Multiple-Input Multiple-Output Two-Way Relaying: A Space-Division Approach

Xiaojun Yuan, *Member, IEEE*, Tao Yang, *Member, IEEE*, and Iain B. Collings, *Senior Member, IEEE*

***Abstract*—We propose a novel space-division-based network-coding scheme for multiple-input multiple-output (MIMO) two-way relay channels (TWRCs), in which two multiantenna users exchange information via a multiantenna relay. In the proposed scheme, the overall signal space at the relay is divided into two subspaces. In one subspace, the spatial streams of the two users have nearly orthogonal directions and are completely decoded at the relay. In the other subspace, the signal directions of the two users are nearly parallel, and linear functions of the spatial streams are computed at the relay, following the principle of physical-layer network coding. Based on the recovered messages and message-functions, the relay generates and forwards network-coded messages to the two users. We show that, at high signal-to-noise ratio, the proposed scheme achieves the asymptotic sum-rate capacity of the MIMO TWRC within $\frac{1}{2}\log(5/4) \approx 0.161$ bits per user-antenna, for any antenna configuration and any channel realization. We perform large-system analysis to derive the average sum-rate of the proposed scheme over Rayleigh-fading MIMO TWRCs. We show that the average asymptotic sum-rate gap to the capacity is at most 0.053 bits per relay-antenna. It is demonstrated that the proposed scheme significantly outperforms the existing schemes.**

***Index Terms*—Multiple-input multiple-output (MIMO), physical-layer network coding, space-division precoding, two-way relay channel (TWRC).**

I. Introduction

DURING the past decade, tremendous progress has been made in the field of network coding [1]. In [2]–[4], the concept of physical-layer network coding (PNC) was introduced and applied to wireless networks. The simplest model for PNC is a two-way relay channel (TWRC), in which two users A and B exchange information via an intermediate relay. Compared with conventional schemes, PNC allows the relay to deliver linear functions of user messages, which can potentially double the network throughput. It has been shown that PNC can achieve the capacity of the Gaussian TWRC within 1/2 bit per user [5], [6], and its gap to the capacity vanishes at high-signal-to-noise-ratio (SNR) regime.

Recently, research effort has been directed toward efficient communications over MIMO TWRCs, in which the two users and the relay are all equipped with multiple antennas. Most existing approaches on MIMO TWRCs focus on classical relaying strategies borrowed from one-way relay channels, such as amplify-and-forward [7]–[10], compress-and-forward [11], [12], and decode-and-forward [13]–[15]. These strategies generally perform well away from the channel capacity, e.g., due to noise amplification and multiplexing loss [15]. Recently, several relaying schemes have been proposed to support PNC in MIMO TWRCs [16]–[22]. The basic idea is to jointly decompose the channel matrices of the two users to create multiple scalar channels, over which multiple PNC streams are transmitted. Let n_A, n_B, and n_R denote the number of antennas of users A and B, and of the relay, respectively. For configurations with $\max(n_A, n_B) \geq n_R$, a generalized singular-value-decomposition (GSVD) scheme was shown to achieve the asymptotic capacity of the MIMO TWRC at high SNR [16], and an eigendirection alignment-based scheme was shown to bring about improved performance at medium-to-low SNR [17]. For configurations with $\max(n_A, n_B) < n_R$, all existing schemes in general perform far away from the capacity. Such configurations, however, are of most practical importance. For example, due to the size limitation of user terminals, it is usually more convenient to implement more antennas at the relay station than at the user ends, as suggested in the standards of next generation wireless networks [23], [24].

In this paper, we propose a new space-division-based PNC scheme for MIMO TWRCs. We first establish a novel joint channel decomposition to characterize the mutual orthogonality of the channel directions of the two users impinging upon the relay. Based on this decomposition, the overall receive-signal space at the relay is divided into two orthogonal subspaces. In one subspace, the channel directions of one user are orthogonal (or close to orthogonal) to those of the other user. In this subspace, the corresponding messages of the two users are completely decoded, hence referred to as *completely decoded messages*. In the other subspace, the channel directions of the two users are parallel or close to parallel. In this subspace, linear functions of the corresponding user messages are directly computed, without completely decoding the individual messages. These linear functions of the user messages are referred to as *PNC messages*. The completely decoded and PNC messages are jointly re-encoded at the relay, and then forwarded to the two users. Following this, each user recovers the message from the other user by extracting their own message.

Manuscript received May 28, 2012; revised May 20, 2013; accepted May 30, 2013. Date of publication June 28, 2013; date of current version September 11, 2013. This work was supported in part by grants from the University Grants Committee of the Hong Kong (Project No. 418712 and AoE/E-02/08) and in part by a CSIRO OCE Postdoctoral Fellowship.

X. Yuan is with the Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong (e-mail: xjyuan@inc.cuhk.edu.hk).

T. Yang was with the CSIRO ICT Centre, Marsfield, N.S.W. 2122, Australia. He is now with the School of Electrical Engineering and Telecommunications, University of New South Wales, Australia (e-mail: yangtom0403@gmail.com).

I. B. Collings is with the CSIRO ICT Centre, Marsfield, N.S.W. 2122, Australia (e-mail: iain.collings@csiro.au).

Communicated by A. Lozano, Associate Editor for Communications.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2013.2271632

We derive the achievable rates of the proposed space-division-based network coding scheme for MIMO TWRCs. We analytically show that, in the high-SNR regime, the proposed scheme can achieve the sum-capacity of the MIMO TWRC within $\min\{n_A, n_B\}\log(5/4)$ bits, or $\frac{1}{2}\log(5/4) \approx 0.161$ bit/user-antenna, for any antenna setup and any channel realization. We further consider a large-system analysis to derive the average achievable sum-rate of the proposed scheme in Rayleigh fading MIMO TWRCs. We show that in the high-SNR regime, the average gap between the achievable sum-rate of our proposed scheme and the cut-set bound (serving as a sum-capacity upper bound) is at most 0.053 bit/relay-antenna, which occurs at the antenna configurations of $n_A = n_B = \frac{1}{2}n_R$. For all other configurations, the proposed scheme performs even closer to the cut-set bound. Particularly, as the ratio n_A/n_R (or n_B/n_R) tends to 0 or 1, the gap to the cut-set bound vanishes. Moreover, numerical results demonstrate that the proposed scheme significantly outperforms the existing schemes across the full SNR range of practical interest.

II. Preliminaries

A. Notation

The following notation is used throughout this paper. Scalars are denoted by lowercase regular letters, vectors by lowercase bold letters, and matrices by uppercase bold letters. For any matrix $\mathbf{M}$, $\mathbf{M}^T$ and $\mathbf{M}^\dagger$ denote the transpose and the Hermitian transpose, respectively; $|\mathbf{M}|$ denotes the determinant; and $\|\mathbf{M}\|_F$ denotes the Frobenius norm; $\mathcal{C}(\mathbf{M})$ denotes the column space. $\mathbb{R}^{n\times m}$ and $\mathbb{C}^{n\times m}$ denote the $(n \times m)$-dimensional real space and complex space, respectively; $\log(\cdot)$ denotes the logarithm with base 2; $I(i)$ denotes the indicator function with $I(i) = 1$ for $i = 1$ and $I(i) = 0$ for $i \neq 1$; $[\cdot]^+$ denotes $\max\{\cdot, 0\}$; $\mathrm{sign}(x)$ denotes the sign of x; and $\mathcal{N}_c(\mu, \sigma^2)$ denotes the circularly symmetric complex Gaussian distribution with mean μ and variable σ^2.

B. System Model

In this paper, we consider a discrete memoryless MIMO TWRC in which users A and B exchange information via a relay, as illustrated in Fig. 1. User m is equipped with n_m antennas, $m \in \{A, B\}$, and the relay with n_R antennas. We assume that there is no direct link between the two users. The channel is assumed to be flat-fading and quasi-static, i.e., the channel coefficients remain unchanged during each round of information exchange. The channel matrix from user m to the relay is denoted by $\mathbf{H}_{mR} \in \mathbb{C}^{n_R\times n_m}$, and the channel matrix from the relay to user m is denoted by $\mathbf{H}_{Rm} \in \mathbb{C}^{n_m\times n_R}, m \in \{A, B\}$. We assume that the channel matrices are always of full column or row rank, whichever is smaller.[1] As in [5], [6], [9], [10], [16], and [17], we also assume that these channel matrices are globally known by both users as well as by the relay. In practice, the channel state information (CSI) is usually not known perfectly; however, accurate CSI can reasonably be estimated at the receiver side via training and channel estimation, while accurate CSI estimates at the transmitter side may be more difficult to acquire due to delay in feeding back CSI or limited feedback. We will briefly discuss the realization of this assumption and the impact of imperfect CSI later in Section V; cf., Remark 8.

[1]It can be shown that this full-rank assumption holds for randomly generated channel matrices with probability one.

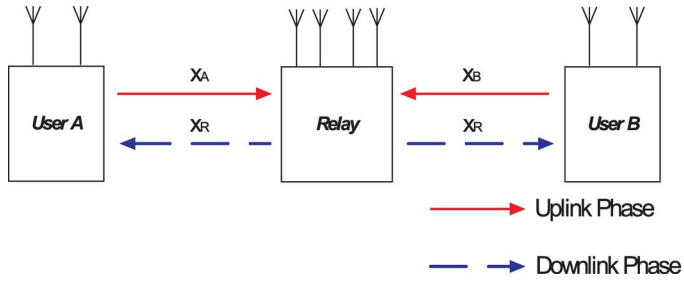

Fig. 1. Configuration of a MIMO TWRC.

The system operates in a half-duplex mode. Two time-slots are employed for each round of information exchange. Following the convention in [16]–[18], we assume that the two time-slots have the same duration. The extension of our results to the case of unequal durations is straightforward.

In the first time-slot (referred to as the *uplink phase*), the two users transmit to the relay simultaneously and the relay remains silent. The transmit signal matrix at user m is denoted by $\mathbf{X}_m \in \mathbb{C}^{n_m\times T}$, $m \in \{A, B\}$, where T is the number of channel uses in one time-slot. Each column of $\mathbf{X}_m$ denotes the signal vector transmitted by the n_m antennas in one channel use. The average power at each user is constrained as $\frac{1}{T}E\left[\|\mathbf{X}_m\|_F^2\right] \leq P_m$, $m \in \{A, B\}$. The received signal at the relay is denoted by $\mathbf{Y}_R \in \mathbb{C}^{n_R\times T}$ with

$$\mathbf{Y}_R = \mathbf{H}_{AR}\mathbf{X}_A + \mathbf{H}_{BR}\mathbf{X}_B + \mathbf{Z}_R, \tag{1}$$

where $\mathbf{Z}_R \in \mathbb{C}^{n_R\times T}$ denotes the additive white Gaussian noise (AWGN) at the relay. We assume that the elements of $\mathbf{Z}_R$ are independent and identically drawn from $\mathcal{N}_c(0, N_0)$. Upon receiving $\mathbf{Y}_R$, the relay generates a signal matrix $\mathbf{X}_R \in \mathbb{C}^{n_R\times T}$. In the second time-slot (referred to as the *downlink phase*), $\mathbf{X}_R$ is broadcast to the two users. The average power at the relay is constrained as $\frac{1}{T}E\left[\|\mathbf{X}_R\|_F^2\right] \leq P_R$. The signal matrix received by user m is denoted by $\mathbf{Y}_m \in \mathbb{C}^{n_m\times T}$, $m \in \{A, B\}$, with

$$\mathbf{Y}_m = \mathbf{H}_{Rm}\mathbf{X}_R + \mathbf{Z}_m, \quad m \in \{A, B\}, \tag{2}$$

where, without loss of generality, $\mathbf{Z}_m$ represents the AWGN matrix at user m, with the elements independently drawn from $\mathcal{N}_c(0, N_0)$.

C. Definition of Achievable Rates

For the system model described above, the message of user m is denoted by $W_m \in \{1, 2, \ldots, 2^{2TR_m}\}$. The cardinality of W_m is given by 2^{2TR_m}, where the factor of $2T$ is because each round of information exchange consists of two length-T time-slots. At user A, the estimated message of user B, denoted by $\hat{W}_B$, is obtained from the received signal $\mathbf{Y}_A$ and the perfect knowledge of the self-message W_A. The decoding operation at user B is likewise. The error probability is defined as $P_e \triangleq \Pr\{\hat{W}_A \neq W_A \text{ or } \hat{W}_B \neq W_B\}$. A rate-pair (R_A, R_B) is achievable if the error probability P_e vanishes as T tends to infinity. The achievable rate-region of a scheme is defined as the closure of all possible achievable rate-pairs.

D. Capacity Upper Bound

This section briefly describes a capacity upper bound of the MIMO TWRC. Let $\mathbf{Q}_m \triangleq \frac{1}{T}E\left[\mathbf{X}_m\mathbf{X}_m^\dagger\right]$, $m \in \{A,B,R\}$, be the input covariance matrices. For given $\{\mathbf{Q}_A,\mathbf{Q}_B,\mathbf{Q}_R\}$ satisfying $\mathrm{tr}(\mathbf{Q}_m) \le P_m,\ m \in \{A,B,R\}$, the achievable rate-pair (R_A,R_B) of the MIMO TWRC is upper bounded as [16]

$$R_A \le \min\left\{R_A^{\mathrm{UL}}(\mathbf{Q}_A), R_A^{\mathrm{DL}}(\mathbf{Q}_R)\right\} \tag{3a}$$

$$R_B \le \min\left\{R_B^{\mathrm{UL}}(\mathbf{Q}_B), R_B^{\mathrm{DL}}(\mathbf{Q}_R)\right\} \tag{3b}$$

where

$$R_m^{\mathrm{UL}}(\mathbf{Q}_m) = \frac{1}{2}\log\left|\mathbf{I}_{n_R} + \frac{1}{N_0}\mathbf{H}_{mR}\mathbf{Q}_m\mathbf{H}_{mR}^\dagger\right|, \quad m\in\{A,B\}, \tag{4a}$$

$$R_A^{\mathrm{DL}}(\mathbf{Q}_R) = \frac{1}{2}\log\left|\mathbf{I}_{n_B} + \frac{1}{N_0}\mathbf{H}_{RB}\mathbf{Q}_R\mathbf{H}_{RB}^\dagger\right|, \tag{4b}$$

$$R_B^{\mathrm{DL}}(\mathbf{Q}_R) = \frac{1}{2}\log\left|\mathbf{I}_{n_A} + \frac{1}{N_0}\mathbf{H}_{RA}\mathbf{Q}_R\mathbf{H}_{RA}^\dagger\right|. \tag{4c}$$

Here, the superscripts "UL" and "DL" represent *uplink* and *downlink*, respectively, and the factor of $\frac{1}{2}$ is due to the two time-slots used for each round of information exchange.

A capacity-region outer bound is defined as the closure of the upper bound rate-pairs in (3). This outer bound can be determined by optimizing $\mathbf{Q}_A$, $\mathbf{Q}_B$, and $\mathbf{Q}_R$, as detailed in [16] and [17]. The goal of this paper is to develop a communication strategy that can approach this outer bound.

III. Relaying Strategies for TWRCs With Single-Antenna Users

In this section, we consider TWRCs with single-antenna users and a multiantenna relay, i.e., $n_A = n_B = 1$ and $n_R \ge 1$. The results developed in this section will be used for the study of general MIMO TWRCs in the later sections.

A. Relaying Strategies: CD Versus PNC

For the case of single-antenna users, the channel model of the uplink phase in (1) can be simplified as

$$\mathbf{Y}_R = \mathbf{h}_{AR}\mathbf{x}_A^T + \mathbf{h}_{BR}\mathbf{x}_B^T + \mathbf{Z}_R \tag{5}$$

where $\mathbf{h}_{mR} \in \mathbb{C}^{n_R\times 1}$ is the reduced version of $\mathbf{H}_{mR}$, and $\mathbf{x}_m \in \mathbb{C}^{T\times 1}$ is the transmit signal vector of user m, with the ith entry of $\mathbf{x}_m$ being the signal transmitted at the ith time interval, $m \in \{A,B\}$.

We have two competing relay strategies described as follows. From (5), the signal direction of user m impinging upon the relay is given by $\mathbf{h}_{mR},\ m \in \{A,B\}$. On one hand, if $\mathbf{h}_{AR}$ and $\mathbf{h}_{BR}$ turn out to be orthogonal, both messages of the two users can be decoded free of interference from each other. The recovered messages of the two users are then network-coded, e.g., using algebraic operations [1], and forwarded to the two users. We refer to this first strategy as a complete decoding (CD) strategy. On the other hand, if $\mathbf{h}_{AR}$ and $\mathbf{h}_{BR}$ turn out to be parallel (i.e., in the same direction), then it is advantageous to compute a linear function of $\mathbf{x}_A$ and $\mathbf{x}_B$, referred to as a network-coded message, instead of completely decoding both $\mathbf{x}_A$ and $\mathbf{x}_B$. We refer to this second strategy as a *PNC* strategy.

In general, the following strategy can be adopted: if $\mathbf{h}_{AR}$ and $\mathbf{h}_{BR}$ tend to be orthogonal, the CD strategy is applied; if $\mathbf{h}_{AR}$ and $\mathbf{h}_{BR}$ tend to be parallel, the PNC strategy is applied. The selection between these two strategies is based on their achievable rates, as described below.

B. Achievable Rates of the CD Strategy

For the CD strategy, the uplink channel in (5) becomes a multiple-access channel (MAC). Let $R_m^{\mathrm{CD}}, m \in \{A,B\}$, be the rate of user m for the CD strategy. Then, the uplink rate-region of the CD strategy, denoted by $\mathcal{R}^{\mathrm{CD}}$, is given by

$$R_A^{\mathrm{CD}} + R_B^{\mathrm{CD}} \le \frac{1}{2}\log\left|\mathbf{I}_{n_R} + \sum_{m\in\{A,B\}}\frac{P_m}{N_0}\mathbf{h}_{mR}\mathbf{h}_{mR}^\dagger\right| \tag{6a}$$

$$R_m^{\mathrm{CD}} \le \frac{1}{2}\log\left(1+\frac{P_m}{N_0}\mathbf{h}_{mR}^\dagger\mathbf{h}_{mR}\right),\ m\in\{A,B\} \tag{6b}$$

which follows from the well-known MAC capacity region [27]–[29].

