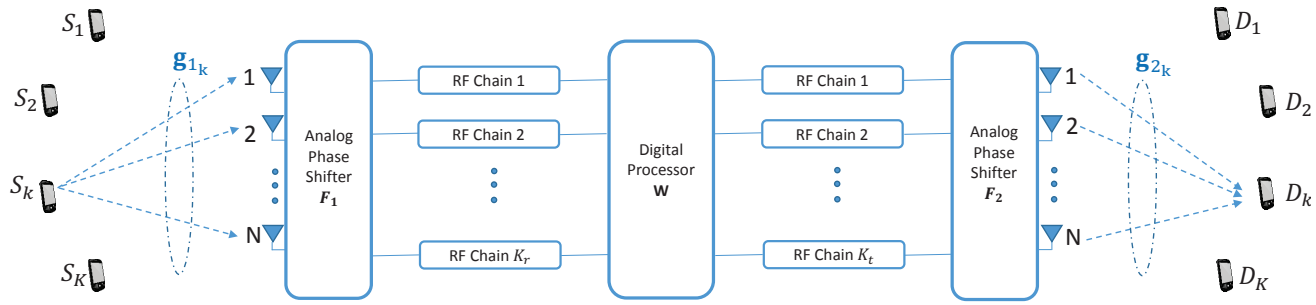

Fig. 1. Block diagram of a massive relay system with a baseband digital processor combined with two analog RF beamformers which are implemented using quantized phase shifters.

flat-fading propagation channel in the same time-frequency block. To keep the implementation cost of this massive MIMO relaying topology at low levels, we consider K_r receive and K_t transmit RF chains at the relay, with $K_r, K_t \ll N$. As mentioned above, by reducing the number of RF paths, we can avail of reduced power consumption (reduced numbers of power amplifiers and analog-to-digital converters) and reduced circuitry. Moreover, to reduce the power dissipation of DSP, we deploy an analog combiner $\mathbf{F}_1 \in \mathbb{C}^{K_r \times N}$ and precoder $\mathbf{F}_2 \in \mathbb{C}^{K_t \times N}$ at the relay station which perform phase matching at a much lower dimension compared to full DSP. Since analog processing alone is not flexible enough, the remaining fraction of signal processing is performed in the digital domain through the smaller dimensional matrix $\mathbf{W} \in \mathbb{C}^{K_t \times K_r}$. Under this model, the received signal at the relay and destination can be mathematically expressed, respectively, as

$$\mathbf{y}_R = \sqrt{P_u}\mathbf{G}_1\mathbf{x} + \mathbf{n}_R \tag{1}$$

$$\mathbf{y}_D = \sqrt{P_u}\mathbf{G}_2^H\mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\mathbf{G}_1\mathbf{x} + \mathbf{G}_2^H\mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\mathbf{n}_R + \mathbf{n}_D \tag{2}$$

where P_u represents the transmitted power of each source, and $\mathbf{x} = [x_1, x_2, \cdots, x_K]^T$ is the zero-mean Gaussian symbol vector such that $\mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right] = \mathbf{I}_K$. Also, the received signal at the destinations is included in $\mathbf{y}_D \in \mathbb{C}^{K\times 1}$, while the N-dimensional vector $\mathbf{n}_R$ and K-dimensional vector $\mathbf{n}_D$ model the additive circularly symmetric complex Gaussian noise such that $\mathbf{n}_R \sim \mathcal{CN}(0, \sigma_{n_R}^2\mathbf{I}_N)$ and $\mathbf{n}_D \sim \mathcal{CN}(0, \sigma_{n_D}^2\mathbf{I}_K)$. Moreover, $\mathbf{G}_1 \in \mathbb{C}^{N\times K}$ and $\mathbf{G}_2 \in \mathbb{C}^{N\times K}$ express the propagation channel between sources and relay, and between relay and destinations, respectively. More precisely, $\mathbf{G}_1 = \mathbf{H}_1\mathbf{D}_1^{\frac{1}{2}}$ and $\mathbf{G}_2 = \mathbf{H}_2\mathbf{D}_2^{\frac{1}{2}}$, where $\mathbf{H}_1$, $\mathbf{H}_2 \in \mathbb{C}^{N\times K}$ refer to small-scale fading channels with independent and identically distributed (i.i.d.) entries, each of them following $\mathcal{CN}(0,1)$. The diagonal matrices $\mathbf{D}_1$ and $\mathbf{D}_2 \in \mathbb{C}^{K\times K}$ include the large-scale fading parameters, where we define $\eta_{1,k} \triangleq \left[\mathbf{D}_1\right]_{k,k}$ and $\eta_{2,k} \triangleq \left[\mathbf{D}_2\right]_{k,k}$. From (2) the received signal at the k-th destination is given by

