



**QUEEN'S
UNIVERSITY
BELFAST**

Multipair Two-Way Half-Duplex DF Relaying with Massive Arrays and Imperfect CSI

Kong, C., Zhong, C., Matthaiou, M., Björnson, E., & Zhang, Z. (2018). Multipair Two-Way Half-Duplex DF Relaying with Massive Arrays and Imperfect CSI. IEEE Transactions on Wireless Communications. DOI: 10.1109/TWC.2018.2808952

Published in:

IEEE Transactions on Wireless Communications

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2017 IEEE.

This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Multipair Two-Way Half-Duplex DF Relaying with Massive Arrays and Imperfect CSI

Chuli Kong, *Student Member, IEEE*, Caijun Zhong, *Senior Member, IEEE*, Michail Matthaiou, *Senior Member, IEEE*, Emil Björnson, *Senior Member, IEEE*, and Zhaoyang Zhang, *Member, IEEE*

Abstract—This paper considers a two-way half-duplex decode-and-forward relaying system where multiple pairs of single-antenna users exchange information via a multiple-antenna relay. Assuming that the channel knowledge is non-ideal and the relay employs maximum ratio processing, we derive a large-scale approximation of the sum spectral efficiency (SE) that is tight when the number of relay antennas, M , becomes very large. Furthermore, we study how the transmit power scales with M to maintain a desired SE. In particular, three special power-scaling cases are discussed and the corresponding asymptotic SE is deduced with clear insights. Our elegant power-scaling laws reveal a tradeoff between the transmit powers of the user/relay and pilot symbol. Finally, we formulate a power allocation problem in terms of maximizing the sum SE and obtain a local optimum by solving a sequence of geometric programming problems.

Index Terms—Decode-and-forward, geometric programming, massive MIMO, power-scaling law, two-way relaying.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) is one of the key technologies for the next-generation wireless communications due to its distinctive features and advantages over conventional MIMO systems. Some key benefits include: 1) very narrow beams and little inter-user interference due to the asymptotic channel orthogonality; 2) low computational complexity since linear signal processing is asymptotically optimal; 3) transmit power can be made extremely low, since it scales down inversely proportional to the number of antennas when perfect channel state information (CSI) is available. In addition to the use of massive MIMO in cellular networks,

Manuscript received August 14, 2017, revised December 19, 2017, accepted February 12, 2018. The work was supported in part by the National Science and Technology Major Project of China 2017ZX03001002–003, the National Natural Science Foundation of China under Grant 61671406, U1709219, 61725104, and the Zhejiang Provincial Natural Science Foundation of China (LR15F010001), and in part by the Open Research Fund of State Key Laboratory of Integrated Services Networks under Grant ISN19-05. The associate editor coordinating the review of this paper and approving it for publication was B. Clerckx. (*Corresponding author: Caijun Zhong.*)

C. Kong, C. Zhong and Z. Zhang are with the Institute of Information and Communication Engineering and also with the Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Zhejiang University, Hangzhou 310027, China, and also with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710126, China (e-mail: kcl_dut@163.com; caijunzhong@zju.edu.cn; sunrise.heaven@gmail.com).

M. Matthaiou is with the Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, Belfast, BT3 9DT, U.K. (email: m.matthaiou@qub.ac.uk).

E. Björnson is with the Department of Electrical Engineering (ISY), Linköping University, Linköping, SE-581 83, Sweden (email: emil.bjornson@liu.se).

the technique has been analyzed in other applications, such as cognitive radio, heterogeneous networks, energy harvesting units, and relaying systems, etc. [1]–[4].

The massive MIMO technique also finds an important application in the multipair relaying system, where multiple pairs of users simultaneously establish communication links with the aid of a shared relay with large antenna arrays. The resulting multipair massive MIMO relaying system has attracted substantial attention from both academia and industry because of its ability to enhance the coverage and service quality for cell edge users [5]–[8]. The initial works on multipair massive MIMO relaying mainly focus on the one-way relaying systems. For one-way multipair massive MIMO amplify-and-forward (AF) relaying systems, the work [9] studied the max-min user selection problem and the power allocation issue. Later in [10], taking into account the semi-blind gain control and relay oscillation, the authors proposed a low-complexity power control scheme. In contrast, for the decode-and-forward (DF) protocol, [5] compared the achievable spectral efficiency (SE) of two linear processing techniques, i.e., zero-forcing (ZF) and maximum-ratio (MR) in Rayleigh fading channels. Then, [11] extended the analysis to the Ricean fading case. The energy efficiency was optimized in [12]. Yet, such one-way mechanism incurs a 50% SE loss. To reduce this loss, two-way relaying becomes a potential solution [13]–[16], where the two communicating nodes execute bidirectional simultaneous data transmission.

Recently, there has been intensive research on two-way massive MIMO relaying. For example, the power-scaling laws of MR and/or ZF processing methods are characterized for half-duplex [17], [18] and full-duplex [19], [20] schemes, respectively. The work [21] considered the maximization of the energy efficiency of the system subject to maximum power and minimum SE constraints. Nevertheless, one major limitation of the above works is that perfect CSI is assumed. Since obtaining perfect CSI is very challenging in the context of massive MIMO, it is important to look into the realistic scenario with imperfect CSI. By employing the minimum mean-square-error (MMSE) estimation at the relay, [22], [23] studied the impact of pilot power on the SE, [24] maximizes the energy efficiency using the max-min approach, while [25] adopted the composite channel estimation method to reduce the pilot overhead by half and demonstrated that this approach outperforms the individual channel estimation scheme in [22], [23] when the coherence interval is smaller. However, all the aforementioned two-way massive MIMO relaying works focus on the AF protocol, and the DF case is largely overlooked.

Unlike AF relaying, DF relaying does not suffer from the problem of noise amplification. Thus, DF two-way relaying may achieve better performance than AF two-way relay, especially at low signal-to-noise ratios (SNRs) [26]. Besides this, DF two-way relaying has the flexibility of performing separate power allocation/precoding for relaying the communication on each direction. Therefore, it is of great interest to study the performance of two-way massive MIMO relaying systems adopting the DF protocol.

Motivated by this, in the current work, we consider a multipair two-way DF relaying system, taking into account channel estimation errors, and present a comprehensive analysis of the achievable SE and power-scaling law of MR processing. Specifically, the main contributions of this paper are summarized as follows:

- We propose a general multipair massive MIMO two-way relaying system employing the DF protocol, and present a large-scale approximation of the SE under the imperfect CSI when the number of relay antennas approaches infinity.
- We characterize new power-scaling laws, which generalize the results presented in [5], [17], [18]. It turns out that there exists a trade-off between the transmit powers of each user, pilot symbol and the relay; in other words, the same SE can be achieved with different combination of power-scaling parameters, which offers great flexibility in the design of practical systems.
- To improve the sum SE, we study the power allocation problem subject to a sum power constraint. Local optimum solutions are obtained by solving a sequence of geometric programming (GP) problems. Our numerical results suggest that the proposed power allocation strategy significantly improves the sum SE.

The remainder of the paper is organized as follows: Section II introduces the multipair two-way half-duplex DF relaying system model. Section III presents a large-scale approximation of the SE, with imperfect CSI, while Section IV studies the power-scaling laws of different system configurations. The power allocation problem is discussed in Section V. The numerical results are verified in Section VI. Finally, Section VII provides some concluding remarks.

Notation: We use bold upper case letters to denote matrices, bold lower case letters to denote vectors and lower case letters to denote scalars. Moreover, $(\cdot)^H$, $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^{-1}$ represent the conjugate transpose operator, the conjugate operator, the transpose operator, and the matrix inverse, respectively. Also, $\|\cdot\|$ is the Euclidian norm, $\|\cdot\|_F$ denotes the Frobenius norm, and $|\cdot|$ is the absolute value. In addition, $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes a circularly symmetric complex Gaussian random vector \mathbf{x} with covariance matrix $\mathbf{\Sigma}$, while \mathbf{I}_k is the identity matrix of size k . Finally, the statistical expectation operator is represented by $E\{\cdot\}$, the variance operator is $\text{Var}\{\cdot\}$, and the notation $\xrightarrow{a.s.}$ means almost sure convergence.

II. SYSTEM MODEL

Consider a multipair two-way relaying system, where N pairs of single-antenna users, denoted as $T_{A,i}$ and $T_{B,i}$, $i =$

$1, \dots, N$, exchange information with each other, under the assistance of a shared relay T_R equipped with M antennas, shown in Fig. 1. We assume that the direct links between $T_{A,i}$ and $T_{B,i}$ do not exist due to shadowing. Also, the relay operates in the half-duplex mode, i.e., it cannot transmit and receive simultaneously.

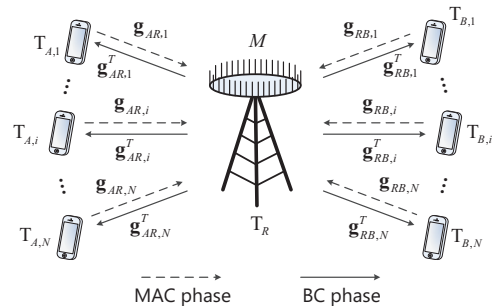


Fig. 1: Illustration of the multipair two-way relaying system.

It is assumed that the system works under a time division duplex (TDD) protocol and channel reciprocity holds. As such, the uplink and downlink channels between $T_{A,i}$ and T_R can be denoted as $\mathbf{g}_{AR,i} \sim \mathcal{CN}(\mathbf{0}, \beta_{AR,i} \mathbf{I}_M)$ and $\mathbf{g}_{AR,i}^T$, respectively. Similarly, the channels between $T_{B,i}$ and T_R are denoted as $\mathbf{g}_{RB,i} \sim \mathcal{CN}(\mathbf{0}, \beta_{RB,i} \mathbf{I}_M)$ and $\mathbf{g}_{RB,i}^T$, $i = 1, \dots, N$, respectively. This model is known as uncorrelated Rayleigh fading, and $\beta_{AR,i}$ and $\beta_{RB,i}$ model the large-scale effect, which are assumed to be constant over many coherence intervals and known a priori. For notational convenience, the channel vectors can be collected together in a matrix form as $\mathbf{G}_{AR} \triangleq [\mathbf{g}_{AR,1}, \dots, \mathbf{g}_{AR,N}] \in \mathbb{C}^{M \times N}$ and $\mathbf{G}_{RB} \triangleq [\mathbf{g}_{RB,1}, \dots, \mathbf{g}_{RB,N}] \in \mathbb{C}^{M \times N}$.

For the considered multipair two-way relaying system, the information transmission process consists of two separate phases. In the first phase, i.e., multiple-access channel (MAC) phase, the N user pairs $T_{A,i}$ and $T_{B,i}$ simultaneously transmit their respective signals to T_R . Thus, the received signal at T_R is given by

$$\mathbf{y}_r = \sum_{i=1}^N (\sqrt{p_{A,i}} \mathbf{g}_{AR,i} x_{A,i} + \sqrt{p_{B,i}} \mathbf{g}_{RB,i} x_{B,i}) + \mathbf{n}_r, \quad (1)$$

where $x_{A,i}$ and $x_{B,i}$ are Gaussian signals with zero mean and unit power transmitted by the i -th user pair, $p_{A,i}$ and $p_{B,i}$ are the average transmit power of $T_{A,i}$ and $T_{B,i}$, respectively, and \mathbf{n}_r is a vector of additive white Gaussian noise (AWGN) at T_R , whose elements are identically and independently distributed (i.i.d.) $\mathcal{CN}(0, 1)$. Note that to keep the notation clean and without loss of generality, we take the noise variance to be 1 in this paper. To reduce the complexity, single-user linear detection is considered at the relay. As such, the transformed signal after linear processing can be expressed as

$$\mathbf{r}^{\text{DF}} = \mathbf{W}^T \mathbf{y}_r, \quad (2)$$

where $\mathbf{W}^T \in \mathbb{C}^{2N \times M}$ is the linear receiver matrix, which will be specified shortly.

In the second phase, i.e., broadcasting (BC) phase, the relay first decodes the received information, and then re-encodes

and broadcasts it to the users. A linear precoding matrix $\mathbf{J} \in \mathbb{C}^{M \times 2N}$ is applied to the decoded signal \mathbf{x} . As such, the transmit signal of T_R is given by

$$\mathbf{y}_t^{\text{DF}} = \rho_{\text{DF}} \mathbf{J} \mathbf{x}, \quad (3)$$

where $\mathbf{x} = [\mathbf{x}_A^T, \mathbf{x}_B^T]^T$ with $\mathbf{x}_A = [x_{A,1}, \dots, x_{A,N}]^T$ and $\mathbf{x}_B = [x_{B,1}, \dots, x_{B,N}]^T$, and ρ_{DF} is the normalization coefficient, which is determined by the average power constraint at the relay, i.e., $\mathbb{E} \{ \|\mathbf{y}_t^{\text{DF}}\|^2 \} = p_r$. Hence, the signals received at $T_{X,i}$ (where $X \in \{A, B\}$) can be expressed as

$$z_{X,i}^{\text{DF}} = \mathbf{g}_{XR,i}^T \mathbf{y}_t^{\text{DF}} + n_{X,i}, \quad (4)$$

where $n_{X,i} \sim \mathcal{CN}(0, 1)$ represents the AWGN at $T_{X,i}$. Please note that, to simplify notation, we introduce $\mathbf{g}_{BR,i}$ which is defined as $\mathbf{g}_{BR,i} \triangleq \mathbf{g}_{RB,i}$ due to the channel reciprocity.

