Received June 24, 2016, accepted July 23, 2016, date of publication July 28, 2016, date of current version September 28, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2595590

Multi-Pair Two-Way Massive MIMO AF Full-Duplex Relaying With Imperfect CSI Over Ricean Fading Channels

XIAOLI SUN[1], KUI XU[1], (Member, IEEE), WENFENG MA[1], YOUYUN XU[1,2], (Senior Member, IEEE), XIAOCHEN XIA[1], AND DONGMEI ZHANG[1]

[1]College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China
[2]Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Corresponding author: K. Xu (lgdxxukui@126.com)

This work was supported in part by the Major Research Plan through the National Natural Science Foundation of China under Grant 91438115, in part by the National Natural Science Foundation of China under Grant 61671472, Grant 61301165 and Grant 61371123, in part by the Jiangsu Province Natural Science Foundation under Grant BK20160079, in part by the Special Financial Grant of the China Postdoctoral Science Foundation under Grant 2015T81079, in part by the China Postdoctoral Science Foundation under Grant 2014M552612, in part by the Jiangsu Postdoctoral Science Foundation under Grant 1401178C, and in part by the Preresearch Project Foundation under Grant 9140C020306130C02007.

ABSTRACT This paper investigates a multi-pair two-way massive MIMO amplify-and-forward full-duplex relay (FDR) system over Ricean fading channels, where multiple pairs of full-duplex users exchange information within pair through a full-duplex relay with a very large number of antennas (M for transmission and M for reception). First, the zero-forcing reception/zero-forcing transmission and maximum-ratio combining/maximum-ratio transmission processing matrices with imperfect channel state information at the relay are presented. Then, the corresponding asymptotic expressions (in M) of the end-to-end signal-to-interference-plus-noise ratio are derived. Finally, the asymptotic spectral efficiency (SE) and energy efficiency (EE) at the general power scaling schemes when the number of the relay antennas tends to infinity is analyzed. Theoretical analyses and simulation results show that, when $M \to \infty$, the effects of channel estimation error, the self-loop of each user as well as the relay, inter-user interferences, the noise at users and the relay can be eliminated, respectively, if the power scaling scheme is properly selected. Besides, the considered multi-pair two-way FDR outperforms the multi-pair two-way half-duplex relaying on SE and EE performances when M is large. Moreover, a low complexity power control scheme is proposed to optimize the SE and we analyze the impact of the number of user pairs on the SE.

INDEX TERMS Massive MIMO, full-duplex, Ricean fading channels, imperfect CSI, power control.

I. INTRODUCTION

Massive multiple input and multiple output (MIMO), is an evolving technology that has been regarded as one of the key potential technologies for the fifth generation (5G) mobile communications, which is robust, secure, spectrum and energy efficient [1].

At the same time, relay networks have generated a wide interest in the past decade since they can improve the network performance by extending the coverage and increasing the network capacity [2]. A large number of existing works considered the half-duplex relaying (HDR), where the relay transmits and receives using the orthogonal frequency or time resources [3]. However, full-duplex relaying (FDR) has recently attracted considerable attention as an approach to double the spectral efficiency (SE) of the traditional HDR due to the pre-log factor 1/2 eliminated in ergodic capacity expression. In FDR, the relay transmits and receives signals simultaneously over the same time-frequency resource block, which including one-way [4], [5] and two-way communication [6]. In one-way FDR, one or more half-duplex (HD) sources send information to one or more HD destinations through a full-duplex (FD) relay. The sources transmit signals to relay and relay transmit signals to destinations at the same time. However, in two-way FDR, FD terminals on two sides exchange messages using shared FD relay, which means all nodes transmit and receive data at the same time and frequency. Thus, two-way FDR can overcome the inherent time and spectrum resource loss in one-way FDR.

The main obstacle faced by FDR is the difficulty in canceling the signal leakage from the relay output to its input [7], [8]. However, notable SE increase still makes FDR attractive. Traditionally, loop interference (LI) suppression is achieved in the antenna domain, which may require sophisticated electronic implementation [9]. Moreover, MIMO processing provides an effective approach to suppress the LI in the spatial domain [9], [10]. Recent works sought to combine FDR and massive MIMO together to suppress the LI in the spatial domain [11], [12].

[13] studied the achievable rate and optimal power allocation scheme to maximize the energy efficiency (EE) for a given SE in a multi-pair decode and forward (DF)-based massive MIMO FDR system. Since the amplify-and-forward (AF) relaying is more attractive for its lower complexity compared with DF relaying, the asymptotic signal-to-interference-plus-noise ratio (SINR) expression of the AF based massive MIMO FDR system and a low complexity power control scheme to optimize the SE are investigated in [14]. However, they only consider the one-way FDR communication.

The asymptotic SE and EE of multi-pair massive MIMO AF two-way relaying system in Rayleigh fading channels are obtained analytically in [15], but all nodes operate in the HD mode. In [16], the asymptotic achievable sum rates when the number of antennas $M \to \infty$ of the multi-pair two way AF massive MIMO FDR system over Rayleigh fading channel with four power scaling schemes are investigated analytically, wherein the perfect channel state information (CSI) is assumed available.

As we have known that Ricean fading channel model is the most important model for representing a received signal comprising of both diffuse scatter components and a line-of-sight (LOS) component. This type of communication channel is typical for micro/macro cellular communication, high altitude platform communication, as well as indoor radio channels [17], [18]. Hence, it is critical to carry out the beamforming (BF) design, performance analysis and optimization of the massive MIMO system over Ricean fading channel. Meanwhile, the channel estimates can hardly be perfect in practice due to the delayed feedback and channel estimation errors, which not only result in the self-interference being incompletely canceled, but also introduce additional noises.

Motivated by the aforementioned considerations, in this paper, we consider a multi-pair two-way massive MIMO amplify-and-forward full-duplex relay (MM-AF-FDR) system over Ricean fading channels with imperfect CSI in detail. Main contributions of this paper can be summarized as follows:

- The low-complexity transceiver designs at the relay based on zero-forcing reception/zero-forcing transmission (ZFR/ZFT) and maximum-ratio combining/ maximum-ratio transmission (MRC/MRT) processing with imperfect CSI are presented.
- The asymptotic expressions (in M) of the end-to-end SINR are derived.
- The asymptotic SE and EE at the general power scaling schemes ($P_{s,k} = E_{s,k}/M^a, P_R = E_R/M^b, 0 < a, b \le 1$, E_S and E_R are fixed), when M tends to infinity, are analyzed.
- A low-complexity power control (PC) scheme is proposed to optimize the SE.
- The Monte-Carlo simulations are conducted to validate the accuracy of the theoretical analysis and compare the performance of these power scaling schemes. We show that, when $M \to \infty$, the effects of channel estimation error, the self-loop of each user as well as the relay, inter-user interferences, the noise at U_k and the relay can be eliminated respectively if the power scaling scheme is properly selected. Besides, the considered multi-pair two-way FDR outperforms the multi-pair two-way HDR on SE and EE performances when M is large. Moreover, the achievable SE is improved significantly by the proposed PC scheme.

Notation: We use $\mathbf{X}^T$, $\mathbf{X}^H$, $\mathbf{X}^*$, $\mathbf{X}^{-1}$ and $Tr(\mathbf{X})$ to denote the transpose, conjugate-transpose, conjugate, inverse and the trace of $\mathbf{X}$, respectively. Moreover, $\mathbf{I}_M$ denotes an $M \times M$ identity matrix. Finally, $\mathbb{E}\{\cdot\}$ is the expectation operator, $\|\cdot\|$ represents the Euclidean norm.

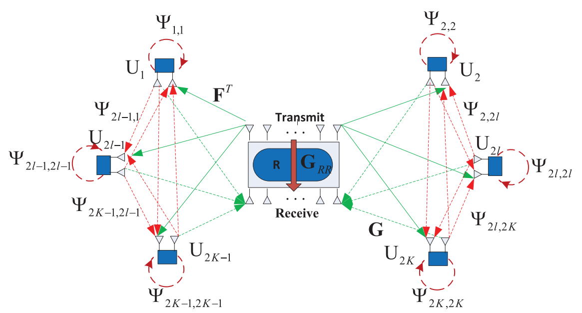


FIGURE 1. Illustration of the multi-pair AF full-duplex two-way relay system.

II. SYSTEM MODEL

Fig. 1 shows the considered K ($K \ge 2$) pairs two-way MM-AF-FDR system, where K pairs full-duplex users (U_{2l-1}, U_{2l}), ($l = 1, \cdots, K$) with two antennas (one for transmission and the other for reception) on two sides try to exchange information within pair through a full-duplex relay (R) with $2M$ antennas (M for transmission and M for reception).

Without loss of generality, a pair of source nodes U_{2l-1} and U_{2l} are too far apart to communicate directly. Meanwhile, one source can be inevitably disturbed by others on the same side.

A. CHANNEL MODEL

We define $\mathbf{G} \in \mathbb{C}^{M \times 2K}$ and $\mathbf{F}^T \in C^{2K \times M}$ as the channel from the transmit antennas of all users to the receive antenna array of R and the channel from the transmit antennas of R to the receive antennas of all users, respectively. More precisely, $\mathbf{G}$ and $\mathbf{F}$ can be expressed as $\mathbf{G} = \mathbf{H}_1\mathbf{D}_1^{1/2}$ and $\mathbf{F} = \mathbf{H}_2\mathbf{D}_2^{1/2}$, where $\mathbf{H}_i \in C^{M\times 2K} (i = 1, 2)$ is the channel matrix representing fast fading and $\mathbf{D}_i \in C^{2K\times 2K}$ is the diagonal matrix representing large-scale fading with $[\mathbf{D}_1]_{kk} = \beta_{g,k}$, $[\mathbf{D}_2]_{kk} = \beta_{f,k}$. Then, the fast fading matrix can be written as $\mathbf{H}_i = \overline{\mathbf{H}}_i\left[\Omega(\Omega+\mathbf{I}_{2K})^{-1}\right]^{1/2} + \mathbf{H}_{i,w}\left[(\Omega+\mathbf{I}_{2K})^{-1}\right]^{1/2}$, where Ω is a $2K\times 2K$ Ricean K-factor diagonal matrix with $[\Omega]_{ii} = \mathrm{K}_i$, $\mathbf{H}_{i,w}$ contains the independent identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ entries. $\overline{\mathbf{H}}_i$ denotes the deterministic component, and we let $\overline{\mathbf{H}}_i$ have an arbitrary rank as [19] $[\overline{\mathbf{H}}_i]_{mk} = e^{-j(m-1)(2\pi d/\lambda)\sin(\theta_k)}$, where θ_k is the arrival angle of the k-th user, λ is the wavelength, and d is the antenna spacing. For convenience, we will set $d = \lambda/2$ in this paper.