$$y_{D_k} = \sqrt{P_u}\mathbf{g}_{2_k}^H\mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\mathbf{g}_{1_k}x_k + \sqrt{P_u}\sum_{i\neq k}^{K}\mathbf{g}_{2_k}^H\mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\mathbf{g}_{1_i}x_i + \mathbf{g}_{2_k}^H\mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\mathbf{n}_R + n_{D_k} \tag{3}$$

where $\mathbf{g}_{1_k}$, and $\mathbf{g}_{2_k}$ denote the k-th column of the matrices $\mathbf{G}_1$ and $\mathbf{G}_2$, respectively. In (3), the first term corresponds to the desired signal, the second term refers to the interpair-interference, while the last two terms correspond to the amplified noise at the relay and noise at the destination, respectively. Thus, the instantaneous end-to-end signal-to-interference-noise ratio (SINR) for the k-th pair is given by

$$\mathrm{SINR}_k = \frac{P_u\left|\mathbf{g}_{2_k}^H\mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\mathbf{g}_{1_k}\right|^2}{P_u\sum_{i\neq k}^{K}\left|\mathbf{g}_{2_k}^H\mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\mathbf{g}_{1_i}\right|^2 + \|\mathbf{g}_{2_k}^H\mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\|^2\sigma_{n_R}^2 + \sigma_{n_D}^2}. \tag{4}$$

Consequently, the average spectral efficiency (bits/s/Hz) of this multipair massive MIMO relaying system can be obtained as

$$R = \frac{1}{2}\sum_{k=1}^{K}\mathbb{E}\Big[\log_2\left(1 + \mathrm{SINR_k}\right)\Big] \tag{5}$$

where the pre-log factor $1/2$ is due to the half-duplex relaying. As mentioned before, the role of the analog combiners is to balance out the phase of the propagation matrices. It is noteworthy that the matrices $\mathbf{F}_1$ and $\mathbf{F}_2$ can only perform analog phase shifting, hence, their elements amplitude are assumed to be fixed by $1/\sqrt{N}$. To this end, we have

$$\begin{aligned} \angle\left[\mathbf{F}_1\right]_{i,j} &= -\angle\left[\mathbf{G}_1\right]_{j,i} \\ \angle\left[\mathbf{F}_2\right]_{i,j} &= -\angle\left[\mathbf{G}_2\right]_{j,i} \\ \left|\left[\mathbf{F}_1\right]_{i,j}\right| &= \left|\left[\mathbf{F}_2\right]_{i,j}\right| = \frac{1}{\sqrt{N}}. \end{aligned} \tag{6}$$

On the other hand, the baseband precoder matrix $\mathbf{W}$ can modify both the amplitude and phase of the incoming vector. Moreover, we introduce the following long-term transmit power constraint for the output of the relay station:

$$\mathrm{Tr}\left(\mathbb{E}\left[\tilde{\mathbf{y}}_R\tilde{\mathbf{y}}_R^H\right]\right) = P_r \tag{7}$$

where, $\tilde{\mathbf{y}}_R = \mathbf{F}_2^H\mathbf{W}\mathbf{F}_1\mathbf{y}_R$ demonstrates the combined relay output signal. In the rest of this paper, we assume MRC to combine received signals at the relay, and also consider MRT to forward the received signals from the relay to the destinations. We recall that MRC/MRT type of processing has been well integrated in the context of massive MIMO, since

it offers a near-optimal performance and can be implemented in a distributed manner [20].

We now define the following symbols that will be used in our subsequent analysis; $\mathbf{A}_1 \triangleq \mathbf{F}_1\mathbf{G}_1$, and $\mathbf{A}_2 \triangleq \mathbf{F}_2\mathbf{G}_2$. Regarding the digital MRC/MRT transformation matrix, $\mathbf{W} \triangleq \alpha\mathbf{A}_2\mathbf{A}_1^H$ where α is a normalization constant that guarantees that the power constraint in (7) is satisfied. Therefore, we can obtain after some mathematical manipulations

$$\alpha = \sqrt{\frac{P_r}{P_u\|\mathbf{F}_2^H\mathbf{A}_2\mathbf{A}_1^H\mathbf{A}_1\|_F^2 + \sigma_{n_R}^2\|\mathbf{F}_2^H\mathbf{A}_2\mathbf{A}_1^H\mathbf{F}_1\|_F^2}}. \quad (8)$$

III. Large N analysis

In this section, we asymptotically analyze the performance of the massive MIMO relay with hybrid processing in two dedicated subsections: section III-A assuming ideal (continuous) phase shifters, and section III-B assuming phase quantization.