A. Channel Estimation

Since we consider a block fading system, the channels \mathbf{G}_{AR} and \mathbf{G}_{RB} are not known and need to be estimated at the relay in every coherence interval. The typical way of doing this in TDD systems is to transmit pilots [27]. To this end, during each coherence interval of length τ_c (in symbols), τ_p symbols are used for channel training. In this case, $T_{A,i}$ and $T_{B,i}$ simultaneously transmit mutually orthogonal pilot sequences to T_R . Thus, the received pilot matrix at T_R is

$$\mathbf{Y}_p = \sqrt{\tau_p p_p} \mathbf{G}_{AR} \Phi_A^T + \sqrt{\tau_p p_p} \mathbf{G}_{RB} \Phi_B^T + \mathbf{N}_p, \quad (5)$$

where p_p is the transmit power of each pilot symbol, \mathbf{N}_p is AWGN matrix including i.i.d. $\mathcal{CN}(0, 1)$ elements, while the i -th columns of $\Phi_A \in \mathbb{C}^{\tau_p \times N}$ and $\Phi_B \in \mathbb{C}^{\tau_p \times N}$ are the pilot sequences transmitted from $T_{A,i}$ and $T_{B,i}$, respectively. Since all pilot sequences are assumed to be mutually orthogonal, $\tau_p \geq 2N$ is required, and we have that $\Phi_A^T \Phi_A = \mathbf{I}_N$, $\Phi_B^T \Phi_B = \mathbf{I}_N$, and $\Phi_A^T \Phi_B = \mathbf{0}_N$.

As in [5], [28], we assume that T_R uses the MMSE estimator to estimate \mathbf{G}_{AR} and \mathbf{G}_{RB} . As such, we have

$$\mathbf{g}_{AR,i} = \hat{\mathbf{g}}_{AR,i} + \mathbf{e}_{AR,i}, \quad (6)$$

$$\mathbf{g}_{RB,i} = \hat{\mathbf{g}}_{RB,i} + \mathbf{e}_{RB,i}, \quad (7)$$

where $\hat{\mathbf{g}}_{AR,i}$, $\hat{\mathbf{g}}_{RB,i}$, $\mathbf{e}_{AR,i}$, and $\mathbf{e}_{RB,i}$ are the i -th columns of the estimated matrices $\hat{\mathbf{G}}_{AR}$, $\hat{\mathbf{G}}_{RB}$, and the estimation error matrices \mathbf{E}_{AR} and \mathbf{E}_{RB} , respectively, which are mutually independent. The elements of $\hat{\mathbf{g}}_{AR,i}$, $\mathbf{e}_{AR,i}$ are Gaussian random variables with zero mean, variance $\sigma_{AR,i}^2$ and $\tilde{\sigma}_{AR,i}^2$, respectively, where $\sigma_{AR,i}^2 \triangleq \frac{\tau_p p_p \beta_{AR,i}^2}{1 + \tau_p p_p \beta_{AR,i}}$ and $\tilde{\sigma}_{AR,i}^2 \triangleq \frac{\beta_{AR,i}}{1 + \tau_p p_p \beta_{AR,i}}$. Similarly, the elements of $\hat{\mathbf{g}}_{RB,i}$, and $\mathbf{e}_{RB,i}$ are complex Gaussian random variables with zero mean, variance $\sigma_{RB,i}^2$ and $\tilde{\sigma}_{RB,i}^2$, respectively, where $\sigma_{RB,i}^2 \triangleq \frac{\tau_p p_p \beta_{RB,i}^2}{1 + \tau_p p_p \beta_{RB,i}}$ and $\tilde{\sigma}_{RB,i}^2 \triangleq \frac{\beta_{RB,i}}{1 + \tau_p p_p \beta_{RB,i}}$.

B. Linear Processing Matrices

To keep the complexity costs at the relay to reasonable levels, a simple linear processing scheme is used at the relay.

With the MR method¹, the processing matrix $\mathbf{W}^T \in \mathbb{C}^{2N \times M}$ and $\mathbf{J} \in \mathbb{C}^{M \times 2N}$ are given by

$$\mathbf{W}^T = \left[\hat{\mathbf{G}}_{AR}, \hat{\mathbf{G}}_{RB} \right]^H, \quad (8)$$

$$\mathbf{J} = \left[\hat{\mathbf{G}}_{RB}, \hat{\mathbf{G}}_{AR} \right]^*, \quad (9)$$

respectively, while ρ_{DF} is given by

$$\rho_{\text{DF}} = \sqrt{\frac{p_r}{\mathbb{E} \{ \|\mathbf{J}\|_{\text{F}}^2 \}}} = \sqrt{\frac{p_r}{M \sum_{n=1}^N (\sigma_{AR,n}^2 + \sigma_{RB,n}^2)}}. \quad (10)$$

III. SPECTRAL EFFICIENCY

In this section, we investigate the SE (in bit/s/Hz) of the two-way half-duplex DF relaying system. In particular, a large-scale approximation of the SE is deduced when $M \rightarrow \infty$.

In the MAC phase, a linear processing matrix \mathbf{W}^T is applied to the received signals prior to signal detection, hence, the post-processing signals at the relay are given by

$$\mathbf{r}^{\text{DF}} = \quad (11)$$

$$\begin{bmatrix} \hat{\mathbf{G}}_{AR}^H \left(\sum_{i=1}^N (\sqrt{p_{A,i}} \mathbf{g}_{AR,i} x_{A,i} + \sqrt{p_{B,i}} \mathbf{g}_{RB,i} x_{B,i}) + \mathbf{n}_R \right) \\ \hat{\mathbf{G}}_{RB}^H \left(\sum_{i=1}^N (\sqrt{p_{A,i}} \mathbf{g}_{AR,i} x_{A,i} + \sqrt{p_{B,i}} \mathbf{g}_{RB,i} x_{B,i}) + \mathbf{n}_R \right) \end{bmatrix},$$

where the top N elements of \mathbf{r}^{DF} stand for the signals from $T_{A,i}$ ($i = 1, \dots, N$), while the bottom N elements of \mathbf{r}^{DF} represent the signals from $T_{B,i}$ ($i = 1, \dots, N$). Without loss of generality, we focus only on the i -th pair of users, i.e., $T_{A,i}$ and $T_{B,i}$, which is given by (12). Since the relay has imperfect CSI, it treats the channel estimates as the true channels to decode the signals. To this end, using a standard lower capacity bound based on the worst-case uncorrelated additive noise [30] yields the sum achievable SE of the i -th user pair in the MAC phase:

$$R_{1,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \times \quad (13)$$

$$\mathbb{E} \left\{ \log_2 \left(1 + \frac{A_i^{\text{DF}} + B_i^{\text{DF}}}{\mathbb{E} \{ (C_i^{\text{DF}} + D_i^{\text{DF}} + E_i^{\text{DF}}) | \hat{\mathbf{G}}_{AR}, \hat{\mathbf{G}}_{RB} \}} \right) \right\},$$

where the inner and outer expectations are taken over the

¹Note that MR is a very attractive linear processing technique in the context of massive MIMO systems due to its low complexity. Most importantly, it can be implemented in a distributed manner [27], [29]. Leveraging on the properties of long Gaussian vectors, the extensions to the ZF and MMSE processing can be easily made by using the same technique as in the MR method. In addition, to simplify the notations, we have assumed that the MAC and BC phases take place in the same coherence interval. However, this is not necessary.

$$\begin{aligned}
r_i^{\text{DF}} = r_i^{\text{DF}} + r_{N+i}^{\text{DF}} &= \underbrace{\sqrt{p_{A,i}} (\hat{\mathbf{g}}_{AR,i}^H \hat{\mathbf{g}}_{AR,i} + \hat{\mathbf{g}}_{RB,i}^H \hat{\mathbf{g}}_{AR,i}) x_{A,i} + \sqrt{p_{B,i}} (\hat{\mathbf{g}}_{AR,i}^H \hat{\mathbf{g}}_{RB,i} + \hat{\mathbf{g}}_{RB,i}^H \hat{\mathbf{g}}_{RB,i}) x_{B,i}}_{\text{desired signal}} \\
&+ \underbrace{\sqrt{p_{A,i}} (\hat{\mathbf{g}}_{AR,i}^H \mathbf{e}_{AR,i} + \hat{\mathbf{g}}_{RB,i}^H \mathbf{e}_{AR,i}) x_{A,i} + \sqrt{p_{B,i}} (\hat{\mathbf{g}}_{AR,i}^H \mathbf{e}_{RB,i} + \hat{\mathbf{g}}_{RB,i}^H \mathbf{e}_{RB,i}) x_{B,i}}_{\text{estimation error}} \\
&+ \underbrace{\sum_{j \neq i} (\sqrt{p_{A,j}} (\hat{\mathbf{g}}_{AR,i}^H \mathbf{g}_{AR,j} + \hat{\mathbf{g}}_{RB,i}^H \mathbf{g}_{AR,j}) x_{A,j} + \sqrt{p_{B,j}} (\hat{\mathbf{g}}_{AR,i}^H \mathbf{g}_{RB,j} + \hat{\mathbf{g}}_{RB,i}^H \mathbf{g}_{RB,j}) x_{B,j})}_{\text{inter-user interference}} + \underbrace{(\hat{\mathbf{g}}_{AR,i}^H + \hat{\mathbf{g}}_{RB,i}^H) \mathbf{n}_R}_{\text{compound noise}}.
\end{aligned} \tag{12}$$

estimation errors and channel estimates, respectively, and

$$A_i^{\text{DF}} = p_{A,i} (|\hat{\mathbf{g}}_{AR,i}^H \hat{\mathbf{g}}_{AR,i}|^2 + |\hat{\mathbf{g}}_{RB,i}^H \hat{\mathbf{g}}_{AR,i}|^2), \tag{14}$$

$$B_i^{\text{DF}} = p_{B,i} (|\hat{\mathbf{g}}_{AR,i}^H \hat{\mathbf{g}}_{RB,i}|^2 + |\hat{\mathbf{g}}_{RB,i}^H \hat{\mathbf{g}}_{RB,i}|^2), \tag{15}$$

$$C_i^{\text{DF}} = p_{A,i} (|\hat{\mathbf{g}}_{AR,i}^H \mathbf{e}_{AR,i}|^2 + |\hat{\mathbf{g}}_{RB,i}^H \mathbf{e}_{AR,i}|^2) \tag{16}$$

$$+ p_{B,i} (|\hat{\mathbf{g}}_{AR,i}^H \mathbf{e}_{RB,i}|^2 + |\hat{\mathbf{g}}_{RB,i}^H \mathbf{e}_{RB,i}|^2),$$

$$D_i^{\text{DF}} = \sum_{j \neq i} p_{A,j} (|\hat{\mathbf{g}}_{AR,i}^H \mathbf{g}_{AR,j}|^2 + |\hat{\mathbf{g}}_{RB,i}^H \mathbf{g}_{AR,j}|^2) \tag{17}$$

$$+ \sum_{j \neq i} p_{B,j} (|\hat{\mathbf{g}}_{AR,i}^H \mathbf{g}_{RB,j}|^2 + |\hat{\mathbf{g}}_{RB,i}^H \mathbf{g}_{RB,j}|^2),$$

$$E_i^{\text{DF}} = \|\hat{\mathbf{g}}_{AR,i}\|^2 + \|\hat{\mathbf{g}}_{RB,i}\|^2. \tag{18}$$

In addition, the SE of the $T_{X,i} \rightarrow T_R$ ($X \in \{A, B\}$) link can be obtained as

$$R_{XR,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \times \tag{19}$$

$$\mathbb{E} \left\{ \log_2 \left(1 + \frac{X_i^{\text{DF}}}{\mathbb{E} \left\{ (C_i^{\text{DF}} + D_i^{\text{DF}} + E_i^{\text{DF}}) |\hat{\mathbf{G}}_{AR}, \hat{\mathbf{G}}_{RB} \right\}} \right) \right\}.$$

In the BC phase, the relay broadcasts to all users using the MR principle; hence the received signal at $T_{X,i}$ is given by

$$\mathbf{z}_{X,i}^{\text{DF}} = \rho_{\text{DF}} \sum_{j=1}^N (\mathbf{g}_{XR,i}^T \hat{\mathbf{g}}_{RB,j}^* \mathbf{x}_{A,j} + \mathbf{g}_{XR,i}^T \hat{\mathbf{g}}_{AR,j}^* \mathbf{x}_{B,j}) + \mathbf{n}_{X,i}. \tag{20}$$

As in [5], [25], [30], we consider the realistic case where the users do not have any instantaneous CSI and instead only have statistical CSI since the acquisition of instantaneous CSI at the user is extremely costly. Note that this type of lower bounds has already been proved to be very good due to the channel hardening [31]². Therefore, after performing the partial self-interference cancellation according to the statistical knowledge of the channel gains, the post-processing signals at $T_{A,i}$ can

be re-expressed as

$$z_{A,i}^{\text{DF}} = z_{A,i}^{\text{DF}} - \rho_{\text{DF}} \mathbb{E} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^* \} x_{A,i} \tag{21}$$

$$= \underbrace{\rho_{\text{DF}} \mathbb{E} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^* \}}_{\text{desired signal}} x_{B,i}$$

$$+ \underbrace{\rho_{\text{DF}} (\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^* - \mathbb{E} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^* \})}_{\text{gain uncertainty}} x_{B,i}$$

$$+ \underbrace{\rho_{\text{DF}} (\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^* - \mathbb{E} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^* \})}_{\text{residual self-interference}} x_{A,i}$$

$$+ \underbrace{\rho_{\text{DF}} \sum_{j \neq i} (\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,j}^* x_{A,j} + \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,j}^* x_{B,j})}_{\text{inter-user interference}} + \underbrace{n_{A,i}}_{\text{noise}}$$

and the post-processing signals at $T_{B,i}$ is obtained by replacing A and B in (21) with B and A .

Therefore, the SE of the $T_R \rightarrow T_{X,i}$ link is expressed as

$$R_{RX,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 (1 + \text{SINR}_{RX,i}^{\text{DF}}), \tag{22}$$

where $\text{SINR}_{RX,i}^{\text{DF}}$ is given by (23) (on the top of the next page).