$\mathbf{G}_{RR} \in C^{M\times M}$ denotes the echo interference (EI) channel matrix between the transmit and receive arrays of R with i.i.d $\mathcal{CN}(0, \sigma_{rr}^2)$ elements. $\Psi_{k,k}$ and $\Psi_{k,i} (i \in S_k, i \neq k)$ represent the self-loop interference coefficient at U_k and the inter-user interference channel coefficient from U_i to U_k, respectively, where $S_k = \{1, 3, \cdots, 2K-1\}$ or $\{2, 4, \cdots, 2K\}$ denotes the set of users on the same side. The element of $\Psi_{k,k}$ and $\Psi_{k,i}$ can be modeled as i.i.d. $\mathcal{CN}(0, \phi_{k,k})$ and $\mathcal{CN}(0, \phi_{k,i})$ random variables.

B. CHANNEL ESTIMATION AND DATA TRANSMISSION

In real scenarios, R can not obtain the complete information of all the channels, the channel matrices should be estimated at R. In this paper, we consider R has estimation of channels $\mathbf{G}$ and $\mathbf{F}$, but no instantaneous knowledge of $\mathbf{G}_{RR}$. For the considered Ricean fading channel model, we assume that both the Ricean K-factor matrix and the deterministic LOS component are perfectly known at both R and the users, so we only need to estimate $\mathbf{G}_w \triangleq \mathbf{H}_{1,w}{\mathbf{D}_1}^{1/2}$ and $\mathbf{F}_w \triangleq \mathbf{H}_{2,w}{\mathbf{D}_2}^{1/2}$. Since $\mathbf{G}_w$ and $\mathbf{F}_w$ are i.i.d, we can use minimum mean square error (MMSE).[1] And for the estimation of arbitrary correlation matrix, the computational complexity of the MMSE is high in massive MIMO system, some new channel estimators need to be used [20], [23]. Let T be the length of the coherence interval and let τ be the number of symbols used for uplink pilots. During the training part of the coherence interval, all users' receive and transmit antennas simultaneously[2] transmit pilot sequences of length τ symbols to R. The received pilots matrices at the R's receive and transmit antenna arrays are given by $\mathbf{Y}_r = \sqrt{\tau P_T}\mathbf{G}\Phi_1 + \sqrt{\tau P_T}\widetilde{\mathbf{F}}\Phi_2 + \mathbf{N}_r$ and $\mathbf{Y}_t = \sqrt{\tau P_T}\widetilde{\mathbf{G}}\Phi_1 + \sqrt{\tau P_T}\mathbf{F}\Phi_2 + \mathbf{N}_t$, where $\Phi_i \in C^{2K\times\tau}$, $(i = 1, 2)$ are the pilot sequences transmitted from all users' transmit antennas and all users' receive antennas, respectively; $\widetilde{\mathbf{G}} \in C^{M\times 2K}$ and $\widetilde{\mathbf{F}} \in C^{M\times 2K}$ are the channel matrices from all users' transmit antennas to the R's transmit antenna array and from all users' receive antennas to the R's receive antenna array, respectively; $\mathbf{N}_r$ and $\mathbf{N}_t$ are $M \times \tau$ additive white Gaussian noise (AWGN) matrices with i.i.d. $\mathcal{CN}(0, \sigma)$ elements, and P_T is the transmit power of each pilot symbol. All pilot sequences are assumed to be pairwisely orthogonal, i.e., $\overline{\Phi}_i\overline{\Phi}_i^H = \mathbf{I}_{2K}$, and $\overline{\Phi}_i\overline{\Phi}_j^H = 0_{2K}, (i \neq j)$, where $\overline{\Phi}_i \triangleq [(\Omega+\mathbf{I}_{2K})^{-1}]^{1/2}\Phi_i$. This requires that $\tau \geq 4K$, in this paper, we set $\tau = 4K$. We can remove the LOS component, which is assumed to be already known, and the remaining terms of the received matrices are $\mathbf{Y}_{r,w} = \sqrt{\tau P_T}\mathbf{G}_w\overline{\Phi}_1 + \sqrt{\tau P_T}\widetilde{\mathbf{F}}_w\overline{\Phi}_2 + \mathbf{N}_r$, and $\mathbf{Y}_{t,w} = \sqrt{\tau P_T}\widetilde{\mathbf{G}}_w\overline{\Phi}_1 + \sqrt{\tau P_T}\mathbf{F}_w\overline{\Phi}_2 + \mathbf{N}_t$. The MMSE estimate of $\mathbf{G}_w$ and $\mathbf{F}_w$ are $\hat{\mathbf{G}}_w = \frac{1}{\sqrt{\tau P_T}}\mathbf{Y}_{r,w}{\overline{\Phi}_1}^H(\mathbf{I}_{2K} + \frac{1}{\sqrt{\tau P_T}}\mathbf{D}_1^{-1})^{-1}$ and $\hat{\mathbf{F}}_w = \frac{1}{\sqrt{\tau P_T}}\mathbf{Y}_{t,w}{\overline{\Phi}_2}^H(\mathbf{I}_{2K} + \frac{1}{\sqrt{P_p}}\mathbf{D}_2^{-1})^{-1}$ [13], [21].

With MMSE estimator, the real channel can be represented as $\mathbf{G} = \hat{\mathbf{G}} + \Delta\mathbf{G}$ and $\mathbf{F} = \hat{\mathbf{F}} + \Delta\mathbf{F}$, where $\hat{\mathbf{G}}$, $\hat{\mathbf{F}}$ and $\Delta\mathbf{G}$, $\Delta\mathbf{F}$ denote the available channel estimate and estimation error, respectively. The elements of the i-th column of $\Delta\mathbf{G}$ and $\Delta\mathbf{F}$ are RVs with zero means and variances $\varepsilon_{g,i}^2 = \frac{\beta_{g,i}\sigma}{(\sigma+P_p\beta_{g,i})(K_i+1)}, \varepsilon_{f,i}^2 = \frac{\beta_{f,i}\sigma}{(\sigma+P_p\beta_{f,i})(K_i+1)}$, where $P_p = \tau P_T$ [21]. Furthermore, owing to the properties of MMSE estimation, $\Delta\mathbf{G}$ is independent of $\hat{\mathbf{G}}$ and $\Delta\mathbf{F}$ is independent of $\hat{\mathbf{F}}$.

At time instant t, all sources transmit their symbols to R and R forwards the amplified signal to destinations. The received signals at the relay and the k-th user are given by

$$\mathbf{y}_R(t) = \mathbf{G}\mathbf{x}(t) + \mathbf{G}_{RR}\mathbf{x}_R(t) + \mathbf{n}_R(t) \tag{1}$$

$$y_k(t) = \mathbf{f}_k^T\mathbf{x}_R(t) + \sum_{i\in S_k}\Psi_{k,i}x_i(t) + n_k(t) \tag{2}$$

respectively, where $\mathbf{x}(t) = [x_1(t), x_2(t), \cdots, x_{2K}(t)]^T$, and $\mathbb{E}\left\{\mathbf{x}(t)\mathbf{x}^H(t)\right\} = diag(P_{s,1}, \cdots, P_{s,2K}) = \Lambda$. Moreover, $\mathbf{x}_R(t)$ denotes the transmit vector of R with power $P_R = Tr\left(\mathbb{E}\left\{\mathbf{x}_R(t)\mathbf{x}_R^H(t)\right\}\right)$. $\mathbf{n}_R(t)$ and $n_k(t)$ represent the noise vector at R and U_k, the elements of $\mathbf{n}_R(t)$ and $n_k(t)$ are assumed to be i.i.d. $\mathcal{CN}(0, \sigma)$ and $\mathcal{CN}(0, \sigma_n)$.

The transmit vector of R at time instant t can be expressed as

$$\mathbf{x}_R(t) = \mathbf{W}\mathbf{y}_R(t-d) \tag{3}$$

where $\mathbf{W} \in C^{M\times M}$ is the beamforming matrix, and d denotes the processing delay at R.

According to [16], after some ingenious loop interference cancellation (LIC) technique, the residual loop interference (RLI) is so weak that it can be regarded as additional noise. In this paper, some LIC methods can be adopted at the relay before carrying out (3), so we can regard the RLI at R as additional noise. As a result, we replace $\mathbf{x}_R(t)$ in the LI term in (1) by a Gaussian noise source $\widetilde{\mathbf{x}}_R(t)$ with the same power limitation to represent the RLI signal. Then (2) can be rewritten as

$$y_k(t) = (\hat{\mathbf{f}}_k^T + \Delta\mathbf{f}_k^T)\mathbf{W}[(\hat{\mathbf{G}} + \Delta\mathbf{G})\mathbf{x}(t-d) + \mathbf{G}_{RR}\widetilde{\mathbf{x}}_R(t-d) + \mathbf{n}_R(t-d)] + \sum_{i\in S_k}\Psi_{k,i}x_i(t) + n_k(t) \tag{4}$$

[1]According to the estimation expression of channels $\mathbf{G}$ and $\mathbf{F}$ at the end of this paragragh, we can see the inversion operation is for a $2K \times 2K$ diagonal matrix when the channels are i.i.d, so the computational complexity is not high.

[2]Using uplink pilots from all users' transmit antennas to R's receive antenna arrays, $\mathbf{G}$ can estimated. $\mathbf{F}$ can estimated using uplink pilots from all users' receive antennas to R's transmit antenna arrays. Relying on channel reciprocity property of time division duplexing (TDD) based systems, the downlink channel matrix is $\mathbf{F}^T$.

$$a_{zf} \xrightarrow[M\to\infty]{a.s.} \sqrt{\frac{P_R}{\frac{1}{M}\sum_{i=1}^{2K} P_{s,i}\eta_{f,i'}^{-1} + \frac{1}{M^2}(\sigma + \sigma_{rr}^2 P_R + \sum_{j=1}^{2K} P_{s,j}\varepsilon_{g,j}^2)\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}}} \tag{8}$$

III. RELAY TRANSCEIVER DESIGN AND ASYMPTOTIC ANALYSIS

In this section, we propose a ZFR/ZFT-based and MRC/MRT-based beamforming design, the asymptotic end-to-end SINR of the proposed transceiver schemes for multi-pair two-way MM-AF-FDR system are analyzed.