A. Ideal (continuous) phase shifters

We now briefly review some asymptotic results that will be particularly useful in our analysis.

Lemma 1. Let $\mathbf{p}$ and $\mathbf{q}$ be two $n \times 1$ mutually independent vectors whose elements are i.i.d RVs with variances σ_p^2 and σ_q^2, respectively. Then, based on the law of large numbers, we have

$$\frac{1}{n}\mathbf{p}^H\mathbf{p} \xrightarrow{a.s.} \sigma_p^2 \quad \text{, and} \quad \frac{1}{n}\mathbf{q}^H\mathbf{q} \xrightarrow{a.s.} \sigma_q^2 \quad \text{, as} \quad n \to \infty \quad (9)$$

where $\xrightarrow{a.s.}$ indicates almost sure convergence. Moreover, based on the Lindeberg–Lévy central limit theorem we can write

$$\frac{1}{\sqrt{n}}\mathbf{p}^H\mathbf{q} \xrightarrow{\text{dist.}} \mathcal{CN}(0, \sigma_p^2\sigma_q^2) \quad (10)$$

where $\xrightarrow{\text{dist.}}$ shows the convergence in distribution.

We can now turn our attention to the analog processing matrices $\mathbf{F}_1$ and $\mathbf{F}_2$ which satisfy the following relationship

Lemma 2. As $N \to \infty$, the matrices $\mathbf{F}_1\mathbf{F}_1^H$ and also $\mathbf{F}_2\mathbf{F}_2^H$ converge pairwise to the identity matrix as follows

$$\begin{aligned} \mathbf{F}_1\mathbf{F}_1^H &\xrightarrow{a.s.} \mathbf{I}_{K_r} \\ \mathbf{F}_2\mathbf{F}_2^H &\xrightarrow{a.s.} \mathbf{I}_{K_t}. \end{aligned} \quad (11)$$

Proof. See Appendix I. □

Lemma 3. As $N \to \infty$, the analog phase shifter, $\mathbf{F}_1$, preserves the distribution of the AWGN noise due to its orthonormal rows.

Lemma 4. Let us define $\mathbf{I}_{a,b,r} \in \mathbb{C}^{a\times b}$ as an $a \times b$ diagonal matrix whose first r elements on the main diagonal are 1, and the rest are 0. Then,

$$\begin{aligned} \mathbf{F}_1\mathbf{H}_1 &\xrightarrow{a.s.} \sqrt{\frac{N\pi}{4}}\mathbf{I}_{K_r,K,r_1} \\ \mathbf{F}_2\mathbf{H}_2 &\xrightarrow{a.s.} \sqrt{\frac{N\pi}{4}}\mathbf{I}_{K_t,K,r_2} \end{aligned} \quad (12)$$

where $r_1 = \min(K_r, K)$ and $r_2 = \min(K_t, K)$.

Proof. See Appendix II. □

Turning now to (2) and using the aforementioned lemmas, when $N \to \infty$, it can be shown that

$$\begin{aligned} \mathbf{y}_D \to &\sqrt{P_u}\alpha\Big(\frac{N\pi}{4}\Big)^2\big(\mathbf{D}_2^{\frac{1}{2}}\big)^H\mathbf{D}_2^{\frac{1}{2}}\big(\mathbf{D}_1^{\frac{1}{2}}\big)^H\mathbf{D}_1^{\frac{1}{2}}\mathbf{x} \\ &+\alpha\Big(\frac{N\pi}{4}\Big)^{\frac{3}{2}}\big(\mathbf{D}_2^{\frac{1}{2}}\big)^H\mathbf{D}_2^{\frac{1}{2}}\big(\mathbf{D}_1^{\frac{1}{2}}\big)^H\mathbf{n}_R + \mathbf{n}_D \end{aligned} \quad (13)$$

which can be simplified for the k-th destination as

$$y_{D_k} \to \sqrt{P_u}\alpha\Big(\frac{N\pi}{4}\Big)^2\eta_{2_k}\eta_{1_k}x_k + \alpha\Big(\frac{N\pi}{4}\Big)^{\frac{3}{2}}\eta_{2_k}\eta_{1_k}^{\frac{1}{2}}n_{R_k} + n_{D_k} \quad (14)$$

where $r = \min(K_r, K_t, K)$ and $k \in \{1, 2, ..., r\}$. Thus, from (4) we can obtain the corresponding SINR for the k-th destination in the case that the number of antennas increases without bound

$$\text{SINR}_k \to \frac{(\frac{N\pi}{4})^4 P_u\alpha^2\eta_{1_k}^2\eta_{2_k}^2}{(\frac{N\pi}{4})^3\sigma_{n_R}^2\alpha^2\eta_{1_k}\eta_{2_k}^2 + \sigma_{n_D}^2}. \quad (15)$$

In the following, we investigate three power scaling strategies and draw very interesting engineering insights. Our analysis can be divided into three main cases, namely, Case 1) fixed NP_u and NP_r while $N \to \infty$; Case 2) fixed NP_u while $N \to \infty$; Case 3) fixed NP_r while $N \to \infty$.