Now, according to [32]–[36], the sum SE of the i -th user pair over both MAC and BC phases for the considered two-way DF relaying can be expressed as

$$R_i^{\text{DF}} = \min (R_{1,i}^{\text{DF}}, R_{2,i}^{\text{DF}}), \tag{24}$$

where $R_{2,i}^{\text{DF}}$ is the sum SE of the i -th user pair in the BC phase, which is given by the sum of the end-to-end SE from $T_{A,i}$ to $T_{B,i}$ (i.e., $\min (R_{AR,i}^{\text{DF}}, R_{RB,i}^{\text{DF}})$) and the end-to-end SE from $T_{B,i}$ to $T_{A,i}$ (i.e., $\min (R_{BR,i}^{\text{DF}}, R_{RA,i}^{\text{DF}})$),

$$R_{2,i}^{\text{DF}} = \min (R_{AR,i}^{\text{DF}}, R_{RB,i}^{\text{DF}}) + \min (R_{BR,i}^{\text{DF}}, R_{RA,i}^{\text{DF}}). \tag{25}$$

Thus, the sum SE of the multipair two-way DF relaying system is

$$R^{\text{DF}} = \sum_{i=1}^N R_i^{\text{DF}}. \tag{26}$$

When T_R employs a very large antenna array, i.e., $M \rightarrow \infty$, a large-scale approximation of the SE of the i -th user pair is presented in the following theorem.

Theorem 1: With the DF protocol, as the number of relay antennas grows to infinity, then we have $R_i^{\text{DF}} - \tilde{R}_i^{\text{DF}} \xrightarrow{M \rightarrow \infty} 0$, where \tilde{R}_i^{DF} is given by

$$\tilde{R}_i^{\text{DF}} \triangleq \min (\tilde{R}_{1,i}^{\text{DF}}, \tilde{R}_{2,i}^{\text{DF}}), \tag{27}$$

²Note that there are two typical ways to obtain the instantaneous CSI, i.e., downlink training and feedback. However, both methods incur huge overheads, hence they are not scalable in the massive MIMO ecosystem.

$$\text{SINR}_{RX,i}^{\text{DF}} = \frac{|\mathbf{E}\{\mathbf{g}_{XR,i}^T \hat{\mathbf{g}}_{XR,i}^*\}|^2}{\text{Var}\{\mathbf{g}_{XR,i}^T \hat{\mathbf{g}}_{XR,i}^*\} + \text{Var}\{\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^*\} + \sum_{j \neq i} \left(\mathbf{E}\{|\mathbf{g}_{XR,i}^T \hat{\mathbf{g}}_{RB,j}^*|^2\} + \mathbf{E}\{|\mathbf{g}_{XR,i}^T \hat{\mathbf{g}}_{AR,j}^*|^2\} \right) + \frac{1}{\rho_{\text{DF}}}}. \quad (23)$$

where

$$\tilde{R}_{1,i}^{\text{DF}} \triangleq \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{t_{A,i} + t_{B,i}}{(\sigma_{AR,i}^2 + \sigma_{RB,i}^2) q_i} \right), \quad (28)$$

$$\tilde{R}_{2,i}^{\text{DF}} \triangleq \min(\tilde{R}_{AR,i}^{\text{DF}}, \tilde{R}_{RB,i}^{\text{DF}}) + \min(\tilde{R}_{BR,i}^{\text{DF}}, \tilde{R}_{RA,i}^{\text{DF}}), \quad (29)$$

with

$$\tilde{R}_{AR,i}^{\text{DF}} \triangleq \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{t_{A,i}}{(\sigma_{AR,i}^2 + \sigma_{RB,i}^2) q_i} \right), \quad (30)$$

$$\tilde{R}_{RA,i}^{\text{DF}} \triangleq \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{p_r M \sigma_{AR,i}^4}{(p_r \beta_{AR,i} + 1) \sum_{j=1}^N (\sigma_{AR,j}^2 + \sigma_{RB,j}^2)} \right), \quad (31)$$

$$t_{A,i} \triangleq p_{A,i} (M \sigma_{AR,i}^4 + \sigma_{AR,i}^2 \sigma_{RB,i}^2), \quad (32)$$

$$t_{B,i} \triangleq p_{B,i} (M \sigma_{RB,i}^4 + \sigma_{AR,i}^2 \sigma_{RB,i}^2), \quad (33)$$

$$q_i \triangleq p_{A,i} \tilde{\sigma}_{AR,i}^2 + p_{B,i} \tilde{\sigma}_{RB,i}^2 + \sum_{j \neq i} (p_{A,j} \beta_{AR,j} + p_{B,j} \beta_{RB,j}) + 1, \quad (34)$$

and $\tilde{R}_{BR,i}^{\text{DF}}$ and $\tilde{R}_{RB,i}^{\text{DF}}$ are obtained by replacing the transmit powers $p_{A,i}$, $p_{B,i}$, and the subscripts ‘‘AR’’, ‘‘RB’’ with the transmit powers $p_{B,i}$, $p_{A,i}$, and the subscripts ‘‘RB’’, ‘‘AR’’ in $\tilde{R}_{AR,i}^{\text{DF}}$ and $\tilde{R}_{RA,i}^{\text{DF}}$, respectively.

Proof: See Appendix A. ■

Theorem 1 provides a large-scale approximation of the i -th user pair’s SE. More specifically, $\tilde{R}_{1,i}^{\text{DF}}$, $\tilde{R}_{AR,i}^{\text{DF}}$, and $\tilde{R}_{BR,i}^{\text{DF}}$ are computed by utilizing Lemma 1 in Appendix A, while $\tilde{R}_{RA,i}^{\text{DF}}$ and $\tilde{R}_{RB,i}^{\text{DF}}$ are the exact expressions for $R_{RA,i}^{\text{DF}}$ and $R_{RB,i}^{\text{DF}}$. We can see that all the signal-to-interference-plus-noise ratio (SINR) of $\tilde{R}_{1,i}^{\text{DF}}$, $\tilde{R}_{AR,i}^{\text{DF}}$, $\tilde{R}_{BR,i}^{\text{DF}}$, $\tilde{R}_{RA,i}^{\text{DF}}$, and $\tilde{R}_{RB,i}^{\text{DF}}$ can be expressed in the form of $\frac{ax+b}{cx+d}$, where $a > 0$, $b \geq 0$, $c > 0$, $d > 0$, and x denotes the transmit power, i.e., $p_{A,i}$, $p_{B,i}$, or p_r . Since $\frac{ax+b}{cx+d}$ increases with x and converges to $\frac{a}{c}$ as $x \rightarrow \infty$, we conclude that 1) \tilde{R}_i^{DF} is an increasing function of $p_{A,i}$, $p_{B,i}$, and p_r ; and 2) when $p_{A,i} \rightarrow \infty$, $p_{B,i} \rightarrow \infty$, and/or $p_r \rightarrow \infty$, \tilde{R}_i^{DF} converges to a non-zero limit, due to strong inter-user interference. Moreover, we observe that \tilde{R}_i^{DF} increases with the number of relay antennas M , indicating the strong advantage of employing massive antenna arrays at the relay, while decreases with the number of user pairs N , which is expected since larger number of users increases the amount of inter-user interference.

IV. POWER-SCALING LAWS

In this section, we pursue a detailed investigation of the power-scaling laws; that is, how the powers can be reduced

with M while retaining a desired SE. In the case where the users have different power levels from the beginning, the differences can be absorbed into the $\sigma_{AR,i}^2$ and $\sigma_{RB,i}^2$ terms without loss of generality. Thus, we assume that all the users have the same transmit power, i.e., $p_{A,i} = p_{B,i} = p_u$. Then, we can characterize the interplay between the relay’s transmit power p_r , the user’s transmit power p_u , and the transmit power of each pilot symbol p_p , as M grows to infinity. More precisely, we consider three different scenarios:

- Scenario A: Fixed p_u and p_r , while $p_p = \frac{E_p}{M^\gamma}$ with $\gamma > 0$, and E_p being a constant. Such a scenario represents the potential of power saving in the channel training stage.
- Scenario B: Fixed p_p , while $p_u = \frac{E_u}{M^\alpha}$, $p_r = \frac{E_r}{M^\beta}$, with $\alpha \geq 0$ and $\beta \geq 0$, and E_u , E_r are constants. Hence, the channel estimation accuracy remains unchanged, and the objective is to study the potential power savings in the data transmission stage.
- Scenario C: This is the most general case where $p_u = \frac{E_u}{M^\alpha}$, $p_r = \frac{E_r}{M^\beta}$, and $p_p = \frac{E_p}{M^\gamma}$, with $\alpha \geq 0$, $\beta \geq 0$, and $\gamma > 0$, E_u , E_r , and E_p are constants.

1) *Scenario A:* We present the following power-scaling law for Scenario A.

Theorem 2: With the DF protocol, for fixed p_u , p_r and E_p , when $p_p = \frac{E_p}{M^\gamma}$ with $\gamma > 0$, as $M \rightarrow \infty$, we have

$$R_i^{\text{DF}} - \min(\bar{R}_{1,i}^{\text{DF}}, \bar{R}_{2,i}^{\text{DF}}) \xrightarrow{M \rightarrow \infty} 0, \quad (35)$$

where

$$\bar{R}_{1,i}^{\text{DF}} \triangleq \frac{\tau_c - \tau_p}{2\tau_c} \times \quad (36)$$

$$\log_2 \left(1 + \frac{p_u \frac{\tau_p E_p}{M^{\gamma-1}} (\beta_{AR,i}^4 + \beta_{RB,i}^4)}{(\beta_{AR,i}^2 + \beta_{RB,i}^2) \left(p_u \sum_{j=1}^N (\beta_{AR,j} + \beta_{RB,j}) + 1 \right)} \right),$$

$$\bar{R}_{2,i}^{\text{DF}} \triangleq \min(\bar{R}_{AR,i}^{\text{DF}}, \bar{R}_{RB,i}^{\text{DF}}) + \min(\bar{R}_{BR,i}^{\text{DF}}, \bar{R}_{RA,i}^{\text{DF}}), \quad (37)$$

with

$$\bar{R}_{AR,i}^{\text{DF}} \triangleq \frac{\tau_c - \tau_p}{2\tau_c} \times \quad (38)$$

$$\log_2 \left(1 + \frac{p_u \frac{\tau_p E_p}{M^{\gamma-1}} \beta_{AR,i}^4}{(\beta_{AR,i}^2 + \beta_{RB,i}^2) \left(p_u \sum_{j=1}^N (\beta_{AR,j} + \beta_{RB,j}) + 1 \right)} \right)$$

$$\bar{R}_{RA,i}^{\text{DF}} \triangleq \frac{\tau_c - \tau_p}{2\tau_c} \times \quad (39)$$

$$\log_2 \left(1 + \frac{p_r \frac{\tau_p E_p}{M^{\gamma-1}} \beta_{AR,i}^4}{(p_r \beta_{AR,i} + 1) \sum_{j=1}^N (\beta_{AR,j}^2 + \beta_{RB,j}^2)} \right),$$

and $\bar{R}_{BR,i}^{\text{DF}}$ and $\bar{R}_{RB,i}^{\text{DF}}$ are obtained by replacing the subscripts “AR”, “RB” in $\bar{R}_{AR,i}^{\text{DF}}$ and $\bar{R}_{RA,i}^{\text{DF}}$ with the subscripts “RB”, “AR”, respectively.

As can be seen, the large-scale approximation of the SE R_i^{DF} in Scenario A depends on the choice of γ . When we cut down p_p too much, i.e., $\gamma > 1$, R_i^{DF} converges to zero. On the other hand, when $0 < \gamma < 1$, R_i^{DF} grows unboundedly. Finally, when $\gamma = 1$, R_i^{DF} converges to a non-zero limit.

2) *Scenario B*: Next, we turn our attention to Scenario B and present the following result.

Theorem 3: With the DF protocol, for fixed p_p , E_u , and E_r , when $p_u = \frac{E_u}{M^\alpha}$, $p_r = \frac{E_r}{M^\beta}$, with $\alpha \geq 0$, $\beta \geq 0$, as $M \rightarrow \infty$, we have

$$R_i^{\text{DF}} - \min(\bar{R}_{1,i}^{\text{DF}}, \bar{R}_{2,i}^{\text{DF}}) \xrightarrow{M \rightarrow \infty} 0, \quad (40)$$

where

$$\bar{R}_{1,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_u}{M^{\alpha-1}} \frac{\sigma_{AR,i}^4 + \sigma_{RB,i}^4}{\sigma_{AR,i}^2 + \sigma_{RB,i}^2} \right), \quad (41)$$

$$\bar{R}_{2,i}^{\text{DF}} = \min(\bar{R}_{AR,i}^{\text{DF}}, \bar{R}_{RB,i}^{\text{DF}}) + \min(\bar{R}_{BR,i}^{\text{DF}}, \bar{R}_{RA,i}^{\text{DF}}), \quad (42)$$

with

$$\bar{R}_{AR,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_u}{M^{\alpha-1}} \frac{\sigma_{AR,i}^4}{\sigma_{AR,i}^2 + \sigma_{RB,i}^2} \right), \quad (43)$$

$$\bar{R}_{RA,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_r}{M^{\beta-1}} \frac{\sigma_{AR,i}^4}{\sum_{j=1}^N (\sigma_{AR,j}^2 + \sigma_{RB,j}^2)} \right), \quad (44)$$

and $\bar{R}_{BR,i}^{\text{DF}}$ and $\bar{R}_{RB,i}^{\text{DF}}$ are obtained by replacing the subscripts “AR”, “RB” in $\bar{R}_{AR,i}^{\text{DF}}$ and $\bar{R}_{RA,i}^{\text{DF}}$ with the subscripts “RB”, “AR”, respectively.

Theorem 3 indicates that when both the transmit power of each user p_u and the transmit power of the relay p_r are scaled down inversely proportional to M (as $M \rightarrow \infty$), the effects of estimation error, residual self-interference, and inter-user interference vanish, and the only remaining impairment comes from the noise at users and the relay. Moreover, when each user’s transmit power is sufficiently large, i.e., $E_u \rightarrow \infty$, the large-scale approximation of R_i^{DF} is determined only by $\bar{R}_{RA,i}^{\text{DF}}$

and $\bar{R}_{RB,i}^{\text{DF}}$, suggesting that the bottleneck of SE appears in the BC phase. In contrast, when the relay’s transmit power becomes large, i.e., $E_r \rightarrow \infty$, the large-scale approximation of R_i^{DF} is determined only by $\bar{R}_{1,i}^{\text{DF}}$, $\bar{R}_{AR,i}^{\text{DF}}$, $\bar{R}_{BR,i}^{\text{DF}}$, indicating that the bottleneck of SE occurs in the MAC phase.