C. Achievable Rates of the PNC Strategy

For the PNC strategy, it is required that the two user-signals lie in the same spatial direction [5], [6]. This is not guaranteed here due to the availability of multiple antennas at the relay. To implement PNC, we employ a projection-based method as follows. We first project the signals from the two users onto a common direction, denoted by a unit vector $\mathbf{p} \in \mathbb{C}^{n_R\times 1}$. The choice of $\mathbf{p}$ will be discussed momentarily. This projection operation creates an aligned scalar channel given by

$$\mathbf{p}^\dagger\mathbf{Y}_R = \mathbf{p}^\dagger\mathbf{h}_{AR}\mathbf{x}_A^T + \mathbf{p}^\dagger\mathbf{h}_{BR}\mathbf{x}_B^T + \mathbf{p}^\dagger\mathbf{Z}_R \tag{7}$$

with the effective channel coefficients given by $\mathbf{p}^\dagger\mathbf{h}_{AR}$ and $\mathbf{p}^\dagger\mathbf{h}_{BR}$. Let R_m^{PNC} be the rate of user m, $m \in \{A,B\}$. It was shown in [5] that, with nested lattice coding [25], the following rates are achievable over the scalar TWRC in (7):

$$R_m^{\mathrm{PNC}} = \frac{1}{2}\left[\log\left(\frac{Q_m\left|\mathbf{p}^\dagger\mathbf{h}_{mR}\right|^2}{Q_A\left|\mathbf{p}^\dagger\mathbf{h}_{AR}\right|^2 + Q_B\left|\mathbf{p}^\dagger\mathbf{h}_{BR}\right|^2 + \dfrac{Q_m\left|\mathbf{p}^\dagger\mathbf{h}_{mR}\right|^2}{N_0}}\right)\right]^+, \ m\in\{A,B\}, \tag{8}$$

where $Q_m = \frac{1}{T}E[\mathbf{x}_m^\dagger\mathbf{x}_m]$ represents the transmission power of user m. Note that the above rates are achieved by decoding the *lattice-modulo* of $\mathbf{p}^\dagger\mathbf{h}_{AR}\mathbf{x}_A^T + \mathbf{p}^\dagger\mathbf{h}_{BR}\mathbf{x}_B^T$ (referred to as the codeword sum), instead of directly decoding the codeword sum [5]. However, the latter is required to support successive interference cancellation (SIC), as in the case of MIMO TWRCs

where multiple spatial-stream pairs need to be successively decoded. It was shown in [16] that, if the codeword sum is decoded, the achievable rates $(R_A^{\rm PNC}, R_B^{\rm PNC})$ are given by

$$R_m^{\rm PNC} = \frac{1}{2}\left[\log\left(\frac{Q_m\left|\mathbf{p}^\dagger \mathbf{h}_{mR}\right|^2}{N_0}\right)\right]^+, \quad m \in \{A,B\}. \tag{9}$$

Notice that (8) and (9) become identical at high SNR.

The uplink rate-region of the PNC scheme is given by

$$\mathcal{R}^{\rm PNC} \triangleq \big\{(R_A,R_B)|R_m \le R_m^{\rm PNC},\ Q_m \le P_m,\ m\in\{A,B\},\ \mathbf{p}^\dagger\mathbf{p}=\mathbf{1}\big\}. \tag{10}$$

The boundary of $\mathcal{R}^{\rm PNC}$ can be found by optimizing Q_A, Q_B, and $\mathbf{p}$, as detailed below.

We start with the rate-pair given in (9). As the achievable rate-region is convex, the boundary points of $\mathcal{R}^{\rm PNC}$ can be determined by solving the weighted sum-rate maximization problem:

$$\begin{aligned} \text{maximize} \quad & w_A R_A^{\rm PNC} + w_B R_B^{\rm PNC} && \text{(11a)}\\ \text{subject to} \quad & \|\mathbf{p}\|=1,\ Q_m \le P_m,\ m\in\{A,B\} && \text{(11b)} \end{aligned}$$

where w_A and w_B are arbitrary nonnegative weighting coefficients.[2] By inspecting (9), the maximum of (11) is achieved at $Q_m = P_m, m \in \{A,B\}$. Thus, we only need to optimize $\mathbf{p}$. The optimal $\mathbf{p}$ to (11) is given in Appendix A. Particularly, for the sum-rate case, i.e., $w_A = w_B = 1$, the optimal projection direction $\widetilde{\mathbf{p}}$ is just the angular bisector of $\widetilde{\mathbf{h}}_{AR}$ and $\widetilde{\mathbf{h}}_{BR}$ if $\widetilde{\mathbf{h}}_{AR}^T\widetilde{\mathbf{h}}_{BR} > 0$, or the angular bisector of $\widetilde{\mathbf{h}}_{AR}$ and $-\widetilde{\mathbf{h}}_{BR}$ if $\widetilde{\mathbf{h}}_{AR}^T\widetilde{\mathbf{h}}_{BR} < 0$. By varying the ratio of w_A/w_B, $\mathcal{R}^{\rm PNC}$ can be determined.

We now briefly discuss the treatment for using the rate-pair (8) in formulating (11). The problem then becomes more complicated due to the coupling between the two rates (since both Q_A and Q_B are present in both rate expressions in (8)). At high SNR, in particular, as (8) and (9) are asymptotically identical, we see that the high-SNR optimal projection direction for the sum-rate case is still the angular bisector of the channel directions of the two users.

D. Overall Scheme

We now present an overall scheme which exploits the benefits of both CD and PNC strategies. Specifically, in the uplink, the proposed scheme selects the CD strategy, the PNC strategy, or their combination by using time sharing. Then, the uplink achievable rate-region, denoted by $\mathcal{R}^{\rm UL}$, is given by the convex hull of $\mathcal{R}^{\rm CD}$ and $\mathcal{R}^{\rm PNC}$. An example of $\mathcal{R}^{\rm UL}$ can be found in Fig. 2.

For the downlink phase, the achievable rate-region is determined as follows. For CD, the relay jointly re-encodes the decoded messages $\mathbf{x}_A$ and $\mathbf{x}_B$, and forwards the resulting codeword to the two users in the downlink. For PNC decoding, the relay forwards the lattice-modulo of $\mathbf{p}^\dagger\mathbf{h}_{AR}\mathbf{x}_A^T + \mathbf{p}^\dagger\mathbf{h}_{BR}\mathbf{x}_B^T$, referred to as the network-coded message, to the two users. Then, each user recovers the message of the other user with the help of perfect knowledge of the self-message. From [16]–[18], the downlink rate-regions for the two strategies are the same given by

$$\mathcal{R}^{\rm DL} \triangleq \big\{(R_A,R_B)|R_A \le R_A^{\rm DL}, R_B \le R_B^{\rm DL}\big\} \tag{12}$$

with

$$R_A^{\rm DL} = \frac{1}{2}\log\left(1+\frac{P_R}{N_0}\mathbf{h}_{RB}^\dagger\mathbf{h}_{RB}\right) \tag{13a}$$

$$R_B^{\rm DL} = \frac{1}{2}\log\left(1+\frac{P_R}{N_0}\mathbf{h}_{RA}^\dagger\mathbf{h}_{RA}\right). \tag{13b}$$

Finally, an achievable rate-region of the overall scheme is the intersection of $\mathcal{R}^{\rm UL}$ and $\mathcal{R}^{\rm DL}$.

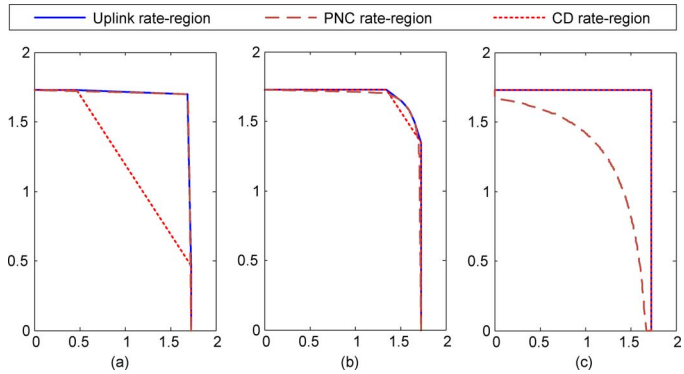

Fig. 2. Uplink rate-regions of the TWRCs with single-antenna users. $\mathbf{h}_A = [1,0]^T$ and $\mathbf{h}_B = [\cos\theta, \sin\theta]^T$. The channel SNR $= 1/N_0 = 10$ dB; the horizontal axes represent the rate of user A; the vertical axes represent the rate of user B; the unit is bit per channel use. It is seen that the PNC strategy outperforms the CD strategy for $\theta = 1°$, and the opposite is true for $\theta = 89°$. The time-sharing combination of the two strategies in general yields a larger rate-region, as seen in the figure of $\theta = 45°$. (a) $\theta = 1°$. (b) $\theta = 45°$.(c) $\theta = 89°$.

IV. Space-Division Approach for MIMO TWRCs

In this section, we consider a general MIMO TWRC with $n_A \ge 1$, $n_B \ge 1$. As aforementioned, the case of $n_R \le \max(n_A, n_B)$ has been well studied in [16], where the asymptotic capacity was established. However, the case of $n_R > \max(n_A, n_B)$ is of more practical importance, and all existing schemes perform well away from the capacity in this case. In what follows, we always assume $n_R \ge \max(n_A, n_B)$ and propose a novel space-division approach which brings about significant performance improvement over the existing schemes.

A. Preliminaries

What motivates the proposed space-division approach is the following property of $\mathbf{H}_{AR}$ and $\mathbf{H}_{BR}$. Denote by $\mathcal{C}(\mathbf{H}_{AR})$ and $\mathcal{C}(\mathbf{H}_{BR})$ the column spaces of the uplink channel matrices $\mathbf{H}_{AR}$ and $\mathbf{H}_{BR}$, respectively. In general, we can partition the

[2] In fact, it is sufficient to describe w_A and w_B using one parameter, since only the ratio $\frac{w_A}{w_B}$ matters in solving (11). However, we keep using both w_A and w_B for notational convenience, as seen later in this paper.

column space $\mathcal{C}(\mathbf{H}_{AR}) \in \mathbb{C}^{n_R}$ as the direct sum[3] of three orthogonal subspaces: a subspace $\mathcal{S}_{A\|B}$ that is parallel to $\mathcal{C}(\mathbf{H}_{BR})$, i.e., any vector in $\mathcal{S}_{A\|B}$ belongs to $\mathcal{C}(\mathbf{H}_{BR})$; a subspace $\mathcal{S}_{A\nparallel B}$ that is neither parallel nor orthogonal to $\mathcal{C}(\mathbf{H}_{BR})$; and a subspace $\mathcal{S}_{A\perp B}$ that is orthogonal to $\mathcal{C}(\mathbf{H}_{BR})$. Similarly, $\mathcal{C}(\mathbf{H}_{BR})$ is the direct sum of three orthogonal subspaces $\mathcal{S}_{B\|A}, \mathcal{S}_{B\nparallel A}$, and $\mathcal{S}_{B\perp A}$. Note that $\mathcal{S}_{A\|B} = \mathcal{S}_{B\|A}$ since both represent the *common space* of $\mathcal{C}(\mathbf{H}_{AR})$ and $\mathcal{C}(\mathbf{H}_{BR})$.

In $\mathcal{S}_{A\|B}$, the signal directions of the two users can be efficiently aligned to a common set of directions, providing a platform to carry out PNC, as in [16]–[18]. On the other hand, in $\mathcal{S}_{A\perp B}$ and $\mathcal{S}_{B\perp A}$, the two users do not interfere with each other; hence, the CD strategy can be employed. The above treatments are similar to those for the case of single-antenna users, as discussed in the preceding section. What remains is the treatment for the signals in $\mathcal{S}_{A\nparallel B}$ and $\mathcal{S}_{B\nparallel A}$ that are neither parallel nor orthogonal. Heuristically, some channel directions in $\mathcal{S}_{A\nparallel B}$ and $\mathcal{S}_{B\nparallel A}$ may be nearly parallel to each other. For these channel directions, the PNC strategy is preferable for the related spatial streams. On the other hand, some channel directions in $\mathcal{S}_{A\nparallel B}$ and $\mathcal{S}_{B\nparallel A}$ may be nearly orthogonal to each other. Then, the CD strategy is preferable. The main challenge lies in how to identify those nearly parallel/orthogonal channel directions. To this end, we next propose a new joint channel decomposition of $\mathbf{H}_{AR}$ and $\mathbf{H}_{BR}$. To begin with, we present some useful lemmas, with the proofs given in Appendix B.

Let the compact singular value decomposition of $\mathbf{H}_{mR}$ be

$$\mathbf{H}_{mR} = \mathbf{U}_m \mathbf{\Delta}_m \mathbf{V}_m^\dagger,\ m \in \{A, B\} \tag{14}$$

where $\mathbf{U}_m$ is an $n_R \times n_m$ matrix with orthonormal columns, i.e., $\mathbf{U}_m^\dagger \mathbf{U}_m = \mathbf{I}_{n_m}$. Denote by λ_i the ith eigenvalue of the matrix $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$, and by $\mathbf{u}_i$ the corresponding eigenvector with unit length. Without loss of generality, we arrange $\{\lambda_i\}$ in the descending order. By definition, we have

$$\left(\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger\right)\mathbf{u}_i = \lambda_i \mathbf{u}_i. \tag{15}$$

The eigenvalues $\{\lambda_i\}$ are valued between 0 and 2, since the eigenvalues of $\mathbf{U}_m\mathbf{U}_m^\dagger$ are either 1 or 0, $m \in \{A, B\}$. We are interested in the following four cases of λ_i: (a) $\lambda_i = 2$; (b) $1 < \lambda_i < 2$; (c) $\lambda_i = 1$; and (d) $0 < \lambda_i < 1$.

For case (a), $\lambda_i = 2$ implies that

$$\mathbf{U}_A\mathbf{U}_A^\dagger \mathbf{u}_i = \mathbf{u}_i \text{ and } \mathbf{U}_B\mathbf{U}_B^\dagger \mathbf{u}_i = \mathbf{u}_i.$$

Thus, $\mathbf{u}_i$ lies in the common space of $\mathcal{C}(\mathbf{U}_A)$ and $\mathcal{C}(\mathbf{U}_B)$, or equivalently, $\mathbf{u}_i$ is in the common space $\mathcal{S}_{A\|B}$.

For case (c), we have the following lemma.

Lemma 1: If $\lambda_i = 1$, then the corresponding $\mathbf{u}_i$ satisfies either

$$\mathbf{U}_A\mathbf{U}_A^\dagger \mathbf{u}_i = \mathbf{u}_i \text{ and } \mathbf{U}_B\mathbf{U}_B^\dagger \mathbf{u}_i = \mathbf{0} \tag{16a}$$

or

$$\mathbf{U}_A\mathbf{U}_A^\dagger \mathbf{u}_i = \mathbf{0} \text{ and } \mathbf{U}_B\mathbf{U}_B^\dagger \mathbf{u}_i = \mathbf{u}_i. \tag{16b}$$

From Lemma 1, $\mathbf{u}_i$ in case (c) is in $\mathcal{S}_{A\perp B}$ (or $\mathcal{S}_{B\perp A}$) which is orthogonal to the space spanned by $\mathbf{H}_{BR}$ (or $\mathbf{H}_{AR}$).

We next show that the eigenvalues in cases (b) and (d) appear in a pairwise manner. Denote

$$\mathbf{l}_{m;i} = \mathbf{U}_m \left(\mathbf{U}_m^\dagger \mathbf{U}_m\right)^{-1} \mathbf{U}_m^\dagger \mathbf{u}_i = \mathbf{U}_m \mathbf{U}_m^\dagger \mathbf{u}_i. \tag{17}$$

Note that $\mathbf{l}_{m;i}$ is the projection of vector $\mathbf{u}_i$ onto the column space of $\mathbf{U}_m$. From (15), we obtain

$$\mathbf{u}_i = \frac{1}{\lambda_i}\left(\mathbf{l}_{A;i} + \mathbf{l}_{B;i}\right). \tag{18}$$

The above implies that $\mathbf{u}_i$, $\mathbf{l}_{A;i}$ and $\mathbf{l}_{B;i}$ lie in the same two-dimension plane (denoted by $\mathcal{S}_i$). We have the following lemmas.

Lemma 2: For any λ_i in case (b), the corresponding $\mathbf{u}_i$ is the angular bisector of $\mathbf{l}_{A;i}$ and $\mathbf{l}_{B;i}$, i.e.,

$$\|\mathbf{l}_{A;i}\|^2 = \mathbf{u}_i^\dagger \mathbf{l}_{A;i} = \mathbf{u}_i^\dagger \mathbf{l}_{B;i} = \|\mathbf{l}_{B;i}\|^2. \tag{19}$$

Lemma 3: For any $\lambda_i \in (1, 2)$ (as in case (b)), $\lambda_i' = 2 - \lambda_i$ is also an eigenvalue of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$, and the corresponding unit-length eigenvector is given by

$$\mathbf{u}_i' = \frac{1}{\sqrt{\lambda_i\lambda_i'}}\left(\mathbf{l}_{A;i} - \mathbf{l}_{B;i}\right). \tag{20}$$

Lemma 4: The subspace $\mathcal{S}_i$ spanned by $\mathbf{l}_{A;i}$ and $\mathbf{l}_{B;i}$ is orthogonal to $\mathcal{S}_j$, for any $j \neq i$.

From Lemma 3, we see that the eigenvalues in case (b) and case (d) indeed appear in a pairwise manner. As a result, the number of eigenvalues in $(1, 2)$ is the same as that in $(0, 1)$.

B. Joint Channel Decomposition

To develop the new joint channel decomposition of $\mathbf{H}_{AR}$ and $\mathbf{H}_{BR}$, we first give an overall picture of the eigenvalues and eigenvectors of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$ based on Lemmas 1–4. Without loss of generality, let k be the number of eigenvalues of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$ equal to 2 [for case (a)]; l be the number of eigenvalues $\in (1, 2)$ [for case (b)]; d_A be the number of eigenvalues equal to 1 with $\{\mathbf{U}_A\mathbf{U}_A^\dagger \mathbf{u}_i = \mathbf{u}_i, \mathbf{U}_B\mathbf{U}_B^\dagger \mathbf{u}_i = \mathbf{0}\}$ [for case (c.1)]; d_B be the number of eigenvalues equal to 1 with $\{\mathbf{U}_A\mathbf{U}_A^\dagger \mathbf{u}_i = \mathbf{0}, \mathbf{U}_B\mathbf{U}_B^\dagger \mathbf{u}_i = \mathbf{u}_i\}$ [for case (c.2)].

Denote the k eigenvalues in case (a) by $\lambda_1, \ldots, \lambda_k$, and the corresponding orthogonal eigenvectors by $\mathbf{u}_1, \ldots, \mathbf{u}_k$. Also denote the l eigenvalues in case (b) by $\lambda_{k+1}, \ldots, \lambda_{k+l}$ in the descending order, and the corresponding eigenvectors by $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_{k+l}$. From Lemma 3, the eigenvalues in cases (b) and (d) appear in a pairwise manner. We denote l eigenvalues in case (d) by $\lambda_{k+1}', \ldots, \lambda_{k+l}'$ in the descending order, and the corresponding eigenvectors by $\mathbf{u}_{k+1}', \ldots, \mathbf{u}_{k+l}'$. Furthermore, we denote the d_A orthogonal eigenvectors in case (c.1) by $\mathbf{u}_{k+l+1}, \cdots, \mathbf{u}_{k+l+d_A}$, and the d_B orthogonal eigenvectors in case (c.2) by $\mathbf{u}_{k+l+d_A+1}, \ldots, \mathbf{u}_{k+l+d_A+d_B}$. Define

$$\mathbf{U} \triangleq \big[\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \mathbf{u}_{k+1}', \ldots, \mathbf{u}_{k+l}, \mathbf{u}_{k+l}', \mathbf{u}_{k+l+1}, \ldots, \mathbf{u}_{k+l+d_A+d_B}\big]. \tag{21}$$

[3]Let $\mathcal{S}$ be a vector space, and let $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$ be subspaces of $\mathcal{S}$. $\mathcal{S}$ is defined to be a direct sum of $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$ when $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$ are mutually orthogonal and for every vector $\mathbf{x}$ in $\mathcal{S}$, there is $\mathbf{x}_i$ in $\mathcal{S}_i$, $i = 1, 2, \ldots, n$, such that $\mathbf{x} = \sum_{i=1}^{n} \mathbf{x}_i$.