Lemma 1: By the law of large numbers, when M is large enough, the inner product of any two columns in the estimate channel matrix $\hat{\mathbf{G}}$ can be found as [21]:

$$\frac{\hat{\mathbf{g}}_n^H \hat{\mathbf{g}}_i}{M} \xrightarrow{a.s.} \begin{cases} \frac{\beta_{g,n}}{\mathrm{K}_n+1}(\mathrm{K}_n + \frac{P_p\beta_{g,n}}{\sigma+P_p\beta_{g,n}}) = \eta_{g,n}, & i=n \\ 0, & i\neq n \end{cases}$$
$$\frac{\hat{\mathbf{f}}_n^H \hat{\mathbf{f}}_i}{M} \xrightarrow{a.s.} \begin{cases} \frac{\beta_{f,n}}{\mathrm{K}_n+1}(\mathrm{K}_n + \frac{P_p\beta_{f,n}}{\sigma+P_p\beta_{f,n}}) = \eta_{f,n}, & i=n \\ 0, & i\neq n \end{cases} \tag{5}$$

According to *Lemma1*, it can be easily obtained that

$$\frac{\hat{\mathbf{G}}^H\hat{\mathbf{G}}}{M} \xrightarrow{a.s.} diag\left\{\eta_{g,1},\eta_{g,2},\cdots,\eta_{g,2K}\right\} = \mathbf{Q}_1$$
$$\frac{\hat{\mathbf{F}}^H\hat{\mathbf{F}}}{M} \xrightarrow[M\to\infty]{a.s.} diag\left\{\eta_{f,1},\eta_{f,2},\cdots,\eta_{f,2K}\right\} = \mathbf{Q}_2 \tag{6}$$

For simplicity, the time labels are omitted in the sequel.

A. ZFR/ZFT BEAMFORMING

The ZFR/ZFT beamforming matrix is [15]

$$\mathbf{W}_{zf} = a_{zf}\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H \tag{7}$$

where a_{zf} is the amplification factor, $\mathbf{P} = diag\{\mathbf{P}_1,\mathbf{P}_2\cdots,\mathbf{P}_K\}$ and $\mathbf{P}_l = [0\ 1;\ 1\ 0]\,, l=1,\cdots,K$.

Lemma 2: As M approaches to infinity, a_{zf} that satisfies the transmit power constraint of R can be expressed as (8), shown at the top of this page, where (i,i') is a pair of user.

Proof: Please see Appendix A.

Since the instantaneous knowledge of $\mathbf{G}_{RR}$ is not available. To deal with this problem, *Lemma 2* is used in the computation of beamforming matrix.

Substituting (7) into (4), the received signal at the k-th user is

$$y_k = a_{zf}x_{k'} + a_{zf}\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{n}_R + \sum_{i\in S_k}\Psi_{k,i}x_i + n_k + y_k^{ee} + y_k^{ei} \tag{9}$$

where $1_{k'}$ represents a $1\times 2K$ vector, in which the k'-th entry is 1 and others are all 0; y_k^{ee} and y_k^{ei} represent the channel estimation errors and relay EI term, which are expressed as

$$\begin{aligned} y_k^{ee} &= a_{zf}\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\Delta\mathbf{G}\mathbf{x} + a_{zf}\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}\mathbf{x} \\ &\quad + a_{zf}\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\Delta\mathbf{G}\mathbf{x} \\ &\quad + a_{zf}\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{n}_R \\ y_k^{ei} &= a_{zf}\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\tilde{\mathbf{x}}_R \\ &\quad + a_{zf}\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\tilde{\mathbf{x}}_R \end{aligned} \tag{10}$$

Theorem 1: Using ZFR/ZFT beamforming matrix with imperfect CSI from MMSE estimation, the end-to-end SINR at the k-th user can be expressed asymptotically (in M) as

$$\gamma_k = \frac{P_{s,k'}a_{zf}^2}{\mathbb{E}[|y_k^{ee}|^2] + \mathbb{E}[|y_k^{ei}|^2] + \frac{a_{zf}^2}{M}\sigma\eta_{g,k'}^{-1} + \sum_{i\in S_k}P_{s,i}\phi_{k,i} + \sigma_n} \tag{11}$$

where a_{zf} is given by *Lemma 2*, the power for channel estimation error and EI can be expressed asymptotically (in M) as

$$\mathbb{E}[|y_k^{ee}|^2] \xrightarrow[M\to\infty]{a.s.} \frac{a_{zf}^2\eta_{g,k'}^{-1}}{M}\sum_{i=1}^{2K}P_{s,i}\varepsilon_{g,i}^2 + \frac{a_{zf}^2\varepsilon_{f,k}^2}{M}\sum_{i=1}^{2K}P_{s,i}\eta_{f,i'}^{-1} + \frac{a_{zf}^2\varepsilon_{f,k}^2}{M^2}\sum_{j=1}^{2K}P_{s,j}\varepsilon_{g,j}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1} + \frac{a_{zf}^2\sigma}{M^2}\varepsilon_{f,k}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1} \tag{12}$$

$$\mathbb{E}[|y_k^{ei}|^2] \xrightarrow{a.s.} \frac{P_R a_{zf}^2}{M^2}\varepsilon_{f,k}^2\sigma_{rr}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1} + \frac{P_R a_{zf}^2}{M}\sigma_{rr}^2\eta_{g,k'}^{-1} \tag{13}$$

Proof: Please see Appendix B.

Corollary 1: Using ZFR/ZFT beamforming matrix with imperfect CSI from MMSE estimation, at the power-scaling ($P_{s,k} = E_{s,k}/M^a, P_R = E_R/M^b,\ 0 < a,b \le 1$), the asymptotic SINR when $M\to\infty$ is (14) at the bottom of this page.

Proof: Substituting (8) and $P_{s,k} = E_{s,k}/M^a$, $P_R = E_R/M^b$ into *Theorem 1*, when $M\to\infty$, by keeping

$$\gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}}{(E_R M^{-b}\sigma_{rr}^2+\sigma)M^{a-1}\eta_{g,k'}^{-1} + (M^{-a}\sum_{i\in S_k}E_{s,i}\phi_{k,i}+\sigma_n)\frac{M^{b-1}}{E_R}\sum_{i=1}^{2K}E_{s,i}\eta_{f,i'}^{-1} + \frac{M^{a+b-2}}{E_R}\sigma\sigma_n\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}} \tag{14}$$

the major terms and omitting the relatively small terms, *corollary 1* can be obtained.

According to *Corollary 1*, we have the following results

$$a=0,\quad b=1,\ \gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}E_R}{(\sigma_n+\sum_{i\in S_k}E_{s,i}\phi_{k,i})\sum_{i=1}^{2K}E_{s,i}\eta_{f,i'}^{-1}}$$
$$0<a<1,\quad b=1,\ \gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}E_R}{\sigma_n\sum_{i=1}^{2K}E_{s,i}\eta_{f,i'}^{-1}}$$
$$a=1,\quad b=0,\ \gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}\eta_{g,k'}}{E_R\sigma_{rr}^2+\sigma}$$
$$a=1,\quad 0<b<1,\ \gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}\eta_{g,k'}}{\sigma}$$
$$a=1,\quad b=1,$$
$$\gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}}{\sigma\eta_{g,k'}^{-1}+\frac{\sigma_n}{E_R}\sum_{i=1}^{2K}E_{s,i}\eta_{f,i'}^{-1}+\frac{\sigma\sigma_n}{E_R}\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}} \tag{15}$$

Equation (15) indicates that, when $M\to\infty$, the effects of channel estimation error, the self-loop of the user and the relay, inter-user interferences, the noise at users and the relay can be eliminated respectively if the power scaling scheme is properly selected. And in the regime of very large M, the SINR of the above cases will saturates to a deterministic values. When $M\to\infty$, for the case $a=0, b=1$, the detrimental effects from R disappear, the inter-user interference and user noise still remain; for the case $0<a<1, b=1$, the noise from the user still remain; for the case $a=1, b=0$, the detrimental effects from the users disappear, the loop interference and the noise of R still remain; for the case $a=1, 0<b<1$, the interference is caused by the noise of R; for the case $a=1, b=1$, the interference is caused by the noise at R and the user. From the case $a=0$ and $a=1$, we can see that, decreasing the power of the users can reduce the user self-loop and inter-user interferences while increase the detrimental effects from the noise at R; the comparison between the case $b=0$ and $b=1$ indicates that diminishing the power of R can reduce the relay loop interference while increase the baneful effects from the noise at U_k. Moreover, with imperfect CSI and increasing M, the value of the deterministic equivalent for γ_k is dependent on both the scaling parameter a, b, the number of user pairs K and the Ricean K-factor.

B. MRC/MRT BEAMFORMING

The MRC/MRT beamforming matrix is [15]

$$\mathbf{W}_{mrc}=a_{mrc}\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H \tag{16}$$

where a_{mrc} is the amplification factor, where $\mathbf{P}=diag\{\mathbf{P}_1,\mathbf{P}_2\cdots,\mathbf{P}_K\}$ and $\mathbf{P}_l=[0\ 1;\ 1\ 0], l=1,\cdots,K$.

Lemma 3: As M approaches to infinity, a_{mrc} that satisfies the transmit power constraint of R can be expressed as (17), shown at the bottom of this page.

Proof: Please see Appendix C.

Since the instantaneous knowledge of $\mathbf{G}_{RR}$ is not available. To deal with this problem, *Lemma 3* is used in the computation of beamforming matrix.

Substituting (16) into (4), the received signal at the k-th user is

$$y_k=a_{mrc}\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\hat{\mathbf{g}}_{k'}x_{k'}+a_{mrc}\sum_{i\neq k,k'}^{2K}\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\hat{\mathbf{g}}_i x_i$$
$$+a_{mrc}\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{n}_R+\sum_{i\in S_k}\Psi_{k,i}x_i+n_k+y_k^{ee}+y_k^{ei} \tag{18}$$

where y_k^{ee} and y_k^{ei} represent the channel estimation errors and relay EI term, which are expressed as

$$y_k^{ee}=a_{mrc}\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\Delta\mathbf{G}\mathbf{x}+a_{mrc}\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\hat{\mathbf{G}}\mathbf{x}$$
$$+a_{mrc}\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\Delta\mathbf{G}\mathbf{x}+a_{mrc}\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{n}_R$$
$$y_k^{ei}=a_{mrc}\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\widetilde{\mathbf{x}}_R+a_{mrc}\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\widetilde{\mathbf{x}}_R \tag{19}$$