Also, when we cut down the transmit powers of the relay and/or of each user too much, namely, 1) $\alpha > 1$, and $\beta \geq 0$, 2) $\alpha \geq 0$, and $\beta > 1$, 3) $\alpha > 1$, and $\beta > 1$, R_i^{DF} converges to zero. On the contrary, when we cut down both the transmit powers of the relay and of each user moderately, i.e., $0 \leq \alpha < 1$ and $0 \leq \beta < 1$, R_i^{DF} grows unboundedly. So the most important task is how to select the parameters α and β to make R_i^{DF} converge to a non-zero finite limit. We discuss this in the following corollaries.

Corollary 1: With the DF protocol, for fixed p_p , E_u , and E_r , when $\alpha = \beta = 1$, namely, $p_u = \frac{E_u}{M}$, $p_r = \frac{E_r}{M}$, as $M \rightarrow \infty$, the SE of the i -th user pair has the limit

$$R_i^{\text{DF}} \rightarrow \min(\bar{R}_{1,i}^{\text{DF}}, \bar{R}_{2,i}^{\text{DF}}), \quad (45)$$

where

$$\bar{R}_{1,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_u (\sigma_{AR,i}^4 + \sigma_{RB,i}^4)}{\sigma_{AR,i}^2 + \sigma_{RB,i}^2} \right), \quad (46)$$

$$\bar{R}_{2,i}^{\text{DF}} = \min(\bar{R}_{AR,i}^{\text{DF}}, \bar{R}_{RB,i}^{\text{DF}}) + \min(\bar{R}_{BR,i}^{\text{DF}}, \bar{R}_{RA,i}^{\text{DF}}), \quad (47)$$

with

$$\bar{R}_{AR,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_u \sigma_{AR,i}^4}{\sigma_{AR,i}^2 + \sigma_{RB,i}^2} \right), \quad (48)$$

$$\bar{R}_{RA,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_r \sigma_{AR,i}^4}{\sum_{j=1}^N (\sigma_{AR,j}^2 + \sigma_{RB,j}^2)} \right), \quad (49)$$

and $\bar{R}_{BR,i}^{\text{DF}}$ and $\bar{R}_{RB,i}^{\text{DF}}$ are obtained by replacing the subscripts “AR”, “RB” in $\bar{R}_{AR,i}^{\text{DF}}$ and $\bar{R}_{RA,i}^{\text{DF}}$ with the subscripts “RB”, “AR”, respectively.

Corollary 1 reveals that when both the transmit powers of the relay and of each user are scaled down with the same speed, i.e., $1/M$, R_i^{DF} converges to a non-zero limit. Moreover, this non-zero limit is an increasing function with respect to E_u and E_r , while a decreasing function with respect to the number of user pairs N .

Corollary 2: With the DF protocol, for fixed p_p , E_u , and E_r , when $\alpha = 1$ and $0 \leq \beta < 1$, namely, $p_u = \frac{E_u}{M}$, $p_r = \frac{E_r}{M^\beta}$, as $M \rightarrow \infty$, the SE of the i -th user pair has the limit

$$R_i^{\text{DF}} \rightarrow \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_u (\sigma_{AR,i}^4 + \sigma_{RB,i}^4)}{\sigma_{AR,i}^2 + \sigma_{RB,i}^2} \right). \quad (50)$$

Corollary 2 suggests that when we cut down the transmit power of each user too much, i.e., $0 \leq \beta < \alpha = 1$, the large-scale approximation of the SE is determined by the performance in the MAC phase, i.e., $\bar{R}_{1,i}^{\text{DF}}$, which depends only on E_u , and is independent of E_r . This result is expected since when the transmit power of each user is much less than the transmit power of the relay, the bottleneck of SE occurs in the MAC phase. On the other hand, when the transmit power of

the relay is cut down more compared with that of each user, i.e., $0 \leq \alpha < \beta = 1$, the bottleneck of SE appears in the BC phase, thus R_i^{DF} is determined by $\bar{R}_{RA,i}^{\text{DF}}$ and $\bar{R}_{RB,i}^{\text{DF}}$ as shown in the following corollary.

Corollary 3: With the DF protocol, for fixed p_p , E_u , and E_r , when $0 \leq \alpha < 1$ and $\beta = 1$, namely, $p_u = \frac{E_u}{M^\alpha}$, $p_r = \frac{E_r}{M}$, as $M \rightarrow \infty$, the SE of the i -th user pair has the limit

$$R_i^{\text{DF}} \rightarrow \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_r \sigma_{AR,i}^4}{\sum_{j=1}^N (\sigma_{AR,j}^2 + \sigma_{RB,j}^2)} \right) + \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{E_r \sigma_{RB,i}^4}{\sum_{j=1}^N (\sigma_{AR,j}^2 + \sigma_{RB,j}^2)} \right). \quad (51)$$

3) *Scenario C:* Finally, a corresponding power-scaling law for Scenario C is obtained as follows.

Theorem 4: With the DF protocol, for fixed E_u , E_r , and E_p , when $p_u = \frac{E_u}{M^\alpha}$, $p_r = \frac{E_r}{M^\beta}$, and $p_p = \frac{E_p}{M^\gamma}$, with $\alpha \geq 0$, $\beta \geq 0$, and $\gamma > 0$, as $M \rightarrow \infty$, we have

$$R_i^{\text{DF}} - \min(\bar{R}_{1,i}^{\text{DF}}, \bar{R}_{2,i}^{\text{DF}}) \xrightarrow{M \rightarrow \infty} 0, \quad (52)$$

where

$$\bar{R}_{1,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_u E_p \beta_{AR,i}^4 + \beta_{RB,i}^4}{M^{\alpha+\gamma-1} \beta_{AR,i}^2 + \beta_{RB,i}^2} \right), \quad (53)$$

$$\bar{R}_{2,i}^{\text{DF}} = \min(\bar{R}_{AR,i}^{\text{DF}}, \bar{R}_{RB,i}^{\text{DF}}) + \min(\bar{R}_{BR,i}^{\text{DF}}, \bar{R}_{RA,i}^{\text{DF}}), \quad (54)$$

with

$$\bar{R}_{AR,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_u E_p \beta_{AR,i}^4}{\beta_{AR,i}^2 + \beta_{RB,i}^2} \right), \quad (55)$$

$$\bar{R}_{RA,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_r E_p \beta_{AR,i}^4}{M^{\beta+\gamma-1} \beta_{AR,i}^2 + \beta_{RB,i}^2} \right), \quad (56)$$

and $\bar{R}_{BR,i}^{\text{DF}}$ and $\bar{R}_{RB,i}^{\text{DF}}$ are obtained by replacing the subscripts ‘‘AR’’, ‘‘RB’’ in $\bar{R}_{AR,i}^{\text{DF}}$ and $\bar{R}_{RA,i}^{\text{DF}}$ with the subscripts ‘‘RB’’, ‘‘AR’’, respectively.

As expected, the large-scale approximation of the SE R_i^{DF} depends on the relationship between α , β , and γ . Moreover, the term $\alpha + \gamma$ determines the SE in the MAC phase, while $\beta + \gamma$ determines the SE in the BC phase, as elaborated in the following corollaries.

Corollary 4: With the DF protocol, for fixed E_u , E_r , and E_p , when $\alpha = \beta > 0$ and $\alpha + \gamma = 1$, namely, $p_u = \frac{E_u}{M^\alpha}$, $p_r = \frac{E_r}{M^\beta}$, and $p_p = \frac{E_p}{M^\gamma}$, with $\gamma > 0$, as $M \rightarrow \infty$, the SE of the i -th user pair has the limit

$$R_i^{\text{DF}} \rightarrow \min(\bar{R}_{1,i}^{\text{DF}}, \bar{R}_{2,i}^{\text{DF}}), \quad (57)$$

where

$$\bar{R}_{1,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_u E_p (\beta_{AR,i}^4 + \beta_{RB,i}^4)}{\beta_{AR,i}^2 + \beta_{RB,i}^2} \right), \quad (58)$$

$$\bar{R}_{2,i}^{\text{DF}} = \min(\bar{R}_{AR,i}^{\text{DF}}, \bar{R}_{RB,i}^{\text{DF}}) + \min(\bar{R}_{BR,i}^{\text{DF}}, \bar{R}_{RA,i}^{\text{DF}}), \quad (59)$$

with

$$\bar{R}_{AR,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_u E_p \beta_{AR,i}^4}{\beta_{AR,i}^2 + \beta_{RB,i}^2} \right), \quad (60)$$

$$\bar{R}_{RA,i}^{\text{DF}} = \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_r E_p \beta_{AR,i}^4}{\sum_{j=1}^N (\beta_{AR,j}^2 + \beta_{RB,j}^2)} \right), \quad (61)$$

and $\bar{R}_{BR,i}^{\text{DF}}$ and $\bar{R}_{RB,i}^{\text{DF}}$ are obtained by replacing the subscripts ‘‘AR’’, ‘‘RB’’ in $\bar{R}_{AR,i}^{\text{DF}}$ and $\bar{R}_{RA,i}^{\text{DF}}$ with the subscripts ‘‘RB’’, ‘‘AR’’, respectively.

Corollary 4 presents a trade-off between the transmit powers of each pilot symbol and of each user/the relay. In other words, if we cut down the transmit power of each pilot symbol too much, which causes poor channel estimation accuracy, the transmit power of each user/the relay should be increased to compensate this imperfection and maintain the same asymptotic SE.

Corollary 5: With the DF protocol, for fixed E_u , E_r , and E_p , when $\alpha > \beta \geq 0$ and $\alpha + \gamma = 1$, namely, $p_u = \frac{E_u}{M^\alpha}$, $p_r = \frac{E_r}{M^\beta}$, and $p_p = \frac{E_p}{M^\gamma}$, with $\gamma > 0$, as $M \rightarrow \infty$, the SE of the i -th user pair has the limit

$$R_i^{\text{DF}} \rightarrow \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_u E_p (\beta_{AR,i}^4 + \beta_{RB,i}^4)}{\beta_{AR,i}^2 + \beta_{RB,i}^2} \right). \quad (62)$$

From Corollary 5, we can see that the limit of R_i^{DF} is an increasing function with respect to E_u and E_p , indicating that we can boost the SE by increasing the transmit power of each user and of each pilot symbol. In addition, the limit of R_i^{DF} is independent of N , indicating that the sum SE of the system is an increasing function with respect to N .

Corollary 6: With the DF protocol, for fixed E_u , E_r , and E_p , when $0 \leq \alpha < \beta$ and $\beta + \gamma = 1$, namely, $p_u = \frac{E_u}{M^\alpha}$, $p_r = \frac{E_r}{M^\beta}$, and $p_p = \frac{E_p}{M^\gamma}$, with $\gamma > 0$, as $M \rightarrow \infty$, the SE of the i -th user pair has the limit

$$R_i^{\text{DF}} \rightarrow \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_r E_p \beta_{AR,i}^4}{\sum_{j=1}^N (\beta_{AR,j}^2 + \beta_{RB,j}^2)} \right) \quad (63)$$

$$+ \frac{\tau_c - \tau_p}{2\tau_c} \log_2 \left(1 + \frac{\tau_p E_r E_p \beta_{RB,i}^4}{\sum_{j=1}^N (\beta_{AR,j}^2 + \beta_{RB,j}^2)} \right).$$

Corollary 6 provides the trade-off between the transmit powers of the relay and of each pilot symbol.

V. POWER ALLOCATION

Power allocation is an effective means to enhance the sum SE of the system. In this section, we assume that the design for channel training stage is done in advance, i.e., the pilot power p_p is determined. We are interested in designing a power allocation algorithm in the data transmission stage maximizing the sum SE subject to a total power constraint, i.e., $\sum_{i=1}^N (p_{A,i} + p_{B,i}) + p_r \leq P$. For notational simplicity, we define $\mathcal{N} \triangleq \{1, \dots, N\}$, $\mathbf{p}_A \triangleq [p_{A,1}, \dots, p_{A,N}]^T$, and $\mathbf{p}_B \triangleq [p_{B,1}, \dots, p_{B,N}]^T$.

For analytical tractability, we use the large-scale approximation (27) in Theorem 1 instead of the exact expression (24). Hence, the power allocation optimization problem is formulated as

$$\begin{aligned} & \underset{\mathbf{p}_A, \mathbf{p}_B, p_r}{\text{maximize}} && \sum_{i=1}^N \tilde{R}_i^{\text{DF}} \\ & \text{subject to} && \sum_{i=1}^N (p_{A,i} + p_{B,i}) + p_r \leq P \\ & && \mathbf{p}_A \geq \mathbf{0}, \mathbf{p}_B \geq \mathbf{0}, p_r \geq 0 \\ & && \tilde{R}_i^{\text{DF}} \geq R_{\min}, i \in \mathcal{N} \end{aligned} \quad (64)$$

where R_{\min} is the minimum SE requirement of the i -th user pair. Since $\log(\cdot)$ is an increasing function, (64) can be equivalently reformulated as $\mathcal{P}_1^{\text{DF}}$:

$$\begin{aligned} & \underset{\substack{\mathbf{p}_A, \mathbf{p}_B, p_r \\ \gamma_i^{\text{DF}}, \gamma_{A,i}^{\text{DF}}, \gamma_{B,i}^{\text{DF}}}}{\text{minimize}} && \prod_{i=1}^N (1 + \gamma_i^{\text{DF}})^{-1} \\ & \text{subject to} && \gamma_i^{\text{DF}} \leq \frac{a_i^{\text{DF}} p_{A,i} + b_i^{\text{DF}} p_{B,i}}{\sum_{j=1}^N (c_{i,j}^{\text{DF}} p_{A,i} + d_{i,j}^{\text{DF}} p_{B,i}) + 1}, i \in \mathcal{N} \\ & && \gamma_i^{\text{DF}} \leq \gamma_{A,i}^{\text{DF}} + \gamma_{B,i}^{\text{DF}} + \gamma_{A,i}^{\text{DF}} \gamma_{B,i}^{\text{DF}}, i \in \mathcal{N} \\ & && \gamma_{A,i}^{\text{DF}} \leq \min \left\{ \frac{a_i^{\text{DF}} p_{A,i}}{g_i}, \frac{p_r}{e_i^{\text{DF}} p_r + f_i^{\text{DF}}} \right\}, i \in \mathcal{N} \\ & && \gamma_{B,i}^{\text{DF}} \leq \min \left\{ \frac{b_i^{\text{DF}} p_{B,i}}{g_i}, \frac{p_r}{\tilde{e}_i^{\text{DF}} p_r + \tilde{f}_i^{\text{DF}}} \right\}, i \in \mathcal{N} \\ & && \sum_{i=1}^N (p_{A,i} + p_{B,i}) + p_r \leq P \\ & && \mathbf{p}_A \geq \mathbf{0}, \mathbf{p}_B \geq \mathbf{0}, p_r \geq 0 \\ & && (\gamma_i^{\text{DF}})^{-1} \left(2^{\frac{2\tau_c R_{\min}}{\tau_c - \tau_p}} - 1 \right) \leq 1, i \in \mathcal{N} \end{aligned}$$