It can be readily verified that $\mathbf{U}$ satisfies $\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}_{k+2l+d_A+d_B}$, i.e., $\mathbf{U}$ has orthonormal columns.

We are now ready to present the joint channel decomposition as follows.

Theorem 1: The channel matrices $\mathbf{H}_{AR}$ and $\mathbf{H}_{BR}$ can be jointly decomposed as

$$\mathbf{H}_{mR} = \mathbf{U}\mathbf{D}_m\mathbf{G}_m,\ m \in \{A, B\} \tag{22}$$

where $\mathbf{G}_m \in \mathbb{C}^{n_m \times n_m}$ is a square matrix, and $\mathbf{D}_m \in \mathbb{C}^{(n_A+n_B-k)\times n_m}, m \in \{A, B\}$, are defined as

$$\mathbf{D}_A = \begin{bmatrix} \mathbf{I}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_A & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{d_A} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \text{ and } \mathbf{D}_B = \begin{bmatrix} \mathbf{I}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_B & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{d_B} \end{bmatrix} \tag{23a}$$

with

$$\mathbf{E}_m = \begin{bmatrix} \mathbf{e}_{m;k+1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_{m;k+2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{e}_{m;k+l} \end{bmatrix} \in \mathbb{C}^{2l\times l}, \tag{23b}$$

$$\mathbf{e}_{A;k+i} = \begin{bmatrix} \sqrt{\frac{\lambda_{k+i}}{2}} \\ \sqrt{\frac{2-\lambda_{k+i}}{2}} \end{bmatrix} \text{ and } \mathbf{e}_{B;k+i} = \begin{bmatrix} \sqrt{\frac{\lambda_{k+i}}{2}} \\ -\sqrt{\frac{2-\lambda_{k+i}}{2}} \end{bmatrix}, \tag{23c}$$

$$1 < \lambda_{k+i} < 2,\ \text{for } i = 1, \ldots, l. \tag{23d}$$

Proof: Let

$$\mathbf{U}'_A = \left[\mathbf{u}_1, \ldots, \mathbf{u}_k, \frac{\mathbf{l}_{A;k+1}}{\|\mathbf{l}_{A;k+1}\|}, \ldots, \frac{\mathbf{l}_{A;k+l}}{\|\mathbf{l}_{A;k+l}\|}, \mathbf{u}_{k+l+1}, \ldots, \mathbf{u}_{k+l+d_A}\right] \tag{24a}$$

and

$$\mathbf{U}'_B = \left[\mathbf{u}_1, \ldots, \mathbf{u}_k, \frac{\mathbf{l}_{B;k+1}}{\|\mathbf{l}_{B;k+1}\|}, \ldots, \frac{\mathbf{l}_{B;k+l}}{\|\mathbf{l}_{B;k+l}\|}, \mathbf{u}_{k+l+d_A+1}, \ldots, \mathbf{u}_{k+l+d_A+d_B}\right]. \tag{24b}$$

In the above, $\mathbf{u}_1, \ldots, \mathbf{u}_k$ are the eigenvectors in case (a); $\mathbf{u}_{k+l+1}, \ldots, \mathbf{u}_{k+l+d_A}$ are the eigenvectors in case (c) satisfying $\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{u}_i = \mathbf{u}_i$, for $i = k+l+1, \ldots, k+l+d_A$; $\mathbf{u}_{k+l+d_A+1}, \ldots, \mathbf{u}_{k+l+d_A+d_B}$ are the eigenvectors in case (c) satisfying $\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i = \mathbf{u}_i$, for $i = k+l+d_A+1, \ldots, k+l+d_A+d_B$. Then, with Lemmas 2 and 3, it can be verified that $\mathbf{D}_m$ in (23a) satisfies

$$\mathbf{U}'_m = \mathbf{U}\mathbf{D}_m, \quad m \in \{A, B\}. \tag{25}$$

Recall that $\mathbf{l}_{m;k+i} \in \mathcal{S}_{k+i}$ for $i = 1, \ldots, l$, where $\mathcal{S}_{k+i}$ is the 2-D subspace spanned by $\mathbf{u}_{k+i}$ and $\mathbf{u}'_{k+i}$. Then, from Lemma 4 and the orthogonality of the eigenvectors, the columns of $\mathbf{U}'_m$ are orthonormal. Together with the fact that all columns of $\mathbf{U}'_m$ lie in the column space of $\mathbf{U}_m$ (and so in the column space of $\mathbf{H}_m$), we see that $\mathbf{H}_m$ and $\mathbf{U}'_m$ share the same column space. Thus, there exists an $n_m \times n_m$ square matrix $\mathbf{G}_m$ such that

$$\mathbf{H}_{mR} = \mathbf{U}'_m\mathbf{G}_m. \tag{26}$$

Combining (25) and (26), we obtain

$$\mathbf{H}_{mR} = \mathbf{U}\mathbf{D}_m\mathbf{G}_m \tag{27}$$

which completes the proof of Theorem 1. ■

TABLE I
SUBSPACES AND THEIR DIMENSIONS

Subspaces	Dim.	Descriptions
$\mathcal{S}_{A\|\|B}$	k	interception of $\mathcal{C}(\mathbf{H}_{AR})$ and $\mathcal{C}(\mathbf{H}_{BR})$
$\mathcal{S}_{A\nparallel B}$	l	not parallel/orthogonal to $\mathcal{C}(\mathbf{H}_{BR})$
$\mathcal{S}_{B\nparallel A}$	l	not parallel/orthogonal to $\mathcal{C}(\mathbf{H}_{AR})$
$\mathcal{S}_{A\perp B}$	d_A	orthogonal to $\mathcal{C}(\mathbf{H}_{BR})$
$\mathcal{S}_{B\perp A}$	d_B	orthogonal to $\mathcal{C}(\mathbf{H}_{AR})$

Remark 1: We interpret (22) as follows. By inspection of (22), we see that $\mathbf{U}\mathbf{D}_m$ specifies the column space of $\mathbf{H}_m$, i.e., $\mathcal{C}(\mathbf{U}\mathbf{D}_m) = \mathcal{C}(\mathbf{H}_m)$. Note that $\mathbf{U}\mathbf{D}_m$ has orthonormal columns, as $\mathbf{D}_m^\dagger\mathbf{U}^\dagger\mathbf{U}\mathbf{D}_m = \mathbf{I}_{n_m}, m \in \{A, B\}$. Therefore, the columns of $\mathbf{U}\mathbf{D}_m$ give an orthogonal basis of $\mathcal{C}(\mathbf{H}_{mR})$, with the coordinates of $\mathbf{H}_{mR}$ specified in $\mathbf{G}_m$. Moreover, from (21), it can be shown

$$\mathcal{C}(\mathbf{U}) = \mathcal{S}_{A\|B} \cup \mathcal{S}_{k+1} \cup \cdots \cup \mathcal{S}_{k+l} \cup \mathcal{S}_{A\perp B} \cup \mathcal{S}_{B\perp A}$$

where $\mathcal{S}_i = \text{span}(\mathbf{l}_{A;i}, \mathbf{l}_{B;i})$ is defined above Lemma 2. Thus, $\mathcal{C}(\mathbf{U})$ is actually the overall column space of the two channel matrices, i.e., $\mathcal{C}(\mathbf{U}) = \mathcal{C}(\mathbf{H}_{AR}) \cup \mathcal{C}(\mathbf{H}_{BR})$.

Remark 2: The column structures of $\mathbf{U}\mathbf{D}_A$ and $\mathbf{U}\mathbf{D}_B$ are explained as follows. In the first place, we note that $\mathbf{U}\mathbf{D}_A$ and $\mathbf{U}\mathbf{D}_B$ share the same first k columns. Thus, the first k columns of $\mathbf{U}\mathbf{D}_A$ span $\mathcal{S}_{A\|B}$, i.e., the common space of $\mathcal{C}(\mathbf{H}_{AR})$ and $\mathcal{C}(\mathbf{H}_{BR})$. Second, from (23a), the last d_A columns of $\mathbf{U}\mathbf{D}_A$ (obtained from multiplying $\mathbf{U}$ with the third block column of $\mathbf{D}_A$) are orthogonal to $\mathbf{U}\mathbf{D}_B$. Hence, these columns of $\mathbf{U}\mathbf{D}_A$ span the subspace $\mathcal{S}_{A\perp B}$, i.e., the subspace orthogonal to $\mathcal{C}(\mathbf{H}_{BR})$. Third, the remaining l columns of $\mathbf{U}\mathbf{D}_A$ span the subspace $\mathcal{S}_{A\nparallel B}$, by noting the facts that $\mathcal{C}(\mathbf{H}_{AR}) = \mathcal{C}(\mathbf{U}\mathbf{D}_A)$ and that $\mathcal{C}(\mathbf{H}_{AR})$ is the direct-sum of three orthogonal subspaces $\mathcal{S}_{A\|B}$, $\mathcal{S}_{A\nparallel B}$, and $\mathcal{S}_{A\perp B}$. Similarly, the first k columns of $\mathbf{U}\mathbf{D}_B$ span $\mathcal{S}_{A\|B}$, the next l columns span $\mathcal{S}_{B\nparallel A}$, and the last d_B columns span $\mathcal{S}_{B\perp A}$. Recall that $\mathcal{C}(\mathbf{H}_{AR})$ is the direct sum of $\mathcal{S}_{A\|B}$, $\mathcal{S}_{A\nparallel B}$, and $\mathcal{S}_{A\perp B}$, and that $\mathcal{C}(\mathbf{H}_{BR})$ is the direct sum of $\mathcal{S}_{B\|A}$, $\mathcal{S}_{B\nparallel A}$, and $\mathcal{S}_{B\perp A}$. Thus, the dimensions of these subspaces have the following relationship:

$$k + l + d_m = n_m, \quad m \in \{A, B\}. \tag{28}$$

We summarize the geometrical meanings of these subspaces and their dimensions in Table I.

The joint channel decomposition in Theorem 1 identifies the channel directions of the two users based on the degree of orthogonality, as detailed below. Let $\mathbf{v}_{m;i}$ be the ith column of

$\mathbf{UD}_m$, $m \in \{A, B\}$. We refer to $\mathbf{v}_{A;i}$ and $\mathbf{v}_{B;i}$ as the ith channel direction pair. Here, $\mathbf{v}_{A;i}^{\dagger}\mathbf{v}_{B;i} = 1$ means that $\mathbf{v}_{A;i}$ and $\mathbf{v}_{B;i}$ are parallel, and $\mathbf{v}_{A;i}^{\dagger}\mathbf{v}_{B;i} = 0$ means that they are orthogonal. Thus, $\mathbf{v}_{A;i}^{\dagger}\mathbf{v}_{B;i}$ can be regarded as a measure of the degree of orthogonality of $\mathbf{v}_{A;i}$ and $\mathbf{v}_{B;i}$. In the following corollary, the degree of orthogonality of each channel direction pair $(\mathbf{v}_{A;i}, \mathbf{v}_{B;i})$ is determined by the magnitude of λ_i, i.e., the ith eigenvalue of $\mathbf{U}_A\mathbf{U}_A^{\dagger} + \mathbf{U}_B\mathbf{U}_B^{\dagger}$.

Corollary 1: For $i = 1, \ldots, k+l$, the degree of orthogonality of the ith channel direction pair $(\mathbf{v}_{A,i}, \mathbf{v}_{B,i})$ is given by $\mathbf{v}_{A,i}^{\dagger}\mathbf{v}_{B,i} = \lambda_i - 1$.

Proof: For $i = 1, \ldots, k$, we see from (23a) that $\lambda_i = 2$ and $\mathbf{v}_{A;i} = \mathbf{v}_{B;i}$, and so $\mathbf{v}_{A;i}^{\dagger}\mathbf{v}_{B;i} = \lambda_i - 1$. For $i = k+1, \ldots, k+l$, from (23a) and (23b), the ith column of $\mathbf{UD}_m$ is given by

$$\mathbf{v}_{m;i} = [\,\widetilde{\mathbf{u}}_{2i-k-1} \quad \widetilde{\mathbf{u}}_{2i-k}\,]\,\mathbf{e}_{m;i}, \quad m \in \{A, B\}\,, \tag{29}$$

where $\widetilde{\mathbf{u}}_i$ represents the ith column of $\mathbf{U}$. Then, we obtain $\mathbf{v}_{A;k+i}^{\dagger}\mathbf{v}_{B;k+i} = \mathbf{e}_{A;k+i}^{\dagger}\mathbf{e}_{B;k+i} = \lambda_i - 1$, where the first step utilizes the fact that $\mathbf{U}$ has orthonormal columns, and the second step follows from (23c). ■

Corollary 2: For $i = k+l+1, \ldots, n_A$, $\mathbf{v}_{A;i}$ is an eigenvector of $\mathbf{U}_A\mathbf{U}_A^{\dagger} + \mathbf{U}_B\mathbf{U}_B^{\dagger}$ corresponding to $\lambda_i = 1$, and is orthogonal to $\mathcal{C}(\mathbf{H}_{BR})$; for $i = k+l+1, \ldots, n_B$, $\mathbf{v}_{B;i}$ is an eigenvector corresponding to $\lambda_i = 1$ and is orthogonal to $\mathcal{C}(\mathbf{H}_{AR})$.

Remark 3: The above corollaries show that the eigenvalue λ_i is an indicator of the degree of orthogonality of the ith direction pair. In particular, $\lambda_i \approx 2$ means that the two channel directions are close to parallel, and $\lambda_i \approx 1$ means that the two channel directions are close to orthogonal.

Remark 4: For the case of $\max(n_A, n_B) = n_R$, we obtain $k = \min(n_A, n_B)$, implying that the eigenvalues satisfy $\lambda_i = 2$ for $i = 1, \ldots, k$, and $\lambda_i = 1$ for $i = k+1, \ldots, n_R$. In this case, efficient channel alignment techniques have been proposed in [16]–[18] for the implementation of PNC. In what follows, we are mainly interested in the case of $\max(n_A, n_B) < n_R$, i.e., there exist λ_i valued between, but not including, 1 and 2.

C. Space-Division for MIMO Two-Way Relaying

Based on the joint channel decomposition in Theorem 1, we now propose a new space-division approach for MIMO two-way relaying. The main idea is to divide the overall signal space $\mathcal{C}([\mathbf{H}_{AR}\ \mathbf{H}_{BR}]) = \mathcal{C}(\mathbf{U})$ into two orthogonal subspaces: 1) $\mathcal{S}^{\mathrm{PNC}}$, in which the channel direction pairs $(\mathbf{v}_{A;i}, \mathbf{v}_{B;i})$ are parallel or close to parallel, for carrying out PNC; 2) $\mathcal{S}^{\mathrm{CD}}$ for carrying out the CD strategy. Let l' be an arbitrary integer between 0 and l. Recall that the channel direction pairs are ordered by the degree of orthogonality as in Corollary 1. That is, the first $k + l'$ direction pairs have lower degree of orthogonality compared to the remaining pairs. Thus, we allocate the first $k + l'$ direction pairs to form a basis of $\mathcal{S}^{\mathrm{PNC}}$. The remaining channel directions give a basis of $\mathcal{S}^{\mathrm{CD}}$. In this section, we assume that l' is given. The details on the optimization of l' will be discussed later in Sections V and VI.

1) Space-Division Operation: Let the RQ decomposition of $\mathbf{G}_m$ be

$$\mathbf{G}_m = \mathbf{R}_m\mathbf{S}_m^{\dagger},\ m \in \{A, B\} \tag{30}$$

where $\mathbf{R}_m \in \mathbb{C}^{n_m \times n_m}$ is an upper triangular matrix given by

$$\mathbf{R}_m = \begin{bmatrix} r_{m;1,1} & r_{m;1,2} & \cdots & r_{m;1,n_m} \\ 0 & r_{m;2,2} & \cdots & r_{m;2,n_m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{m;n_m,n_m} \end{bmatrix}, \tag{31}$$

and $\mathbf{S}_m \in \mathbb{C}^{n_m \times n_m}$ is unitary. Together with (22), the channel matrices can be jointly decomposed as

$$\mathbf{H}_m = \mathbf{UD}_m\mathbf{R}_m\mathbf{S}_m^{\dagger},\ m \in \{A, B\}\,. \tag{32}$$

Then, the received signal at the relay, after left-multiplying $\mathbf{U}^{\dagger}$, can be represented as

$$\mathbf{Y}_R' = \mathbf{U}^{\dagger}\mathbf{Y}_R = \mathbf{D}_A\mathbf{R}_A\mathbf{X}_A' + \mathbf{D}_B\mathbf{R}_B\mathbf{X}_B' + \mathbf{Z}_R', \tag{33}$$

where $\mathbf{X}_m' = \mathbf{T}_m^{\dagger}\mathbf{X}_m, m \in \{A, B\}$, and $\mathbf{Z}_R' = \mathbf{U}^{\dagger}\mathbf{Z}_R$ with i.i.d. elements $\sim \mathcal{N}_c(0, N_0)$.

We partition $\mathbf{D}_m$ and $\mathbf{R}_m$, respectively, as

$$\mathbf{D}_m = \begin{bmatrix} \mathbf{D}_{m;1,1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{m;2,2} \end{bmatrix} \tag{34a}$$

$$\mathbf{R}_m = \begin{bmatrix} \mathbf{R}_{m;1,1} & \mathbf{R}_{m;1,2} \\ \mathbf{0} & \mathbf{R}_{m;2,2} \end{bmatrix}, \quad m \in \{A, B\} \tag{34b}$$

where $\mathbf{D}_{m;1,1} \in \mathbb{C}^{(k+2l')\times(k+l')}$ and $\mathbf{D}_{m;2,2} \in \mathbb{C}^{(n_R-k-2l')\times(l-l'+d_m)}$, $m \in \{A, B\}$, are block-diagonal matrices, and $\mathbf{R}_{m;1,1} \in \mathbb{C}^{(k+l')\times(k+l')}$ and $\mathbf{R}_{m;2,2} \in \mathbb{C}^{(l-l'+d_m)\times(l-l'+d_m)}$, $m \in \{A, B\}$, are upper triangular matrices. Then, (33) can be rewritten as (35), given at the bottom of the page, where $\mathbf{Y}_R'$, $\mathbf{X}_m'$, and $\mathbf{Z}_R'$ are correspondingly partitioned as

$$\mathbf{Y}_R' = \begin{bmatrix} \mathbf{Y}_R^{\mathrm{PNC}} \\ \mathbf{Y}_R^{\mathrm{CD}} \end{bmatrix}, \mathbf{X}_m' = \begin{bmatrix} \mathbf{X}_m^{\mathrm{PNC}} \\ \mathbf{X}_m^{\mathrm{CD}} \end{bmatrix}, \text{ and } \mathbf{Z}_R' = \begin{bmatrix} \mathbf{Z}_R^{\mathrm{PNC}} \\ \mathbf{Z}_R^{\mathrm{CD}} \end{bmatrix}. \tag{36}$$

$$\begin{bmatrix} \mathbf{Y}_R^{\mathrm{PNC}} \\ \mathbf{Y}_R^{\mathrm{CD}} \end{bmatrix} = \sum_{m\in\{A,B\}} \begin{bmatrix} \mathbf{D}_{m;1,1}\mathbf{R}_{m;1,1} & \mathbf{D}_{m;1,1}\mathbf{R}_{m;1,2} \\ \mathbf{0} & \mathbf{D}_{m;2,2}\mathbf{R}_{m;2,2} \end{bmatrix} \begin{bmatrix} \mathbf{X}_m^{\mathrm{PNC}} \\ \mathbf{X}_m^{\mathrm{CD}} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_R^{\mathrm{PNC}} \\ \mathbf{Z}_R^{\mathrm{CD}} \end{bmatrix} \tag{35}$$

Here, the superscript "PNC" (or "CD") represents the PNC (or CD) strategy.