Theorem 2: Using MRC/MRT beamforming matrix with imperfect CSI from MMSE estimation, the end-to-end SINR at the k-th user can be expressed asymptotically (in M) as (20), shown at the bottom of this page, where a_{mrc} is given by Lemma 3, the power for channel estimation error and EI can be expressed asymptotically (in M) as

$$\mathbb{E}[|y_k^{ee}|^2]\xrightarrow[M\to\infty]{a.s.}M^3a_{mrc}^2\eta_{f,k}^2\eta_{g,k'}\sum_{i=1}^{2K}P_{s,i}\varepsilon_{g,i}^2$$
$$+M^2a_{mrc}^2\varepsilon_{f,k}^2\sum_{j=1}^{2K}P_{s,j}\varepsilon_{g,j}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}$$
$$+M^3a_{mrc}^2\varepsilon_{f,k}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}^2P_{s,i'}+M^2a_{mrc}^2\sigma\varepsilon_{f,k}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'} \tag{21}$$

$$a_{mrc}\xrightarrow{a.s.}\sqrt{\frac{P_R}{M^3\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}^2P_{s,i'}+M^2(P_R\sigma_{rr}^2+\sigma+\sum_{j=1}^{2K}P_{s,j}\varepsilon_{g,j}^2)\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}}} \tag{17}$$

$$\gamma_k=\frac{P_{s,k'}\eta_{f,k}^2\eta_{g,k'}^2a_{mrc}^2M^4}{\mathbb{E}[|y_k^{ee}|^2]+\mathbb{E}[|y_k^{ei}|^2]+a_{mrc}^2M^3\sigma\eta_{f,k}^2\eta_{g,k'}+\sum_{i\in S_k}P_{s,i}\phi_{k,i}+\sigma_n} \tag{20}$$

$$\gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}\eta_{f,k}^2\eta_{g,k'}^2}{(E_R M^{-b}\sigma_{rr}^2+\sigma)M^{a-1}\eta_{f,k}^2\eta_{g,k'}+(M^{-a}\sum_{i\in S_k}E_{s,i}\phi_{k,i}+\sigma_n)\frac{M^{b-1}}{E_R}\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}^2E_{s,i'}+\frac{M^{a+b-2}}{E_R}\sigma\sigma_n\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}} \tag{23}$$

$$\mathbb{E}[\left|y_k^{ei}\right|^2] \xrightarrow[M\to\infty]{a.s.} P_R M^3 a_{mrc}^2\sigma_{rr}^2\eta_{f,k}^2\eta_{g,k'} + P_R M^2 a_{mrc}^2\varepsilon_{f,k}^2\sigma_{rr}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'} \tag{22}$$

Proof: Please see Appendix D.

Corollary 2: Using MRC/MRT beamforming matrix with imperfect CSI from MMSE estimation, at the power-scaling ($P_{s,k} = E_{s,k}/M^a, P_R = E_R/M^b$, $0 < a, b \leq 1$), the asymptotic SINR when $M \to \infty$ is (23) at the top of this page.

Proof: Substituting (17) and $P_{s,k} = E_{s,k}/M^a, P_R = E_R/M^b$ into *Theorem 2*, when $M \to \infty$, by keeping the major terms and omitting the relatively small terms, *corollary 2* can be obtained.

According to *Corollary 2*, we have the following results

$$a=0,\quad b=1,$$
$$\gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}E_R\eta_{f,k}^2\eta_{g,k'}^2}{(\sigma_n+\sum_{i\in S_k}E_{s,i}\phi_{k,i})\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}^2E_{s,i'}}$$
$$0<a<1,\quad b=1,\ \gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}E_R\eta_{f,k}^2\eta_{g,k'}^2}{\sigma_n\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}^2E_{s,i'}}$$
$$a=1,\quad b=0,\ \gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}\eta_{g,k'}}{E_R\sigma_{rr}^2+\sigma}$$
$$a=1,\quad 0<b<1,\ \gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}\eta_{g,k'}}{\sigma}$$
$$a=1,\quad b=1,$$
$$\gamma_k \xrightarrow[M\to\infty]{a.s.} \frac{E_{s,k'}\eta_{f,k}^2\eta_{g,k'}^2}{\sigma\eta_{f,k}^2\eta_{g,k'}+\frac{\sigma_n}{E_R}\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}^2E_{s,i'}+\frac{\sigma\sigma_n}{E_R}\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}} \tag{24}$$

Equation (24) indicates that, similar to ZFR/ZFT, when $M \to \infty$, some baneful effects can be eliminated if the power scaling scheme is properly selected. Moreover, with imperfect CSI and increasing M, the value of the deterministic equivalent for γ_k is also dependent on both the scaling parameter a, b, the number of user pairs K and the Ricean K-factor. The main factors that affects SINR of MRC/MRT are the same as ZFR/ZFT. For the case $a = 1, b = 0$ and the case $a = 1, 0 < b < 1$, MRC/MRT and ZFR/ZFT scheme have the same asymptotic γ_k. Specially, when $\eta_{f,i} = \eta_{g,i} = \eta$, the deterministic equivalent for γ_k of ZFR/ZFT becomes the same as that of MRC/MRT in all of the above cases.

IV. SPECTRAL EFFICIENCY OPTIMIZATION

This section presents a low-complexity PC scheme to optimize the SE subject to the maximum power constrains $P_{s,k}^{\max}$ at U_k and $P_R^{\max}$ at R. The SE is defined as $SE = \frac{T-\tau}{T}[\sum_{i=1}^{2K}\log_2(1+\gamma_i)]$.

To facilitate the analysis, we rewrite the end-to-end SINR in *Theorem 1, 2* under ZFR/ZFT and MRC/MRT beamforming schemes as an unified expression (25) at the bottom of this page. For ZFR/ZFT beamforming scheme, we have

$$\begin{aligned}
A_{k,i} &= \frac{\eta_{g,k'}^{-1}\varepsilon_{g,i}^2}{M}+\frac{\varepsilon_{f,k}^2\eta_{f,i'}^{-1}}{M}+\frac{\varepsilon_{f,k}^2\varepsilon_{g,i}^2}{M^2}\sum_{j=1}^{2K}\eta_{f,j}^{-1}\eta_{g,j'}^{-1}\\
a_{k,i} &= \frac{\phi_{k,i}\sigma_{rr}^2\sum_{j=1}^{2K}\eta_{f,j}^{-1}\eta_{g,j'}^{-1}}{M^2}\\
B_k &= \frac{\varepsilon_{f,k}^2\sigma_{rr}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}}{M^2}+\frac{\sigma_{rr}^2\eta_{g,k'}^{-1}}{M}\\
c_i &= \frac{\sigma_n\eta_{f,i'}^{-1}}{M}+\frac{\varepsilon_{g,i}^2\sigma_n\sum_{j=1}^{2K}\eta_{f,j}^{-1}\eta_{g,j'}^{-1}}{M^2}\\
c_{k,i} &= \frac{\phi_{k,i}\sigma\sum_{j=1}^{2K}\eta_{f,j}^{-1}\eta_{g,j'}^{-1}}{M^2}\\
d_i &= \frac{\eta_{f,i'}^{-1}}{M}+\frac{\varepsilon_{g,i}^2}{M^2}\sum_{ii=1}^{2K}\eta_{f,ii}^{-1}\eta_{g,ii'}^{-1}\\
e_k &= \frac{\sigma_n\sigma\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}}{M^2}\\
f_k &= \frac{\sigma}{M^2}\varepsilon_{f,k}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}+\frac{\sigma_n\sigma_{rr}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}}{M^2}+\frac{\sigma\eta_{g,k'}^{-1}}{M}
\end{aligned} \tag{26}$$

$$\gamma_k = \frac{P_{s,k'}}{\sum_{i=1}^{2K}A_{k,i}P_{s,i}+\sum_{i\in S_k}a_{k,i}P_{s,i}+B_kP_R+\sum_{i=1}^{2K}c_i\frac{P_{s,i}}{P_R}+\sum_{i\in S_k}c_{k,i}\frac{P_{s,i}}{P_R}+\sum_{j\in S_k}\phi_{k,j}P_{s,j}\sum_{i=1}^{2K}d_i\frac{P_{s,i}}{P_R}+e_kP_R^{-1}+f_k} \tag{25}$$

For MRC/MRT beamforming scheme, we have

$$A_{k,i} = \frac{\eta_{g,k'}^{-1}\varepsilon_{g,i}^2}{M} + \frac{\eta_{f,i'}\varepsilon_{f,k}^2\eta_{g,i}^2}{\eta_{f,k}^2\eta_{g,k'}^2 M} + \frac{\varepsilon_{f,k}^2\varepsilon_{g,i}^2\sum_{j=1}^{2K}\eta_{f,j}\eta_{g,j'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2}$$

$$a_{k,i} = \frac{\phi_{k,i}\sigma_{rr}^2\sum_{j=1}^{2K}\eta_{f,j}\eta_{g,j'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2}$$

$$B_k = \frac{\varepsilon_{f,k}^2\sigma_{rr}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2} + \frac{\sigma_{rr}^2}{M\eta_{g,k'}}$$

$$C_i = \frac{\sigma_n\eta_{f,i'}\eta_{g,i}^2}{\eta_{f,k}^2\eta_{g,k'}^2 M} + \frac{\varepsilon_{g,i}^2\sigma_n\sum_{j=1}^{2K}\eta_{f,j}\eta_{g,j'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2}$$

$$c_{ki} = \frac{\phi_{k,i}\sigma\sum_{j=1}^{2K}\eta_{f,j}\eta_{g,j'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2}$$

$$d_i = \frac{\eta_{f,i'}\eta_{g,i}^2}{\eta_{f,k}^2\eta_{g,k'}^2 M} + \frac{\varepsilon_{g,i}^2\sum_{ii=1}^{2K}\eta_{f,ii}\eta_{g,ii'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2}$$

$$e_k = \frac{\sigma_n\sigma\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2}$$

$$f_k = \frac{\sigma\varepsilon_{f,k}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2} + \frac{\sigma}{M\eta_{g,k'}} + \frac{\sigma_n\sigma_{rr}^2\sum_{j=1}^{2K}\eta_{f,j}\eta_{g,j'}}{\eta_{f,k}^2\eta_{g,k'}^2 M^2} \tag{27}$$

Algorithm 1 The PC Strategy by GP

Input: The initial value $\hat{\gamma}_k^{(0)}$, a tolerance ε, $P_{s,k}^{\max}$, $P_R^{\max}$, $i = 1$.
Export: The optimal $\hat{\gamma}_k$ and its corresponding $P_{s,k}$, P_R.

while $i <$ Limited Iteration Number **do**
- solve the GP problem:
  $\min_{P_{s,1},\cdots P_{s,2K};P_R} \prod_{k=1}^{2K}\theta_k^{-1}\gamma_k^{-w_k}$ $s.t.$ C_1, C_2, C_3, C_4,
  where $w_k = \hat{\gamma}_k^{(i-1)}/(1+\hat{\gamma}_k^{(i-1)})$, $\theta_k = (\hat{\gamma}_k^{(i-1)})^{-w_k}(1+\hat{\gamma}_k^{(i-1)})$.

  if $\max_{k=1,\cdots,2K}\left|\hat{\gamma}_k^{(i)} - \hat{\gamma}_k^{(i-1)}\right| < \varepsilon$ **then**
    break
  else
    $i = i+1$
  end if

end while

The SE optimization problem can be written as (28) at the bottom of this page.

By introducing the auxiliary variables $v_k \le 1+\gamma_k$, (28) can be reformulated as (29) at the bottom of this page.

