Forty-Eighth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 29 - October 1, 2010

Asymptotic Capacity of the Separated MIMO Two-Way Relay Channel with Linear Precoding

Hyun Jong Yang*, Joohwan Chun*, and Arogyaswami Paulraj†
*Dept. of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea
Email: {hjyang, chun}@sclab.kaist.ac.kr
†Dept. of Electrical Engineering, Stanford University, Stanford, CA
Email: apaulraj@stanford.edu

***Abstract*—A multiple-input multiple-output two-way relay channel consisting of two communication nodes and a full-duplex relay node is considered, assuming that no direct link exists between the two communication nodes. We propose an achievable scheme using lattice codes combined with generalized singular value decomposition-based precoding for the first phase and vertically encoded structural bining for the second phase. We show that the proposed scheme achieves the cut-set bound asymptotically as the signal-to-noise ratio tends to infinity.**

***Index Terms*—Capacity of two-way relay channels, lattice coding, generalized singular value decomposition, linear precoding**

I. Introduction

Two-way relaying (TWR) was first considered and defined by Shannon [1], and has since been intensively studied in wireless channels [2]–[5], with the expectation that it will achieve greatly increased spectral efficiency. Generally, in TWR, the relay node simultaneously receives and transmits signals between the two communication nodes. In this way, the number of phases (e.g., time slots or frequency bands) required can be halved, compared to one-way relaying, which results in the potential benefit of a doubled spectral efficiency. Based on the process performed by the relay node, TWR schemes can be roughly classified into three categories: amplify-and-forward (AF), compress-and-forward (CF), or decode-and-forward (DF) TWR.

In AF TWR, the relay node simply amplifies and retransmits the received signal without any decoding or detection, and hence, the computational or implemental complexity is minimal, compared with the other types of TWR. On the other hand, AF TWR suffers from an achievable rate loss due to the noise propagation at the relay node. Many AF TWR schemes have been proposed for MIMO TWR channels (TWRC) as well as for single-input single-output (SISO) TWRCs [6]–[12]. Optimal precoding methods have been proposed for multiple-input single-output (MISO) and MIMO TWRC in [6], [7], [12].

In DF TWR, the relay node can decode the two codewords from the two communication nodes either separately or jointly. If not specified particularly, DF TWR normally refers to the scheme based on the separate decoding of the two codewords in the first phase. The separate decoding is not much different from that used in a multiple access channel (MAC) with two users. For this separate decoding-based DF TWR, there have been several conclusive studies [4], [13]. Under certain conditions of channel coefficients, it was shown that the separate decoding-based TWR can be superior to all other types of TWR [4], [14]. Nevertheless, separate decoding is basically subject to a loss of multiplexing gain [15]. The joint information decoding-based TWR was inspired by the fact that the relay node does not necessarily need to retrieve the two messages separately, but only needs to obtain the joint information of those messages that is sufficient for each node to obtain the message of the other node, having the message it sent in the first phase as side information. In the binary phase shift keying (BPSK) case, this joint information can be the XORed codeword of the two codewords, under the assumption that the two codewords are encoded by the same binary linear encoder [2].

If we confine ourselves to an additive white Gaussian noise (AWGN) SISO TWRC, it was shown that partial DF or CF TWR can asymptotically achieve the cut-set bound in terms of the achievable rate within a constant gap for the symmetric signal-to-noise ratio (SNR) TWRC [16]–[18]. In [19], it was shown that XOR decoding-based TWR is information-theoretically optimal for the symmetric SNR TWRC in a low-SNR regime. It was further shown in [19] that, for higher order pulse amplitude modulations, the cut-set bound can be achieved with a difference of 1.53 dB (shaping loss) using a $GF(q)$-low-density parity-check (LDPC) code and the modulo-q operation, where q denotes the modulation size. In [20] and [21], Nam *et al.* showed that the cut-set bound can be achieved within a bit, even for an asymmetric SNR

TWRC, by using nested lattice coding formed from a lattice chain.

For a fading SISO TWRC, a combination of separate DF and CF TWR was shown to achieve nearly the cut-set bound for a significantly asymmetric SNR TWRC [22]. Moreover, in [23], [24], a hybrid scheme that selects either a simple AF TWR or a superposition coding strategy was shown to achieve the cut-set bound within three bits. Recently, these results for the TWRC employing one-pair of users have been extended in [25]–[27] to a TWRC employing two or more pairs of users. Furthermore, a method to optimize the denoising strategy for given fading coefficients was proposed in [28], which improves the BER performance compared to the XOR decoding-based TWR.

For the MIMO TWRC, only a few of studies addressed the optimality of the TWR scheme in terms of the achievable rate. In [29] and [30], it was shown that a CF TWR scheme achieves the optimal diversity-multiplexing tradeoff (DMT). Note, however, that the optimal DMT does not necessarily guarantee the close-achievability to the cut-set bound of capacity. In [31] and [32], a MIMO power allocation problem was solved for a TWRC employing separate decoding; therefore, the scheme is subject to a loss of multiplexing gain. In [33], the authors showed that the cut-set bound can be asymptotically achieved by using AF relays if the number of relay nodes is sufficiently large. In the scheme proposed in [33], each spatial stream at each communication node is dedicated to different relay nodes disjointly, and each relay node receives and retransmits the dedicated spatial stream via maximum ratio combining and beamforming. As a consequence, each spatial stream can be delivered to the destination node with full spatial diversity obtained by the infinite number of relay nodes, which results in the total sum-rate becoming close to the cut-set bound.