Based on the signal model in (35), the proposed space-division based relaying strategy is described as follows. At user m, two groups of spatial streams are generated: one group, referred to as the CD spatial streams, forms the codeword matrix $\mathbf{X}_m^{\mathrm{CD}}$, and the other group, referred to as the PNC spatial streams, forms the codeword matrix $\mathbf{X}_m^{\mathrm{PNC}}$, $m \in \{A, B\}$.

2) CD Spatial Streams: Due to the block upper triangular structure of the channel matrices in (35), the relay can completely decode the spatial streams $\mathbf{X}_A^{\mathrm{CD}}$ and $\mathbf{X}_B^{\mathrm{CD}}$ free of interference from the PNC spatial streams. Specifically, the relay completely decodes both $\mathbf{X}_A^{\mathrm{CD}}$ and $\mathbf{X}_B^{\mathrm{CD}}$ based on

$$\mathbf{Y}_R^{\mathrm{CD}} = \sum_{m\in\{A,B\}} \mathbf{D}_{m;2,2}\mathbf{R}_{m;2,2}\mathbf{X}_m^{\mathrm{CD}} + \mathbf{Z}_R^{\mathrm{CD}}. \quad (37)$$

Then, $\mathbf{X}_A^{\mathrm{CD}}$ and $\mathbf{X}_B^{\mathrm{CD}}$ are canceled from the received signal in (35).

3) PNC Spatial Streams: After the cancellation of $\mathbf{X}_A^{\mathrm{CD}}$ and $\mathbf{X}_B^{\mathrm{CD}}$, the system model for the PNC spatial streams is given by

$$\mathbf{Y}_R^{\mathrm{PNC}} = \sum_{m\in\{A,B\}} \mathbf{D}_{m;1,1}\mathbf{R}_{m;1,1}\mathbf{X}_m^{\mathrm{PNC}} + \mathbf{Z}_R^{\mathrm{PNC}}. \quad (38)$$

From (23a), the first k columns of $\mathbf{D}_{A;1,1}$ and $\mathbf{D}_{B;1,1}$ are identical, meaning that the corresponding signal directions are already aligned; however, for $i = k+1, \ldots, k+l'$, the ith columns of $\mathbf{D}_{A;1,1}$ and $\mathbf{D}_{B;1,1}$ are not identical, meaning that the signals are in different directions. To carry out PNC for these non-aligned signal directions, we project each column pair of $\mathbf{D}_{A;1,1}$ and $\mathbf{D}_{B;1,1}$ onto a common direction. The treatment is similar to what described in Section III, as detailed below.

By inspection, the only difference between the ith columns of $\mathbf{D}_{A;1,1}$ and $\mathbf{D}_{B;1,1}$ is given by the 2×1 vectors $\mathbf{e}_{A;i}$ and $\mathbf{e}_{B;i}$, for $i = k+1, \ldots, k+l'$. Without loss of generality, denote by $\mathbf{p}_i$ a 2×1 unit vector representing the projection direction of $\mathbf{e}_{A;i}$ and $\mathbf{e}_{B;i}$. The choice of $\mathbf{p}_i$ is similar to that described in Section III and will be detailed in the next section.

Now the overall projection process can be described using the following projection matrix:

$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_k & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{p}_{k+1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{p}_{k+l'} \end{bmatrix} \in \mathbb{C}^{(k+2l')\times(k+l')}. \quad (39)$$

After the projection, the resulting signal model is given by

$$\widetilde{\mathbf{Y}}_R^{\mathrm{PNC}} = \mathbf{P}^T\mathbf{Y}_R^{\mathrm{PNC}} = \sum_{m\in\{A,B\}} \widetilde{\mathbf{H}}_m^{\mathrm{PNC}}\mathbf{X}_m^{\mathrm{PNC}} + \widetilde{\mathbf{Z}}_R^{\mathrm{PNC}} \quad (40)$$

with

$$\widetilde{\mathbf{H}}_m^{\mathrm{PNC}} = \mathbf{P}^T\mathbf{D}_{m;1,1}\mathbf{R}_{m;1,1} = \tilde{\mathbf{D}}_{m;1,1}\mathbf{R}_{m;1,1} \quad (41a)$$

$$\tilde{\mathbf{D}}_{m;1,1} = \mathrm{diag}\left\{1, \ldots, 1, \mathbf{p}_{k+1}^T\mathbf{e}_{m;k+1}, \ldots, \mathbf{p}_{k+l'}^T\mathbf{e}_{m;k+l'}\right\} \quad (41b)$$

where $\widetilde{\mathbf{Z}}_R^{\mathrm{PNC}} = \mathbf{P}^T\mathbf{Z}_R^{\mathrm{PNC}} \sim \mathcal{N}_c(\mathbf{0}, N_0\mathbf{I}_{n_R-l'})$. Note that the equivalent channel matrices $\widetilde{\mathbf{H}}_A^{\mathrm{PNC}}$ and $\widetilde{\mathbf{H}}_B^{\mathrm{PNC}}$ are $(k+l')$-by-$(k+l')$ square matrices. For such an equivalent MIMO TWRC, efficient techniques exist to align the signal directions of the two users into a common set of $k+l'$ directions. Based on that, we carry out k PNC streams in these directions. We will analyze the achievable rates of the proposed scheme in the next section.

V. Achievable Rate-Region of MIMO TWRC

In this section, we derive an achievable rate-pair of the proposed space-division based network coding scheme. Based on that, we optimize the system parameters to determine the achievable rate-region.

A. Achievable Rate-Pairs

1) CD Spatial Streams: The equivalent channel model seen by the CD spatial streams is given in (37), with the equivalent channel matrices given by $\mathbf{D}_{m,2,2}\mathbf{R}_{m;2,2}$, $m \in \{A, B\}$.

The signal model in (37) is a standard MIMO MAC channel. Let $\mathbf{Q}_m^{\mathrm{CD}} = \frac{1}{T}E\left[\mathbf{X}_m^{\mathrm{CD}}(\mathbf{X}_m^{\mathrm{CD}})^\dagger\right]$ be the input covariance matrix of the CD spatial steams of user m. Then, the achievable rate-pair of the CD spatial streams satisfies (42), at the bottom of the page; see [29]. It is worth noting that (42a) is redundant when the CD signals of the two users are orthogonal to each other, i.e., $\mathcal{C}(\mathbf{D}_{A;2,2}\mathbf{R}_{A;2,2})$ is orthogonal to $\mathcal{C}(\mathbf{D}_{B;2,2}\mathbf{R}_{B;2,2})$, implying that there is no cross-interference between the two users' CD signals.

2) PNC Spatial Streams: The equivalent channel seen by the PNC streams is given in (40). Recall that $\widetilde{\mathbf{H}}_m^{\mathrm{PNC}}$ is a $(k+l')\times(k+l')$ square matrix, and the efficient design of PNC for this case has been studied in [16]–[18]. Here, we follow the GSVD-based approach in [16], as briefly described below.

$$R_A^{\mathrm{CD}} + R_B^{\mathrm{CD}} \le \frac{1}{2}\log\left|\mathbf{I} + \frac{1}{N_0}\sum_{m\in\{A,B\}} \mathbf{D}_{m;2,2}\mathbf{R}_{m;2,2}\mathbf{Q}_m^{\mathrm{CD}}\mathbf{R}_{m;2,2}^\dagger\mathbf{D}_{m;2,2}^\dagger\right| \quad (42a)$$

$$R_m^{\mathrm{CD}} \le \frac{1}{2}\log\left|\mathbf{I} + \frac{1}{N_0}\mathbf{D}_{m;2,2}\mathbf{R}_{m;2,2}\mathbf{Q}_m^{\mathrm{CD}}\mathbf{R}_{m;2,2}^\dagger\mathbf{D}_{m;2,2}^\dagger\right|, m \in \{A, B\} \quad (42b)$$

Applying the GSVD [26] to $\widetilde{\mathbf{H}}_m^{\mathrm{PNC}}$, we obtain

$$\widetilde{\mathbf{H}}_m^{\mathrm{PNC}} = \mathbf{B}\mathbf{\Sigma}_m\widetilde{\mathbf{S}}_m^{\dagger}, \quad m \in \{A, B\}, \tag{43}$$

where $\mathbf{B} \in \mathbb{C}^{(k+l')\times(k+l')}$ is a nonsingular matrix, $\widetilde{\mathbf{S}}_m \in \mathbb{C}^{(k+l')\times(k+l')}$ is an orthogonal matrix, $m \in \{A, B\}$, and $\mathbf{\Sigma}_m \in \mathbb{C}^{(k+l')\times(k+l')}$ is a diagonal matrix with the ith diagonal element denoted by $\sigma_{m;i}$. We further take the QR decomposition to the matrix $\mathbf{B}$, yielding

$$\widetilde{\mathbf{H}}_m^{\mathrm{PNC}} = \mathbf{Q}\widetilde{\mathbf{R}}\mathbf{\Sigma}_m\widetilde{\mathbf{S}}_m^{\dagger}, \quad m \in \{A, B\}, \tag{44a}$$

where $\widetilde{\mathbf{R}} \in \mathbb{C}^{(k+l')\times(k+l')}$ is an upper triangular matrix given by

$$\widetilde{\mathbf{R}} = \begin{bmatrix} \tilde{r}_{1,1} & \tilde{r}_{1,2} & \cdots & \tilde{r}_{1,k+l'} \\ 0 & \tilde{r}_{2,2} & \cdots & \tilde{r}_{2,k+l'} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \tilde{r}_{k+l',k+l'} \end{bmatrix}. \tag{44b}$$

The transmit signal $\mathbf{X}_m^{\mathrm{PNC}}$ in (40) is designed as

$$\mathbf{X}_m^{\mathrm{PNC}} = \widetilde{\mathbf{S}}_m\mathbf{\Psi}_m^{1/2}\mathbf{C}_m^{\mathrm{PNC}}, \quad m \in \{A, B\}, \tag{45}$$

where $\mathbf{\Psi}_m^{1/2} = \mathrm{diag}\left\{\sqrt{\psi_{m;1}}, \sqrt{\psi_{m;2}}, \ldots, \sqrt{\psi_{m;k+l'}}\right\}$ is a diagonal matrix with $\psi_{m;i} \geq 0, i = 1, 2, \ldots, k + l'$, and $\mathbf{C}_m^{\mathrm{PNC}} \in \mathbb{C}^{(k+l')\times T}$ is the codeword matrix with unit power (i.e., the elements of $\mathbf{C}_m^{\mathrm{PNC}}$ have zero mean and unit variance). Note that the total power consumption of the PNC streams of user m is given by $\sum_{i=1}^{k+l'} \psi_{m;i}, m \in \{A, B\}$.

By left-multiplying $\mathbf{Q}^{\dagger}$ to the received signal in (40), the signal component induced by the PNC spatial streams is given by $\widetilde{\mathbf{R}}\mathbf{\Sigma}_A\mathbf{\Psi}_A^{1/2}\mathbf{C}_A^{\mathrm{PNC}} + \widetilde{\mathbf{R}}\mathbf{\Sigma}_B\mathbf{\Psi}_B^{1/2}\mathbf{C}_B^{\mathrm{PNC}}$. Note that each column of $\widetilde{\mathbf{R}}\mathbf{\Sigma}_A\mathbf{\Psi}_A^{1/2}$ is proportional to the corresponding column of $\widetilde{\mathbf{R}}\mathbf{\Sigma}_B\mathbf{\Psi}_B^{1/2}$. This allows us to successively decode the linear combinations on the main diagonal by peeling off the interference from the other PNC spatial streams [16]. Specifically, let $(\mathbf{c}_{m,i}^{\mathrm{PNC}})^T$ be the ith row of $\mathbf{C}_m^{\mathrm{PNC}}$, and $\tilde{r}_{i,i}$ be the (i, i)th element of $\widetilde{\mathbf{R}}$. With the upper triangular form of the equivalent channel matrices, the PNC spatial streams, i.e., $\tilde{r}_{i,i}\sigma_{A;i}\psi_{A;i}^{1/2}\mathbf{c}_{A,i}^{\mathrm{PNC}} + \tilde{r}_{i,i}\sigma_{B;i}\psi_{B;i}^{1/2}\mathbf{c}_{B,i}^{\mathrm{PNC}}$, $i = k + l', k + l' - 1, \ldots, 1$, are successively network-decoded and canceled from the received signal. Let R_m^{PNC} be the total rate of the PNC spatial streams of user m. From [16, Th. 1], the achievable rate-pair is given by

$$R_m^{\mathrm{PNC}} = \sum_{i=1}^{k+l'} \frac{1}{2}\left[\log\left(\frac{I(i)\sigma_{m;i}^2\psi_{m;i}}{\sigma_{A;i}^2\psi_{A;i} + \sigma_{B;i}^2\psi_{B;i}} + \frac{\tilde{r}_{i,i}^2\sigma_{m;i}^2\psi_{m;i}}{N_0}\right)\right]^+ \quad \text{for } m \in \{A, B\} \tag{46}$$

where $I(i)$ is the indicator function with $I(i) = 1$ for $i = 1$ and $I(i) = 0$ for $i \neq 1$.

3) Overall Achievable Rate-Pair: We now consider the overall achievable rate-pair of the proposed space-division based network coding scheme. Before going into details, we note that the power constraint of user m, i.e., $\frac{1}{T}E\left[\|\mathbf{X}_m\|_F^2\right] \leq P_m, \ m \in \{A, B\}$, can be equivalently expressed as

$$\mathrm{tr}\{\mathbf{Q}_m^{\mathrm{CD}}\} + \sum_{i=1}^{k+l'} \psi_{m;i} \leq P_m, m \in \{A, B\}. \tag{47}$$

We are now ready to present the following theorem on the achievable rates of the proposed scheme. The proof can be found in Appendix C.

Theorem 2: For given $\mathbf{Q}_m^{\mathrm{CD}}$, $\mathbf{\Psi}_m$, and $\mathbf{Q}_R$ satisfying (47) and $\mathrm{tr}(\mathbf{Q}_R) \leq P_R$, a rate-pair (R_A, R_B) for the MIMO TWRC is achievable if

$$R_m \leq \min\{R_m^{\mathrm{CD}} + R_m^{\mathrm{PNC}}, R_m^{\mathrm{DL}}\}, \quad m \in \{A, B\}, \tag{48}$$

where R_A^{CD} and R_B^{CD} satisfy (42), R_m^{PNC} is given by (46), and R_m^{DL} is given by (4).

Remark 5: Here, we briefly explain the difference between our proposed scheme and the existing GSVD-based scheme [16]. For the case of $n_R = \max(n_A, n_B)$, we have $\mathcal{C}(\mathbf{H}_{AR}) \subseteq \mathcal{C}(\mathbf{H}_{BR})$ or $\mathcal{C}(\mathbf{H}_{BR}) \subseteq \mathcal{C}(\mathbf{H}_{AR})$ or both, i.e., the two users share a common dimension-$\min(n_A, n_B)$ signal space. Then, the GSVD approach in [16] is equivalent to our space-division approach, and both can identify $k = \min(n_A, n_B)$ aligned spatial directions for PNC. On the other hand, for the case of $n_R > \max(n_A, n_B)$, any of $\mathcal{C}(\mathbf{H}_{AR})$ and $\mathcal{C}(\mathbf{H}_{BR})$ generally does not contain the other. As a consequence, the GSVD approach can only identify k aligned PNC directions in the intersection of $\mathcal{C}(\mathbf{H}_{AR})$ and $\mathcal{C}(\mathbf{H}_{BR})$. In contrast, in the proposed space-division approach, we first project the PNC signals into a properly chosen common subspace to obtain the equivalent channel matrices $\widetilde{\mathbf{H}}_A^{\mathrm{PNC}}$ and $\widetilde{\mathbf{H}}_B^{\mathrm{PNC}}$ satisfying $\mathcal{C}(\widetilde{\mathbf{H}}_A^{\mathrm{PNC}}) = \mathcal{C}(\widetilde{\mathbf{H}}_B^{\mathrm{PNC}}) = \mathbb{C}^{k+l'}$. Then, we use GSVD to identify $k + l'$ ($\geq k$) aligned spatial directions for $k + l'$ streams of PNC. The remaining portions of signal directions are projected to the space orthogonal to the PNC streams, where CD can achieve improved performance. This difference eventually leads to the performance advantage of the space-division approach over the GSVD approach for the case of $n_R > \max(n_A, n_B)$.

Remark 6: The proposed space-division scheme involves jointly re-encoding the completely decoded messages and network-coded messages at the relay. This re-encoding operation includes symbolwise algebraic operations for network-coding and conventional channel coding. These two operations are generally required in all regenerative relaying schemes utilizing network coding [2], [5], [16], [17], [35]. In particular, the algebraic operations involve simple additions and multiplications in finite fields. Thus, the overall complexity of the re-encoding process is dominated by channel coding, for which efficient encoding algorithms have been well studied in the literature. Compared with conventional regenerative relaying, the proposed space-division scheme requires a marginal extra overhead for the joint re-encoding process at the relay.

Remark 7: We note that identical received power between the two users is not required in our proposed scheme. Unlike the original PNC scheme proposed in [2], PNC with nested lattice coding allows unequal reception powers of the two users and is capacity-achieving for SISO TWRCs in the high-SNR regime [5]. Our work in this paper borrows the result in [5] for

PNC spatial streams. Specifically, our SD approach employs a new precoding technique to decompose a MIMO TWRC into multiple scalar SISO TWRCs, and then nested lattice coding is applied to each of those SISO TWRCs. Therefore, pre-equalization with identical received powers is not required in the proposed space-division scheme.