Since the target function is not in the posynomial form, the similar method in [22] for solving geometric programming (GP) can be used. Specifically, for any $\gamma_k > 0$, $1+\gamma_k$ can be approximated by a posynomial function $\theta_k(\gamma_k)^{w_k}$ near $\hat{\gamma}_k$, where $w_k = \hat{\gamma}_k/(1+\hat{\gamma}_k)$ and $\theta_k = \hat{\gamma}_k^{-w_k}(1+\hat{\gamma}_k)$. Then the the SE optimization problem can be solved by using several GPs (Algorithm 1 is presented).

$$\max_{P_{s,1},\cdots P_{s,2K};P_R}\prod_{k=1}^{2K}(1+\gamma_k) = \min_{P_{s,1},\cdots P_{s,2K};P_R}\prod_{k=1}^{2K}\frac{1}{1+\gamma_k}$$

$$s.t.\begin{cases} C_1: \dfrac{P_{s,k'}}{\sum_{i=1}^{2K}A_{k,i}P_{s,i} + \sum_{i\in S_k}a_{k,i}P_{s,i} + B_kP_R + \sum_{i=1}^{2K}c_i\frac{P_{s,i}}{P_R} + \sum_{i\in S_k}c_{k,i}\frac{P_{s,i}}{P_R} + \sum_{j\in S_k}\phi_{k,j}P_{s,j}\sum_{i=1}^{2K}d_i\frac{P_{s,i}}{P_R} + e_kP_R^{-1} + f_k} \ge \gamma_k, \\ \quad k = 1,2,\cdots,2K \\ C_2: 0 \le P_{s,k} \le P_{s,k}^{\max}, \\ \quad k = 1,2,\cdots,2K \\ C_3: 0 \le P_R \le P_R^{\max} \end{cases} \tag{28}$$

$$\min_{P_{s,1},\cdots P_{s,2K};P_R}\prod_{k=1}^{2K}\frac{1}{v_k}$$

$$s.t.\begin{cases} C_1: v_k \le 1+\gamma_k, \quad k = 1,2,\cdots,2K \\ C_2: \sum_{i=1}^{2K}A_{k,i}P_{s,i}P_{s,k'}^{-1}\gamma_k + \sum_{i\in S_k}a_{k,i}P_{s,i}P_{s,k'}^{-1}\gamma_k + B_kP_{s,k'}^{-1}P_R\gamma_k + \sum_{i=1}^{2K}c_iP_{s,i}P_{s,k'}^{-1}P_R^{-1}\gamma_k + \sum_{i\in S_k}c_{k,i}P_{s,i}P_{s,k'}^{-1}P_R^{-1}\gamma_k \\ \quad + \sum_{j\in S_k}\phi_{k,j}P_{s,j}\sum_{i=1}^{2K}d_iP_{s,i}P_{s,k'}^{-1}P_R^{-1}\gamma_k + e_kP_{s,k'}^{-1}P_R^{-1}\gamma_k + f_kP_{s,k'}^{-1}\gamma_k \le 1 \\ C_3: 0 \le P_{s,k} \le P_{s,k}^{\max}, \quad k = 1,2,\cdots,2K \\ C_4: 0 \le P_R \le P_R^{\max} \end{cases} \tag{29}$$

V. SIMULATION RESULTS

In this section, we examine the SE and EE of the multi-pair two-way MM-AF-FDR system. The EE is defined as $EE = \frac{SE}{\sum_{k=1}^{2K} P_{s,k}+P_R}$. The performance of the HDR considered in [15] are also simulated for comparison. To make a fair comparison, the transmit powers in HDR are set two times of the FDR for all nodes transmit only half of the time in HDR. Without loss of generality, we assume $\sigma = \sigma_n = \phi_{k,i} = 1$, $\mathbf{D}_i = \mathbf{I}_{2K}$, and all users have the same Ricean K-factor K_i. Note that, under the above assumption, the asymptotic SE and EE of MRC/MRT and ZFR/ZFT are equal. For CSI estimation with uplink pilots, the length of pilot sequences is set to $\tau = 4K$, $P_T = -3.8db$, and T is set to 196.

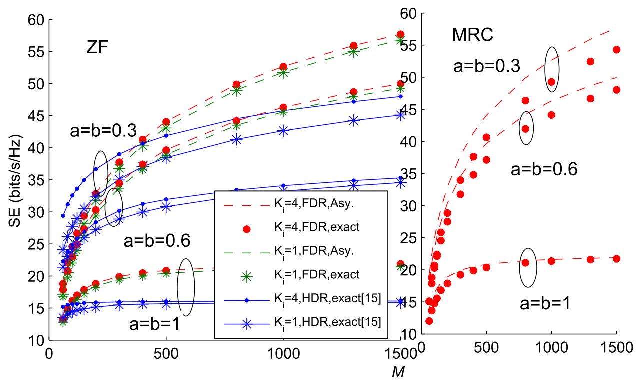


FIGURE 2. The SE of multi-pair FDR and HDR v.s. M, where $K = 5$, $E_{s,k} = 10db$, $E_R = 2KE_{s,k}$, $\sigma_{rr}^2 = 1$.

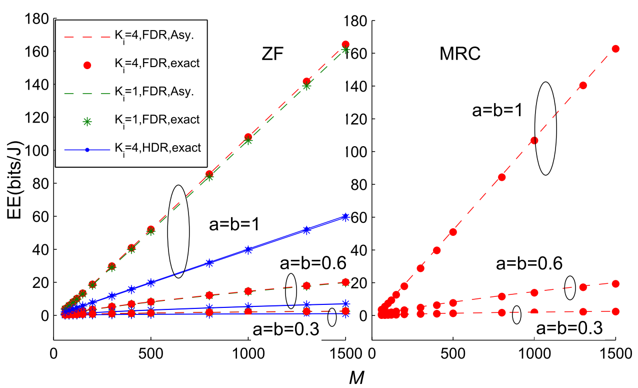


FIGURE 3. The EE of multi-pair FDR and HDR v.s. M, where $K = 5$, $E_{s,k} = 10db$, $E_R = 2KE_{s,k}$, $\sigma_{rr}^2 = 1$.

Fig. 2 and Fig. 3 show the SE and EE of multi-pair FDR and HDR v.s. M. Clearly, the asymptotic expression derived in *Theorem 1, 2* can predict the performance of multi-pair FDR precisely with the increasing of M. Besides, it is seen that the FDR outperforms HDR for the large M scenario. As a and b increasing, the required M for FDR outperforming HDR decreases. This is because that the effect of EI in FDR becomes smaller in this case. The gaps between the asymptotic expression and the corresponding simulated SE with MRC/MRT processing is larger than that with ZFR/ZFT processing, This is because ZFR/ZFT processing can completely eliminate the inter-pair interferences, but the inter-pair interferences can only be eliminated when $M \to \infty$ with MRC/MRT processing. Again, we can see that the FDR outperforms HDR in EE for the large M scenario. Moreover, the power scaling scheme which results in the best SE is not necessary to be optimal for the EE. Specifically, the power scaling scheme with $a = b = 0.3$ which is optimal for SE achieves the worst EE. Moreover, both the SE and EE depend on the Ricean K-factor.

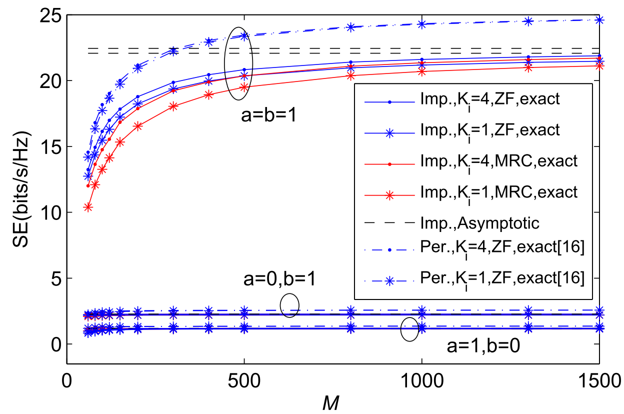


FIGURE 4. Comparing the SE of imperfect CSI with perfect CSI [16], where $K = 5$, $E_{s,k} = 10db$, $E_R = 2KE_{s,k}$, $\sigma_{rr}^2 = 1$.

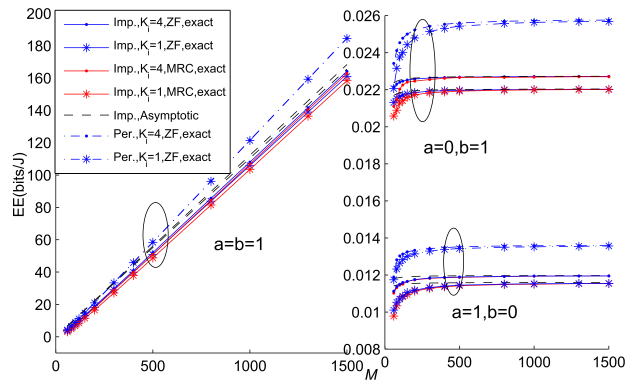


FIGURE 5. Comparing the EE of imperfect CSI with perfect CSI [16], where $K = 5$, $E_{s,k} = 10db$, $E_R = 2KE_{s,k}$, $\sigma_{rr}^2 = 1$.

Fig. 4 and Fig. 5 describe the simulated SE, EE and the presented analytical asymptotic results when $M \to \infty$ in *Corollary 1, 2* v.s. M. Clearly, as the number of antennas increases, the SE increase and saturate to the same value. Furthermore, the ZFR/ZFT beamforming can derive greater SE and EE than the MRC/MRT beamforming in the regime of finite M, since the system is interference-limited under the simulation condition, and ZFR/ZFT can eliminate the inter-pair interference. In addition, as M increases, the EE increases linearly with M in Case $a = 0, b = 1$. For other two cases, as M increases, the EE increase and saturate to a constant value. Moreover, as the number of antennas increases, the SE and EE increase and saturate to a constant value which is independent of the Ricean K-factor in the case of perfect CSI [16]. However, for imperfect CSI, the asymptotic result of SE and EE depend on both the scaling parameter a, b, the number of user pairs K and the Ricean K-factor.

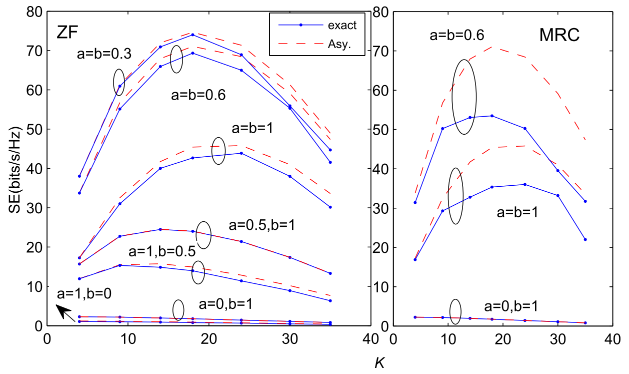


FIGURE 6. The SE of multi-pair FDR v.s. K, where $M = 500$, $E_{s,k} = 10db$, $E_R = 2KE_{s,k}$, $\sigma_{rr}^2 = 1$, $\mathrm{K}_i = 4$.