To the best of the authors' knowledge, however, no conclusive scheme that achieves the cut-set bound for a three-node MIMO TWRC has been known, and the capacity region is open. Several schemes have been proposed to handle the channel matrices effectively in a three-node MIMO TWRC, however, their maximum achievable rates are far below the cut-set bound [31], [32], [34], [35].

In this paper, we propose an achievable scheme for the three-node separated MIMO TWRC, where no direct link between the two communication nodes is assumed and the relay operates in the full duplex mode. In the first phase, linear precoding based on the generalized singular value-decomposition (GSVD) is used to make the effective channel matrices upper-triangular. Therefore, in the proposed scheme, the global channel state information (CSI) is assumed. Furthermore, we decode the added codeword instead of the two individual codewords, and thus, the multiplexing gain can be preserved. To decode the added codeword, lattice codes formed from a lattice chain are considered in the first phase, that are useful because of their group property. At the relay node, the added codeword is decoded in a successive interference cancelation (SIC) manner. In the second phase, vertically encoded structural bining of the added message is considered. We show that the proposed scheme achieves the cut-set bound asymptotically as the SNRs in the first phase tend to infinity.

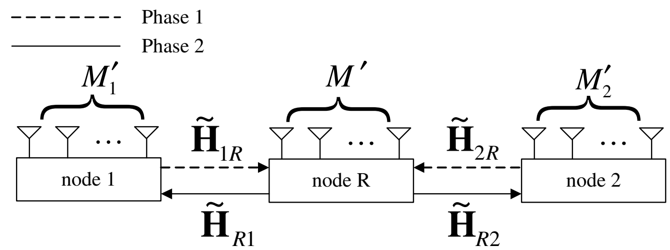

Fig. 1. $(M_1' \times M' \times M_2')$-three-node separated MIMO two-way relay channel model

The rest of the paper is organized as follows: Section II presents the separated MIMO TWRC model. Section III describes the proposed TWR scheme and characterizes the achievable rate region. Section IV concludes the paper.

Throughout the paper, we shall use the following notations: $\mathbf{A}^{\mathrm{T}}$ denotes the transpose of a matrix $\mathbf{A}$. $\Re(x)$ and $\Im(x)$ denote the real and imaginary components, respectively, of x. $\mathcal{N}(\mu, \sigma^2)$ and $\mathcal{CN}(\mu, \sigma^2)$ denote the real and complex Gaussian distributions, respectively, with mean μ and variance σ^2. $\mathbb{R}^{n\times m}$ denotes the $(n \times m)$-dimensional real space. The diagonal matrix with diagonal elements $(a_1, a_2, \ldots, a_M)$ is denoted by $\mathrm{diag}(a_1, a_2, \ldots, a_M)$. $\mathbf{I}_M$ denotes the $M \times M$ identity matrix.

II. System Model

We consider the discrete memoryless separated MIMO TWRC as depicted in Fig. 1. The two communication nodes, nodes i, $i = 1, 2$, are assumed to have M_i' antennas, respectively, and the relay node, node R, is assumed to have M' antennas. Although the proposed scheme can be applied without loss of achievable rate if $M_1', M_2' \geq M'$, we assume that $M_1' = M_2' = M'$ due to the space-limit. No direct link between node 1 and node 2 is assumed. It is further assumed that each channel is frequency-flat and quasi-static, and hence, the fading coefficients remain unchanged for the duration of a codeword that is denoted by T. Subsequently, the $(M' \times M')$-dimensional channel matrices can be represented with the subscript denoting the transmit and receive nodes; specifically, the channel matrix from node i to node R is denoted as $\tilde{\mathbf{H}}_{i\mathrm{R}}$, and the channel matrix from node R to node i is denoted by $\tilde{\mathbf{H}}_{\mathrm{R}i}$, $i \in \{1, 2\}$. The $(M' \times T)$-dimensional transmit signal matrix at node i is denoted by $\tilde{\mathbf{S}}_i$, and the maximum average power at node i is denoted by P_i such that $\frac{1}{T}E\{\|\tilde{\mathbf{S}}_i\|_{\mathrm{F}}^2\} \leq P_i$, $i = 1, 2$. The signal received at node R in Phase 1 can be written as

$$\tilde{\mathbf{Y}}_{\mathrm{R}} = \tilde{\mathbf{H}}_{1\mathrm{R}}\tilde{\mathbf{S}}_1 + \tilde{\mathbf{H}}_{2\mathrm{R}}\tilde{\mathbf{S}}_2 + \tilde{\mathbf{Z}}_{\mathrm{R}}, \tag{1}$$

where $\tilde{\mathbf{Z}}_{\mathrm{R}}$ denotes the AWGN at node R, each element of which is i.i.d. according to $\mathcal{CN}(0, 2)$.