Remark 8: Global CSI is assumed in developing the proposed space-division scheme. However, acquiring *global* CSI may pose a heavy burden on the network backhaul in practice. We now describe an efficient way to distribute CSI in a MIMO TWRC, so as to realize the proposed scheme. We assume that channel estimation is conducted at the relay node, with the consideration that the relay node is usually more computationally powerful than the user nodes, such as in a cellular network. Upon acquiring the CSI, the relay computes necessary information for implementing the proposed scheme, including the precoding matrices of both users, the number of PNC spatial streams (that is the same for both users), and the information rates of the PNC/CD spatial stream for both users. Then, the relay sends the information to the two users via backhaul links. Note that it is sufficient for each user to know its *individual information* including its own precoding matrix, the number of its PNC spatial streams, and the rates of its PNC and CD spatial streams. Generally speaking, the overhead incurred by sending the individual information to a user is much smaller than that by directly sending the global CSI (including the channel matrices of the two user-relay links) to this user. Furthermore, the assumption of *perfect* CSI is usually difficult to be exactly met in practice, e.g., due to the Doppler effect or due to capacity limitation of the backhaul links. In what follows, we briefly discuss the robustness of the proposed scheme against uncertainties in the CSI. With perfect CSI, the signal directions of the PNC spatial streams (as described in Section IV-B and in this subsection) impinging upon the relay are perfectly aligned, so that the PNC streams are successively decoded free of interference. However, if the users only partially know CSI, the PNC streams cannot be perfectly aligned, which causes interference between the PNC streams (that cannot be handled using the SIC technique in the proposed scheme). In the low-SNR regime, this interference leakage has little impact on the system performance, as the interference is overwhelmed by the channel noise; however, in the high-SNR regime, this leakage in general impairs the achievable degrees of freedom (DoF) of the PNC streams. We will return to this issue in Section VI-A.

B. Determining Achievable Rate-Region

Now we consider determining the boundary of the achievable rate-region. From (48), the downlink achievable rates are the same as the capacity upper bound in (4). Here, we focus on the uplink rate-region.

The boundary of the uplink rate-region can be determined by solving the weighted-sum-rate maximization problem defined in (49), given at the bottom of the page. This problem involves the optimization of the number of PNC streams l', the projection matrix $\mathbf{P}$, the covariance matrices $\mathbf{Q}_A^{\mathrm{CD}}$ and $\mathbf{Q}_B^{\mathrm{CD}}$ for the CD streams, and the power allocation coefficients $\{\psi_{A;i}\}_{i=1}^{k+l'}$ and $\{\psi_{B;i}\}_{i=1}^{k+l'}$. This problem is in general difficult to solve. We next propose a suboptimal solution, as detailed below.

1) Determining the Projection Directions: The optimization of the projection matrix $\mathbf{P}$ to maximize the weighted sum-rate is in general difficult to solve. To simplify the problem, we consider the high-SNR regime, with the weighted sum-rate given by

$$\begin{aligned}
&w_A R_A^{\mathrm{PNC}} + w_B R_B^{\mathrm{PNC}} \\
&\overset{(a)}{\approx} \frac{1}{2} \sum_{m\in\{A,B\}} w_m \left(\sum_{i=1}^{k+l'} \log\left(\frac{\tilde{r}_{i,i}^2 \sigma_{m;i}^2 \psi_{m;i}}{N_0} \right) \right) \\
&\overset{(b)}{=} \frac{1}{2} \sum_{m\in\{A,B\}} w_m \log \left| \frac{P_m}{N_0 n_m} \tilde{\mathbf{R}} \mathbf{\Sigma}_m \mathbf{\Sigma}_m^{\dagger} \tilde{\mathbf{R}}^{\dagger} \right| \\
&\overset{(c)}{=} \frac{1}{2} \sum_{m\in\{A,B\}} w_m \log \left| \frac{P_m}{N_0 n_m} \tilde{\mathbf{D}}_{m;1,1} \mathbf{R}_{m;1,1} \mathbf{R}_{m;1,1}^{\dagger} \tilde{\mathbf{D}}_{m;1,1}^{\dagger} \right| \\
&\overset{(d)}{=} \frac{1}{2} \sum_{m\in\{A,B\}} \left(w_m \log \left| \frac{P_m}{N_0 n_m} \mathbf{R}_{m;1,1} \mathbf{R}_{m;1,1}^{\dagger} \right| \right. \\
&\qquad \left. + w_m \log \left| \tilde{\mathbf{D}}_{m;1,1} \tilde{\mathbf{D}}_{m;1,1}^{\dagger} \right| \right) \qquad (50)
\end{aligned}$$

$$\text{maximize} \quad \sum_{m\in\{A,B\}} w_m \left(R_m^{\mathrm{CD}} + R_m^{\mathrm{PNC}} \right) \qquad (49a)$$

$$\text{subject to} \quad \sum_{j=1}^{k+l'} \psi_{m;j} + \mathrm{tr}\{\mathbf{Q}_m^{\mathrm{CD}}\} \le P_m, \mathbf{Q}_m^{\mathrm{CD}} \succeq \mathbf{0}, \psi_{m;i} \ge 0, \text{for } i = 1, \ldots, k+l', \qquad (49b)$$

$$R_m^{\mathrm{PNC}} = \sum_{i=1}^{k+l'} \frac{1}{2} \left[\log \left(\frac{I(i)\sigma_{m;i}^2 \psi_{m;i}}{\sigma_{A;i}^2 \psi_{A;i} + \sigma_{B;i}^2 \psi_{B;i}} + \frac{\tilde{r}_{i,i}^2 \sigma_{m;i}^2 \psi_{m;i}}{N_0} \right) \right]^+, \qquad (49c)$$

$$R_A^{\mathrm{CD}} + R_B^{\mathrm{CD}} \le \frac{1}{2} \log \left| \mathbf{I} + \frac{1}{N_0} \sum_{m\in\{A,B\}} \mathbf{D}_{m;2,2} \mathbf{R}_{m;2,2} \mathbf{Q}_m^{\mathrm{CD}} \mathbf{R}_{m;2,2}^{\dagger} \mathbf{D}_{m;2,2}^{\dagger} \right|, \qquad (49d)$$

$$R_m^{\mathrm{CD}} \le \frac{1}{2} \log \left| \mathbf{I} + \frac{1}{N_0} \mathbf{D}_{m;2,2} \mathbf{R}_{m;2,2} \mathbf{Q}_m^{\mathrm{CD}} \mathbf{R}_{m;2,2}^{\dagger} \mathbf{D}_{m;2,2}^{\dagger} \right|, m \in \{A, B\} \qquad (49e)$$

where step (*a*) follows from substituting (46), step (*b*) from the facts that $\tilde{\mathbf{R}}$ is upper triangular and that equal power allocation is asymptotically optimal (i.e., $\psi_{m;i}=\frac{P_m}{n_m}$),[4] step (*c*) by noting $\tilde{\mathbf{D}}_{m;1,1}\mathbf{R}_{m;1,1}=\tilde{\mathbf{H}}_m^{\mathrm{PNC}}=\mathbf{Q}\tilde{\mathbf{R}}\mathbf{\Sigma}_m\tilde{\mathbf{S}}_m^{\dagger}$ [cf., (40) and (44a)], and step (d) from the fact that $\tilde{\mathbf{D}}_{m;1,1}$ is diagonal. In the above, $\log\left|\tilde{\mathbf{D}}_{m;1,1}\tilde{\mathbf{D}}_{m;1,1}^{\dagger}\right|$ is the only term related to $\mathbf{P}$. Recall from (41b) that $\tilde{\mathbf{D}}_{m;1,1}=\mathbf{P}^T\mathbf{D}_{m;1,1}$ with $\mathbf{D}_{m;1,1}$ being the principle submatrix of $\mathbf{D}_m$ in (34a). Thus, the weighted sum-rate maximization problem over $\mathbf{P}$ can be decoupled into l' independent subproblems as

$$\max_{\|\mathbf{p}_i\|=1} w_A\log\left(\left|\mathbf{p}_i^{\dagger}\mathbf{e}_{A;i}\right|^2\right)+w_B\log\left(\left|\mathbf{p}_i^{\dagger}\mathbf{e}_{B;i}\right|^2\right),\quad \text{for } i=k+1,\ldots,k+l', \tag{51}$$

where $\mathbf{e}_{A;i}$ and $\mathbf{e}_{B;i}$ are given in (23c). From (71) and the discussions therein, the optimal $\mathbf{p}_i$ to maximize the weighted sum-rate is a real vector given by

$$\mathbf{p}_i=\gamma_i\left(\mathbf{e}_{A;i}+\beta_i\mathbf{e}_{B;i}\right),\ \text{for } i=k+1,\ldots,k+l', \tag{52}$$

where γ_i is a scaling factor to ensure $\|\mathbf{p}_i\|=1$, and

$$\beta_i\equiv\frac{1}{2}\left(\sqrt{t_i^2+4\frac{w_B}{w_A}}-t_i\right) \tag{53}$$

with

$$t_i\equiv(\lambda_i-1)\left(1-\frac{w_B}{w_A}\right). \tag{54}$$

[4]At high SNR, the sum-rate maximization problem in (49) reduces to maximize $\sum_{m\in\{A,B\}}w_m\left(\sum_{i=1}^{k+l'}\log\psi_{m;i}+\log|\mathbf{Q}_m^{\mathrm{CD}}|\right)$ subject to (49b). This formulation is similar to the power allocation problem for point-to-point MIMO, and it is known that equal power allocation is asymptotically optimal at high SNR [33].

2) Determining $\mathbf{Q}_A^{\mathrm{CD}}$ and $\mathbf{Q}_B^{\mathrm{CD}}$: Given $\{\mathbf{p}_i\}$ in (52), the optimization problem in (49) can be decoupled into two separate problems by predetermining the power allocated to the two signal subspaces. Let P_m^{CD} be the power of user m used for the CD spatial streams. Then, the power for the PNC streams is given by $P_m^{\mathrm{PNC}}=P_m-P_m^{\mathrm{CD}}, m\in\{A,B\}$. For given P_A^{CD} and P_B^{CD}, the optimal $\mathbf{Q}_A^{\mathrm{CD}}$ and $\mathbf{Q}_B^{\mathrm{CD}}$ to (49) can be found by solving the problem in (55), given at the bottom of the page. This is a standard weighted sum-rate maximization problem for a MIMO multiple-access channel with two users [34]. The optimal solution can be computed using convex programming [30] or the iterative water-filling technique in [34].

3) Determining Power Allocation for PNC Streams: Now we consider the optimization of $\{\psi_{A;i}\}_{i=1}^{k+l'}$ and $\{\psi_{B;i}\}_{i=1}^{k+l'}$. Given P_A^{PNC} and P_B^{PNC}, the optimal $\{\psi_{A;i}\}_{i=1}^{k+l'}$ and $\{\psi_{B;i}\}_{i=1}^{k+l'}$ can be determined by solving the problem in (56), given at the bottom of the page. A similar problem has been considered in [16], and the optimal solution can be obtained by solving the Karush–Kuhn–Tuchker (KKT) conditions. The involved complexity is negligible. We refer interested readers to [16] for more details.

4) Overall Algorithm: We summarize the above suboptimal solution to (49) as follows:

Algorithm

Step1: Calculate $\mathbf{p}_i$ for $i=k+1,\ldots,k+l$ using (52);

Step2:

For $\{l',P_m^{\mathrm{CD}},P_m^{\mathrm{PNC}}\mid m\in\{A,B\}\}$ in set $\mathfrak{S}$

Calculate $\left\{\mathbf{Q}_A^{\mathrm{CD}},\mathbf{Q}_B^{\mathrm{CD}},\{\psi_{A;i}\}_{i=1}^{k+l'},\{\psi_{B;i}\}_{i=1}^{k+l'}\right\}$ by solving (55) and (56);

End

Step3: Find the optimal $\{\mathbf{Q}_m^{\mathrm{CD}},\{\psi_{m;i}\}_{i=1}^{k+l'},l',P_m^{\mathrm{CD}},P_m^{\mathrm{PNC}}|m\in\{A,B\}\}$ among the results of Step 2.

$$\begin{aligned}
\text{maximize}\quad & w_AR_A^{\mathrm{CD}}+w_AR_A^{\mathrm{CD}} && \text{(55a)}\\
\text{subject to}\quad & \mathrm{tr}\{\mathbf{Q}_m^{\mathrm{CD}}\}\le P_m^{\mathrm{CD}},\mathbf{Q}_m^{\mathrm{CD}}\succeq\mathbf{0}, && \text{(55b)}\\
& R_A^{\mathrm{CD}}+R_B^{\mathrm{CD}}\le\frac{1}{2}\log\left|\mathbf{I}+\frac{1}{N_0}\sum_{m\in\{A,B\}}\mathbf{D}_{m,2,2}\mathbf{R}_{m;2,2}\mathbf{Q}_m^{\mathrm{CD}}\mathbf{R}_{m;2,2}^{\dagger}\mathbf{D}_{m,2,2}^{\dagger}\right| && \text{(55c)}\\
& R_m^{\mathrm{CD}}\le\frac{1}{2}\log\left|\mathbf{I}+\frac{1}{N_0}\mathbf{D}_{m,2,2}\mathbf{R}_{m;2,2}\mathbf{Q}_m^{\mathrm{CD}}\mathbf{R}_{m;2,2}^{\dagger}\mathbf{D}_{m,2,2}^{\dagger}\right|,m\in\{A,B\} && \text{(55d)}
\end{aligned}$$

$$\begin{aligned}
\text{maximize}\quad & \sum_{m\in\{A,B\}}w_m\left(\sum_{i=1}^{k+l'}\frac{1}{2}\left[\log\left(\frac{I(i)\sigma_{m;i}^2\psi_{m;i}}{\sigma_{A;i}^2\psi_{A;i}+\sigma_{B;i}^2\psi_{B;i}}+\frac{\tilde{r}_{i,i}^2\sigma_{m;i}^2\psi_{m;i}}{N_0}\right)\right]^+\right) && \text{(56a)}\\
\text{subject to}\quad & \sum_{i=1}^{k+l'}\psi_{m;i}\le P_m^{\mathrm{PNC}},\ \psi_{m;i}\ge0,\text{for } i=1,\ldots,k+l' && \text{(56b)}
\end{aligned}$$

Remark 9: In the above algorithm, the projection directions are fixed in (52) and are independent of the other optimization parameters. This treatment is in general not optimal, but allows a simple solution to (49) as described in the above algorithm. We will show in the next section that the proposed scheme (with this suboptimal treatment of the projection directions) has guaranteed sum-rate performance in both high- and low-SNR regimes. However, no guaranteed performance can be promised in the medium-SNR regime, though our empirical experience suggests that this treatment can also provide good performance at medium SNR.

Remark 10: In Step 2 of the above algorithm, the exhaustive search over all possible choices of l', P_m^{CD}, and P_m^{PNC}, $m \in \{A, B\}$ (denoted by the set $\mathfrak{S}$) is detailed as follows. First note that enumerating l' is straightforward, as l' is an integer between 0 and l. In addition, denote by $\alpha_m \in [0, 1]$ the portion of power allocated for the PNC spatial streams of user m, $m \in \{A, B\}$, i.e., $P_m^{\mathrm{CD}} = \alpha_m P_m$ and $P_m^{\mathrm{PNC}} = (1 - \alpha_m)P_m$, $m \in \{A, B\}$. Then, we only need to conduct a bounded 2-D exhaustive search, i.e., to search α_A and α_B over a discretized set of $[0, 1] \times [0, 1]$ up to a certain precision level.

Remark 11: The computational complexity of the above algorithm is dominated by Step 2, which involves solving two decoupled problems (55) and (56) for every possible choice of l', P_m^{CD}, and P_m^{PNC}, $m \in \{A, B\}$. Note that (56) can be done by solving KKT conditions with marginal complexity. Thus, the complexity is dominated by (55) which is a weighted sum-rate maximization problem for a vector MAC and is solvable using the iterative water-filling technique in [34].

VI. Asymptotic Sum-Rate Analysis

In the preceding section, we have shown the achievable rates of the proposed space-division-based network-coding scheme for MIMO TWRCs. In general, it is difficult to represent the achievable rate of the optimized space-division scheme in a closed form. Thus, we turn to the asymptotic regimes to assess the performance of our proposed scheme.

In the low-SNR regime, our proposed scheme asymptotically achieves the capacity of MIMO TWRC as SNR tends to zero.[5] This is because the proposed scheme reduces to the CD scheme, and the optimal power allocation for each user is to allocate the entire power to the signal stream that has the highest channel gain as SNR tends to zero. In this way, our scheme achieves the capacity at low SNR, as similarly to water-filling for a point-to-point MIMO system [17].

In what follows, we focus on the high-SNR analysis. We derive a closed-form expression for the asymptotic sum-rate of the proposed scheme in the high-SNR regime.

A. Asymptotic Sum-Rate as $\mathrm{SNR} \to \infty$

We start with analyzing the uplink achievable sum-rate

$$R^{\mathrm{SD}} = \sum_{m\in\{A,B\}} R_m^{\mathrm{CD}} + R_m^{\mathrm{PNC}} \tag{57}$$

[5]On the contrary, the GSVD approach in [16] performs away from the capacity at low SNR. This is because PNC with nested lattice coding generally fails to achieve the capacity of TWRC, as we see from (8) and (9) that there is no "1+" term in the logarithm of the rate formulas.

as the SNRs, i.e., $\frac{P_A}{N_0}$ and $\frac{P_B}{N_0}$, tend to infinity. In the high-SNR regime, equal power allocation is asymptotically optimal; cf., Footnote 4. Then, the upper bound of the uplink sum-rate of the MIMO TWRC is given by [cf., (4)]

$$R^{\mathrm{UL}} \approx \frac{1}{2} \sum_{m\in\{A,B\}} \log \left| \mathbf{I}_{n_R} + \frac{P_m}{N_0 n_m} \mathbf{H}_m \mathbf{H}_m^{\dagger} \right| \tag{58}$$

where "$x \approx y$" means

$$\lim_{N_0 \longrightarrow 0} (x - y) = 0.$$

We now present the following theorem on the asymptotic sum-rate of the proposed scheme. Denote by R^{SD} the uplink achievable sum-rate of the proposed space-division scheme.

Theorem 3: For a given l', the uplink achievable sum-rate of the proposed space-division scheme satisfies

$$\lim_{N_0 \longrightarrow 0} R^{\mathrm{UL}} - R^{\mathrm{SD}} = \Delta^{\mathrm{SD}} \tag{59a}$$

where

$$\Delta^{\mathrm{SD}} \triangleq -\log \prod_{i=k+1}^{k+l'} \frac{\lambda_i}{2} - \log \prod_{i=k+l'+1}^{k+l} \sqrt{\lambda_i(2-\lambda_i)} \geq 0. \tag{59b}$$

The proof of Theorem 3 can be found in Appendix D. Notice that the first term in (59b), i.e., $-\log \prod_{i=k+1}^{k+l'} \frac{\lambda_i}{2}$, is the rate loss incurred by the PNC spatial streams, and the second term, i.e., $\log \prod_{i=k+l'+1}^{k+l} \sqrt{\lambda_i(2-\lambda_i)}$, is that incurred by the CD spatial streams.

Remark 12: For the case of $\max(n_A, n_B) = n_R$, we have $l = 0$ and $\lambda_i = 2$ for $i = 1, \ldots, k$; see Remark 4. Then, from (59b), we have $\Delta^{\mathrm{SD}} = 0$, which means that the scheme is asymptotically optimal. This agrees with the fact that our proposed space-division scheme reduces to the GSVD scheme which is indeed asymptotically optimal in the high-SNR regime [16].