In Fig. 6, the simulated SE and the asymptotic SE using the expression derived in *Theorem 1, 2* v.s. K are plotted under MRC/MRT and ZFR/ZFT beamforming scheme in the case of imperfect CSI. It is clearly shown that, when $a = 0, b = 1$ and $a = 1, b = 0$, the SE tends to zero as increasing K. For other cases, the gaps between the asymptotic results and the corresponding simulated SE increase with the increasing of K for the additional inter-pair and inter-user interferences. In addition, we also see that there is an optimal K^* for many cases. This is because when K is small, the multiplexing gain is greater than the effect of interference, but as K gets larger, the multiplexing gain can not compensate the effect of interference. So with the increasing of K, the SE first grows then decreases. Besides, when M is finite, the gaps between the asymptotic expression and the corresponding simulated SE with MRC/MRT processing is larger than that

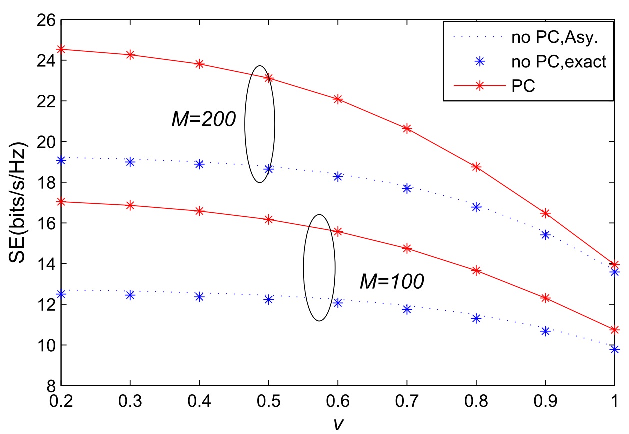


FIGURE 7. The SE of multi-pair FDR v.s. v, where $K = 5$, $E_{s,k} = 10db$, $E_R = 2KE_{s,k}$, $\mathrm{K}_i = 4$, $P_{s,k}^{max} = E_{s,k}/M^v$, $P_R^{max} = 2KP_{s,k}^{max}$, $\sigma_{rr}^2 = 5$.

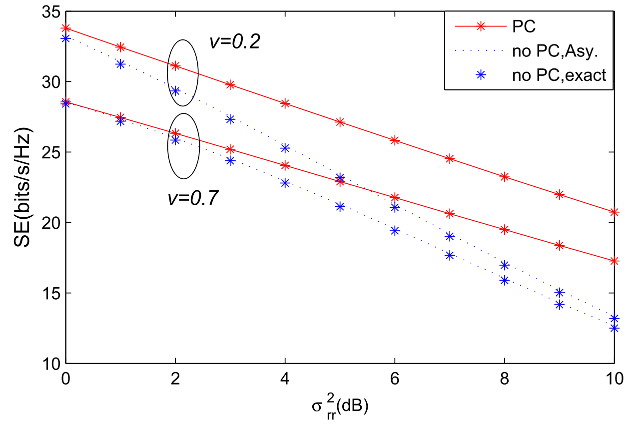


FIGURE 8. The SE of multi-pair FDR v.s. σ_{rr}^2, where $K = 5$, $E_{s,k} = 10db$, $E_R = 2KE_{s,k}$, $\mathrm{K}_i = 4$, $P_{s,k}^{max} = E_{s,k}/M^v$, $P_R^{max} = 2KP_{s,k}^{max}$, $M = 200$.

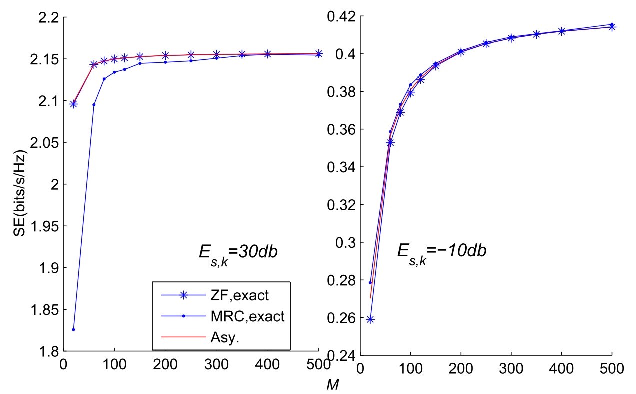


FIGURE 9. The SE of multi-pair FDR v.s. M, where $K = 2$, $E_R = 2KE_{s,k}$, $\sigma_{rr}^2 = 1$, $a = 0$, $b = 1$.

with ZFR/ZFT processing. This is because ZFR/ZFT processing can completely eliminate the inter-pair interferences, but the inter-pair interferences can only be eliminated when $M \to \infty$ with MRC/MRT processing. Fig. 7 and Fig. 8 show that the achievable SE is improved significantly by the proposed PC scheme. However, when v is large or the power of relay EI is small, the advantage of the PC scheme is mitigated. From Fig. 9, we can find ZFR/ZFT outperforms MRC/MRT when the scenario is interference-limited, i.e., SNR = 30 dB; MRC/MRT outperforms ZFR/ZFT when the scenario is noise-limited, i.e., SNR = −10 dB.

VI. CONCLUSION

In this paper, we investigate a multi-pair two-way massive MIMO amplify-and-forward full-duplex relay system over Ricean fading channels in the case of imperfect CSI. Our analysis incorporated both MRC/MRT and ZFR/ZFT. The low-complexity processing matrices based on the two schemes at the relay are presented and the asymptotic expressions in M of the end-to-end SINR are derived analytically. Theoretical analyses and simulation results show that, when $M \to \infty$, at the general power scaling schemes ($P_{s,k} = E_{s,k}/M^a$, $P_R = E_R/M^b$, $0 < a, b \leq 1$), the sufficient condition that the effect of relay EI is eliminated is $b > 0$, and the effect of the user self-loop and inter-user interferences

$$a_{zf}^2 = \frac{P_R}{Tr[\Lambda\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}] + \sigma Tr[(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}] + \frac{P_R}{M}\Delta\mathbf{I} + \Delta\mathbf{I}_1} \tag{30}$$

$$a_{zf}^2 \xrightarrow{a.s.} \frac{P_R}{\frac{Tr[\Lambda\mathbf{P}(\mathbf{Q}_2)^{-1}\mathbf{P}]}{M} + (\sigma + \sigma_{rr}^2 P_R + \sum_{j=1}^{2K} P_{s,j}\varepsilon_{g,j}^2)\frac{Tr[(\mathbf{Q}_1)^{-1}\mathbf{P}(\mathbf{Q}_2)^{-1}\mathbf{P}]}{M^2}} \tag{31}$$

is eliminated is $a > 0$. Besides, at some power scaling schemes, as the number of antennas increases, the SE increase and saturate to a constant value which is independent of the Ricean K-factor in the case of perfect CSI. However, for imperfect CSI, the asymptotic result of SE depends on both the scaling parameter a, b, the number of user pairs K and the Ricean K-factor. Moreover, it is shown that the multi-pair FDR outperforms the multi-pair HDR for the large M scenario. And the achievable SE is improved significantly by the proposed PC scheme.

APPENDIX

A. PROOF OF (8) IN LEMMA 2

Using (1), (3), (7), the property $Tr(\mathbf{AB}) = Tr(\mathbf{BA})$ and $P_R = Tr\left(\mathbb{E}\left\{\mathbf{x}_R\mathbf{x}_R^H\right\}\right)$, we have (30), as shown at the top of this page, where $\Delta\mathbf{I} = Tr[(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\mathbf{G}_{RR}^H\hat{\mathbf{G}}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{P} \times (\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}]$ and $\Delta\mathbf{I}_1 = Tr[(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\Delta\mathbf{G}\Lambda\Delta\mathbf{G}^H\hat{\mathbf{G}} \times (\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}]$. Substituting (5), (6) into (30), we have (31), as shown at the top of this page.

Since

$$Tr[\Lambda\mathbf{P}(\mathbf{Q}_2)^{-1}\mathbf{P}] = \sum_{i=1}^{2K} P_{s,i}\eta_{f,i'}^{-1} \tag{32}$$

$$Tr[(\mathbf{Q}_1)^{-1}\mathbf{P}(\mathbf{Q}_2)^{-1}\mathbf{P}] = \sum_{i=1}^{2K} \eta_{f,i}^{-1}\eta_{g,i'}^{-1} \tag{33}$$

then (8) can be easily obtained.

B. PROOF OF THEOREM 1

1) Compute $\mathbb{E}[\left|\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{n}_R\right|^2]$: According to (5), (6) and $\mathbb{E}\left\{\mathbf{n}_R\mathbf{n}_R^H\right\} = \sigma\mathbf{I}_M$, we have

$$\begin{aligned}\mathbb{E}[\left|\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{n}_R\right|^2] &= \sigma\mathbb{E}[\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\hat{\mathbf{G}} \\ &\quad\times(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{1}_{k'}^T] = \sigma\mathbb{E}[\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{1}_{k'}^T] \\ &\xrightarrow[M\to\infty]{a.s.} \frac{\sigma}{M}\mathbf{1}_{k'}(\mathbf{Q}_1)^{-1}\mathbf{1}_{k'}^T = \frac{\sigma}{M}\eta_{g,k'}^{-1}\end{aligned} \tag{34}$$

2) Compute $\mathbb{E}[\left|y_k^{ee}\right|^2]$:

Since $\hat{\mathbf{G}}$ and $\Delta\mathbf{G}$ are independent, we obtain

$$\begin{aligned}&\mathbb{E}[\left|\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\Delta\mathbf{G}\mathbf{x}\right|^2] \\ &= \mathbb{E}[\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\Delta\mathbf{G}\Lambda\Delta\mathbf{G}^H\hat{\mathbf{G}}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{1}_{k'}^T] \\ &= \sum_{i=1}^{2K} P_{s,i}\varepsilon_{g,i}^2[\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{1}_{k'}^T] \\ &\xrightarrow[M\to\infty]{a.s.} \frac{\sum_{i=1}^{2K} P_{s,i}\varepsilon_{g,i}^2}{M}\mathbf{1}_{k'}(\mathbf{Q}_1)^{-1}\mathbf{1}_{k'}^T = \frac{\eta_{g,k'}^{-1}}{M}\sum_{i=1}^{2K} P_{s,i}\varepsilon_{g,i}^2\end{aligned} \tag{35}$$