For later convenience of notation and use of the real-valued lattice code, we only consider the real-valued signals by introducing the equivalent expression

$$\mathbf{Y}_{\mathrm{R}} = \mathbf{H}_{1\mathrm{R}}\mathbf{S}_1 + \mathbf{H}_{2\mathrm{R}}\mathbf{S}_2 + \mathbf{Z}_{\mathrm{R}}, \tag{2}$$

where, for $i = 1, 2$, $\mathbf{Y}_{\mathrm{R}} = \left[\Re(\tilde{\mathbf{Y}}_{\mathrm{R}})^{\mathrm{T}}, \Im(\tilde{\mathbf{Y}}_{\mathrm{R}})^{\mathrm{T}}\right]^{\mathrm{T}}$, $\mathbf{S}_i = \left[\Re(\tilde{\mathbf{S}}_i)^{\mathrm{T}}, \Im(\tilde{\mathbf{S}}_i)^{\mathrm{T}}\right]^{\mathrm{T}}$, $\mathbf{Z}_{\mathrm{R}} = \left[\Re(\tilde{\mathbf{Z}}_{\mathrm{R}})^{\mathrm{T}}, \Im(\tilde{\mathbf{Z}}_{\mathrm{R}})^{\mathrm{T}}\right]^{\mathrm{T}}$,

$$\mathbf{H}_{i\mathrm{R}} = \begin{bmatrix} \Re(\tilde{\mathbf{H}}_{i\mathrm{R}}) & -\Im(\tilde{\mathbf{H}}_{i\mathrm{R}}) \\ \Im(\tilde{\mathbf{H}}_{i\mathrm{R}}) & \Re(\tilde{\mathbf{H}}_{i\mathrm{R}}) \end{bmatrix}, \tag{3}$$

and

$$\mathbf{H}_{\mathrm{R}i} = \begin{bmatrix} \Re(\tilde{\mathbf{H}}_{\mathrm{R}i}) & -\Im(\tilde{\mathbf{H}}_{\mathrm{R}i}) \\ \Im(\tilde{\mathbf{H}}_{\mathrm{R}i}) & \Re(\tilde{\mathbf{H}}_{\mathrm{R}i}) \end{bmatrix}. \tag{4}$$

The message at node i is denoted by W_i, and the corresponding message set is denoted by $\mathcal{W}_i$ such that $W_i \in \mathcal{W}_i = \{1, 2, \ldots, 2^{R_i}\}$, $i = 1, 2$. Let us define M such that $M \triangleq 2M'$. The $(M \times T)$-dimensional encoded symbol matrix at node i, denoted by $\mathbf{X}_i \in \mathcal{X}_i$, is a function of W_i, $i = 1, 2$. Moreover, $\mathbf{S}_i \in \mathcal{S}_i$ is a function of $\mathbf{X}_i$, and we thus have

$$f_i : \mathcal{W}_i \mapsto \mathcal{X}_i, \;\; g_i : \mathcal{X}_i \mapsto \mathcal{S}_i, \;\; i = 1, 2, \tag{5}$$

where f_i is the encoding function and g_i is the MIMO preprocessing function. In Phase 2, with the received signal $\mathbf{Y}_{\mathrm{R}} \in \mathcal{Y}_{\mathrm{R}}$, node R broadcasts the $(M \times T)$-dimensional signal matrix $\mathbf{X}_{\mathrm{R}} \in \mathcal{X}_{\mathrm{R}}$ that is obtained by the encoding function f_{R} such that

$$f_{\mathrm{R}} : \mathcal{Y}_{\mathrm{R}} \mapsto \mathcal{X}_{\mathrm{R}}. \tag{6}$$

The maximum average power at node R is denoted by P_{R} such that $\frac{1}{T}E\{\|\mathbf{X}_{\mathrm{R}}\|_{\mathrm{F}}^2\} \leq P_{\mathrm{R}}$. Subsequently, the received signal at node i in Phase 2 can be represented as

$$\mathbf{Y}_i = \mathbf{H}_{\mathrm{R}i}\mathbf{X}_{\mathrm{R}} + \mathbf{Z}_i, \;\; i = 1, 2, \tag{7}$$

where $\mathbf{Z}_i$ denotes the AWGN at node i, each element of which is i.i.d according to $\mathcal{N}(0, 1)$. With the received signal $\mathbf{Y}_i \in \mathcal{Y}_i$ and the side information W_i, node i estimates the message W_{3-i} using the decoding function d_i such that

$$d_i : \mathcal{Y}_i \times \mathcal{W}_i \mapsto \mathcal{W}_{3-i}, \quad i = 1, 2. \tag{8}$$

The average probability of error is defined as

$$P_e \triangleq \Pr\{d_1(\mathbf{Y}_1, W_1) \neq W_2 \text{ or } d_2(\mathbf{Y}_2, W_2) \neq W_1\}. \tag{9}$$

A rate pair (R_1, R_2) is said to be achievable if there exist encoding and decoding functions as well as MIMO preprocessing functions such that the error probability vanishes as $T \to \infty$.

III. Asymptotic Capacity of the Separated MIMO TWRC

In this section, we provide achievable rate regions for the separated MIMO TWRC. In the proposed achievable scheme, lattice coding based on a lattice chain is used for Phase 1. Moreover, a GSVD-based precoding method is proposed to make the effective channel matrices in Phase 1 upper-triangular, which enables us to decode the added codeword in an SIC manner. Thus, the horizontal encoding and decoding is considered for Phase 1. In Phase 2, we use the vertically encoded structural bining, based on the one-to-one mapping from the estimated added codewords in Phase 1 to the Gaussian random codewords.