Corollary 3: The proposed space-division scheme and the capacity upper bound have a common high-SNR rate slope (referred to as the DoF) given by

$$\lim_{N_0 \longrightarrow 0} \frac{R^{\mathrm{SD}}}{\log(1/N_0)} = \frac{1}{2} \min(n_A + n_B, 2n_R). \tag{60}$$

Proof: From (59a), we see that the proposed space-division scheme and the capacity upper bound have the same rate slope, since their rate gap is a constant. From (58), the capacity upper bound is the sum of the achievable rates of two point-to-point MIMO channels. Note that the DoF of a point-to-point MIMO channel is the minimum of the number of transmit antennas and that of the receive antennas. Thus, the DoF of the capacity upper bound is $\frac{1}{2}\min(n_A + n_B, 2n_R)$, which concludes the proof. ■

Remark 13: The DoF of the CD scheme is limited by the MAC in (1), and is given by $\frac{1}{2}\min(n_A + n_B, n_R)$. From Corollary 3, our proposed scheme achieves a higher DoF than the CD scheme when $2n_R < n_A + n_B$. However, as aforementioned,

the PNC spatial streams lose DoF when channel uncertainty exists. In that case, the DoF of our proposed scheme reduces to that of the CD scheme. We note that such a loss of DoF is a common problem for all existing signal-alignment based techniques [16], [17], [36].

Corollary 4: The optimal l' that minimizes Δ^{SD} in (59b) satisfies

$$2 > \lambda_{k+1} \geq \cdots \geq \lambda_{k+l'} \geq \frac{8}{5} > \lambda_{k+l'+1} \geq \cdots \geq \lambda_{k+l} > 1. \tag{61}$$

With this choice of l', the asymptotic rate gap Δ^{SD} is at most $l\log(5/4)$ bits, which occurs when $\lambda_{k+1} = \lambda_{k+2} = \cdots = \lambda_{k+l} = \frac{8}{5}$.

Proof: The proof follows straightforwardly from (59b) by noting that the minimum of $\max(\frac{x}{2}, \sqrt{x(2-x)})$, $x \in [1,2]$, is $\frac{4}{5}$ achieved at $x = \frac{8}{5}$. ■

Corollary 5: At high SNR, the proposed scheme is at most $\min(n_A, n_B)\log(5/4)$ bits, or $\frac{1}{2}\log(5/4) \approx 0.161$ bits per user-antenna away from the sum-capacity; the corresponding asymptotic SNR gap to the sum-capacity upper bound is at most $\frac{10\log(5/4)}{\log 10} \approx 0.969$ dB.

Proof: From Corollary 4, the asymptotic gap to the sum-capacity upper bound is $l\log(5/4)$ bits for the worst case. Thus, the first half of the corollary is proven by noting $l \leq \min(n_A, n_B)$. For the second half, we see from Corollary 3 that the rate gap and the corresponding SNR gap in dB (denoted by Δ_{SNR}) satisfies

$$\lim_{N_0 \longrightarrow 0} \frac{\Delta^{\mathrm{SD}}}{\Delta_{SNR}} = \frac{\log 10}{10} \cdot \frac{n_A + n_B}{2}. \tag{62}$$

Substituting $\Delta^{\mathrm{SD}} = \min(n_A, n_B)\log(5/4)$, we obtain

$$\Delta_{SNR} = \frac{10\log(5/4)}{\log 10} \cdot \frac{\min(n_A, n_B)}{\frac{n_A+n_B}{2}} \leq \frac{10\log(5/4)}{\log 10} \approx 0.969. \tag{63}$$

■

B. Average Sum-Rate via Large-System Analysis

In this subsection, we investigate the statistical average of the rate gap Δ^{SD} in fading channels. To this end, the distribution of $\{\lambda_i\}$, i.e., the eigenvalues of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$, is required. However, such a distribution is difficult to obtain in general. Here, we employ the large-system analysis to find an approximation of the distribution of $\{\lambda_i\}$. The distribution obtained in this way becomes exact as the number of antennas in the system is large.

We assume Rayleigh fading, in which the channel coefficients are i.i.d. circularly symmetric complex Gaussian random variables. Then, the matrices $\mathbf{U}_A$ and $\mathbf{U}_B$ in (14) are truncated uniformly distributed unitary matrices or, alternatively, are asymptotically free random matrices [31]. Thus, we can use the theory of free probability to derive the asymptotic eigenvalue distribution (a.e.d.) of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$ as n_R tends to infinity, with the result given in the lemma below. Define $\eta_m \triangleq \frac{n_m}{n_R}, m \in \{A, B\}$.

Lemma 5: As $n_R \to \infty$ with $\frac{n_A}{n_R} \to \eta_A$ and $\frac{n_B}{n_R} \to \eta_B$, the a.e.d. of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$ is given by (64), at the bottom of the page, where $\delta(\cdot)$ is a Dirac delta function and $\mathrm{Im}[\cdot]$ is the imaginary part of a complex number.

The proof of the above lemma can be found in Appendix E. As $n_R \to \infty$, we see that for $\eta_A + \eta_B \geq 1$, the portion of eigenvalues $\{\lambda_i\}$ equal to 2 is given by $\eta_A + \eta_B - 1$. This portion corresponds to the dimension of the common space $\mathcal{S}_{A\|B}$ of $\mathbf{H}_{AR}$ and $\mathbf{H}_{BR}$. In addition, for $\eta_A \neq \eta_B$, the portion of eigenvalues $\{\lambda_i\}$ equal to 1 is given by $|\eta_A - \eta_B|$. This portion corresponds to the dimension of $\mathcal{S}_{A\perp B}$ if $\eta_A \geq \eta_B$ or the dimension of $\mathcal{S}_{B\perp A}$ if $\eta_A < \eta_B$.

We are now ready to present the following asymptotic result. Define $r^{\mathrm{SD}} \equiv \lim_{n_R\to\infty} \frac{\Delta^{\mathrm{SD}}}{n_R}$. Then, we have the following.

Theorem 4: As $n_R \to \infty$ with $\frac{n_A}{n_R} \to \eta_A$ and $\frac{n_B}{n_R} \to \eta_B$, the gap to the capacity upper bound satisfies

$$r^{\mathrm{SD}} = -\left(\int_1^{\frac{8}{5}} \log\sqrt{\lambda(2-\lambda)} + \int_{\frac{8}{5}}^2 \log\frac{\lambda}{2}\right)\mathcal{F}(\lambda; \eta_A, \eta_B)d\lambda. \tag{65}$$

Proof: The a.e.d. of λ_i is given by Lemma 3. Then, letting n_R tends to infinity in (59b), we immediately obtain the theorem. ■

Let $\overline{R}^{\mathrm{UL}}$ be the average sum-capacity upper bound. Then, for a large n_R, the average sum-rate of the proposed SD scheme can be first-order approximated as

$$\overline{R}^{\mathrm{SD}} = \overline{R}^{\mathrm{UL}} - n_R r^{\mathrm{SD}} \tag{66}$$

with r^{SD} given in (65).

We next study the symmetric case that the two users have the same number of antennas, i.e., $\eta_A = \eta_B = \eta$.

Corollary 6: For $0 \leq \eta \leq \frac{1}{10}$,

$$r^{\mathrm{SD}} = -\int_1^{\lambda^*(\eta)} \log\sqrt{\lambda(2-\lambda)}\mathcal{G}(\lambda; \eta)d\lambda; \tag{67a}$$

$$\mathcal{F}(\lambda; \eta_A, \eta_B) = [1 - \eta_A - \eta_B]^+ \delta(\lambda) + |\eta_A - \eta_B|\,\delta(\lambda - 1)\,[\eta_A + \eta_B - 1]^+ \delta(\lambda - 2) + \frac{1}{\pi}\mathrm{Im}\left[\sqrt{\left(\frac{1 - \eta_A - \eta_B}{2\lambda - \lambda^2}\right)^2 - \frac{1 - \left(\frac{\eta_A - \eta_B}{\lambda - 1}\right)^2}{(2\lambda - \lambda^2)}}\right] \tag{64}$$

for $\frac{1}{10} < \eta \leq 1$,

$$r^{\mathrm{SD}} = -\left(\int_1^{\frac{8}{5}} \log\sqrt{\lambda(2-\lambda)} + \int_{\frac{8}{5}}^{\lambda^*(\eta)} \log\frac{\lambda}{2}\right)\mathcal{G}(\lambda;\eta)d\lambda \tag{67b}$$

where $\lambda^*(\eta) = 1 + \sqrt{1-(1-2\eta)^2}$ and

$$\mathcal{G}(\lambda;\eta) = \frac{1}{\pi}\frac{\sqrt{(2\lambda-\lambda^2)-(1-2\eta)^2}}{2\lambda-\lambda^2}. \tag{67c}$$

Proof: Letting $\eta_A = \eta_B = \eta$, we obtain that $\mathcal{F}(\lambda;\eta_A,\eta_B) = \mathcal{G}(\lambda;\eta)$ for $1 < \lambda < \lambda^*$, and $\mathcal{F}(\lambda;\eta_A,\eta_B) = 0$ for $\lambda^* < \lambda < 2$. In addition, $\lambda^*(\eta) = \frac{8}{5}$ implies $\eta = \frac{1}{10}$. Based on these facts and Theorem 4, we obtain the corollary. ■

Remark 14: From the above, we see that, if $\eta \leq \frac{1}{10}$, the probability of $\lambda_i > \frac{8}{5}$ approaches zero as $n_R \to \infty$, implying that CD achieves a higher rate than PNC for all spatial streams.

Corollary 7: The asymptotic normalized rate gap r^{SD} in (65) is maximized at $\eta_A = \eta_B = 1/2$, with the maximum given by

$$-\frac{1}{\pi}\left(\int_1^{\frac{8}{5}} \frac{\log\sqrt{\lambda(2-\lambda)}}{\sqrt{2\lambda-\lambda^2}}d\lambda + \int_{\frac{8}{5}}^{2} \frac{\log\frac{\lambda}{2}}{\sqrt{2\lambda-\lambda^2}}d\lambda\right) \approx 0.053 \text{ bit}. \tag{68}$$

Proof: We first consider optimizing η_A and η_B under the constraint of $\eta_A + \eta_B = 2\eta$. From (64), we see that, for any $\lambda \in (1,2)$, $\mathcal{F}(\lambda;\eta_A,\eta_B)$ is maximized at $\eta_A = \eta_B = \eta$, and so is r^{SD}.

What remains is to optimize η. From (67c), $\mathcal{G}(\lambda;\eta)$ is maximized at $\eta = 1/2$. Therefore, r^{SD} is maximized at $\eta = 1/2$, which completes the proof. ■

Remark 15: We can represent the above rate gap in terms of the equivalent SNR gap as follows. From (62), together with $\Delta^{\mathrm{SD}} = r^{\mathrm{SD}} n_R$ and $\eta = n_m/n_R = 1/2, m \in \{A,B\}$, we obtain the corresponding SNR gap as

$$\Delta_{SNR} = \frac{10}{\log 10}\frac{r^{\mathrm{SD}}}{\eta} \approx 0.319 \text{ dB}. \tag{69}$$

Fig. 3 illustrates the function of the normalized asymptotic rate gap r^{SD} against η. From Fig. 3, this rate gap is maximized to be about 0.053 bit at $\eta = 1/2$, which verifies Corollary 7. Also, this rate gap vanishes as η tends to 0, implying that, for any fixed $n_A = n_B$, the proposed space-division scheme can achieve the asymptotic capacity as n_R tends to infinity. Moreover, this rate gap vanishes as η tends to 1. This agrees with the fact that, for $\eta = 1$, or equivalently, $n_A = n_B = n_R$, the proposed space-division scheme achieves the asymptotic capacity as it reduces to the GSVD scheme in [16].

VII. Numerical Results

In this section, we provide numerical results to evaluate the performance of the proposed space-division based network-coding strategy for MIMO TWRCs. The results presented here are obtained by averaging over 10 000 random channel realizations. Rayleigh-fading is always assumed, i.e., the coefficients in the channel matrices are independently and identically drawn from $\mathcal{N}_c(0,1)$.

We first present the numerical results for the MIMO TWRC with $n_A = n_B = 2$ and $n_R = 4$ in Fig. 4. The sum-capacity

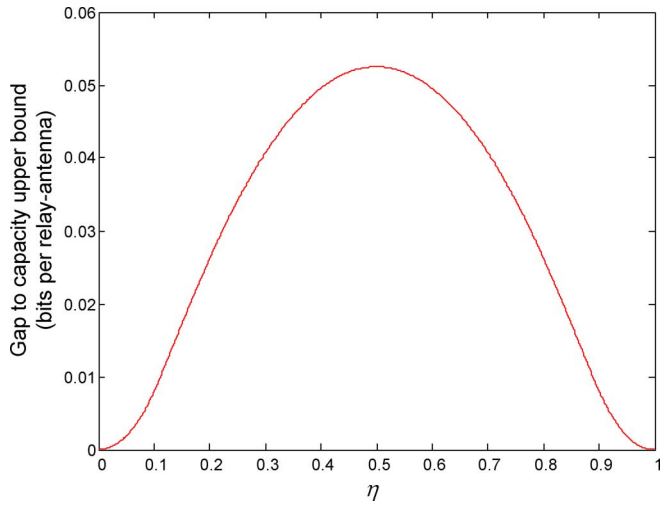

Fig. 3. Function of the average normalized gap r^{SD} in (65) against η.

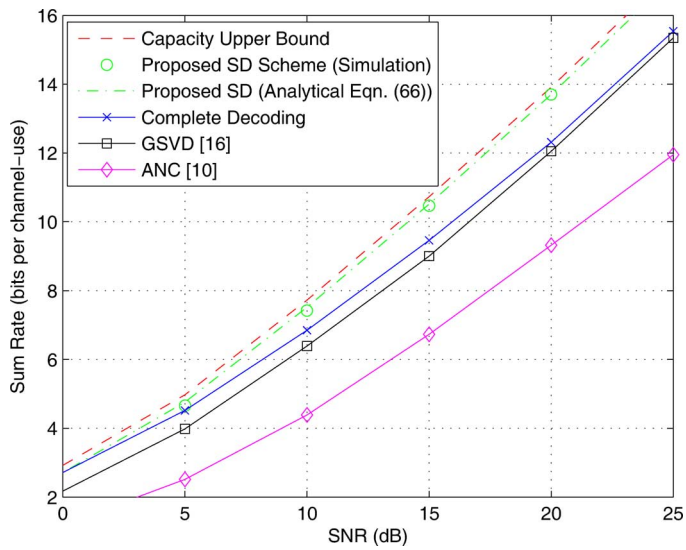

Fig. 4. Average achievable sum-rates of various schemes for the Rayleigh fading MIMO TWRC with $n_A = n_B = 2$ and $n_R = 4$.

upper bound (UB), the proposed space-division (SD) scheme, the GSVD scheme in [16] and the CD scheme in [17] are included for comparison. We see that, at a relatively high SNR, e.g., SNR = 25 dB, the rate gap between the proposed SD scheme and the sum-capacity upper bound is about 0.15 bit/channel-use, which is almost unnoticeable. We also plot the high-SNR analytical result in (66) of the proposed SD scheme. We observe that our analytical result are very tight for SNRs greater than 10 dB. From this figure, it is clear that the proposed SD scheme significantly outperforms the other schemes in the entire SNR range of interest. For example, at the rate of 14 bits per channel use, the proposed SD scheme outperforms the CD and GSVD schemes by more than 2.4 dB. The slope of the achievable sum-rate curve is parallel to that of the capacity upper bound, which implies that the proposed SD scheme achieves full multiplexing gain. Fig. 4 also includes the performance curve for the analog network coding (ANC) approach in [10]. We see that the proposed SD scheme outperforms ANC by a significant power gap of over 6 dB throughout the SNR range.

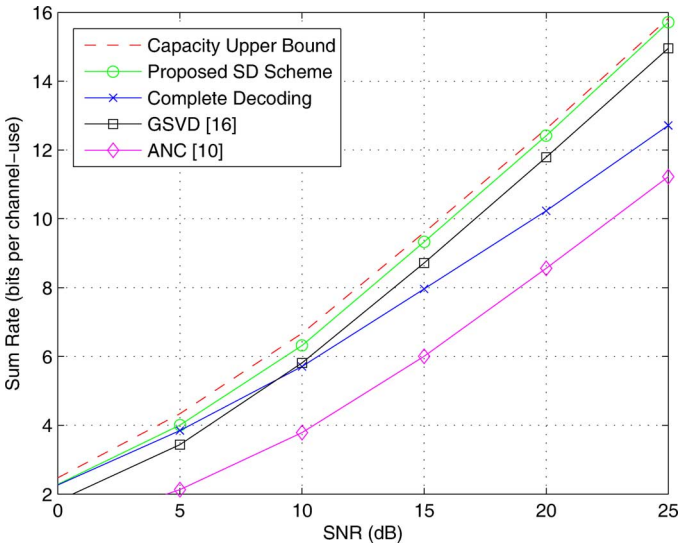

Fig. 5. Average achievable sum-rates of various schemes for the Rayleigh fading MIMO TWRC with $n_A = n_B = 2$ and $n_R = 3$.

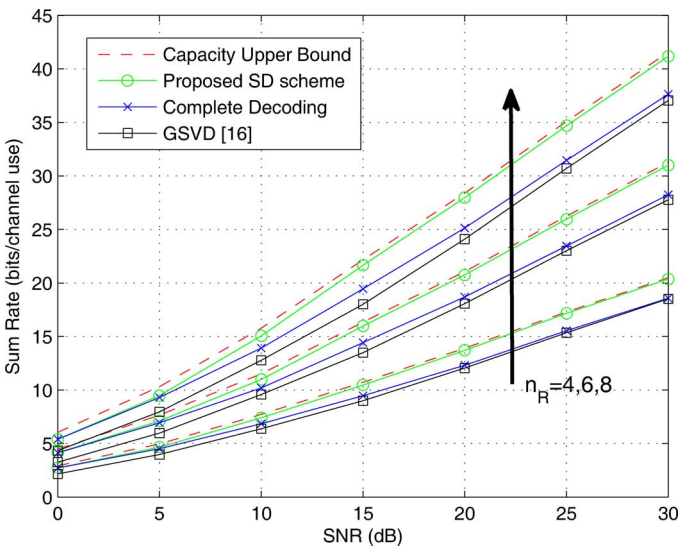

Fig. 6. Scaling effect of the average sum-rates of various schemes for the Rayleigh fading MIMO TWRCs with $\eta_A = \eta_B = 1/2$.