1) Let $\lim_{N\to\infty} NP_u = E_u$ and $\lim_{N\to\infty} NP_r = E_r$ where both E_u and E_r are finite constants. Then, from (8) we can get

$$N^3\alpha^2 \to \frac{E_r}{(\frac{\pi}{4})^3 E_t \sum_{i=1}^{r}\eta_{1_i}^2\eta_{2_i} + (\frac{\pi}{4})^2\sigma_{n_R}^2\sum_{i=1}^{r}\eta_{1_i}\eta_{2_i}} \quad (16)$$

which finally yields (17) shown at the top of next page. As a consequence, under a full CSI assumption we can reduce the transmitted power and also relay power proportionally to $\frac{1}{N}$ if the number of relay antennas grows without bound. This result is consistent with [20].

2) Let $\lim_{N\to\infty} NP_u = E_u$, where E_u is a finite constant. Then, returning to (17) and after a few simplifications we obtain

$$\text{SINR}_k \to \frac{\pi}{4}\frac{E_u\eta_{1_k}}{\sigma_{n_R}^2} \quad (18)$$

which is associated with the following average spectral efficiency

$$R_2 \to \frac{1}{2}\sum_{k=1}^{r}\log_2\Big(1 + \frac{\pi}{4}\frac{E_u\eta_{1_k}}{\sigma_{n_R}^2}\Big). \quad (19)$$

The above result is quite intuitive. It shows that if the number of RF chains is, at least, equal to the number of users, i.e. $\min(K_r, K_t) \geq K$ or equivalently $r = K$, we can enjoy full multiplexing gain and boost the achievable spectral efficiency. Moreover, in comparison with a single-input single-output (SISO) system without

$$\mathrm{SINR}_k \to \frac{\left(\frac{\pi}{4}\right)^2 E_u E_r \eta_{1_k}^2 \eta_{2_k}^2}{\left(\frac{\pi}{4}\right) E_r \sigma_{n_R}^2 \eta_{1_k} \eta_{2_k}^2 + \left(\frac{\pi}{4}\right) E_u \sigma_{n_D}^2 \sum\limits_{i=1}^{r} \eta_{1_i}^2 \eta_{2_i} + \sigma_{n_R}^2 \sigma_{n_D}^2 \sum\limits_{i=1}^{r} \eta_{1_i} \eta_{2_i}} \quad k \in \{1, 2, \cdots, r\} \tag{17}$$

any intra-cell interference, our system model only suffers a $\frac{\pi}{4}$-fold reduction on the power gain due to the analog processing. All in all, this power gain penalty is quite acceptable as we have eliminated many relay RF chains, and consequently, we have substantially reduced the circuitry complexity and power consumption. Similar to Case 1, we can infer that we can scale down the transmit power analogously to the number of relay antennas and, still, maintain a non-zero spectral efficiency.

3) Let $\lim\limits_{N\to\infty} NP_r = E_r$, where E_r is a finite constant. Then, we can find out the average spectral efficiency in the same way as pointed out in Case 2 to get

$$R_3 \to \frac{1}{2} \sum_{k=1}^{r} \log_2 \left(1 + \frac{\pi}{4} \frac{E_r \eta_{1_k}^2 \eta_{2_k}^2}{\sigma_{n_D}^2 \sum\limits_{i=1}^{r} \eta_{1_i}^2 \eta_{2_i}} \right). \tag{20}$$

It is noteworthy that if we ignore large-scale fading effects, we get the same results in Case 2 and 3. However, Case 3 converges faster than Case 2 to its own asymptotic result. This can be observed from (17). In Case 2, we can ignore the constant term $E_r \sigma_{n_R}^2 \eta_{1_k} \eta_{2_k}^2$ in comparison with $E_u \sigma_{n_D}^2 \sum_{i=1}^{r} \eta_{1_i}^2 \eta_{2_i}$ even for moderate number of antennas. In contrast, in Case 2, a much higher number of antennas is required to ignore the constant term $E_u \sigma_{n_D}^2 \sum_{i=1}^{r} \eta_{1_i}^2 \eta_{2_i}$ vs. the scaled term $E_r \sigma_{n_R}^2 \eta_{1_k} \eta_{2_k}^2$ in (17).