Our main result can be characterized by the following Theorem:

Theorem 1: For the separated MIMO TWRC considered in Section II, the set of all rate pairs (R_1, R_2) is achievable, if

$$\begin{aligned} R_1 \;\; &\leq \;\; \min\left\{\left[\frac{1}{2}\log\det\left(\frac{P_1}{M}\mathbf{H}_{1\mathrm{R}}\mathbf{H}_{1\mathrm{R}}^{\mathrm{T}}\right)\right]^+,\right. \\ &\qquad\qquad \left.\frac{1}{2}\log\det\left(\mathbf{I} + \frac{P_{\mathrm{R}}}{M}\mathbf{H}_{\mathrm{R}2}\mathbf{H}_{\mathrm{R}2}\right)\right\}, \end{aligned} \tag{10}$$

$$\begin{aligned} R_2 \;\; &\leq \;\; \min\left\{\left[\frac{1}{2}\log\det\left(\frac{P_2}{M}\mathbf{H}_{2\mathrm{R}}\mathbf{H}_{2\mathrm{R}}^{\mathrm{T}}\right)\right]^+,\right. \\ &\qquad\qquad \left.\frac{1}{2}\log\det\left(\mathbf{I} + \frac{P_{\mathrm{R}}}{M}\mathbf{H}_{\mathrm{R}1}\mathbf{H}_{\mathrm{R}1}\right)\right\}, \end{aligned} \tag{11}$$

where $[a]^+ \triangleq \max\{a, 0\}$ for any a. As a corollary, this region asymptotically achieves the outer-bound as the SNRs of the channels from node i, $i = 1, 2$, to node R in Phase 1 tend to infinity.

In the rest of this section, Theorem 1 is proved.

A. Phase 1

In Phase 1, horizontal lattice coding is used. The message set $\mathcal{W}_i$ is divided into M independent sets such that

$$\mathcal{W}_i = \mathcal{W}_{i,1} \times \mathcal{W}_{i,2} \times \cdots \times \mathcal{W}_{i,M}, \;\; i = 1, 2, \qquad (12)$$

where $\mathcal{W}_{i,m} = \{1, 2, \ldots, 2^{TR_{i,m}}\}$ denotes the message set for the mth spatial stream at node i; thus, the message for the mth spatial stream at node i is denoted as $w_{i,m} \in \mathcal{W}_{i,m}$. Because $|\mathcal{W}_i| = \prod_{m=1}^{M} |\mathcal{W}_{i,m}|$, the rate R_i can be written as

$$R_i = \sum_{m=1}^{M} R_{i,m}, \quad i = 1, 2. \qquad (13)$$

The $(1 \times T)$-dimensional encoded codeword for the mth spatial stream at node i is denoted as $\mathbf{x}_{i,m}$, $i = 1, 2$, $m = 1, \ldots, M$. Subsequently, the $(M \times T)$-dimensional encoded symbol matrix $\mathbf{X}_i$ is formed as

$$\mathbf{X}_i = \left[\mathbf{x}_{i,1}{}^{\mathrm{T}}, \mathbf{x}_{i,2}{}^{\mathrm{T}}, \ldots, \mathbf{x}_{i,M}{}^{\mathrm{T}}\right]^{\mathrm{T}}, \quad i = 1, 2. \qquad (14)$$

At this point, we consider the linear precoding strategy based on the GSVD. The GSVD for $\mathbf{H}_{\mathrm{R}1}$ and $\mathbf{H}_{\mathrm{R}2}$ can be represented as [36]

$$\mathbf{H}_{1\mathrm{R}} = \mathbf{B}\boldsymbol{\Sigma}_1\mathbf{V}_1^{\mathrm{T}}, \quad \mathbf{H}_{2\mathrm{R}} = \mathbf{B}\boldsymbol{\Sigma}_2\mathbf{V}_2^{\mathrm{T}}, \qquad (15)$$

where $\mathbf{B} \in \mathbb{R}^{M\times M}$ is a square matrix, and $\mathbf{V}_i \in \mathbb{R}^{M\times M}$ is an orthogonal matrix, $i \in \{1, 2\}$. Furthermore, $\boldsymbol{\Sigma}_i = \mathrm{diag}(\sigma_{i,1}, \sigma_{i,2}, \cdots, \sigma_{i,M})$, where $\sigma_{i,m}$ denotes the mth generalized singular value of $\mathbf{H}_{i\mathrm{R}}$, $i = 1, 2$, $m = 1, 2, \ldots, M$. We further take the QR-decomposition [36] to the matrix $\mathbf{B}$, to get

$$\mathbf{H}_{1\mathrm{R}} = \mathbf{Q}\mathbf{R}\boldsymbol{\Sigma}_1\mathbf{V}_1^{\mathrm{T}}, \quad \mathbf{H}_{2\mathrm{R}} = \mathbf{Q}\mathbf{R}\boldsymbol{\Sigma}_2\mathbf{V}_2^{\mathrm{T}}, \qquad (16)$$

where $\mathbf{Q} \in \mathbb{R}^{M\times M}$ is an orthogonal matrix, and $\mathbf{R} \in \mathbb{R}^{M\times M}$ is an upper-triangular matrix such that

$$\mathbf{R} = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,M} \\ 0 & r_{2,2} & \cdots & r_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{M,M} \end{bmatrix}. \qquad (17)$$

Subsequently, the transmit signal matrces are obtained from

$$\mathbf{S}_i = \mathbf{V}_i\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i, \;\; i = 1, 2. \qquad (18)$$