where $\gamma^{\text{DF}} \triangleq [\gamma_1^{\text{DF}}, \dots, \gamma_N^{\text{DF}}]^T$, $\gamma_A^{\text{DF}} \triangleq [\gamma_{A,1}^{\text{DF}}, \dots, \gamma_{A,N}^{\text{DF}}]^T$, $\gamma_B^{\text{DF}} \triangleq [\gamma_{B,1}^{\text{DF}}, \dots, \gamma_{B,N}^{\text{DF}}]^T$, $g_i = \sum_{j=1}^N (c_{i,j}^{\text{DF}} p_{A,i} + d_{i,j}^{\text{DF}} p_{B,i}) + 1$, $a_i^{\text{DF}} = \frac{M\sigma_{AR,i}^4 + \sigma_{AR,i}^2 \sigma_{RB,i}^2}{\sigma_{AR,i}^2 + \sigma_{RB,i}^2}$, $b_i^{\text{DF}} = \frac{M\sigma_{RB,i}^4 + \sigma_{AR,i}^2 \sigma_{RB,i}^2}{\sigma_{AR,i}^2 + \sigma_{RB,i}^2}$, $c_{i,j}^{\text{DF}} = \begin{cases} \tilde{\sigma}_{AR,i}^2, & j = i, \\ \beta_{AR,j}, & j \neq i, \end{cases}$, $e_i^{\text{DF}} = \frac{\beta_{RB,i}}{M\sigma_{RB,i}^4} \sum_{j=1}^N (\sigma_{AR,j}^2 + \sigma_{RB,j}^2)$, $f_i^{\text{DF}} =$

$\frac{1}{M\sigma_{RB,i}^4} \sum_{j=1}^N (\sigma_{AR,j}^2 + \sigma_{RB,j}^2)$, and $d_{i,j}^{\text{DF}}$, \tilde{e}_i^{DF} , \tilde{f}_i^{DF} are obtained by replacing the subscripts ‘‘AR’’, ‘‘RB’’ with ‘‘RB’’, ‘‘AR’’ in $c_{i,j}^{\text{DF}}$, e_i^{DF} , f_i^{DF} , respectively.

The above problem $\mathcal{P}_1^{\text{DF}}$ is identified as a complementary geometric programming (CGP) problem, which is nonconvex. Also, γ_i^{DF} , $\gamma_{A,i}^{\text{DF}}$, and $\gamma_{B,i}^{\text{DF}}$ are considered as the SINR of $\min(R_{1,i}^{\text{DF}}, R_{2,i}^{\text{DF}})$, $\min(R_{AR,i}^{\text{DF}}, R_{RB,i}^{\text{DF}})$, and $\min(R_{BR,i}^{\text{DF}}, R_{RA,i}^{\text{DF}})$, respectively. In addition, we have replaced the equality ‘‘=’’ with ‘‘ \leq ’’ in the first four constraints of problem $\mathcal{P}_1^{\text{DF}}$; however, this does not change or relax the original problem (64), since the objective function is decreasing with γ_i^{DF} . Therefore, we can guarantee that these four constraints must be active at any optimal solution of $\mathcal{P}_1^{\text{DF}}$.

Although the CGP problem is nonconvex, we can obtain its local optimum solution by jointly solving a sequence of convex GP problems; this is a technique that has been widely used in the resource allocation literature, such as [5], [22], [37]–[40]. Next, we are dedicated to transforming $\mathcal{P}_1^{\text{DF}}$ into a standard GP problem that can be solved efficiently with standard optimization tools such as CVX or gglab. Since the objective function $1 + \gamma_i^{\text{DF}}$ can be approximated by a monomial function $\omega_i^{\text{DF}} (\gamma_i^{\text{DF}})^{\mu_i^{\text{DF}}}$, where $\mu_i^{\text{DF}} = \frac{\hat{\gamma}_i^{\text{DF}}}{1 + \hat{\gamma}_i^{\text{DF}}}$ and $\omega_i^{\text{DF}} = (\hat{\gamma}_i^{\text{DF}})^{-\mu_i^{\text{DF}}} (1 + \hat{\gamma}_i^{\text{DF}})$ (which are obtained by guaranteeing that both the values and the first gradients of the approximations and of the original functions are equal at the same specific points, i.e., by solving the equations $\omega_i^{\text{DF}} (\hat{\gamma}_i^{\text{DF}})^{\mu_i^{\text{DF}}} = 1 + \hat{\gamma}_i^{\text{DF}}$ and $\omega_i^{\text{DF}} \mu_i^{\text{DF}} (\hat{\gamma}_i^{\text{DF}})^{\mu_i^{\text{DF}} - 1} = 1$), the main challenge to convert $\mathcal{P}_1^{\text{DF}}$ into a GP problem is to transform the first two inequality constraints into the form of posynomials. According to [40], [41], the geometric mean is no larger than the arithmetic mean for any set of positive numbers; thus we have

$$a_i^{\text{DF}} p_{A,i} + b_i^{\text{DF}} p_{B,i} \geq \left(\frac{a_i^{\text{DF}} p_{A,i}}{\nu_{A,i}^{\text{DF}}} \right)^{\nu_{A,i}^{\text{DF}}} \left(\frac{b_i^{\text{DF}} p_{B,i}}{\nu_{B,i}^{\text{DF}}} \right)^{\nu_{B,i}^{\text{DF}}}, \quad (65)$$

where $\nu_{A,i}^{\text{DF}} = \frac{a_i^{\text{DF}} \hat{p}_{A,i}}{a_i^{\text{DF}} \hat{p}_{A,i} + b_i^{\text{DF}} \hat{p}_{B,i}}$, $\nu_{B,i}^{\text{DF}} = \frac{b_i^{\text{DF}} \hat{p}_{B,i}}{a_i^{\text{DF}} \hat{p}_{A,i} + b_i^{\text{DF}} \hat{p}_{B,i}}$, and $\hat{p}_{A,i}$, $\hat{p}_{B,i}$ are the initialization values.

As a result, the first inequality constraint in $\mathcal{P}_1^{\text{DF}}$ can be approximated by [40]

$$\gamma_i^{\text{DF}} \leq \frac{\left(\frac{a_i^{\text{DF}} p_{A,i}}{\nu_{A,i}^{\text{DF}}} \right)^{\nu_{A,i}^{\text{DF}}} \left(\frac{b_i^{\text{DF}} p_{B,i}}{\nu_{B,i}^{\text{DF}}} \right)^{\nu_{B,i}^{\text{DF}}}}{\sum_{j=1}^N (c_{i,j}^{\text{DF}} p_{A,i} + d_{i,j}^{\text{DF}} p_{B,i}) + 1}, i \in \mathcal{N}. \quad (66)$$

Now, we focus on the approximation of the second inequality constraint. Following the idea proposed in [37, Lemma 1], we use a monomial function $g(x, y) = \eta x^{\lambda_1} y^{\lambda_2}$ to approximate $f(x, y) = x + y + xy$ near an arbitrary point $\hat{x}, \hat{y} > 0$. To make the approximation accurate, we need to ensure that the following equations hold and then obtain its solution:

$$\begin{cases} \hat{x} + \hat{y} + \hat{x}\hat{y} = \eta \hat{x}^{\lambda_1} \hat{y}^{\lambda_2} \\ 1 + \hat{y} = \eta \lambda_1 \hat{x}^{\lambda_1 - 1} \hat{y}^{\lambda_2} \\ 1 + \hat{x} = \eta \lambda_2 \hat{x}^{\lambda_1} \hat{y}^{\lambda_2 - 1} \end{cases} \Rightarrow \begin{cases} \lambda_1 = \frac{\hat{x}(1+\hat{y})}{\hat{x}+\hat{y}+\hat{x}\hat{y}}, \\ \lambda_2 = \frac{\hat{y}(1+\hat{x})}{\hat{x}+\hat{y}+\hat{x}\hat{y}}, \\ \eta = (\hat{x} + \hat{y} + \hat{x}\hat{y}) \hat{x}^{-\lambda_1} \hat{y}^{-\lambda_2}. \end{cases}$$

To this end, the second inequality constraint in $\mathcal{P}_1^{\text{DF}}$ can be approximated by

$$\gamma_i^{\text{DF}} \leq \eta_i^{\text{DF}} (\gamma_{A,i}^{\text{DF}})^{\lambda_{A,i}^{\text{DF}}} (\gamma_{B,i}^{\text{DF}})^{\lambda_{B,i}^{\text{DF}}}, i \in \mathcal{N}, \quad (67)$$

where $\eta_i^{\text{DF}} = (\hat{\gamma}_{A,i}^{\text{DF}} + \hat{\gamma}_{B,i}^{\text{DF}} + \hat{\gamma}_{A,i}^{\text{DF}} \hat{\gamma}_{B,i}^{\text{DF}}) (\hat{\gamma}_{A,i}^{\text{DF}})^{-\lambda_{A,i}^{\text{DF}}} (\hat{\gamma}_{B,i}^{\text{DF}})^{-\lambda_{B,i}^{\text{DF}}}$,
 $\lambda_{A,i}^{\text{DF}} = \frac{\hat{\gamma}_{A,i}^{\text{DF}} (1 + \hat{\gamma}_{B,i}^{\text{DF}})}{\hat{\gamma}_{A,i}^{\text{DF}} + \hat{\gamma}_{B,i}^{\text{DF}} + \hat{\gamma}_{A,i}^{\text{DF}} \hat{\gamma}_{B,i}^{\text{DF}}}$,
 $\lambda_{B,i}^{\text{DF}} = \frac{\hat{\gamma}_{B,i}^{\text{DF}} (1 + \hat{\gamma}_{A,i}^{\text{DF}})}{\hat{\gamma}_{A,i}^{\text{DF}} + \hat{\gamma}_{B,i}^{\text{DF}} + \hat{\gamma}_{A,i}^{\text{DF}} \hat{\gamma}_{B,i}^{\text{DF}}}$, and $\hat{\gamma}_{A,i}^{\text{DF}}, \hat{\gamma}_{B,i}^{\text{DF}}$ are the initialization values.

We now outline the steps to solve the original problem $\mathcal{P}_1^{\text{DF}}$ in Algorithm 1.

Note that we have removed ω_i^{DF} in the objective function, since they do not affect the optimization problem. Also, five extra inequalities as trust region constraints are included, which limit how much the variables are allowed to differ from the current guess $\hat{\gamma}_i^{\text{DF}}, \hat{\gamma}_{A,i}^{\text{DF}}$, and $\hat{\gamma}_{B,i}^{\text{DF}}$. The limit of any convergent sequence generated by Algorithm 1 is a Karush-Kuhn-Tucker point, and the detailed proof can be found in [42]. The result holds provided Slater's constraint qualification condition holds. The parameter $\theta > 1$ controls the desired accuracy. More precisely, when θ is close to 1 it provides good accuracy for the monomial approximation but with slower convergence speed, and vice versa if θ is large. As discussed in [41], $\theta = 1.1$ offers a good tradeoff between the accuracy and convergence speed.

Algorithm 1 focuses on the case where each user transmits with a different power, and yields a local optimum of the original problem $\mathcal{P}_1^{\text{DF}}$ by solving a sequence of GPs. Now, we turn our attention to the scenario where all the users transmit with the same power, i.e., $p_{A,i} = p_{B,i} = p_u$ and R_{\min} is very low such that the constraint $\tilde{R}_i^{\text{DF}} \geq R_{\min}, i \in \mathcal{N}$ can be neglected; hence, the problem (64) reduces to the following special case:

$$\begin{aligned} \mathcal{P}_3^{\text{DF}} : \text{maximize}_{p_u, p_r} \quad & \sum_{i=1}^N \tilde{R}_i^{\text{DF}} \\ \text{subject to} \quad & 2Np_u + p_r \leq P, p_u \geq 0, p_r \geq 0. \end{aligned}$$

Theorem 5: $\mathcal{P}_3^{\text{DF}}$ is a convex optimization problem.

Proof: See Appendix B. \blacksquare

Since the optimization problem $\mathcal{P}_3^{\text{DF}}$ is convex, the optimal solutions $p_u^{\text{DF,opt}} \in (0, \frac{P}{2N}]$ and $p_r^{\text{DF,opt}} \in (0, P]$ maximizing the sum SE can be obtained efficiently by adopting some standard techniques, such as the bisection method with respect to P , due to the convexity of the optimization problem $\mathcal{P}_3^{\text{DF}}$.