B. Phase Quantization

Until now, we have assumed ideal analog phase shifters (beamformers) which generate any required phases. However, the implementation of such shifters with continuous phase is not feasible or, at least, is quite expensive due to hardware limitations [12]–[15]. Most importantly, quantized analog beamformers are more attractive in limited feedback systems [16], [21]. In the rest of this paper, the system performance will be assessed under quantized phases. Thus, the phase of each entry of $\mathbf{F}_1$ and $\mathbf{F}_2$ is chosen from a codebook Ψ based on the closest Euclidean distance.

$$\Psi = \left\{ 0, \pm\left(\frac{2\pi}{2^\beta}\right), \pm 2\left(\frac{2\pi}{2^\beta}\right), \cdots, \pm 2^{\beta-1}\left(\frac{2\pi}{2^\beta}\right) \right\} \tag{21}$$

where, β denotes the number of quantization bits. As pointed out previously, the channel coefficients $\left[\mathbf{G}_1\right]_{m,n}$ and $\left[\mathbf{G}_2\right]_{m,n}$ all have uniform phase between 0 and 2π, such that $\angle\left[\mathbf{G}_i\right]_{m,n} = \phi_{m,n} \sim U(0, 2\pi)$, for $i = 1, 2$. Let us define $\epsilon_{m,n}$ as the error between the unquantized phase $\phi_{m,n}$ and quantized phase $\hat{\phi}_{m,n}$ chosen from the codebook

$$\epsilon_{m,n} \triangleq \phi_{m,n} - \hat{\phi}_{m,n}. \tag{22}$$

Due to the uniform distribution of phase, we can easily conclude that the error is an uniform RV, i.e. $\epsilon_{m,n} \sim U[-\delta, +\delta)$, where we define $\delta \triangleq \frac{\pi}{2^\beta}$. This error affects Lemma 4, and in turn, the average spectral efficiency. For this reason, we provide the following lemma to account for phase quantization.[1]

Lemma 5. Let $\hat{\mathbf{F}}_1$ and $\hat{\mathbf{F}}_2$ denote the analog detector and precoder, respectively. Then,

$$\begin{aligned} \hat{\mathbf{F}}_1 \mathbf{H}_1 &\xrightarrow{a.s.} \sqrt{\frac{N\pi}{4}} \operatorname{sinc}(\delta) \mathbf{I}_{K_r, K, r_1} \\ \hat{\mathbf{F}}_2 \mathbf{H}_2 &\xrightarrow{a.s.} \sqrt{\frac{N\pi}{4}} \operatorname{sinc}(\delta) \mathbf{I}_{K_t, K, r_2} \end{aligned} \tag{23}$$

where we define $\operatorname{sinc}(\delta) \triangleq \frac{\sin(\delta)}{\delta}$.

Proof. The results follow trivially by using the methodology outlined in Appendix II. □

Now, we incorporate Lemma 5 into the system model and signal description. The modified normalization factor $\hat{\alpha}$ can be found at (24) on the top of next page. Furthermore, the received signal for the k-th destination can be obtained from the following formula

$$\begin{aligned} \hat{y}_{D_k} \to & \sqrt{P_u} \operatorname{sinc}^4(\delta) \hat{\alpha} \left(\frac{N\pi}{4}\right)^2 \eta_{2_k} \eta_{1_k} x_k \\ & + \left(\frac{N\pi}{4}\right)^{\frac{3}{2}} \operatorname{sinc}^3(\delta) \hat{\alpha}\, \eta_{2_k} \eta_{1_k}^{\frac{1}{2}} n_{R_k} + n_{D_k}. \end{aligned} \tag{25}$$

Phase quantization also affects the power scaling strategies considered in Cases 1–3 above. The corresponding results for these three cases under quantized analog processing can be modified as shown in (26) (on the top of next page), (27) and (28), respectively.

$$\hat{R}_2 \to \frac{1}{2} \sum_{k=1}^{r} \log_2 \left(1 + \frac{\pi}{4} \frac{E_u \eta_{1_k}}{\sigma_{n_R}^2} \operatorname{sinc}^2(\delta) \right). \tag{27}$$

$$\hat{R}_3 \to \frac{1}{2} \sum_{k=1}^{r} \log_2 \left(1 + \frac{\pi}{4} \frac{E_r \eta_{1_k}^2 \eta_{2_k}^2}{\sigma_{n_D}^2 \sum\limits_{i=1}^{r} \eta_{1_i}^2 \eta_{2_i}} \operatorname{sinc}^2(\delta) \right). \tag{28}$$