Moreover, to satisfy the power constraint at node i, it should be satisfied that

$$\frac{1}{T}E\left\{\|\mathbf{S}_i\|_{\mathrm{F}}^2\right\} = \frac{1}{T}E\left\{\|\boldsymbol{\Sigma}_i{}^{-1}\mathbf{X}_i\|_{\mathrm{F}}^2\right\} = \frac{1}{T}\sum_{m=1}^{M}(\sigma_{i,m})^{-2}E\left\{\|\mathbf{x}_{i,m}\|^2\right\} \le P_i. \qquad (19)$$

We thus choose the average power for the mth codewords at node i, $\gamma_{1,m}$, such that

$$\gamma_{i,m} = \frac{1}{T}E\{\|\mathbf{x}_{i,m}\|^2\} = \frac{P_i\sigma_{i,m}{}^2}{M}, \qquad (20)$$

$i = 1, 2$, $m = 1, \ldots, M$. Let us assume, without loss of generality, that $\gamma_{1,m} \ge \gamma_{2,m}$. In this case, by [21, Theorem 2], there exists a T-dimensional lattice chain composed of $\Lambda_{i,m}$, $i = 1, 2$, and $\Lambda_{C,m}$ such that $\Lambda_{1,m} \subseteq \Lambda_{2,m} \subseteq \Lambda_{C,m}$ and $V(\Lambda_{i,m})$ is arbitrarily close to $\gamma_{i,m}$, where $\Lambda_{1,m}$ and $\Lambda_{2,m}$ are simultaneously Rogers-good and Poltyrev-good while $\Lambda_{C,m}$ is Poltyrev-good (See [37] for lattice goodness). Here, $V(\cdot)$ denotes the second-order moment of a lattice per dimension [37]; thus, the power constraint given in (20) can be satisfied. Subsequently, we can consider a lattice code that is defined by the set of coset leaders as $\mathcal{C}_{i,m} = \{\Lambda_{C,m} \cup \mathcal{R}_{i,m}\}$, $i = 1, 2$, where $\mathcal{R}_{i,m}$ is the Voronoi region of $\Lambda_{i,m}$, and $|\mathcal{C}_{i,m}| = 2^{TR_{i,m}}$.

The message $w_{i,m} \in \{1, \ldots, 2^{TR_{i,m}}\}$ is one-to-one mapped to $\mathbf{x}_{i,m} \in \mathcal{C}_{i,m}$, and hence, the codeword $\mathbf{x}_{i,m}$ itself represents the message. Consequently, the received signal at node R in Phase 1 can be written as

$$\begin{aligned}\mathbf{Y}_{\mathrm{R}} &= \mathbf{H}_{1\mathrm{R}}\mathbf{S}_1 + \mathbf{H}_{2\mathrm{R}}\mathbf{S}_2 + \mathbf{Z}_{\mathrm{R}} \\ &= \mathbf{Q}\mathbf{R}\left(\mathbf{X}_1 + \mathbf{X}_2\right) + \mathbf{Z}_{\mathrm{R}}. \end{aligned} \qquad (21)$$

Further, without loss of optimality, we can equalize the orthogonal matrix $\mathbf{Q}$ and the diagonal elements of $\mathbf{R}$ as

$$\begin{aligned}\bar{\mathbf{Y}}_{\mathrm{R}} &= \left[\bar{\mathbf{y}}_{\mathrm{R},1}^{\mathrm{T}}, \ldots, \bar{\mathbf{y}}_{\mathrm{R},M}^{\mathrm{T}}\right]^{\mathrm{T}} = \mathbf{R}_d^{-1}\mathbf{Q}^{\mathrm{T}}\mathbf{Y}_{\mathrm{R}} \\ &= \mathbf{X}_1 + \mathbf{X}_2 + \left(\mathbf{R}_d^{-1}\mathbf{R} - \mathbf{I}\right)\left(\mathbf{X}_1 + \mathbf{X}_2\right) + \bar{\mathbf{Z}}_{\mathrm{R}}, \end{aligned} \qquad (22)$$

where $\mathbf{R}_d \triangleq \mathrm{diag}(r_{1,1}, r_{2,2}, \ldots, r_{M,M})$ and $\bar{\mathbf{Z}}_{\mathrm{R}} = \left[\bar{\mathbf{z}}_{\mathrm{R},1}^{\mathrm{T}}, \ldots, \bar{\mathbf{z}}_{\mathrm{R},M}^{\mathrm{T}}\right]^{\mathrm{T}} \triangleq \mathbf{R}_d^{-1}\mathbf{Q}^{\mathrm{T}}\mathbf{Z}_{\mathrm{R}}$. Because, $\left(\mathbf{R}_d^{-1}\mathbf{R} - \mathbf{I}\right)$ is upper-triangular with zero diagonal elements, the self-interference term $\left(\mathbf{R}_d^{-1}\mathbf{R} - \mathbf{I}\right)\left(\mathbf{X}_1 + \mathbf{X}_2\right)$ can be successively canceled from the Mth spatial stream to the first spatial stream.