According to (5), (6), (32) and the property $\Delta\mathbf{f}$ is independent of $\hat{\mathbf{F}}$, we derive

$$\begin{aligned}&\mathbb{E}[\left|\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}\mathbf{x}\right|^2] \\ &= \mathbb{E}[\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}\Lambda\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\hat{\mathbf{F}}^T\Delta\mathbf{f}_k^*] \\ &\xrightarrow[M\to\infty]{a.s.} \frac{\varepsilon_{f,k}^2}{M}Tr[\Lambda\mathbf{P}(\mathbf{Q}_2)^{-1}\mathbf{P}] = \frac{\varepsilon_{f,k}^2}{M}\sum_{i=1}^{2K} P_{s,i}\eta_{f,i'}^{-1}\end{aligned} \tag{36}$$

According to (5), (6), (33) and the property $\Delta\mathbf{f}$, $\hat{\mathbf{F}}$, $\hat{\mathbf{G}}$ and $\Delta\mathbf{G}$ are independent, we obtain

$$\begin{aligned}&\mathbb{E}[\left|\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\Delta\mathbf{G}\mathbf{x}\right|^2] \\ &= \mathbb{E}[\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\Delta\mathbf{G}\Lambda \\ &\quad\times\Delta\mathbf{G}^H\hat{\mathbf{G}}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\hat{\mathbf{F}}^T\Delta\mathbf{f}_k^*] \\ &\xrightarrow[M\to\infty]{a.s.} \frac{\varepsilon_{f,k}^2}{M^2}\sum_{j=1}^{2K} P_{s,j}\varepsilon_{g,j}^2 Tr[(\mathbf{Q}_1)^{-1}\mathbf{P}(\mathbf{Q}_2)^{-1}\mathbf{P}] \\ &= \frac{\varepsilon_{f,k}^2}{M^2}\sum_{j=1}^{2K} P_{s,j}\varepsilon_{g,j}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}\end{aligned} \tag{37}$$

$$\begin{aligned}&\mathbb{E}[\left|\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{n}_R\right|^2] \\ &= \sigma\mathbb{E}[\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\hat{\mathbf{G}}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{P} \\ &\quad\times(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\hat{\mathbf{F}}^T\Delta\mathbf{f}_k^*] \\ &\xrightarrow[M\to\infty]{a.s.} \frac{\sigma}{M^2}\varepsilon_{f,k}^2 Tr[(\mathbf{Q}_1)^{-1}\mathbf{P}(\mathbf{Q}_2)^{-1}\mathbf{P}] \\ &= \frac{\sigma}{M^2}\varepsilon_{f,k}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}\end{aligned} \tag{38}$$

3) Compute $\mathbb{E}[\left|y_k^{ei}\right|^2]$:

According to (5), (6) and the property $\hat{\mathbf{G}}$ is independent of $\mathbf{G}_{RR}$, we derive

$$\begin{aligned}&\mathbb{E}[\left|\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\tilde{\mathbf{x}}_R\right|^2] \\ &= \frac{P_R}{M}\mathbb{E}[\mathbf{1}_{k'}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\mathbf{G}_{RR}^H\hat{\mathbf{G}}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{1}_{k'}^T] \\ &\xrightarrow[M\to\infty]{a.s.} \frac{P_R\sigma_{rr}^2}{M}\mathbf{1}_{k'}(\mathbf{Q}_1)^{-1}\mathbf{1}_{k'}^T = \frac{P_R}{M}\sigma_{rr}^2\eta_{g,k'}^{-1}\end{aligned} \tag{39}$$

$$a_{mrc}^2 = \frac{P_R}{Tr[(\hat{\mathbf{G}}^H\hat{\mathbf{G}})\Lambda(\hat{\mathbf{G}}^H\hat{\mathbf{G}})\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)\mathbf{P} + \frac{P_R}{M}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\mathbf{G}_{RR}^H\hat{\mathbf{G}}\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)\mathbf{P} + \sigma(\hat{\mathbf{G}}^H\hat{\mathbf{G}})\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)\mathbf{P} + \Delta\mathbf{I}]} \tag{41}$$

$$a_{mrc}^2 \xrightarrow[M\to\infty]{a.s.} \frac{P_R}{P_s M^3 Tr[\mathbf{Q}_1\Lambda\mathbf{Q}_1\mathbf{P}\mathbf{Q}_2\mathbf{P}] + (P_R\sigma_{rr}^2 + \sigma + \sum_{j=1}^{2K} P_{s,j}\varepsilon_{g,j}^2)M^2 Tr[\mathbf{Q}_1\mathbf{P}\mathbf{Q}_2\mathbf{P}]} \tag{42}$$

According to (5), (6), (33) and the property $\Delta\mathbf{f}$, $\hat{\mathbf{F}}$, $\hat{\mathbf{G}}$ and $\mathbf{G}_{RR}$ are independent, we obtain

$$\begin{aligned}
&\mathbb{E}[\left|\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\tilde{\mathbf{x}}_R\right|^2] \\
&= \frac{P_R}{M}\mathbb{E}[\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\mathbf{G}_{RR}^H \\
&\quad\times\hat{\mathbf{G}}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})^{-1}\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)^{-1}\hat{\mathbf{F}}^T\Delta\mathbf{f}_k^*] \\
&\xrightarrow[M\to\infty]{a.s.} \frac{P_R}{M^2}\varepsilon_{f,k}^2\sigma_{rr}^2 Tr[(\mathbf{Q}_1)^{-1}\mathbf{P}(\mathbf{Q}_2)^{-1}\mathbf{P}] \\
&= \frac{P_R}{M^2}\varepsilon_{f,k}^2\sigma_{rr}^2\sum_{i=1}^{2K}\eta_{f,i}^{-1}\eta_{g,i'}^{-1}
\end{aligned} \tag{40}$$

By using (34)-(40), we obtain (11).

C. PROOF OF (17) IN LEMMA 3

Using (1), (3), (16), the property $Tr(\mathbf{AB}) = Tr(\mathbf{BA})$ and $P_R = Tr\left(\mathbb{E}\left\{\mathbf{x}_R\mathbf{x}_R^H\right\}\right)$, we have (41) at the top of this page, where $\Delta\mathbf{I} = \hat{\mathbf{G}}^H\Delta\mathbf{G}\Lambda\Delta\mathbf{G}^H\hat{\mathbf{G}}\mathbf{P}(\hat{\mathbf{F}}^T\hat{\mathbf{F}}^*)\mathbf{P}$. Substituting (5), (6) into (41), we have (42), shown at the top of this page.

Since

$$Tr[\mathbf{Q}_1\Lambda\mathbf{Q}_1\mathbf{P}\mathbf{Q}_2\mathbf{P}] = \sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}^2 P_{s,i'} \tag{43}$$

$$Tr[\mathbf{Q}_1\mathbf{P}\mathbf{Q}_2\mathbf{P}] = \sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'} \tag{44}$$

then (17) can be easily obtained.

D. PROOF OF THEOREM 2:

1) Compute $\mathbb{E}[\left|\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\hat{\mathbf{g}}_i\right|^2]$ and $\mathbb{E}[\left|\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{n}_R\right|^2]$:

According to (5), (6) and $\mathbb{E}\left\{\mathbf{n}_R\mathbf{n}_R^H\right\} = \sigma\mathbf{I}_M$, we have

$$\begin{aligned}
&\mathbb{E}[\left|\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\hat{\mathbf{g}}_i\right|^2] = \mathbb{E}[\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\hat{\mathbf{g}}_i\hat{\mathbf{g}}_i^H\hat{\mathbf{G}}\mathbf{P}\hat{\mathbf{F}}^T\hat{\mathbf{f}}_k^*] \\
&\xrightarrow[M\to\infty]{a.s.} M^4\eta_{f,k}\mathbf{1}_{k'}\eta_{g,i}\mathbf{1}_i^T\eta_{g,i}\mathbf{1}_{i'}\eta_{f,k}\mathbf{1}_k^T \\
&= \eta_{f,k}^2\eta_{g,k'}^2 M^4\delta_{k',i}
\end{aligned} \tag{45}$$

$$\begin{aligned}
&\mathbb{E}[\left|\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{n}_R\right|^2] = \mathbb{E}[\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{n}_R\mathbf{n}_R^H\hat{\mathbf{G}}\mathbf{P}\hat{\mathbf{F}}^T\hat{\mathbf{f}}_k^*] \\
&\xrightarrow[M\to\infty]{a.s.} M^3\sigma\eta_{f,k}\mathbf{1}_{k'}\mathbf{Q}_1\eta_{f,k}\mathbf{1}_{k'}^T = M^3\sigma\eta_{f,k}^2\eta_{g,k'}
\end{aligned} \tag{46}$$

2) Compute $\mathbb{E}[\left|y_k^{ee}\right|^2]$:

According to (5), (6), (43), (44) and the property $\Delta\mathbf{f}$, $\hat{\mathbf{F}}$, $\hat{\mathbf{G}}$ and $\Delta\mathbf{G}$ are independent, we obtain the following results:

$$\begin{aligned}
&\mathbb{E}[\left|\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\Delta\mathbf{G}\mathbf{x}\right|^2] = \mathbb{E}[\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\Delta\mathbf{G}\Lambda\Delta\mathbf{G}^H\hat{\mathbf{G}}\mathbf{P}\hat{\mathbf{F}}^T\hat{\mathbf{f}}_k^*] \\
&\xrightarrow[M\to\infty]{a.s.} M^3\sum_{i=1}^{2K}P_{s,i}\varepsilon_{g,i}^2\eta_{f,k}\mathbf{1}_{k'}\mathbf{Q}_1\eta_{f,k}\mathbf{1}_{k'}^T \\
&= M^3\eta_{f,k}^2\eta_{g,k'}\sum_{i=1}^{2K}P_{s,i}\varepsilon_{g,i}^2
\end{aligned} \tag{47}$$

$$\begin{aligned}
&\mathbb{E}[\left|\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\hat{\mathbf{G}}\mathbf{x}\right|^2] \\
&= \mathbb{E}[\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}(\hat{\mathbf{G}}^H\hat{\mathbf{G}})\Lambda(\hat{\mathbf{G}}^H\hat{\mathbf{G}})\mathbf{P}\hat{\mathbf{F}}^T\Delta\mathbf{f}_k^*] \\
&\xrightarrow[M\to\infty]{a.s.} M^3\varepsilon_{f,k}^2 Tr[\mathbf{Q}_1\Lambda\mathbf{Q}_1\mathbf{P}\mathbf{Q}_2\mathbf{P}] \\
&= M^3\varepsilon_{f,k}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}^2 P_{s,i'}
\end{aligned} \tag{48}$$

$$\begin{aligned}
&\mathbb{E}[\left|\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\Delta\mathbf{G}\mathbf{x}\right|^2] = \mathbb{E}[\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\Delta\mathbf{G}\Lambda\Delta\mathbf{G}^H \\
&\quad\times\hat{\mathbf{G}}\mathbf{P}\hat{\mathbf{F}}^T\Delta\mathbf{f}_k^*] \xrightarrow[M\to\infty]{a.s.} M^2\varepsilon_{f,k}^2\sum_{j=1}^{2K}P_{s,j}\varepsilon_{g,j}^2 Tr[\mathbf{Q}_1\mathbf{P}\mathbf{Q}_2\mathbf{P}] \\
&= M^2\varepsilon_{f,k}^2\sum_{j=1}^{2K}P_{s,j}\varepsilon_{g,j}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}
\end{aligned} \tag{49}$$

$$\begin{aligned}
&\mathbb{E}[\left|\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{n}_R\right|^2] = \mathbb{E}[\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{n}_R\mathbf{n}_R^H\hat{\mathbf{G}}\mathbf{P}\hat{\mathbf{F}}^T\Delta\mathbf{f}_k^*] \\
&\xrightarrow[M\to\infty]{a.s.} M^2\sigma\varepsilon_{f,k}^2 Tr[\mathbf{Q}_1\mathbf{P}\mathbf{Q}_2\mathbf{P}] = M^2\sigma\varepsilon_{f,k}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'}
\end{aligned} \tag{50}$$