Taken together, these results indicate a penalty function associated with quantized processing. Roughly speaking, $\operatorname{sinc}^2(\delta)$ is a good approximation of this power gain penalty. In a worst case, where we have only one quantization bit $\beta = 1$, the SINR will be reduced by a factor of $\operatorname{sinc}^2(\frac{\pi}{2}) = \frac{4}{\pi^2} \approx 40\%$. As pointed out in [12], a reasonable rule-of-thumb is to add 1 bit resolution while the number of antennas doubles, since beam width is inversely relative to the number of antennas.

[1] Hereafter, we use a hat sign for the variables that are associated with the quantized beamforming assumption.

$$\hat{\alpha} \to \sqrt{\frac{P_r}{P_u \operatorname{sinc}^6(\delta)\|\mathbf{F}_2^H \mathbf{A}_2 \mathbf{A}_1^H \mathbf{A}_1\|_F^2 + \sigma_{n_R}^2 \operatorname{sinc}^4(\delta)\|\mathbf{F}_2^H \mathbf{A}_2 \mathbf{A}_1^H \mathbf{F}_1\|_F^2}}. \tag{24}$$

$$\hat{R}_1 \to \frac{1}{2}\log_2\left(1 + \frac{(\frac{\pi}{4})^2 \operatorname{sinc}^8(\delta) E_u E_r \eta_{1_k}^2 \eta_{2_k}^2}{(\frac{\pi}{4}) \operatorname{sinc}^6(\delta) E_r \sigma_{n_R}^2 \eta_{1_k} \eta_{2_k}^2 + (\frac{\pi}{4}) \operatorname{sinc}^6(\delta) E_u \sigma_{n_D}^2 \sum_{i=1}^{r} \eta_{1_i}^2 \eta_{2_i} + \operatorname{sinc}^4(\delta) \sigma_{n_R}^2 \sigma_{n_D}^2 \sum_{i=1}^{r} \eta_{1_i} \eta_{2_i}}\right). \tag{26}$$

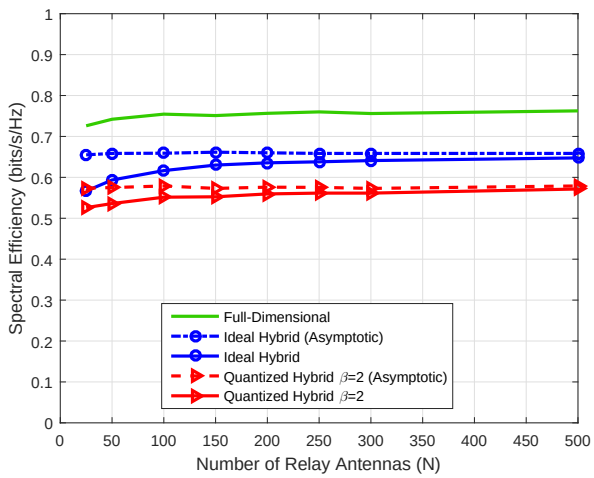

Fig. 2. Average spectral efficiency in Case 1 $(E_u = E_r = 13\text{ dB})$.

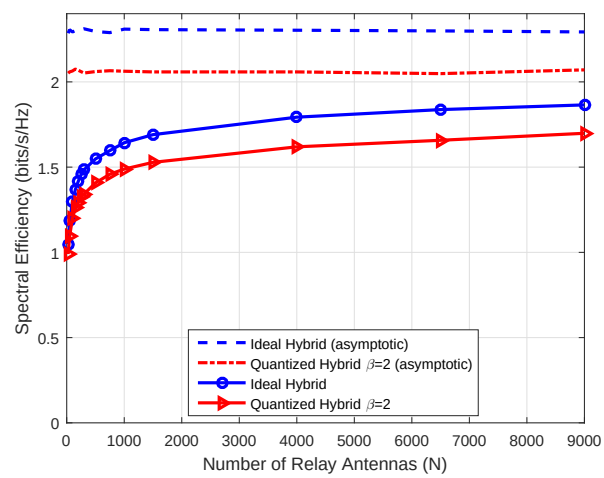

Fig. 3. Average spectral efficiency in Case 2 $(E_u = 13\text{ dB}, P_r = 13\text{ dB})$.