1) Mth spatial stream: Because the Mth spatial is free from the self-interference induced by $\left(\mathbf{R}_d^{-1}\mathbf{R} - \mathbf{I}\right)(\mathbf{X}_1 + \mathbf{X}_2)$, the received signal for the Mth spatial stream can be written as

$$\bar{\mathbf{y}}_{\mathrm{R},M} = (\mathbf{x}_{1,M} + \mathbf{x}_{2,M}) + \bar{\mathbf{z}}_{\mathrm{R},M}. \qquad (23)$$

Note that $\mathbf{x}_{1,m}, \mathbf{x}_{2,m} \in \Lambda_{C,m}$; thus, $\mathbf{x}_{1,m} + \mathbf{x}_{2,m} \in \Lambda_{C,m}$ by the group property. Therefore, node R retrieves the added codeword $\mathbf{x}_{1,m} + \mathbf{x}_{2,m}$ instead of individual $\mathbf{x}_{1,m}$ and $\mathbf{x}_{2,m}$. We apply the ambiguity decoder [38] (See also [39]) that finds the lattice point in the region $\mathcal{E} \subset \mathbb{R}^T$ if and only if there exists only one lattice point in that region. Specifically, denoting the estimate of $\mathbf{x}_{1,m} + \mathbf{x}_{2,m}$

by $\hat{\mathbf{x}}_{\text{added},m}$, the decoder outputs $\hat{\mathbf{x}}_{\text{added},M} \in \Lambda_{C,M}$ if $\bar{\mathbf{y}}_{\text{R},M} \in \hat{\mathbf{x}}_{\text{added,M}} + \mathcal{E}$ and there exist no other lattice point $\mathbf{x}' \in \Lambda_{C,M}$ in that region. If $\bar{\mathbf{y}}_{\text{R},M} \in \{\mathcal{E}+\mathbf{x}\} \cap \{\mathcal{E}+\mathbf{x}'\}$ for some lattice points $\mathbf{x}, \mathbf{x}' \in \Lambda_{C,M}$ such that $\mathbf{x} \neq \mathbf{x}'$, we say the ambiguity event $\mathcal{A}$ occurs. Therefore, for a given $\mathcal{E}$ and $\Lambda_{C,M}$, the probability of error is given by

$$p_e^{(M)}(\mathcal{E}|\Lambda_{C,M}) = \Pr\{\hat{\mathbf{x}}_{\text{added},m} \neq (\mathbf{x}_{1,m} + \mathbf{x}_{2,m})\}$$
$$= \Pr\{(\bar{\mathbf{z}}_{\text{R},M} \notin \mathcal{E}) \text{ or } \mathcal{A}\}. \quad (24)$$

For a sufficiently large T and $\epsilon, \delta, \delta' > 0$, the average probability error $\bar{p}_e^{(M)}$ is upper-bounded as

$$\bar{p}_e^{(M)} \leq \epsilon/2 + 2^{-\frac{T}{2}\log_2\left(\gamma_{i,M} \cdot r_{M,M}{}^2\right) + TR_{i,M}} \times 2^{\log_2\left((1+\delta)(1+\delta')^{\text{T}}\right)} \quad (25)$$

where $\epsilon, \delta, \delta' \to 0$ as $T \to \infty$. The derivation of (25) is presented in Appendix A. Therefore, $\bar{p}_e^{(M)}$ tends to zero as $T \to \infty$ if $R_{i,M} \leq R^*_{i,M}$, where

$$R^*_{i,M} \triangleq \frac{1}{2}\log\left(\gamma_{i,M} \cdot r_{M,M}{}^2\right) = \frac{1}{2}\log\left(\frac{P_i}{M}\sigma_{i,M}{}^2 \cdot r_{M,M}{}^2\right). \quad (26)$$

2) mth spatial stream ($m \leq M-1$): Let us assume that $\mathbf{x}_{1,l} + \mathbf{x}_{2,l}$, $l = m+1, \ldots, M$, have been retrieved without error; namely, $\hat{\mathbf{x}}_{\text{added},l} = \mathbf{x}_{1,l} + \mathbf{x}_{2,l}$. Consequently, the retrieved codewords can be subtracted from the mth spatial signal, $\bar{\mathbf{y}}_{\text{R},m}$, as

$$\bar{\mathbf{y}}_{\text{R},m} - \sum_{l=m+1}^{M}\left(\frac{r_{m,l}}{r_{m,m}}\hat{\mathbf{x}}_{\text{added},l}\right) = \mathbf{x}_{1,m} + \mathbf{x}_{2,m} + \bar{\mathbf{z}}_{\text{R},m}, \quad (27)$$

where $\bar{\mathbf{z}}_{\text{R},m} \sim \mathcal{N}(\mathbf{0}, (r_{m,m})^{-2} \cdot \mathbf{I}_T)$. Using the interference-suppressed signal (27), it can be shown, by the analogous proof for the Mth spatial stream, that the probability of error for the mth spatial stream at node i, $\bar{p}_e^{(m)}$, vanishes for any $R_{i,m} \leq R^*_{i,m}$ as $T \to \infty$, where

$$R^*_{i,m} \triangleq \frac{1}{2}\log\left(\frac{P_i}{M}\sigma_{i,m}{}^2 r_{m,m}{}^2\right). \quad (28)$$