3) Compute $\mathbb{E}[\left|y_k^{ei}\right|^2]$:

According to (5), (6) and the property $\hat{\mathbf{F}}$, $\hat{\mathbf{G}}$ and $\mathbf{G}_{RR}$ are independent, we derive

$$\begin{aligned}
&\mathbb{E}[\left|\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\tilde{\mathbf{x}}_R\right|^2] \\
&= \frac{P_R}{M}\mathbb{E}[\hat{\mathbf{f}}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\mathbf{G}_{RR}^H\hat{\mathbf{G}}\mathbf{P}\hat{\mathbf{F}}^T\hat{\mathbf{f}}_k^*] \\
&\xrightarrow[M\to\infty]{a.s.} P_R M^3\sigma_{rr}^2\eta_{f,k}\mathbf{1}_{k'}\mathbf{Q}_1\eta_{f,k}\mathbf{1}_{k'}^T \\
&= P_R M^3\sigma_{rr}^2\eta_{f,k}^2\eta_{g,k'}
\end{aligned} \tag{51}$$

According to (5), (6), (44) and the property $\Delta\mathbf{f}$, $\hat{\mathbf{F}}$, $\hat{\mathbf{G}}$ and $\mathbf{G}_{RR}$ are independent, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\widetilde{\mathbf{x}}_R\Big|^2\Big] \\
&= \frac{P_R}{M}\mathbb{E}[\Delta\mathbf{f}_k^T\hat{\mathbf{F}}^*\mathbf{P}\hat{\mathbf{G}}^H\mathbf{G}_{RR}\mathbf{G}_{RR}^H\hat{\mathbf{G}}\mathbf{P}\hat{\mathbf{F}}^T\hat{\mathbf{f}}_k^*] \\
&\xrightarrow[M\to\infty]{a.s.} P_R M^2\varepsilon_{f,k}^2\sigma_{rr}^2 Tr[\mathbf{Q}_1\mathbf{P}\mathbf{Q}_2\mathbf{P}] \\
&= P_R M^2\varepsilon_{f,k}^2\sigma_{rr}^2\sum_{i=1}^{2K}\eta_{f,i}\eta_{g,i'} \qquad (52)
\end{aligned}$$

By using (45)-(52), we obtain (20).

REFERENCES

[1] 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.

[2] J. Yang, P. Fan, T. Q. Duong, and X. Lei, "Exact performance of two-way AF relaying in Nakagami-m fading environment," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 980–987, Mar. 2011.

[3] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions and performance comparison of half duplex bi-directional relaying protocols," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6405–6418, Oct. 2011.

[4] L. J. Rodriguez, N. H. Tran, and T. Le-Ngoc, "Optimal power allocation and capacity of full-duplex AF relaying under residual self-interference," *IEEE Wireless Commun. Lett.*, vol. 3, no. 2, pp. 233–236, Apr. 2014.

[5] L. J. Rodríguez, N. H. Tran, and T. Le-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.

[6] D. Choi and J. H. Lee, "Outage probability of two-way full-duplex relaying with imperfect channel state information," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 933–936, Jun. 2014.

[7] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-Aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038–4050, Aug. 2013.

[8] X. Xia, K. Xu, D. Zhang, and Y. Xu, "Low-complexity transceiver design and antenna subset selection for cooperative half- and full-duplex relaying systems," in *Proc. IEEE Globecom*, Dec. 2014, pp. 3314–3319.

[9] 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.

[10] H. A. Suraweera, I. Krikidis, and C. Yuen, "Antenna selection in the full-duplex multi-antenna relay channel," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Budapest, Hungary, Jun. 2013, pp. 4823–4828.

[11] H. Q. Ngo, H. A. Suraweerat, M. Matthaiou, and E. G. Larsson, "Multipair massive MIMO full-duplex relaying with MRC/MRT processing," in *Proc. IEEE ICC*, Jun. 2014, pp. 4807–4813.

[12] X. Xia, W. Xie, D. Zhang, K. Xu, and Y. Xu, "Multi-pair full-duplex amplify-and-forward relaying with very large antenna arrays," in *Proc. IEEE WCNC*, Mar. 2015, pp. 304–309.

[13] 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, Sep. 2014.

[14] 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, Jul. 2016, doi: 10.1109/TVT.2015.2467211.

[15] H. Cui, L. Song, and B. Jiao, "Multi-pair two-way amplify-and-forward relaying with very large number of relay antennas," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2636–2645, May 2014.

[16] Z. Zhang, Z. Chen, M. Shen, B. Xia, and L. Luo, "Achievable rate analysis for multi-pair two-way massive MIMO full-duplex relay systems," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2015, pp. 2598–2602.

[17] M. R. Bhatnagar and M. K. Arti, "Performance analysis of AF based hybrid satellite-terrestrial cooperative network over generalized fading channels," *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1912–1915, Oct. 2013.

[18] J. D. Sanchez-Heredia, J. F. Valenzuela-Valdes, A. M. Martinez-Gonzalez, and D. A. Sanchez-Hernandez, "Emulation of MIMO Rician-fading environments with mode-stirred reverberation chambers," *IEEE Trans. Antennas Propag.*, vol. 59, no. 2, pp. 654–660, Feb. 2011.

[19] N. Ravindran, N. Jindal, and H. C. Huang, "Beamforming with finite rate feedback for LOS MIMO downlink channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Nov. 2007, pp. 4200–4204.

[20] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 815–830, Oct. 2014.

[21] Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.

[22] P. C. Weeraddana, M. Codreanu, M. Latva-Aho, and A. Ephremides, "Resource allocation for cross-layer utility maximization in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2790–2809, Jul. 2011.

[23] M. K. Arti, "A simple scheme of channel estimation in large MIMO systems," in *Proc. 83rd Veh. Technol. Conf. (VTC Spring)*, Nanjing, China, May 2016, pp. 15–18.

XIAOLI SUN was born in 1992. She received the B.E. degree from the PLA University of Science and Technology in 2014, where she is currently pursuing the B.S. degree with the Institution of Communications Engineering. Her research interests include MIMO techniques, relaying network, full-duplex communication, and network coding.

KUI XU (M'13) was born in 1982. He received the B.S. degree in wireless communications and the Ph.D. degree in software defined radio from the PLA University of Science and Technology (PLAUST), Nanjing, China, in 2004 and 2009, respectively. He is currently an Associate Professor with the College of Communications Engineering, PLAUST. Since 2013, he has been a Post-Doctoral Fellow with PLAUST. His research interests include broadband wireless communications, signal processing for communications, network coding, and wireless communication networks. He has authored about 100 papers in refereed journals and conference proceedings and holds five patents in China.

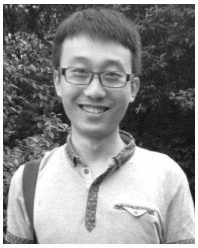
Prof. Xu served on the Technical Program Committee of the IEEE WCSP 2014, WCSP 2015, ICSPDM 2015, and SIRS 2015 TPC. He received the URSI Young Scientists Award in 2014 and the 2010 Ten Excellent Doctor Degree Dissertation Award of PLAUST. He serves as the Reviewer of the IEEE Transactions on Wireless Communication, the IEEE Transactions on Vehicle Technology, the IEEE Communications Letter, and the IEEE Signal Processing Letters.

WENFENG MA was born in 1974. He received the B.S. degree in microwave communication and the M.S. degree in communications and information systems from the Institute of Communications Engineering, Nanjing, China, in 1995 and 1998, respectively, and the Ph.D. degree in communications and information system from the PLA University of Science and Technology, Nanjing, in 2002. From 2004 to 2006, he was a Post-Doctoral Fellow with Shanghai Jiao Tong University, Shanghai, China.

He is currently an Associate Professor with the PLA University of Science and Technology. His research interests include wireless communication networks and broadband wireless communications.

YOUYUN XU (M'02–SM'11) was born in 1966. He received the Ph.D. degree in information and communication engineering from Shanghai Jiao Tong University (SJTU), China, in 1999.

He is currently a Professor with the Nanjing University of Posts and Telecommunications and the PLA university of Science and Technology. He is also a part-time Professor with the Institute of Wireless Communication Technology, SJTU. He has more than 20 years of professional experience in teaching and researching in communication theory and engineering. His research interests are focusing on new generation wireless mobile communication systems (IMT-advanced and related), advanced channel coding and modulation techniques, multiuser information theory and radio resource management, wireless sensor networks, and cognitive radio networks. He is a Senior Member of the Chinese Institute of Electronics.



XIAOCHEN XIA received the B.E. degree in electronic science and technology from Tianjin University in 2010, and the M.S. degree in communication and information system from the PLA University of Science and Technology in 2013, where he is currently pursuing the Ph.D. degree with the Institution of Communications Engineering. His research interests include relaying network, full-duplex communication, network coding, and MIMO techniques. He received the 2013 Excellent Master Degree Dissertation Award of Jiangsu Province, China.

DONGMEI ZHANG was born in 1972. She received the B.S. and M.S. degrees in communication engineering in 1993 and 2005, respectively, and the Ph.D. degree in communications and information system from the PLA University of Science and Technology, Nanjing, China, in 2013.

She is currently an Associate Professor with the Institution of Communications Engineering, PLA University of Science and Technology. Her research interests include new generation wireless mobile communication system, radio resource management, and network coding in wireless communication.