IV. Simulation Results

In this section, Monte-Carlo simulations are provided to assess the validity of the average spectral efficiency of a multipair relay system. We assume that the relay covers a circular area with a radius of 1000 meters. Users are located with a uniform random distribution around the relay with a guard zone of $r_g = 100$ meters. We consider a Rayleigh flat fading channel for small-scale fading effects. Also, the large-scale fading is modeled via a log-normal RV, with standard deviation σ_{sh}, which is multiplied by $\left(\frac{r_k}{r_g}\right)^{\nu}$ to model the path-loss as well. Here, r_k is the distance between the k-th user and the relay, and also ν denotes the path loss exponent. Without loss of generality, we set $\sigma_{n_R}^2 = \sigma_{n_D}^2 = 1$, $\nu = 3.8$, $K_r = K_t = K = 10$ and $\sigma_{sh} = 8$ dB for all simulations.

Figure 2 compares the performance of full-dimensional topology, where all amount of detection/precoding is performed in the digital domain, against that of hybrid topology with continuous and quantized analog processing. A full-dimensional massive relay is equipped by N RF chains which seems to be infeasible in practice, while this number is reduced to only $K = 10$ in the hybrid structure. Moreover, a hybrid relay deploys two inexpensive beamformers which can be actually implemented in the analog domain with phase shifters. It can be also observed that the hybrid scheme performs very close to the conventional scheme, with about a 10% reduction in spectral efficiency but substantially reduced complexity. However, this reduction in spectral efficiency can be compensated by deploying more antennas at the relay without any additional RF chains. Hence, this promising idea seems to be a viable alternative to conventional relaying topologies. Moreover, this figure examines a more restricted case, where there is a severe phase control on beamformers with only 2 bit resolution. Results confirm that the proposed method suffers a negligible reduction. Figures 3 and 4 demonstrate similar results for Case 2 and 3, respectively. Clearly, as the number of relay antennas increases, the average spectral efficiency approaches to the saturation value which is expected by our power scaling laws. Note also that the curve scales much slower in Case 2 in comparison with Case 1 and 3.

V. Conclusion

Massive MIMO is a major candidate for the next generation of wireless systems. This technique combined with relays can enhance the cell coverage while it enjoys a simple signal processing at the relay. On the other hand, the high cost and power consumption of RF chains can be prohibitive due to the large number of mixers and power amplifiers. For this reason, we used an analog/digital (hybrid) structure at the relay and also reduced the number of RF chains analogously to the number of users while the system still enjoys a full multiplexing gain. Finally, we analytically quantified the system

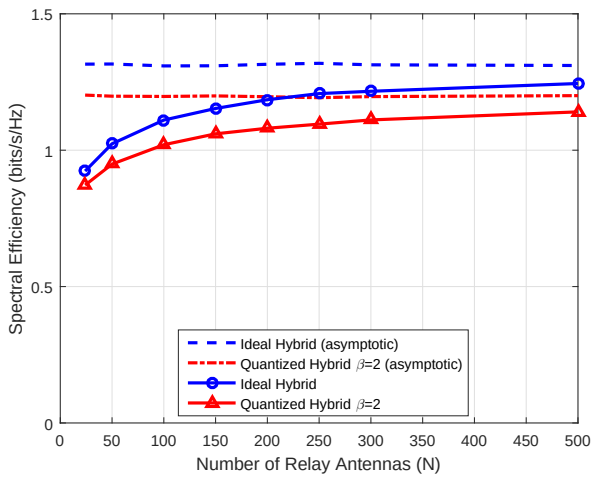

Fig. 4. Average spectral efficiency in Case 3 $\left(P_u = 13\text{ dB}, E_r = 13\text{ dB}\right)$.

spectral efficiency and demonstrated a great performance of the proposed configuration even under coarse quantization.

Appendix I
Proof of Lemma 2

Having discussed how to construct $\mathbf{F}_1 \in \mathbb{C}^{K_r \times N}$, we can write each entry of this matrix as $\frac{1}{\sqrt{N}} \exp\left(j\theta_{m,n}\right)$, where $\theta_{m,n}$ is a uniform RV, i.e. $\theta_{m,n} \sim U\left[0, 2\pi\right)$. Now, let the vectors $\mathbf{f}_{1_p}$ and $\mathbf{f}_{1_q}$ denote the p-th and q-th rows of matrix $\mathbf{F}_1$, respectively. Then, $\mathbf{f}_{1_p}\mathbf{f}_{1_p}^H = 1$ since the phases cancel out each other. On the other hand, if $N \to \infty$, due to the central limit theorem for any $p \neq q$ we have

$$\mathbf{f}_{1_p}\mathbf{f}_{1_q}^H = \frac{1}{N}\sum_{l=1}^{N} \mathrm{e}^{j(\theta_{p,l}-\theta_{q,l})} \to \mathbb{E}\left[\mathrm{e}^{j\theta_p}\right]\mathbb{E}\left[\mathrm{e}^{-j\theta_q}\right] = 0 \quad (29)$$

where the distribution of θ_p and θ_q are defined similar to $\theta_{m,n}$. Likewise, we can prove the second part.