Subsequently, let us define R_i^*, $i = 1, 2$, as

$$\begin{aligned}
R_i^* &\triangleq \sum_{m=1}^{M} R^*_{i,m} = \frac{1}{2}\sum_{m=1}^{M}\log\left(\frac{P_i}{M}\sigma_{i,m}{}^2 r_{m,m}{}^2\right) \\
&= \frac{1}{2}\log\left(\prod_{m=1}^{M}\frac{P_i}{M}\sigma_{i,m}{}^2 r_{m,m}{}^2\right) \\
&= \frac{1}{2}\log\det\left(\frac{P_1}{M}\mathbf{R}\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_i\mathbf{R}^{\text{T}}\right) && (29) \\
&= \frac{1}{2}\log\det\left(\frac{P_1}{M}\mathbf{Q}\mathbf{R}\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_i\mathbf{R}^{\text{T}}\mathbf{Q}^{\text{T}}\right) && (30) \\
&= \frac{1}{2}\log\det\left(\frac{P_1}{M}\mathbf{H}_{i\text{R}}\mathbf{H}_{i\text{R}}{}^{\text{T}}\right), && (31)
\end{aligned}$$

where (29) follows from the fact that the determinant of a triangular matrix is identical to the product of its diagonal elements, and (30) follows because the determinant of a matrix is unchanged by the right- or left-multiplication of orthogonal matrices. For any $R_i \leq R_i^*$, let us denote R_i as

$$R_i = \mu R_i^*, \ \ 0 \leq \mu \leq 1, \quad (32)$$

and choose the rate of each spatial stream such that $R_{i,m} = \mu R^*_{i,m} \leq R^*_{i,m}$, $m = 1, \ldots, M$. As a consequence, probability of error for each spatial stream vanishes as $T \to \infty$, and therefore the probability of error in Phase 1 tends to zero by the union bound argument of probability.

B. Phase 2

In fact, the rate region for Phase 2 and its derivation are not much different from those in [40] (See Theorem 1 therein), except that the added codeword, instead of the two individual codewords, is decoded and broadcasted in the proposed scheme. Therefore, we briefly derive the rate region for Phase 2 to avoid duplication. Denoting the estimated added codeword in Phase 1 as $\hat{\mathbf{X}}_{\text{added}} = \left[\hat{\mathbf{x}}^{\text{T}}_{\text{added},1}, \ldots, \hat{\mathbf{x}}^{\text{T}}_{\text{added},M}\right]^{\text{T}}$, node R maps $\hat{\mathbf{X}}_{\text{added}}$ to Gaussian codewords. Let us denote the set of the added codewords by $\mathcal{C}_{\text{added}}$ such that

$$\mathcal{C}_{\text{added}} = \left\{\mathbf{X}_1 + \mathbf{X}_2 = \left[(\mathbf{x}_{1,1} + \mathbf{x}_{2,1})^{\text{T}}, \ldots, (\mathbf{x}_{1,M} + \mathbf{x}_{2,M})^{\text{T}}\right]^{\text{T}} : \mathbf{x}_{i,m} \in \mathcal{C}_{i,m}, \forall i, m\right\}, \quad (33)$$

and denote the corresponding rate by $R_{\text{added}} \triangleq \frac{1}{T}\log_2|\mathcal{C}_{\text{added}}|$. Because $|\mathcal{C}_{\text{added}}| \leq \prod_{m=1}^{M}|\mathcal{C}_{1,m}| \cdot |\mathcal{C}_{2,m}|$, it can be readily shown that

$$R_{\text{added}} \leq \sum_{m=1}^{M}(R_{1,m} + R_{2,m}) = R_1 + R_2. \quad (34)$$

First, we generate $2^{MT\cdot R_{\text{added}}}$ independent codewords of length MT, each element of which is chosen independently and identically according to $\mathcal{N}(0, P_{\text{R}}/M)$, to form a codebook $\mathcal{C}_{\text{R}}$, where P_{R} is the power constant at node R. Here, for simplicity, we assume that the transmit power at each antenna of node R is the same. The estimated $\hat{\mathbf{X}}_{\text{added}}$ is then one-to-one mapped to one of the sequences in $\mathcal{C}_{\text{R}}$, and the mapped sequence is vertically reshaped to $(M \times T)$-dimensional matrix $\mathbf{X}_{\text{R}}$. The received signal at node i can be written as (7). Using the side information $\mathbf{X}_i$, node i obtains $\hat{\mathbf{X}}_{3-i}$, the estimate of $\mathbf{X}_{3-i}$, and thereby estimates the message by demapping $\hat{\mathbf{X}}_{3-i}$ to one of the message set $\mathcal{W}_{3-i}$ if there exists a unique codeword in $\mathcal{C}_{\text{R},3-i}$ such that $(\hat{\mathbf{X}}_{3-i}\mathbf{Y}_i)$ are jointly typical, where

$$\mathcal{C}_{\text{R},1} \triangleq \Big\{\mathbf{X}_{\text{R}}(\mathbf{T}_1) : \mathbf{T}_1 = \left[\mathbf{t}_{1,1}{}^{\text{T}}, \ldots, \mathbf{t}_{1,M}{}^{\text{T}}\right]^{\text{T}} + \mathbf{X}_2, \quad \mathbf{t}_{1,m} \in \mathcal{C}_{1,m}, m = 1, \ldots, M\Big\}, \tag{35}$$

$$\mathcal{C}_{\text{R},2} \triangleq \Big\{\mathbf{X}_{\text{R}}(\mathbf{T}_2) : \mathbf{T}_2 = \mathbf{X}_1 + \left[\mathbf{t}_{2,1}{}^{\text{T}}, \ldots, \mathbf{t}_{2,M}{}^{\text{T}}\right]^{\text{T}}, \quad \mathbf{t}_{2,m} \in \mathcal{C}_{2,m}, m = 1, \ldots, M\Big\}. \tag{36}$$

Because $|\mathcal{C}_{\text{R},i}| = 2^{TR_i}$, from the jointly typical decoding argument [41], the probability error at node i in Phase 2 vanishes as $T \to \infty$ if

$$R_{3-i} \leq \frac{1}{2}\log\det\left(\mathbf{I} + \frac{P_R}{M}\mathbf{H}_{\text{R}i}\mathbf{H}_{\text{R}i}\right), \quad i = 1, 2. \tag{37}$$

Therefore, from the results in (31) and (37), the probability of error vanishes if a rate pair (R_1, R_2) satisfies (10) and (11) by the union bound argument of probability.