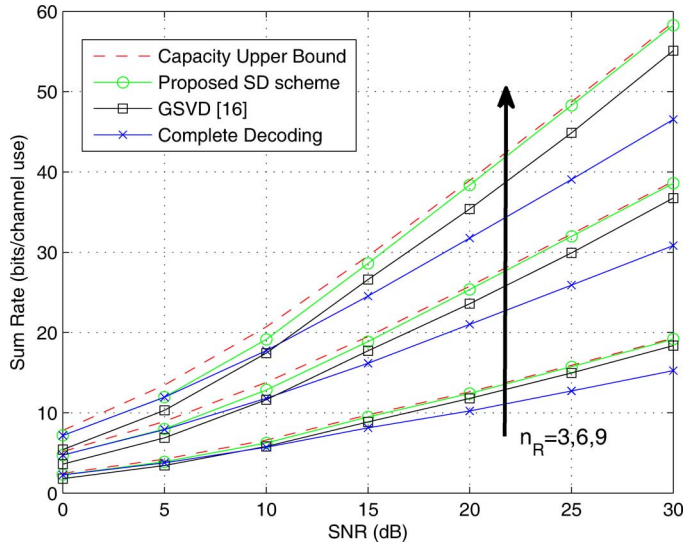

Fig. 7. Scaling effect of the average sum-rates of various schemes for the Rayleigh fading MIMO TWRCs with $\eta_A = \eta_B = 2/3$.

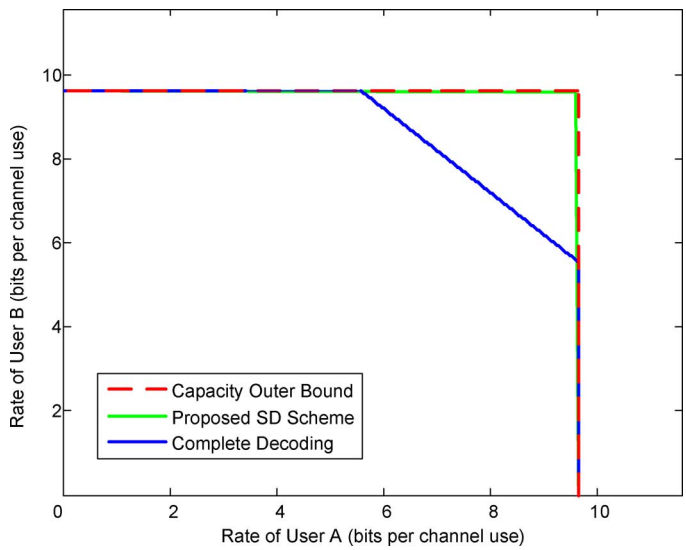

Fig. 8. Average achievable rate-regions for the Rayleigh fading MIMO TWRC with $n_A = n_B = 2$ and $n_R = 3$. The average SNRs for all the channel links are set to 30 dB.

In Fig. 5, we present the numerical results for the MIMO TWRC with $n_A = n_B = 2$ and $n_R = 3$. The same set of rate curves as in Fig. 4 are included for comparison. Again, we see that the gap between the sum-rate of the proposed SD scheme and the sum-capacity upper bound is almost unnoticeable at a relatively high SNR, say, $\mathrm{SNR} \geq 15$ dB. Again, we observe that the proposed SD scheme significantly outperforms its counterparts throughout the SNR range of interest.

In Figs. 6 and 7, we show the scaling effect of the antennas on the average achievable sum-rates. We see that the asymptotic rate gap between the proposed SD scheme and the sum-capacity upper bound increases linearly as the increase of n_R for fixed η_A and η_B. For example, for the case of $\eta_A = \eta_B = 1/2$ in Fig. 6, the rate gap at $\mathrm{SNR} = 25$ dB is 0.14 bits per channel use for $n_R = 4$; 0.29 bits per channel use for $n_R = 6$; and 0.40 bits per channel use for $n_R = 8$. These numerical results agree well with the asymptotic results in Corollaries 6 and 7.

In Fig. 8, we show the achievable rate-region of the proposed SD scheme. The capacity-region outer bound and the achievable rate-region of the CD scheme are also included for comparison. From Fig. 8, the difference between the achievable rate-region of the proposed SD scheme and the capacity region outer bound is negligible for a relatively high SNR. We also see that the proposed SD scheme can achieve rate-pairs that cannot be achieved by the CD scheme.

VIII. Conclusion

In this paper, we developed a new joint channel decomposition for the uplink of MIMO TWRCs. Based on that, we proposed a space-division-based network-coding scheme that is able to approach the sum-rate capacity of the MIMO

TWRC within $\frac{1}{2}\log(5/4) \approx 0.161$ bit per user-antenna in the high-SNR regime. We also showed that, for a large-system Rayleigh-fading MIMO TWRC, the average gap between the achievable sum-rate of the proposed scheme and the capacity is no more than 0.053 bit per relay-antenna in the high-SNR regime.

We conclude this paper by pointing out some promising research avenues for future work. First, the capacity of the MIMO TWRC still remains an open problem. There are possibilities to further reduce the gap toward the capacity, e.g., by designing more advanced multidimensional PNC relaying strategies, or by developing capacity upper bounds tighter than the cut-set bound. Second, global perfect CSI was assumed in this study. The relay design for the MIMO TWRC with imperfect CSI is an interesting topic for future research. For example, as mentioned in Remark 8, imperfect CSI generally impairs the achievable DoF of the proposed space-division scheme. Thereby, it is of pressing interest to develop a smarter signal-alignment technique that can avoid this DoF loss.

Appendix A
Optimal Solution to Problem (11)

Suppose that $R_A^{\mathrm{PNC}} = 0$ (or $R_B^{\mathrm{PNC}} = 0$). Then, from (9), the optimal $\mathbf{p}$ is trivially taken as $\frac{\mathbf{h}_{BR}}{\|\mathbf{h}_{BR}\|}$ (or $\frac{\mathbf{h}_{AR}}{\|\mathbf{h}_{AR}\|}$). Thus, we focus on the case of $R_m^{\mathrm{PNC}} > 0, m \in \{A, B\}$. In this case, this weighted sum-rate maximization problem is equivalent to maximizing

$$\max_{\|\mathbf{p}\|=1} w_A \log\left(\left|\mathbf{p}^\dagger \mathbf{h}_{AR}\right|^2\right) + w_B \log\left(\left|\mathbf{p}^\dagger \mathbf{h}_{BR}\right|^2\right) \quad (70)$$

or equivalently

$$\max_{\|\widetilde{\mathbf{p}}\|=1} \left|\widetilde{\mathbf{h}}_{AR}^T \widetilde{\mathbf{p}}\right|^{2w_A} \left|\widetilde{\mathbf{h}}_{BR}^T \widetilde{\mathbf{p}}\right|^{2w_B}$$

where $\widetilde{\mathbf{p}} = [\mathrm{Re}[\mathbf{p}^T], \mathrm{Im}[\mathbf{p}^T]]^T$ and $\widetilde{\mathbf{h}}_{mR} = [\mathrm{Re}[\mathbf{h}_{mR}^T], \mathrm{Im}[\mathbf{h}_{mR}^T]]^T$, $m \in \{A, B\}$. By setting the derivative of the Lagrangian with respect to $\mathbf{p}$ to zero, the optimal $\mathbf{p}$ satisfies

$$\alpha\widetilde{\mathbf{p}} = \frac{w_A\widetilde{\mathbf{h}}_{AR}}{\widetilde{\mathbf{p}}^T\widetilde{\mathbf{h}}_{AR}} + \frac{w_B\widetilde{\mathbf{h}}_{BR}}{\widetilde{\mathbf{p}}^T\widetilde{\mathbf{h}}_{BR}},$$

where α is a scaling factor. Then, with some straightforward algebra, we obtain the optimal projection direction given by

$$\widetilde{\mathbf{p}}^{\mathrm{opt}} = \gamma\left(\frac{\widetilde{\mathbf{h}}_{AR}}{\left\|\widetilde{\mathbf{h}}_{AR}\right\|} + \beta\frac{\widetilde{\mathbf{h}}_{BR}}{\left\|\widetilde{\mathbf{h}}_{BR}\right\|}\right), \quad (71)$$

where β is given in (72), at the bottom of the page, and γ is a scaling factor to ensure $\|\widetilde{\mathbf{p}}^{opt}\| = 1$.

Appendix B
Proof of Lemmas 1–4

1) Proof of Lemma 1: If $\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i = \mathbf{0}$, then (15) implies $\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{u}_i = \mathbf{u}_i$. Thus, (16a) holds, which proves the lemma. Therefore, it suffices to consider the situation of $\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i \neq \mathbf{0}$. We multiply both sides of (15) by $\mathbf{U}_A\mathbf{U}_A^\dagger$. Then, after some straightforward manipulations, we obtain

$$(\lambda_i - 1)\,\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{u}_i = \mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i. \quad (73)$$

Noting $\lambda_i = 1$, we obtain

$$\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i = \mathbf{0}. \quad (74)$$

Recall that $\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i \neq \mathbf{0}$. Thus, (74) implies that the vector $\mathbf{a} = \mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i \neq \mathbf{0}$ is in the null space of the matrix $\mathbf{U}_A\mathbf{U}_A^\dagger$, i.e., $\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{a} = \mathbf{0}$. On the other hand, we have $\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{a} = \mathbf{a}$. Therefore, $\mathbf{a}$ is actually an eigenvector of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$ satisfies (16b), which concludes the proof.

2) Proof of Lemma 2: Similar to (74), we have

$$(\lambda_i - 1)\,\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i = \mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{u}_i. \quad (75)$$

Then

$$\mathbf{u}_i^\dagger\mathbf{l}_{A;i} \overset{(a)}{=} \mathbf{u}_i^\dagger\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{u}_i \overset{(b)}{=} \frac{1}{\lambda_i - 1}\mathbf{u}_i^\dagger\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i$$
$$\overset{(c)}{=} \frac{1}{\lambda_i - 1}\mathbf{u}_i^\dagger\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{u}_i \overset{(d)}{=} \mathbf{u}_i^\dagger\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i \overset{(e)}{=} \mathbf{u}_i^\dagger\mathbf{l}_{B;i}$$

where step (a) follows from (17), (b) from (74), (c) from the fact that the Hermitian transpose of a real-valued scalar is itself, (d) from (75), and (e) again from (17). From (17), the projection of $\mathbf{u}_i$ onto $\mathbf{l}_{m;i}$ is just $\mathbf{l}_{m;i}$. Thus, $\|\mathbf{l}_{m;i}\|^2 = \mathbf{u}_i^\dagger\mathbf{l}_{m;i}$, which completes the proof.

$$\beta = \frac{\mathrm{sign}(\widetilde{\mathbf{h}}_{AR}^T\widetilde{\mathbf{h}}_{BR})}{2}\left(\sqrt{\left(\frac{\widetilde{\mathbf{h}}_{AR}^T\widetilde{\mathbf{h}}_{BR}\left(1 - \frac{w_B}{w_A}\right)}{\left\|\widetilde{\mathbf{h}}_{AR}\right\|\left\|\widetilde{\mathbf{h}}_{BR}\right\|}\right)^2 + 4\frac{w_B}{w_A}} - \frac{\left|\widetilde{\mathbf{h}}_{AR}^T\widetilde{\mathbf{h}}_{BR}\right|\left(1 - \frac{w_B}{w_A}\right)}{\left\|\widetilde{\mathbf{h}}_{AR}\right\|\left\|\widetilde{\mathbf{h}}_{BR}\right\|}\right) \quad (72)$$

3) Proof of Lemma 3: By definition, we have

$$\begin{aligned}
&\left(\mathbf{U}_A\mathbf{U}_A^\dagger+\mathbf{U}_B\mathbf{U}_B^\dagger\right)\mathbf{u}_i' \\
&\overset{(a)}{=}\frac{1}{\sqrt{\lambda_i\lambda_i'}}\left(\mathbf{U}_A\mathbf{U}_A^\dagger+\mathbf{U}_B\mathbf{U}_B^\dagger\right)(\mathbf{l}_{A;i}-\mathbf{l}_{B;i}) \\
&\overset{(b)}{=}\frac{1}{\sqrt{\lambda_i\lambda_i'}}\Big(\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{u}_i+(\lambda_i-1)\,\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i \\
&\qquad\qquad -(\lambda_i-1)\,\mathbf{U}_A\mathbf{U}_A^\dagger\mathbf{u}_i-\mathbf{U}_B\mathbf{U}_B^\dagger\mathbf{u}_i\Big) \\
&=\frac{\lambda_i'}{\sqrt{\lambda_i\lambda_i'}}(\mathbf{l}_{A;i}-\mathbf{l}_{B;i})=\lambda_i'\mathbf{u}_i' \qquad (76)
\end{aligned}$$

where step (a) follows from (20), and step (b) from (17), (74) and (75).

What remains is to show that $\|\mathbf{u}_i'\|=1$. To see this, we left-multiply both sides of (15) by $\mathbf{u}_i^\dagger$, yielding

$$\|\mathbf{l}_{A;i}\|^2+\|\mathbf{l}_{B;i}\|^2=\lambda_i. \qquad (77)$$

Together with (19), we obtain

$$\|\mathbf{l}_{A;i}\|^2=\|\mathbf{l}_{B;i}\|^2=\frac{\lambda_i}{2}. \qquad (78)$$

Moreover, left-multiplying (74) and (75), respectively, by $\mathbf{u}_i^\dagger$ and plugging in (17), we obtain

$$\mathbf{l}_{A;i}^\dagger\mathbf{l}_{B;i}=(\lambda_i-1)\,\|\mathbf{l}_{A;i}\|^2=(\lambda_i-1)\,\|\mathbf{l}_{B;i}\|^2=\mathbf{l}_{B;i}^\dagger\mathbf{l}_{A;i}. \qquad (79)$$

Then

$$\begin{aligned}
\mathbf{u}_i'^\dagger\mathbf{u}_i' &= \frac{1}{\lambda_i\lambda_i'}(\mathbf{l}_{A;i}-\mathbf{l}_{B;i})^\dagger(\mathbf{l}_{A;i}-\mathbf{l}_{B;i}) \\
&= \frac{1}{\lambda_i\lambda_i'}\left(\|\mathbf{l}_{A;i}\|^2-\mathbf{l}_{A;i}^\dagger\mathbf{l}_{B;i}-\mathbf{l}_{B;i}^\dagger\mathbf{l}_{A;i}+\|\mathbf{l}_{B;i}\|^2\right) \\
&\overset{(a)}{=} \frac{1}{\lambda_i\lambda_i'}\left(\lambda_i-(\lambda_i-1)\,\|\mathbf{l}_{A;i}\|^2-(\lambda_i-1)\,\|\mathbf{l}_{B;i}\|^2\right) \\
&\overset{(b)}{=} 1 \qquad (80)
\end{aligned}$$

where step (a) follows from (78) and (79), and step (b) from (78) and the fact of $\lambda_i'=2-\lambda_i$. This completes the proof of Lemma 3.

4) Proof of Lemma 4: From (18) and (20), we see that both $\mathbf{u}_i$ and $\mathbf{u}_i'$ lie on the plane $\mathcal{S}_i$. As $\mathbf{u}_i$ and $\mathbf{u}_i'$ are orthogonal to each other, $\mathcal{S}_i$ is spanned by $\mathbf{u}_i$ and $\mathbf{u}_i'$. Then, the lemma holds straightforwardly by noting the orthogonality between the eigenvectors.

Appendix C
Proof of Theorem 2

Here, we provide a sketch of the proof of Theorem 2. The overall encoding and decoding process for the proposed scheme is described as follows. The messages of the user m are doubly indexed as $(W_m^{\mathrm{CD}},W_m^{\mathrm{PNC}})$, with $W_m^{\mathrm{CD}}\in\{1,2,\ldots,2^{2TR_m^{\mathrm{CD}}}\}$ for the CD spatial streams, and $W_m^{\mathrm{PNC}}\in\{1,2,\ldots,2^{2TR_m^{\mathrm{PNC}}}\}$ for the PNC spatial streams. Each W_m^{CD} is one-to-one mapped to $\mathbf{X}_m^{\mathrm{CD}}$ in (35), and each W_m^{PNC} is one-to-one mapped to $\mathbf{C}_m^{\mathrm{PNC}}$ in (45). In the uplink phase, $\mathbf{X}_m^{\mathrm{CD}}$ and $\mathbf{X}_m^{\mathrm{PNC}}=\widetilde{\mathbf{S}}_m\mathbf{\Psi}_m^{1/2}\mathbf{C}_m^{\mathrm{PNC}}$ are transmitted via the channel in (35), with the transmit power limited by (47).

Upon receiving $\mathbf{Y}_R$, the relay first completely decodes $\mathbf{X}_A^{\mathrm{CD}}$ and $\mathbf{X}_B^{\mathrm{CD}}$ based on $\mathbf{Y}_R^{\mathrm{CD}}$ in (37), with the achievable rate-pair given in (42). The decoded $\mathbf{X}_A^{\mathrm{CD}}$ and $\mathbf{X}_B^{\mathrm{CD}}$ are subtracted from $\mathbf{Y}_R^{\mathrm{PNC}}$. Then, the network-coded PNC spatial streams, i.e., $\tilde{r}_{i,i}\sigma_{A;i}\psi_{A;i}^{1/2}\mathbf{c}_{A,i}^{\mathrm{PNC}}+\tilde{r}_{i,i}\sigma_{B;i}\psi_{B;i}^{1/2}\mathbf{c}_{B,i}^{\mathrm{PNC}}$, $i=k+l',k+l'-1,\ldots,1$, are successively recovered and canceled from the received signal, with the achievable rate-pair given in (46). The decoded messages from the CD streams, together with the network-coded messages from the PNC streams, are then jointly encoded. The new codeword is forwarded to the two users in the downlink phase, under the transmit power constraint of $\mathrm{tr}(\mathbf{Q}_R)\le P_R$. Following the discussions in [16]–[18], the achievable rate-pair of the downlink phase is given by $(R_A^{\mathrm{DL}},R_B^{\mathrm{DL}})$ in (4). This completes the proof.

Appendix D
Proof of Theorem 3

We first consider the sum-rate upper bound:

$$\begin{aligned}
R^{\mathrm{UL}} &\overset{(a)}{\approx}\frac{1}{2}\sum_{m\in\{A,B\}}\log\left|\mathbf{I}_{n_R}+\frac{P_m}{N_0n_m}\mathbf{U}\mathbf{D}_m\mathbf{R}_m\mathbf{R}_m^\dagger\mathbf{D}_m^\dagger\mathbf{U}^\dagger\right| \\
&\overset{(b)}{=}\frac{1}{2}\sum_{m\in\{A,B\}}\log\left|\mathbf{I}_{n_m}+\frac{P_m}{N_0n_m}\mathbf{R}_m\mathbf{R}_m^\dagger\right| \\
&\overset{(c)}{\approx}\frac{1}{2}\sum_{m\in\{A,B\}}\log\left|\frac{P_m}{N_0n_m}\mathbf{R}_m\mathbf{R}_m^\dagger\right| \qquad (81)
\end{aligned}$$

where step (a) follows by substituting (32) into (58), step (b) follows from the facts that $\mathbf{D}_m^\dagger\mathbf{U}^\dagger\mathbf{U}\mathbf{D}_m=\mathbf{I}_{n_m}$ and $|\mathbf{I}+\mathbf{A}\mathbf{B}|=|\mathbf{I}+\mathbf{B}\mathbf{A}|$, and step (c) utilizes the fact that $\mathbf{R}_m$ is a square matrix.