VI. NUMERICAL RESULTS

We now present numerical results to validate the above analytical results. Unless otherwise specified, the following set of parameters are used in simulation. The length of the coherence interval is $\tau_c = 196$ (symbols), chosen by the LTE standard. The different large-scale fading parameters are arbitrarily generated by $\beta_{AR,i} = z_i (r_{AR,i}/r_0)^\alpha$ and $\beta_{RB,i} = z_i (r_{RB,i}/r_0)^\alpha$, where z_i is a log-normal random variable with standard deviation 8 dB, $r_{AR,i}$ and $r_{RB,i}$ are the locations of $T_{AR,i}$ and $T_{RB,i}$ from the relay, $\alpha = 3.8$ is the path loss exponent, and r_0 denotes the guard interval

Algorithm 1 Successive approximation algorithm for $\mathcal{P}_1^{\text{DF}}$

1) *Initialization.* Define a tolerance ϵ and parameter θ . Set $k = 1$, the initial values of $\hat{\gamma}_i^{\text{DF}}, \hat{\gamma}_{A,i}^{\text{DF}}$ and $\hat{\gamma}_{B,i}^{\text{DF}}$ are chosen according to the SINR in Theorem 1. Also, we set $\hat{p}_{A,i} = \hat{p}_{B,i} = \frac{P}{4N}$,
 2) *Iteration* k . Compute $\mu_i^{\text{DF}} = \frac{\hat{\gamma}_i^{\text{DF}}}{1 + \hat{\gamma}_i^{\text{DF}}}$,
 $\nu_{A,i}^{\text{DF}} = \frac{a_i^{\text{DF}} \hat{p}_{A,i}}{a_i^{\text{DF}} \hat{p}_{A,i} + b_i^{\text{DF}} \hat{p}_{B,i}}$, $\nu_{B,i}^{\text{DF}} = \frac{b_i^{\text{DF}} \hat{p}_{B,i}}{a_i^{\text{DF}} \hat{p}_{A,i} + b_i^{\text{DF}} \hat{p}_{B,i}}$,
 $\eta_i^{\text{DF}} = (\hat{\gamma}_{A,i}^{\text{DF}} + \hat{\gamma}_{B,i}^{\text{DF}} + \hat{\gamma}_{A,i}^{\text{DF}} \hat{\gamma}_{B,i}^{\text{DF}}) (\hat{\gamma}_{A,i}^{\text{DF}})^{-\lambda_{A,i}^{\text{DF}}} (\hat{\gamma}_{B,i}^{\text{DF}})^{-\lambda_{B,i}^{\text{DF}}}$,
 $\lambda_{A,i}^{\text{DF}} = \frac{\hat{\gamma}_{A,i}^{\text{DF}} (1 + \hat{\gamma}_{B,i}^{\text{DF}})}{\hat{\gamma}_{A,i}^{\text{DF}} + \hat{\gamma}_{B,i}^{\text{DF}} + \hat{\gamma}_{A,i}^{\text{DF}} \hat{\gamma}_{B,i}^{\text{DF}}}$, $\lambda_{B,i}^{\text{DF}} = \frac{\hat{\gamma}_{B,i}^{\text{DF}} (1 + \hat{\gamma}_{A,i}^{\text{DF}})}{\hat{\gamma}_{A,i}^{\text{DF}} + \hat{\gamma}_{B,i}^{\text{DF}} + \hat{\gamma}_{A,i}^{\text{DF}} \hat{\gamma}_{B,i}^{\text{DF}}}$. Then, solve the following GP problem $\mathcal{P}_2^{\text{DF}}$:

$$\begin{aligned} \text{minimize}_{\mathbf{p}_A, \mathbf{p}_B, p_r, \gamma_i^{\text{DF}}, \gamma_{A,i}^{\text{DF}}, \gamma_{B,i}^{\text{DF}}} \quad & \prod_{i=1}^N (\gamma_i^{\text{DF}})^{-\mu_i^{\text{DF}}} \\ \text{subject to} \quad & \theta^{-1} \hat{p}_{A,i}^{\text{DF}} \leq p_{A,i}^{\text{DF}} \leq \theta \hat{p}_{A,i}^{\text{DF}}, i \in \mathcal{N} \\ & \theta^{-1} \hat{p}_{B,i}^{\text{DF}} \leq p_{B,i}^{\text{DF}} \leq \theta \hat{p}_{B,i}^{\text{DF}}, i \in \mathcal{N} \\ & \theta^{-1} \hat{\gamma}_i^{\text{DF}} \leq \gamma_i^{\text{DF}} \leq \theta \hat{\gamma}_i^{\text{DF}}, i \in \mathcal{N} \\ & \theta^{-1} \hat{\gamma}_{A,i}^{\text{DF}} \leq \gamma_{A,i}^{\text{DF}} \leq \theta \hat{\gamma}_{A,i}^{\text{DF}}, i \in \mathcal{N} \\ & \theta^{-1} \hat{\gamma}_{B,i}^{\text{DF}} \leq \gamma_{B,i}^{\text{DF}} \leq \theta \hat{\gamma}_{B,i}^{\text{DF}}, i \in \mathcal{N} \\ & \gamma_i^{\text{DF}} (u_{A,i})^{-\nu_{A,i}^{\text{DF}}} \left(\frac{b_i^{\text{DF}} p_{B,i}}{\nu_{B,i}^{\text{DF}}} \right)^{-\nu_{B,i}^{\text{DF}}} g_i \leq 1, i \in \mathcal{N} \\ & \gamma_i^{\text{DF}} (\eta_i^{\text{DF}})^{-1} (\gamma_{A,i}^{\text{DF}})^{-\lambda_{A,i}^{\text{DF}}} (\gamma_{B,i}^{\text{DF}})^{-\lambda_{B,i}^{\text{DF}}} \leq 1, i \in \mathcal{N} \\ & \gamma_{A,i}^{\text{DF}} (a_i^{\text{DF}})^{-1} p_{A,i}^{-1} g_i \leq 1, i \in \mathcal{N} \\ & \gamma_{B,i}^{\text{DF}} (b_i^{\text{DF}})^{-1} p_{B,i}^{-1} g_i \leq 1, i \in \mathcal{N} \\ & \gamma_{A,i}^{\text{DF}} p_r^{-1} (e_i^{\text{DF}} p_r + f_i^{\text{DF}}) \leq 1, i \in \mathcal{N} \\ & \gamma_{B,i}^{\text{DF}} p_r^{-1} (\tilde{e}_i^{\text{DF}} p_r + \tilde{f}_i^{\text{DF}}) \leq 1, i \in \mathcal{N} \\ & \sum_{i=1}^N (p_{A,i} + p_{B,i}) + p_r \leq P \\ & \mathbf{p}_A \geq \mathbf{0}, \mathbf{p}_B \geq \mathbf{0}, p_r \geq 0 \\ & (\gamma_i^{\text{DF}})^{-1} \left(2^{\frac{2\tau_c R_{\min}}{\tau_c - \tau_p}} - 1 \right) \leq 1, i \in \mathcal{N} \end{aligned}$$

where $u_{A,i} = \frac{a_i^{\text{DF}} p_{A,i}}{\nu_{A,i}^{\text{DF}}}$.

Denote the optimal solutions by $p_{A,i}^{(k),\text{DF}}, p_{B,i}^{(k),\text{DF}}, \gamma_i^{(k),\text{DF}}, \gamma_{A,i}^{(k),\text{DF}}, \gamma_{B,i}^{(k),\text{DF}}, i \in \mathcal{N}$.

3) *Stopping criterion.* If $\max_i |p_{A,i}^{(k),\text{DF}} - \hat{p}_{A,i}^{\text{DF}}| < \epsilon$ and/or $\max_i |p_{B,i}^{(k),\text{DF}} - \hat{p}_{B,i}^{\text{DF}}| < \epsilon$ and/or $\max_i |\gamma_i^{(k),\text{DF}} - \hat{\gamma}_i^{\text{DF}}| < \epsilon$ and/or $\max_i |\gamma_{A,i}^{(k),\text{DF}} - \hat{\gamma}_{A,i}^{\text{DF}}| < \epsilon$ and/or $\max_i |\gamma_{B,i}^{(k),\text{DF}} - \hat{\gamma}_{B,i}^{\text{DF}}| < \epsilon$, stop; otherwise, go to step 4).

4) *Update initial values.* Set $\hat{p}_{A,i}^{\text{DF}} = p_{A,i}^{(k),\text{DF}}, \hat{p}_{B,i}^{\text{DF}} = p_{B,i}^{(k),\text{DF}}, \hat{\gamma}_i^{\text{DF}} = \gamma_i^{(k),\text{DF}}, \hat{\gamma}_{A,i}^{\text{DF}} = \gamma_{A,i}^{(k),\text{DF}}, \hat{\gamma}_{B,i}^{\text{DF}} = \gamma_{B,i}^{(k),\text{DF}}$, and $k = k + 1$. Go to step 2).

which specifies the nearest distance between the users and the relay. The relay is located at the center of a cell with a radius of 1000 meters and $r_0 = 100$ meters. We choose $N = 5$, $\beta_{AR} = [0.2688, 0.0368, 0.00025, 0.1398, 0.0047]$, and $\beta_{RB} = [0.0003, 0.00025, 0.0050, 0.0794, 0.0001]$. The length

of the pilot sequences is $\tau_p = 2N$ which is the minimum requirement. We assume that each user has the same transmit power, i.e., $p_{A,i} = p_{B,i} = p_u$.

A. Validation of analytical expressions and evaluating the impact of τ_p

We assume that $p_p = p_u$, and that the total transmit power of the N user pairs is equal to the transmit power of the relay, i.e., $p_r = 2Np_u$.

Fig. 2(a) shows the sum SE versus the transmit power of each user p_u for different number of relay antennas. Note that the ‘‘Approximations’’ curves are obtained by using (27), and the ‘‘Numerical results’’ curves are generated according to (24) by averaging over 10^4 independent channel realizations. As can be readily observed, the large-scale approximations are very accurate, especially for large antenna arrays. Moreover, we can see that increasing the number of relay antennas significantly yields higher SE, as expected.

Fig. 2(b) studies the impact of τ_p on the sum SE. We can see that at moderate and low p_p , there is an optimal τ_p maximizing the sum SE. Also, the optimal length of pilot sequence decreases with the transmit power of pilot sequence p_p . In contrast, at high p_p , i.e., $p_p = 10$ dB, the sum SE is a decreasing function with respect to τ_p , which means that the optimal τ_p is equal to $2N$.

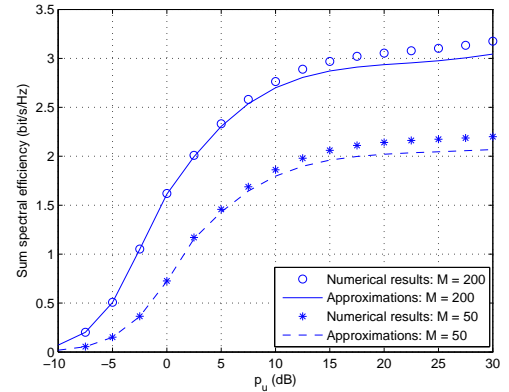
B. Power-scaling laws

In this subsection, we provide numerical simulation results to verify the power-scaling laws presented in the previous subsections, and investigate the potential for power saving when employing large number of antennas at the relay. Since our goal is to show the general power saving behavior and it is unnecessary to pay much attention to one particular user’s location, here we set $\beta_{AR,i} = \beta_{RB,i} = 1$. Note that the curves labeled as ‘‘Approximations’’ are obtained according to Theorems 1.

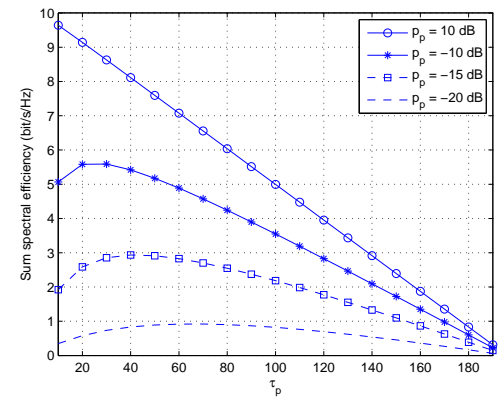
1) *Scenario A*: Fig. 3 verifies the analytical results for Scenario A. The curves labelled as ‘‘Scenario A’’, are plotted according to Theorem 2. It can be readily observed that in the large M regime, the asymptotic curves converge to the exact curves, demonstrating the accuracy of the asymptotic analysis. In addition, when $\gamma > 1$, i.e., $\gamma = 2$, the SE gradually approaches zero. In contrast, when $0 < \gamma < 1$, i.e., $\gamma = 0.8$, the SE of both schemes grows unbounded. Finally, when $\gamma = 1$, the SE converges to a non-zero limit for both schemes.

2) *Scenario B*: Fig. 4 investigates how the scaling of the transmit power of each user $p_u = \frac{E_u}{M^\alpha}$ and the transmit power of the relay $p_r = \frac{E_r}{M^\beta}$ affects the achievable SE. Note that the curves labelled as ‘‘Scenario B’’ are generated by using Theorem 3, while the curves labelled as ‘‘Scenario B-Case X’’ with $X \in \{I, II, III\}$ are plotted according to Corollaries 1–3, respectively.

Fig. 4(a) studies three different cases according to the values of α and β . In agreement with Corollaries 1–3, the sum SE saturates in the asymptotical large M regime for all the three



(a) Validation of analytical expressions



(b) Impact of τ_p

Fig. 2: Validation of analytical expressions and impact of τ_p for $N = 5$, $p_p = p_u$ and $p_r = 2Np_u$.

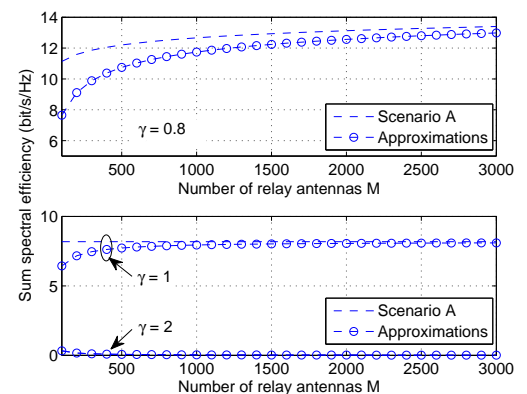


Fig. 3: Sum SE versus the number of relay antennas M for $N = 5$, $p_u = 10$ dB, $p_r = 20$ dB, and $p_p = E_p/M^\gamma$ with $E_p = 10$ dB.

cases. As readily observed, Case I and Case II achieve the same performance due to the setting of $E_r = 2NE_u$.

Fig. 4(b) illustrates the other two extreme scenarios where the transmit power down-scaling is either too aggressive or too moderate. For the former scenario, three different cases are studied, i.e., $\alpha > 1, \beta \geq 0$, $\alpha \geq 0, \beta > 1$, and $\alpha > 1, \beta > 1$.