Appendix II
Proof of Lemma 4

Let us rewrite the (m,n)-th entry of matrix $\mathbf{H}_1 \in \mathbb{C}^{N\times K}$ like $r_{m,n}\mathrm{e}^{j\phi_{m,n}}$, where the amplitude and phase have a Rayleigh and uniform distribution, respectively. In other words, $r_{m,n} \sim R\left(\sqrt{\frac{1}{2}}\right)$, and $\phi_{m,n} \sim U\left[0, 2\pi\right)$. Now, let the vectors $\mathbf{f}_{1_p}$ and $\mathbf{h}_{1_p}$ denote the p-th row of matrix $\mathbf{F}_1$ and p-th column of matrix $\mathbf{H}_1$, respectively. Then, since phases cancel out each other, for any $p \leq r_1$ we have that

$$\mathbf{f}_{1_p}\mathbf{h}_{1_p} = \frac{1}{\sqrt{N}}\sum_{l=1}^{N} r_{p,l} \overset{(a)}{\to} \sqrt{N}\mathbb{E}\left[r_p\right] = \sqrt{\frac{N\pi}{4}} \quad (30)$$

where we have used the central limit theorem in (a), and the fact that r_p is a Rayleigh RV with parameter $\sqrt{\frac{1}{2}}$. We can also prove the second part in a similar way.

Acknowledgment

This work of S. Jin was supported by the National Natural Science Foundation of China under Grants $(61531011, 61450110445)$.

References

[1] H. Q. Ngo, H. 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, Sept. 2014.

[2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[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. Suraweera, H. Q. Ngo, T. Q. Duong, C. Yuen, and E. G. Larsson, "Multi-pair amplify-and-forward relaying with very large antenna arrays," in *Proc. IEEE ICC*, June 2013, pp. 4635–4640.

[5] W. Zhang, X. Ma, B. Gestner, and D. V. Anderson, "Designing low-complexity equalizers for wireless systems," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 56–62, Jan. 2009.

[6] M. Dohler and Y. Li, *Cooperative communications: hardware, channel and PHY*. John Wiley & Sons, 2010.

[7] S. Jin, X. Liang, K.-K. Wong, X. Gao, and Q. Zhu, "Ergodic rate analysis for multipair massive MIMO two-way relay networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1480–1491, Mar. 2015.

[8] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4837–4851, Dec. 2009.

[9] K. Phan, T. Le-Ngoc, S. Vorobyov, and C. Tellambura, "Power allocation in wireless multi-user relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2535–2545, May 2009.

[10] J. Cao, Z. Zhong, and F. Wang, "Regenerative multi-way relaying: Relay precoding and ordered MMSE-SIC receiver," in *Proc. IEEE VTC*, May 2012.

[11] H. Q. Ngo and E. G. Larsson, "Spectral efficiency of the multipair two-way relay channel with massive arrays," in *Proc. IEEE ASILOMAR*, Nov. 2013, pp. 275–279.

[12] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[13] W. Ni, X. Dong, and W.-S. Lu, "Near-optimal hybrid processing for massive MIMO systems via matrix decomposition," [Online]. Available: http://arxiv.org/pdf/1504.03777v1.pdf.

[14] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653–656, Dec. 2014.

[15] D. Ying, F. W. Vook, T. A. Thomas, and D. J. Love, "Hybrid structure in massive MIMO: Achieving large sum rate with fewer RF chains," in *Proc. IEEE ICC*, June 2015, pp. 2344–2349.

[16] A. Alkhateeb, O. El Ayach, G. Leus, and R. Heath, "Hybrid precoding for millimeter wave cellular systems with partial channel knowledge," in *Proc. IEEE ITA*, Feb. 2013.

[17] A. F. Molisch and X. Zhang, "FFT-based hybrid antenna selection schemes for spatially correlated MIMO channels," *IEEE Commun. Lett.*, vol. 8, no. 1, pp. 36–38, Jan. 2004.

[18] J. Lee and Y. H. Lee, "AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing," in *Proc. IEEE ICC*, June 2014, pp. 5838–5842.

[19] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[20] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[21] J. C. Roh and B. D. Rao, "Design and analysis of MIMO spatial multiplexing systems with quantized feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 2874–2886, Aug. 2006.