Now we consider the achievable sum-rate of the proposed space-division scheme. For notational simplicity, let $\mathbf{H}_m^{\mathrm{CD}}=\mathbf{D}_{m;2,2}\mathbf{R}_{m;2,2}$, $m\in\{A,B\}$. From (42), the sum-rate of the CD spatial streams is given by (82), at the bottom of the next page, where (82a) utilizes the fact that equal power allocation is asymptotically optimal, and (82d) follows by substituting $\mathbf{H}_B^{\mathrm{CD}}=\mathbf{D}_{B;2,2}\mathbf{R}_{B;2,2}$. Applying the matrix inversion lemma to $\left(\mathbf{I}+\frac{P_A}{N_0n_A}\mathbf{H}_A^{\mathrm{CD}}(\mathbf{H}_A^{\mathrm{CD}})^\dagger\right)^{-1}$, we further obtain (83), given at the bottom of the next page, where (83b) follows by noting $\mathbf{I}+\frac{P_A}{N_0n_A}(\mathbf{H}_A^{\mathrm{CD}})^\dagger\mathbf{H}_A^{\mathrm{CD}}\approx\frac{P_A}{N_0n_A}(\mathbf{H}_A^{\mathrm{CD}})^\dagger\mathbf{H}_A^{\mathrm{CD}}$ and $\mathbf{H}_A^{\mathrm{CD}}=\mathbf{D}_{A;2,2}\mathbf{R}_{A;2,2}$, $m\in\{A,B\}$, and (83c) utilizes the definitions

in (23a) and (34a). Moreover, letting $w_A = w_B = 1$ in (50), we obtain the sum-rate of the PNC spatial streams as

$$R_A^{\mathrm{PNC}} + R_B^{\mathrm{PNC}} = \frac{1}{2} \sum_{m\in\{A,B\}} \log \left| \frac{P_m}{N_0 n_m} \mathbf{R}_{m;1,1} \mathbf{R}_{m;1,1}^{\dagger} \right| + \frac{1}{2} \sum_{m\in\{A,B\}} \log \left| \tilde{\mathbf{D}}_{m;1,1} \tilde{\mathbf{D}}_{m;1,1}^{\dagger} \right|. \tag{84}$$

From (52), $\mathbf{p}_i$ is the angular bisection of $\mathbf{e}_{A;i}$ and $\mathbf{e}_{B;i}$, or equivalently, $\mathbf{p}_i = [1, 0]^T$, for the sum-rate case of $w_A = w_B = 1$. Then, using the definition in (41b), we obtain

$$\log \left| \tilde{\mathbf{D}}_{m;1,1} \tilde{\mathbf{D}}_{m;1,1}^{\dagger} \right| = \sum_{i=k+1}^{k+l} \log \frac{\lambda_i}{2}. \tag{85}$$

Combining (81)–(85), we complete the proof of Theorem 3.

Appendix E
Proof of Lemma 5

We prove by using the theory of free probability [32]. The a.e.d. of $\mathbf{U}_m \mathbf{U}_m^{\dagger}$ is given by

$$p_m(\lambda) = \eta_m \delta(\lambda - 1) + (1 - \eta_m)\, \delta(\lambda), \; m \in \{A, B\}.$$

Let X_m be a random variable with PDF $p_m(\lambda)$. Its Stieltjes transform is given by (cf., [31, (2.40)])

$$S_{X_m}(z) = E\left[\frac{1}{X_m - z}\right] = \frac{\eta_m}{1 - z} - \frac{1 - \eta_m}{z}.$$

Then, the inverse function of $S_{X_m}(z)$ is given by

$$S_{X_m}^{-1}(s) = \frac{-(1 - s) \pm \sqrt{(1 - s)^2 - 4s(\eta_m - 1)}}{2s}.$$

Using the relation between Stieltjes transform and R-transform (cf., [31, (2.72)]), we obtain the R-transform of X_m as

$$R_{X_m}(z) = S_{X_m}^{-1}(-z) - \frac{1}{z} = \frac{z - 1 \mp \sqrt{(z - 1)^2 + 4\eta_m z}}{2z}.$$

From [31, Th. 2.64], as $\mathbf{U}_A \mathbf{U}_A^{\dagger}$ and $\mathbf{U}_B \mathbf{U}_B^{\dagger}$ are asymptotically free random matrices, the R-transform of the a.e.d. of $\mathbf{U}_A \mathbf{U}_A^{\dagger} + \mathbf{U}_B \mathbf{U}_B^{\dagger}$ is given by

$$R_{AB}(z) = R_{X_A}(z) + R_{X_B}(z) = \sum_{m\in\{A,B\}} \frac{z - 1 \mp \sqrt{(z - 1)^2 + 4\eta_m z}}{2z}.$$

$$R_A^{\mathrm{CD}} + R_B^{\mathrm{CD}} = \frac{1}{2} \log \left| \mathbf{I} + \sum_{m\in\{A,B\}} \frac{P_m}{N_0 n_m} \mathbf{H}_m^{\mathrm{CD}} (\mathbf{H}_m^{\mathrm{CD}})^{\dagger} \right| \tag{82a}$$

$$= \frac{1}{2} \log \left| \mathbf{I} + \frac{P_A}{N_0 n_A} \mathbf{H}_A^{\mathrm{CD}} (\mathbf{H}_A^{\mathrm{CD}})^{\dagger} \right| + \frac{1}{2} \log \frac{\left| \mathbf{I} + \sum_{m\in\{A,B\}} \frac{P_m}{N_0 n_m} \mathbf{H}_m^{\mathrm{CD}} (\mathbf{H}_m^{\mathrm{CD}})^{\dagger} \right|}{\left| \mathbf{I} + \frac{P_A}{N_0 n_A} \mathbf{H}_A^{\mathrm{CD}} (\mathbf{H}_A^{\mathrm{CD}})^{\dagger} \right|} \tag{82b}$$

$$\approx \frac{1}{2} \log \left| \frac{P_A}{N_0 n_A} \mathbf{R}_{A;2,2} \mathbf{R}_{A;2,2}^{\dagger} \right| + \frac{1}{2} \log \left| \frac{P_B}{N_0 n_B} (\mathbf{H}_B^{\mathrm{CD}})^{\dagger} \left(\mathbf{I} + \frac{P_A}{N_0 n_A} \mathbf{H}_A^{\mathrm{CD}} (\mathbf{H}_A^{\mathrm{CD}})^{\dagger} \right)^{-1} \mathbf{H}_B^{\mathrm{CD}} \right| \tag{82c}$$

$$= \sum_{m\in\{A,B\}} \frac{1}{2} \log \left| \frac{P_m}{N_0 n_m} \mathbf{R}_{m;2,2} \mathbf{R}_{m;2,2}^{\dagger} \right| + \frac{1}{2} \log \left| \mathbf{D}_{B;2,2}^{\dagger} \left(\mathbf{I} + \frac{P_A}{N_0 n_A} \mathbf{H}_A^{\mathrm{CD}} (\mathbf{H}_A^{\mathrm{CD}})^{\dagger} \right)^{-1} \mathbf{D}_{B;2,2} \right| \tag{82d}$$

$$R_A^{\mathrm{CD}} + R_B^{\mathrm{CD}} = \sum_{m\in\{A,B\}} \frac{1}{2} \log \left| \frac{P_m}{N_0 n_m} \mathbf{R}_{m;2,2} \mathbf{R}_{m;2,2}^{\dagger} \right| + \frac{1}{2} \log \left| \mathbf{I} - \mathbf{D}_{B;2,2}^{\dagger} \mathbf{H}_A^{\mathrm{CD}} \left(\frac{N_0 n_A}{P_A} \mathbf{I} + (\mathbf{H}_A^{\mathrm{CD}})^{\dagger} \mathbf{H}_A^{\mathrm{CD}} \right)^{-1} (\mathbf{H}_A^{\mathrm{CD}})^{\dagger} \mathbf{D}_{B;2,2} \right| \tag{83a}$$

$$\approx \sum_{m\in\{A,B\}} \frac{1}{2} \log \left| \frac{P_m}{N_0 n_m} \mathbf{R}_{m;2,2} \mathbf{R}_{m;2,2}^{\dagger} \right| + \frac{1}{2} \log \left| \mathbf{I} - \mathbf{D}_{B;2,2}^{\dagger} \mathbf{D}_{A;2,2} \mathbf{D}_{A;2,2}^{\dagger} \mathbf{D}_{B;2,2} \right| \tag{83b}$$

$$= \sum_{m\in\{A,B\}} \frac{1}{2} \log \left| \frac{P_m}{N_0 n_m} \mathbf{R}_{m;2,2} \mathbf{R}_{m;2,2}^{\dagger} \right| + \frac{1}{2} \log \prod_{i=k+l+1}^{k+l} \lambda_i (2 - \lambda_i) \tag{83c}$$

Then, the Stieltjes transform of the a.e.d. of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$ satisfies

$$S_{AB}^{-1}(-z) = 1 \mp \sum_{m\in\{A,B\}} \frac{\sqrt{(z-1)^2 + 4\eta_m z}}{2z}.$$

Letting $y = S_{AB}^{-1}(-z)$, we obtain

$$\sum_{m\in\{A,B\}} \sqrt{(z-1)^2 + 4\eta_m z} = \mp 2z(y-1).$$

Multiplying $\sqrt{(z-1)^2 + 4\eta_A z} - \sqrt{(z-1)^2 + 4\eta_B z}$ on both sides, we have

$$\sqrt{(z-1)^2 + 4\eta_A z} - \sqrt{(z-1)^2 + 4\eta_B z} = \mp\frac{2(\eta_A - \eta_B)}{y-1}.$$

Adding the above two equations and taking the square, we further obtain

$$(z-1)^2 + 4\eta_A z = \left(z(y-1) + \frac{\eta_A - \eta_B}{y-1}\right)^2.$$

Solving z, we obtain

$$S_{AB}(z) = \pm\frac{\sqrt{(1-\eta_A-\eta_B)^2 + (2z-z^2)\left(\left(\frac{\eta_A-\eta_B}{z-1}\right)^2 - 1\right)}}{2z-z^2} - \frac{1-\eta_A-\eta_B}{2z-z^2}.$$

From [31, (2.45)], the a.e.d. of $\mathbf{U}_A\mathbf{U}_A^\dagger + \mathbf{U}_B\mathbf{U}_B^\dagger$ is given by

$$\mathcal{F}(\lambda) = \lim_{\omega\to 0^+} \frac{1}{\pi}\operatorname{Im}\left[S_{AB}(\lambda + j\omega)\right].$$

Thus, for $0 < \lambda < 1$ and $1 < \lambda < 2$, we obtain

$$\mathcal{F}(\lambda) = \frac{1}{\pi}\operatorname{Im}\left[\frac{\sqrt{(1-\eta_A-\eta_B)^2 + (2\lambda-\lambda^2)\left(\left(\frac{\eta_A-\eta_B}{\lambda-1}\right)^2 - 1\right)}}{2\lambda-\lambda^2}\right]. \tag{86}$$

In addition, for a randomly generated pair of $\mathbf{U}_A$ and $\mathbf{U}_B$, there are $n_A + n_B - n_R$ orthogonal eigenvectors for $\lambda_i = 2$, $|n_A - n_B|$ orthogonal eigenvectors for $\lambda_i = 1$, and $n_R - n_A - n_B$ orthogonal eigenvectors for $\lambda_i = 0$. Thus, as n_R tends to infinity, the PDF $\mathcal{F}(\lambda)$ at $\lambda = 2$ is given by $[\eta_A + \eta_B - 1]^+\delta(\lambda - 2)$; that at $\lambda = 1$ is given by $|\eta_A - \eta_B|\,\delta(\lambda - 1)$; and that at $\lambda = 0$ is given by $[1 - \eta_A - \eta_B]^+\delta(\lambda)$. This concludes the proof of the lemma.

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Oct. 2000.

[2] S. Zhang, S. Liew, and P. P. Lam, "Hot topic: Physical-layer network coding," presented at the 12th Ann. Int. Conf. Mobile Comput. Netw., Los Angeles, CA, USA, Sep. 2006.

[3] B. Nazer and M. Gastpar, "The case for structured random codes in network capacity theorems," *Eur. Trans. Telecommn.*, vol. 19, no. 1, Jan./Feb. 2008.

[4] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," presented at the IEEE Int. Conf. Commun., Glasgow, U.K., Jun. 2007.

[5] W. Nam, S. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within 1/2 bit," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

[6] M. P. Wilson, K. Narayanan, H. D. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5641–5654, Nov. 2010.

[7] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007.

[8] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," presented at the ACM SIGCOMM, Kyoto, Japan, Aug. 2007.

[9] 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.

[10] S. Xu and Y. Hua, "Optimal design of spatial source-and-relay matrices for a non-regenerative two-way MIMO relay system," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1645–1655, May 2011.

[11] M. Aleksic, P. Razaghi, and W. Yu, "Capacity of a class of modulo-sum relay channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 921–930, Mar. 2009.

[12] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.

[13] 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, Sep. 2005.

[14] A. El Gamal, N. Hassanpour, and J. Mammen, "Relay networks with delays," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3413–3431, Oct. 2007.

[15] D. Gunduz, A. Goldsmith, and H. V. Poor, "MIMO two-way relay channel: Diversity-multiplexing tradeoff analysis," in *Proc. IEEE Asilomar Conf. Signals Syst. Comput.*, Oct. 2008, pp. 1474–1478.

[16] H. J. Yang, J. Chun, and A. Paulraj, "Asymptotic capacity of the separated MIMO two-way relay channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7542–7554, Nov. 2011.

[17] T. Yang, X. Yuan, P. Li, I. B. Collings, and J. Yuan, "A new physical-layer network coding scheme with eigen-direction alignment precoding for MIMO two-way relaying," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 973–986, Mar. 2013.

[18] A. Khina, Y. Kochman, and U. Erez, "Physical-layer MIMO relaying," presented at the IEEE Int. Symp. Inf. Theory, Saint Petersburg, Russia, Jul./Aug. 2011.

[19] T. Yang, X. Yuan, and I. B. Collings, "Reduced-dimension cooperative precoding for MIMO two-way relay channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4150–4160, Nov. 2012.

[20] C. Y. Leow, Z. Ding, and K. K. Leung, "Joint beamforming and power management for nonregenerative MIMO two-way relaying channels," *EEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4374–4383, Nov. 2011.

[21] F. Wang, X. Yuan, S. C. Liew, and D. Guo, "Wireless MIMO switching: Weighted sum mean square error and sum rate optimization," *IEEE Trans. Inf. Theory*, to be published, to be published.

[22] B. Nazer and M. Gasper, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, pp. 6463–6486, Oct. 2011.

[23] 3GPP, "Overview of 3GPP Release 10," Release 10 V0.1.1 (2011-6) Jun. 2011 [Online]. Available: http://www.3gpp.org/Release-10

[24] *IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Broadband Wireless Access Systems Amendment 3: Advanced Air Interface*, IEEE Std 802.16m-2011, May 2011.

[25] U. Erez and R. Zamir, "Achieving $1/2\log(1+\mathrm{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[26] G. H. Golub and C. F. V. Loan, *Matrix Computation*, 3rd ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1997.

[27] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, no. 5, pp. 805–814, Oct. 1974.

[28] H. Liao, "Multiple access channels," Ph.D. dissertation, Dept. Electr. Eng., Univ. Hawaii, Honolulu, HI, USA, 1972.
[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.
[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[31] A. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Delft, The Netherlands: Now Publishers, 2004.
[32] D. Voiculescu, "The analogues of entropy and of Fisher's information measure in free probability theory, I," *Commun. Math. Phys.*, vol. 155, no. 1, pp. 71–92, Jul. 1993.
[33] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[34] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative waterfilling for Gaussian vector multiple access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–151, Jan. 2004.
[35] T. Huang, T. Yang, J. Yuan, and I. Land, "Design of irregular repeat-accumulate coded physical-layer network coding for Gaussian two-way relay channels," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 897–909, Mar. 2013.
[36] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

Xiaojun Yuan (S'04–M'09) received the Ph.D. degree in electrical engineering from the Department of Electronic Engineering, City University of Hong Kong, 2008. During 2009–2010, he was working as a post-doc research fellow at the Department of Electrical Engineering, University of Hawaii at Manoa. He is now a research assistant professor at the Institute of Network Coding, The Chinese University of Hong Kong. His research interests are in the general areas of wireless communications, signal processing, and information theory, including physical-layer network coding, cooperative communications, massive MIMO communications, channel coding and coded modulation, etc. He has now published over 40 research papers in the leading journals and conferences in the related areas, and has served as the Technical Program Committee (TPC) members for several international conferences including IEEE International Conf. on Communications (ICC), IEEE Wireless Communications and Networking Conf. (WCNC), etc.

Tao Yang (S'07–M'10) received B.Sc. degree in electronic engineering in 2003 from Beijing University of Aeronautics and Astronautics (Beihang University), Beijing, China. He received Master by research and Ph.D. degrees in electrical engineering from the University of New South Wales, Sydney, Australia, in 2006 and 2010, respectively. He was an OCE Postdoc Research Fellow in the Wireless and Networking Technologies Laboratory (WNTL) at Commonwealth Scientific and Industrial Research Organization (CSIRO), Sydney, Australia. He is now with School of electrical engineering and telecommunications, the University of New South Wales, Australia. His research expertise and interests include multi-user and MIMO communications, error-control coding, iterative signal processing and decoding, physical-layer network coding and network information theory. He has published more than 30 research papers in the top IEEE journals and conferences. He was the recipient of Australian Postgraduate Award (APA), NICTA research project award (NRPA) and Supplementary Engineering Award (SEA) from the University of New South Wales.

Iain B. Collings (S'92–M'95–SM'02) received the B.E. degree in Electrical and Electronic Engineering from the University of Melbourne in 1992, and the Ph.D. degree in Systems Engineering from the Australian National University in 1995. Currently he is the Research Director of the Wireless and Networking Technologies Laboratory at the Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO). Prior to this he was an Associate Professor at the University of Sydney (1999–2005); a Lecturer at the University of Melbourne (1996–1999); and a Research Fellow in the Australian Cooperative Research Centre for Sensor Signal and Information Processing (1995). He has published over 260 research papers in the area of wireless digital communications. In 2009 he was awarded the Engineers Australia IREE Neville Thiele Award for outstanding achievements in engineering, and in 2011 he was a recipient of the IEEE CommSoc Stephen O. Rice Best Paper Award for IEEE TRANSACTIONS ON COMMUNICATIONS.

Dr. Collings served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2002–2009), and the Elsevier Physical Communication Journal PHYCOM (2008–2012). He has served as a Co/Vice-Chair of the Conference Technical Program Committees for IEEE International Conf. on Communications (ICC) 2013, IEEE Vehicular Technology Conf. (VTC) Spring 2011, IEEE Wireless Communications and Networking Conf. (WCNC) 2010, and IEEE VTC Spring 2006. He is a founding organizer of the Australian Communication Theory Workshops 2000–2013. He also served as the Chair of the Joint Communications & Signal Processing Chapter in the IEEE NSW Section (2008–2010), and as Secretary of the IEEE NSW Section (2010).