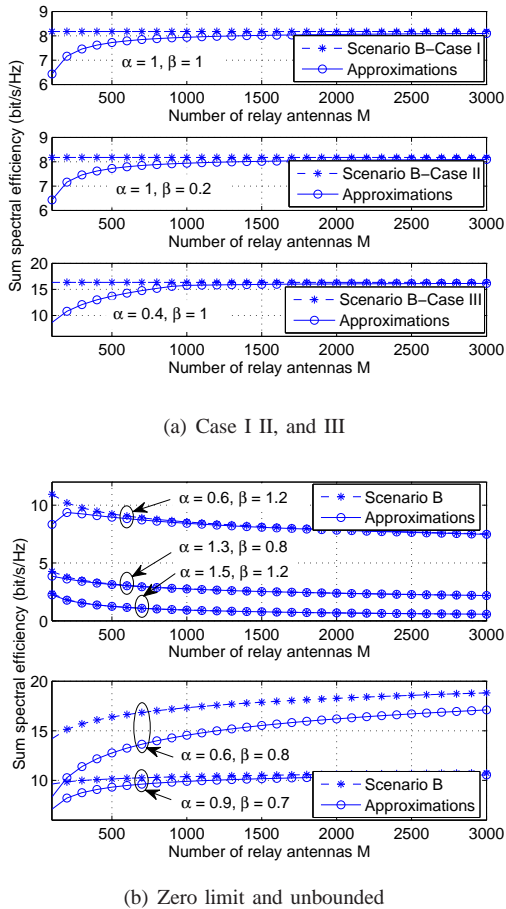


Fig. 4: Sum SE versus the number of relay antennas M for $N = 5$, $p_p = 10$ dB, $p_u = E_u/M^\alpha$ with $E_u = 10$ dB, and $p_r = E_r/M^\beta$ with $E_u = 20$ dB.

As expected, when the number of relay antennas increases, the sum SE gradually reduces to zero. However, the speed of reduction varies significantly depending on the scaling parameters. The larger the scaling parameters, the faster the decay SE. On the other hand, when we cut down the transmit powers of each user and of the relay moderately, the sum SE grows unboundedly.

3) *Scenario C*: Fig. 5 demonstrates the tradeoff between the user/relay power and the pilot symbol power. For illustration purposes, two extreme scenarios where the transmit power down-scaling is either too aggressive or too moderate are considered. For the former scenario, two sets of curves are drawn according to $\alpha = 1.3$, $\beta = 1.1$, $\gamma = 0.5$ and $\alpha = 0.8$, $\beta = 0.6$, $\gamma = 1$, which satisfy $\alpha + \gamma = 1.8$ and $\beta + \gamma = 1.6$. When the number of relay antennas grows large, the sum SE of all system configurations smoothly converges to zero, as predicted. Moreover, the gaps between the two sets of curves reduce with M and eventually vanish. This indicates that as long as $\alpha + \gamma$ and $\beta + \gamma$ are the same, the asymptotic sum SE remains unchanged. Now, let us focus on the two curves associated with $N = 5$. Interestingly, we see that the curve associated with $\gamma = 0.5$ yields better sum SE in the finite antenna regime, despite the fact that the user or relay power is

over-reduced compared to the $\gamma = 1$ case, which suggests that it is of crucial importance to improve the channel estimation accuracy. The same behavior appears for the unbounded SE scenario where $\alpha + \gamma = 0.6$ and $\beta + \gamma = 0.7$.

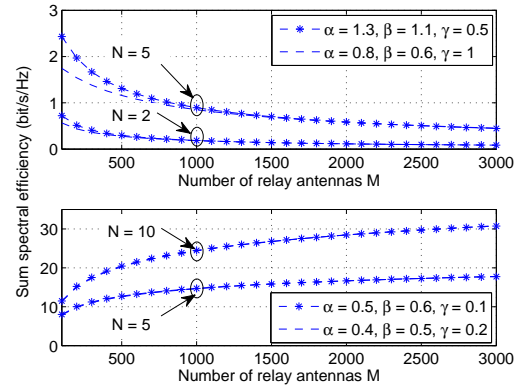


Fig. 5: Sum SE versus the number of relay antennas M for $p_u = E_u/M^\alpha$ with $E_u = 10$ dB, $p_r = E_r/M^\beta$ with $E_r = 15$ dB, and $p_p = E_p/M^\gamma$ with $E_p = 0$ dB.

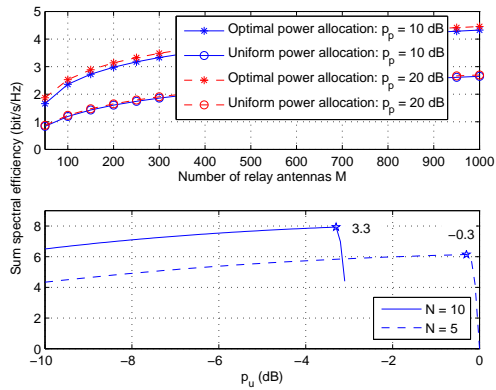
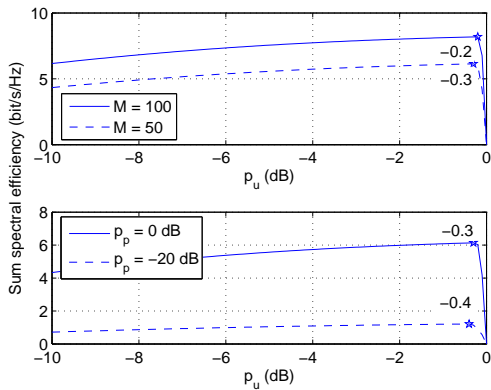
C. Power allocation

Fig. 6(a) illustrates the impact of the optimal power allocation scheme on the sum SE when all users' large-scale fading are different. The optimal power allocation curves are generated by Algorithm 1. As a benchmark scheme for comparison, we also plot the sum SE with uniform power allocation, i.e., the relay transmit power equals to the sum user transmit power. As can be observed, the optimal power allocation policy respectively provides 79.11% and 82.67% SE enhancement for $p_p = 10$ dB and $p_p = 20$ dB when $M = 300$, indicating that high channel estimation accuracy slightly boosts the power allocation efficiency. Moreover, by focusing on the case where every user has the same transmit power, i.e., $p_{A,i} = p_{B,i} = p_u$, we can see that the optimal user transmit power is a decreasing function with respect to the number of user pairs N , for a given power budget $P = 10$ dB.

Fig. 6(b) further examines the impact of key system parameters such as M , and p_p on the optimal power allocation scheme. As can be seen, with fixed number of user pairs $N = 5$, the optimal user transmit power $p_u^{\text{DF,opt}}$ with $M = 100$ is larger than that with $M = 50$, suggesting that we should increase the transmit power of each user when the number of relay antennas is large. In addition, when the pilot training power increases, i.e., from $p_p = -20$ dB to $p_p = 0$ dB, the optimal user transmit power $p_u^{\text{DF,opt}}$ also increases, indicating that when the channel estimation accuracy is improved, we need to use a higher transmit power for each user.

VII. CONCLUSION

The paper studied the sum SE of a multipair two-way DF relaying system with the MR processing by taking realistic CSI assumption into account. In particular, a closed-form

(a) Benefits of optimal power allocation and impact of N (b) Impact of M and p_p Fig. 6: Sum SE for a given power budget $P = 10$ dB.

large-scale approximation of the SE was obtained that is tight when the number of relay antennas is large. Based on the approximation, the system's power-scaling laws were characterized, which demonstrated that the transmit powers of the users, relay, and pilot symbol can be substantially reduced to maintain the desired SE. In addition, it was revealed that there exists a tradeoff between the user/relay transmit power and pilot symbol power, which provides great flexibility for the design of practical systems. Finally, the transmit powers of each user and the relay were jointly optimized in terms of maximizing the sum SE, which substantially enhance the performance compared to the uniform power allocation scheme. For the special case where all users have the same transmit power, it turns out that the users should decrease their transmit power when the number of user pairs becomes large. On the other hand, the users should increase their transmit power when the number of relay antennas increases or when the channel estimation accuracy improves.

APPENDIX A PROOF OF THEOREM 1

Here we only present the detailed derivation for $\tilde{R}_{1,i}^{\text{DF}}$ and $\tilde{R}_{RA,i}^{\text{DF}}$, since $\tilde{R}_{AR,i}^{\text{DF}}$ and $\tilde{R}_{BR,i}^{\text{DF}}$ can be obtained in a straightforward way, while $\tilde{R}_{RB,i}^{\text{DF}}$ can be derived in the same fashion as $\tilde{R}_{RA,i}^{\text{DF}}$.

First, we focus on (13), which consists of five terms: 1) desired signal power of $T_{B,i} A_i^{\text{DF}}$; 2) desired signal power of $T_{A,i} B_i^{\text{DF}}$; 3) estimation error C_i^{DF} ; 4) inter-user interference D_i^{DF} ; 5) compound noise E_i^{DF} . For each of these five terms, we will subsequently derive a deterministic equivalent expression.

Before proceeding, we first review some useful results, which are given in the following lemma.

Lemma 1: Let $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \sigma_x^2 \mathbf{I}_M)$ and $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \sigma_y^2 \mathbf{I}_M)$. Assume that \mathbf{x} and \mathbf{y} are mutually independent. Then, we have

$$\frac{1}{M} \mathbf{x}^\dagger \mathbf{x} \xrightarrow{a.s.} \sigma_x^2, \quad M \rightarrow \infty, \quad (68)$$

$$\frac{1}{M} \mathbf{x}^\dagger \mathbf{y} \xrightarrow{a.s.} 0, \quad M \rightarrow \infty, \quad (69)$$

$$\frac{1}{M^2} |\mathbf{x}^\dagger \mathbf{y}|^2 - \frac{1}{M} \sigma_x^2 \sigma_y^2 \xrightarrow{a.s.} 0, \quad M \rightarrow \infty. \quad (70)$$

Now, we compute the five terms one by one.

1) Deterministic equivalent for A_i^{DF} : we have $\frac{1}{M^2} A_i^{\text{DF}} = \frac{p_{A,i}}{M^2} (|\hat{\mathbf{g}}_{AR,i}^H \hat{\mathbf{g}}_{AR,i}|^2 + |\hat{\mathbf{g}}_{RB,i}^H \hat{\mathbf{g}}_{AR,i}|^2)$.

Then, by invoking Lemma 1, we have

$$\frac{1}{M^2} A_i^{\text{DF}} - p_{A,i} \left(\sigma_{AR,i}^4 + \frac{1}{M} \sigma_{AR,i}^2 \sigma_{RB,i}^2 \right) \xrightarrow{a.s.} 0. \quad (71)$$

2) Deterministic equivalents for B_i^{DF} , C_i^{DF} , D_i^{DF} , and E_i^{DF} : Similarly, we obtain

$$\frac{1}{M^2} B_i^{\text{DF}} - p_{B,i} \left(\sigma_{RB,i}^4 + \frac{1}{M} \sigma_{AR,i}^2 \sigma_{RB,i}^2 \right) \xrightarrow{a.s.} 0, \quad (72)$$

$$\frac{1}{M^2} C_i^{\text{DF}} - \frac{1}{M} \left(\sigma_{AR,i}^2 + \sigma_{RB,i}^2 \right) \left(p_{A,i} \tilde{\sigma}_{AR,i}^2 + p_{B,i} \tilde{\sigma}_{RB,i}^2 \right) \xrightarrow{a.s.} 0, \quad (73)$$

$$\frac{1}{M^2} D_i^{\text{DF}} - \frac{1}{M} \sum_{j \neq i} (p_{A,j} \beta_{AR,j} + p_{B,j} \beta_{RB,j}) \left(\sigma_{AR,i}^2 + \sigma_{RB,i}^2 \right) \xrightarrow{a.s.} 0, \quad (74)$$

$$\frac{1}{M^2} E_i^{\text{DF}} - \frac{1}{M} \left(\sigma_{AR,i}^2 + \sigma_{RB,i}^2 \right) \xrightarrow{a.s.} 0. \quad (75)$$

Substituting (71), (72), (73), (74), and (75) into (13) and (19), and after some algebraic manipulations, we obtain $\tilde{R}_{1,i}^{\text{DF}}$, $\tilde{R}_{AR,i}^{\text{DF}}$, $\tilde{R}_{BR,i}^{\text{DF}}$.

Now, we turn our attention to derive $\tilde{R}_{RA,i}^{\text{DF}}$.

1) Compute $\mathbb{E} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^* \}$:

$$\mathbb{E} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^* \} = \mathbb{E} \{ \|\hat{\mathbf{g}}_{AR,i}\|^2 \} + \mathbb{E} \{ \mathbf{e}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^* \} \quad (76)$$

$$= M \sigma_{AR,i}^2.$$

2) Compute $\text{Var} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^* \}$:

$$\text{Var} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^* \} = \mathbb{E} \{ \|\hat{\mathbf{g}}_{AR,i}\|^4 \} + \mathbb{E} \{ |\mathbf{e}_{AR,i}^T \hat{\mathbf{g}}_{AR,i}^*|^2 \} - M^2 \sigma_{AR,i}^4 \quad (77)$$

$$= M \sigma_{AR,i}^2 \beta_{AR,i}.$$

3) Compute $\text{Var} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^* \}$:

$$\text{Var} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^* \} = \mathbb{E} \{ |\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^*|^2 \} - |\mathbb{E} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^* \}|^2 = M \sigma_{RB,i}^2 \beta_{AR,i}. \quad (78)$$

$$= \mathbb{E} \{ |\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^*|^2 \} - |\mathbb{E} \{ \mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,i}^* \}|^2 = M \sigma_{RB,i}^2 \beta_{AR,i}.$$

4) Compute $\sum_{j \neq i} (\mathbb{E} \{ |\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,j}^*|^2 \} + \mathbb{E} \{ |\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,j}^*|^2 \})$:

For $j \neq i$, we obtain

$$\mathbb{E} \left\{ |\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,j}^*|^2 \right\} = M \sigma_{RB,j}^2 \beta_{AR,i}, \quad (79)$$

$$\mathbb{E} \left\{ |\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,j}^*|^2 \right\} = M \sigma_{AR,j}^2 \beta_{AR,i}. \quad (80)$$

Thus, we have

$$\begin{aligned} & \sum_{j \neq i} \left(\mathbb{E} \left\{ |\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{RB,j}^*|^2 \right\} + \mathbb{E} \left\{ |\mathbf{g}_{AR,i}^T \hat{\mathbf{g}}_{AR,j}^*|^2 \right\} \right) \\ &= M \beta_{AR,i} \sum_{j \neq i} \left(\sigma_{AR,j}^2 + \sigma_{RB,j}^2 \right). \end{aligned} \quad (81)$$

Combining (76), (77), (78), and (81) completes the proof.

APPENDIX B PROOF OF THEOREM 5

Using the same argument as in the proof of Theorem 5, it can be proved that the objective function in $\mathcal{P}_3^{\text{DF}}$ is maximized when $2Np_u + p_r = P$.

Before studying the properties of \tilde{R}_i^{DF} , we first present the following useful lemma.

Lemma 2: The functions $g_1(x) = \log_2 \left(1 + \frac{a_1 x}{b_1 x + c_1} \right)$ and $g_2(x) = \log_2 \left(1 + \frac{a_2 (d_2 - x)}{b_2 (d_2 - x) + c_2} \right)$ are all strictly concave with respect to x when $a_1, b_1, c_1, a_2, b_2, c_2, d_2 > 0$.

Now, focusing on \tilde{R}_i^{DF} and substituting $2Np_u + p_r = P$ into \tilde{R}_i^{DF} , it is easy to show that $\tilde{R}_{1,i}^{\text{DF}}$, $\tilde{R}_{AR,i}^{\text{DF}}$, and $\tilde{R}_{BR,i}^{\text{DF}}$ can be reformulated as $g_1(p_u)$, while $\tilde{R}_{RA,i}^{\text{DF}}$ and $\tilde{R}_{RB,i}^{\text{DF}}$ can be reformulated as $g_2(p_u)$, hence, $\tilde{R}_{1,i}^{\text{DF}}$, $\tilde{R}_{AR,i}^{\text{DF}}$, $\tilde{R}_{BR,i}^{\text{DF}}$, $\tilde{R}_{RA,i}^{\text{DF}}$, and $\tilde{R}_{RB,i}^{\text{DF}}$ are all concave functions with respect to p_u .

Due to the convexity preservation property of pointwise maximum and nonnegative weighted sums operations [43], $\tilde{R}_{2,i}^{\text{DF}} = \min \left(\tilde{R}_{AR,i}^{\text{DF}}, \tilde{R}_{RB,i}^{\text{DF}} \right) + \min \left(\tilde{R}_{BR,i}^{\text{DF}}, \tilde{R}_{RA,i}^{\text{DF}} \right)$ is also a concave function with respect to p_u . Therefore, $\min \left(\tilde{R}_{1,i}^{\text{DF}}, \tilde{R}_{2,i}^{\text{DF}} \right)$ is a concave function, which completes the proof.

REFERENCES

- [1] H. Xie, B. Wang, F. Gao, and S. Jin, "A full-space spectrum-sharing strategy for massive MIMO cognitive radio," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2537–2549, Oct. 2016.
- [2] H. Xie, F. Gao, S. Zhang, S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model" *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3170–3184, Apr. 2017.
- [3] R. Mochaourab, E. Björnson, and M. Bengtsson, "Adaptive pilot clustering in heterogeneous massive MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5555–5568, Aug. 2016.
- [4] G. Yang, C. K. Ho, R. Zhang, and Y. L. Guan, "Throughput optimization for massive MIMO systems powered by wireless energy transfer," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1640–1650, Aug. 2015.
- [5] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Oct. 2014.
- [6] H. A. 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.
- [7] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.
- [8] Q. Wang and Y. Jing, "Performance analysis and scaling law of MR-C/MRT relaying with CSI error in multi-pair massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5882–5896, Sept. 2017.
- [9] H. Gao, T. Lv, X. Su, H. Yang, and J. M. Cioffi, "Energy-efficient resource allocation for massive MIMO amplify-and-forward relay systems," *IEEE Access*, vol. 4, pp. 2771–2787, May 2016.
- [10] X. Xia, Y. Xu, K. Xu, D. Zhang, and W. Ma, "Full-duplex massive MIMO AF relaying with semiblind gain control," *IEEE Trans. Veh. Technol.* vol. 65, no. 7, pp. 5797–5804, July 2016.
- [11] S. Wang, Y. Liu, W. Zhang, and H. Zhang, "Achievable rates of full-duplex massive MIMO relay systems over Rician fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 9825–9837, Nov. 2017.
- [12] F. Tan, T. Lv, and S. Yang, "Power allocation optimization for energy-efficient massive MIMO aided multi-pair decode-and-forward relay systems," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2368–2381, Jun. 2017.
- [13] R. Zhang, Y-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699–712, Jun. 2009.
- [14] K.-J. Lee, H. Sung, E. Park, and I. Lee, "Joint optimization for one and two-way MIMO AF multiple-relay systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3671–3681, Dec. 2010.
- [15] G. Amaraturiya, C. Tellambura, and M. Ardakani, "Two-way amplify-and-forward multiple-input multiple-output relay networks with antenna selection," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1513–1529, Sept. 2012.
- [16] R. Vaze and R.W. Heath Jr., "On the capacity and diversity-multiplexing tradeoff of the two-way relay channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4219–4234, July 2011.
- [17] 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.
- [18] 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.
- [19] J. Feng, S. Ma, G. Yang, and B. Xia, "Power scaling of full-duplex two-way massive MIMO relay systems with correlated antennas and MRC/MRT processing," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4738–4753, July 2017.
- [20] Z. Zhang, Z. Chen, M. Shen, and B. Xia, "Spectral and energy efficiency of multipair two-way full-duplex relay systems with massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 848–863, Apr. 2016.
- [21] M. Yemini, A. Zappone, E. Jorswieck, and A. Leshem, "Energy efficient bidirectional massive MIMO relay beamforming," *IEEE Signal Process. Lett.* vol. 24, no. 7, pp. 1010–1014, July 2017.
- [22] Y. Dai and X. Dong, "Power allocation for multi-pair massive MIMO two-way AF relaying with linear processing," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 5932–5946, Sept. 2016.
- [23] C. Kong, C. Zhong, M. Matthaiou, E. Björnson, and Z. Zhang, "Multi-pair Two-way AF Relaying Systems with Massive Arrays and Imperfect CSI," in *Proc. IEEE ICASSP*, Mar. 2016, pp. 3651–3655.
- [24] E. Sharma, R. Budhiraja, K. Vasudevan, and L. Hanzo, "Full-duplex massive MIMO multi-pair two-way AF relaying: Energy efficiency optimization", Oct. 2017. [Online] Available: <https://arxiv.org/pdf/1705.09043.pdf>
- [25] Z. Zhang, Z. Chen, M. Shen, B. Xia, W. Xie, and Y. Zhao, "Performance analysis for training-based multipair two-way full-duplex relaying with massive antennas," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6130–6145, July 2017.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [31] G. Caire, "On the ergodic rate lower bounds with applications to massive MIMO," [Online] Available: <https://arxiv.org/pdf/1705.03577.pdf>
- [32] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

- [33] J. Gao, S. A. Vorobyov, H. Jiang, J. Zhang, and M. Haardt, "Sum-rate maximization with minimum power consumption for MIMO DF two-way relaying—part I: Relay optimization," *IEEE Trans. Signal Process.*, vol. 61, no. 14, pp. 3563–3577, July 2013.
- [34] —, "Sum-rate maximization with minimum power consumption for MIMO DF two-way relaying—part II: Network optimization," *IEEE Trans. Signal Process.*, vol. 61, no. 14, pp. 3578–3591, July 2013.
- [35] I. Hammerström, M. Kuhn, C. Eşli, J. Zhao, A. Wittneben, and G. Bauch, "MIMO two-way relaying with transmit CSI at the relay", in *Proc. IEEE SPAWC*, June 2007.
- [36] A. Alsharoa, F. Bader, and M.-S. Alouini, "Relay selection and resource allocation for two-way DF-AF cognitive radio networks," *IEEE Wireless Commun. Lett.*, vol. 2, no. 4, pp. 427–430, Aug. 2013.
- [37] 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, July 2011.
- [38] K. Guo, Y. Guo, and G. Ascheid, "Security-constrained power allocation in MU-massive-MIMO with distributed antennas," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8139–8153, Dec. 2016.
- [39] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, July 2007.
- [40] C. He, G. Y. Li, F. Zheng, and X. Hu, "Power allocation criteria for distributed antenna systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5083–5090, Nov. 2015.
- [41] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optim. Eng.*, vol. 8, no. 1, pp. 67–127, Apr. 2007.
- [42] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, pp. 681–683, Jul./Aug. 1978.
- [43] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.



Michail Matthaiou (S'05–M'08–SM'13) was born in Thessaloniki, Greece in 1981. He obtained the Diploma degree (5 years) in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece in 2004. He then received the M.Sc. (with distinction) in Communication Systems and Signal Processing from the University of Bristol, U.K. and Ph.D. degrees from the University of Edinburgh, U.K. in 2005 and 2008, respectively. From September 2008 through May 2010, he was with the Institute for Circuit Theory and Signal Processing, Munich University of Technology (TUM), Germany working as a Postdoctoral Research Associate. He is currently a Reader (equivalent to Associate Professor) in Multiple-Antenna Systems at Queen's University Belfast, U.K. after holding an Assistant Professor position at Chalmers University of Technology, Sweden. His research interests span signal processing for wireless communications, massive MIMO, hardware-constrained communications, and performance analysis of fading channels.

Dr. Matthaiou and his coauthors received the 2017 IEEE Communications Society Leonard G. Abraham Prize. He was the recipient of the 2011 IEEE ComSoc Best Young Researcher Award for the Europe, Middle East and Africa Region and a co-recipient of the 2006 IEEE Communications Chapter Project Prize for the best M.Sc. dissertation in the area of communications. He was co-recipient of the Best Paper Award at the 2014 IEEE International Conference on Communications (ICC) and was an Exemplary Reviewer for IEEE COMMUNICATIONS LETTERS for 2010. In 2014, he received the Research Fund for International Young Scientists from the National Natural Science Foundation of China. In the past, he was an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, Associate Editor/Senior Editor for IEEE COMMUNICATIONS LETTERS and was the Lead Guest Editor of the special issue on "Large-scale multiple antenna wireless systems" of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He was a co-chair of the Wireless Communications Symposium (WCS) at IEEE GLOBECOM 2016.



Chuili Kong (S'15) received the B.S. degree in electronics and information engineering from the Dalian University of Technology, Dalian, China, in 2013. She is currently working towards the Ph.D. degree in information and communication engineering at the Zhejiang University. From October 2016 to October 2017, she was a visiting scholar at University of California, Irvine. She was an exemplary reviewer for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS in 2017. Her research interests include massive MIMO systems and relaying communica-

tions.



Caijun Zhong (S'07–M'10–SM'14) received the B.S. degree in Information Engineering from the Xi'an Jiaotong University, Xi'an, China, in 2004, and the M.S. degree in Information Security in 2006, Ph.D. degree in Telecommunications in 2010, both from University College London, London, United Kingdom. From September 2009 to September 2011, he was a research fellow at the Institute for Electronics, Communications and Information Technologies (ECIT), Queens University Belfast, Belfast, UK. Since September 2011, he has been with Zhejiang

University, Hangzhou, China, where he is currently an associate professor. His research interests include massive MIMO systems, full-duplex communications, wireless power transfer and physical layer security.

Dr. Zhong is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, EURASIP JOURNAL ON WIRELESS COMMUNICATIONS AND NETWORKING, and JOURNAL OF COMMUNICATIONS AND NETWORKS. He is the recipient of the 2013 IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award. He and his coauthors has been awarded a Best Paper Award at the WCSP 2013. He was an Exemplary Reviewer for IEEE TRANSACTIONS ON COMMUNICATIONS in 2014.



Emil Björnson (S'07, M'12, SM'17) received the M.S. degree in Engineering Mathematics from Lund University, Sweden, in 2007. He received the Ph.D. degree in Telecommunications from KTH Royal Institute of Technology, Sweden, in 2011. From 2012 to mid 2014, he was a joint postdoc at the Alcatel-Lucent Chair on Flexible Radio, SUPELEC, France, and at KTH. He joined Linköping University, Sweden, in 2014 and is currently Associate Professor and Docent at the Division of Communication Systems.

He performs research on multi-antenna communications, Massive MIMO, radio resource allocation, energy-efficient communications, and network design. He is on the editorial board of the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING. He is the first author of the textbooks *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency* (2017) and *Optimal Resource Allocation in Coordinated Multi-Cell Systems* (2013). He is dedicated to reproducible research and has made a large amount of simulation code publicly available.

Dr. Björnson has performed MIMO research for more than ten years and has filed more than ten related patent applications. He received the 2016 Best PhD Award from EURASIP, the 2015 Ingvar Carlsson Award, and the 2014 Outstanding Young Researcher Award from IEEE ComSoc EMEA. He has co-authored papers that received best paper awards at WCSP 2017, IEEE ICC 2015, IEEE WCNC 2014, IEEE SAM 2014, IEEE CAMSAP 2011, and WCSP 2009.



Zhaoyang Zhang (M'00) received his Ph.D. degree in communication and information systems from Zhejiang University, Hangzhou, China, in 1998. Since then he has been with the College of Information Science and Electronic Engineering, Zhejiang University, where he became a full professor in 2005. He has a wide variety of research interests including information theory and coding theory, signal processing for communications and in networks, computation-and-communication theoretic analysis, etc. He was a co-recipient of four international conference Best Paper / Best Student Paper Awards. He is currently serving as Editor for *IEEE Transactions on Communications*, *IET Communications* and some other international journals. He served as General Chair, TPC Co-Chair or Symposium Co-Chair for many international conferences and workshops like ChinaCOM 2008, ICUFN 2011/2012/2013, WCSP 2013, Globecom 2014 Wireless Communications Symposium, and VTC-Fall 2017 Workshop HMWC, etc.