C. Numerical Example

In Fig. 2, the average achievable rate regions of several schemes and the outer cut-set bound are depicted for the separated MIMO TWRC with $M' = 2$, where the SNRs of the all channels are assumed to be the same as 20dB.

The outer cut-set bound of the separated MIMO TWRC can be readily obtained by the cut-set theorem [41] as, for $i = 1, 2$,

$$\begin{aligned} R_i &\leq \min\left\{\frac{1}{T}I(\mathbf{X}_i; \mathbf{Y}_{\text{R}}, \mathbf{Y}_{3-i}|\mathbf{X}_{\text{R}}, \mathbf{X}_{3-i}), \frac{1}{T}I(\mathbf{X}_i, \mathbf{X}_{\text{R}}; \mathbf{Y}_{3-i}|\mathbf{X}_{3-i})\right\} \\ &= \min\left\{\frac{1}{2}\log\det\left(\mathbf{I} + \mathbf{H}_{i\text{R}}\mathbf{H}_{i\text{R}}{}^{\text{T}}\right), \frac{1}{2}\log\det\left(\mathbf{I} + \mathbf{H}_{\text{R}(3-i)}\mathbf{H}_{\text{R}(3-i)}{}^{\text{T}}\right)\right\}. \end{aligned} \tag{38}$$

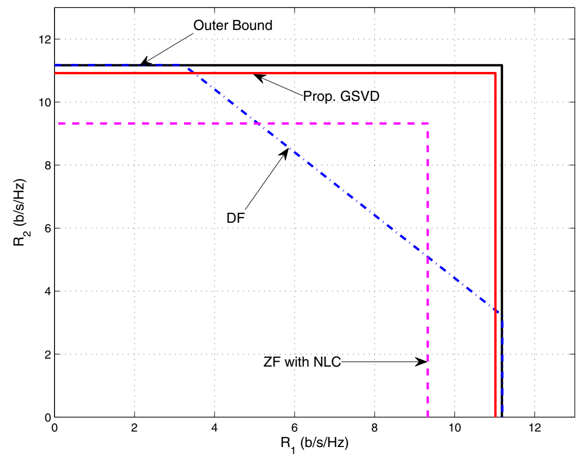

Fig. 2. Average achievable rate regions and outer bound for the separated MIMO TWRC with $M' = 2$. The average SNRs for all channels are assumed to be the same as 20 dB.

In addition, as a baseline scheme, the straightforward zero-forcing (ZF) precoding-based scheme is considered, in which the transmit signal matrix in Phase 1 is designed such that $\mathbf{S}_i = \alpha_i \mathbf{H}_{i\text{R}}^{-1}\mathbf{X}_i$, where the power constant α_i is given by

$$\alpha_i = \sqrt{\frac{M}{\text{tr}\left\{\mathbf{H}_{i\text{R}}^{-1}\left(\mathbf{H}_{i\text{R}}{}^{-1}\right)^{\text{T}}\right\}}} \quad i = 1, 2, \tag{39}$$

where $\text{tr}\{\cdot\}$ is the trace operation. This ZF-based scheme is equivalent to the M independent AWGN SISO TWRCs with asymmetric SNRs as considered in [21]. Thus, the achievable rate region for the ZF precoding-based scheme with nested lattice code (NLC) can be written as (See [21, Theorem 1]), for $i = 1, 2$,

$$R_i \leq \min\left\{\left[\frac{M}{2}\log\left(\frac{\alpha_i{}^2}{\alpha_1{}^2 + \alpha_2{}^2} + \alpha_i{}^2\right)\right]^+, \frac{1}{2}\log\det\left(\mathbf{I} + \mathbf{H}_{\text{R}(3-i)}\mathbf{H}_{\text{R}(3-i)}{}^{\text{T}}\right)\right\}. \tag{40}$$

Due to some instances in which the channel matrices are ill-conditioned, the achievable rate of the ZF-based scheme, denoted by 'ZF with NLC', is far below the cut-set bound, as depicted in Fig. 2. It can be seen that the proposed GSVD precoding-based scheme achieves rate pairs that cannot be achieved by the DF TWR scheme or the ZF precoding-based scheme, and the difference between the proposed GSVD-based scheme and the outer cut-set bound is smaller than a half bit per spatial stream, which is negligible in a high-SNR regime.

IV. Conclusion and Remarks

We have derived an achievable rate region for the quasi-static separated MIMO TWRC, where each node is equipped with the same number of antennas. Using generalized singular value decomposition-based precoding together with the successive interference cancelation decoding in Phase 1, we have shown that the achievable rate region of the proposed scheme asymptotically achieves the cut-set bound within a half bit per real spatial stream that is negligible in a high SNR regime. In our future work, we plan to generalize the antenna configuration and consider an optimal power allocation method at each node.

Appendix A
Derivation of (25)

The probability of error in (24) is upper-bounded by the union bound as

$$p_e^{(M)}(\mathcal{E}|\Lambda_{1,M}) = \Pr\{(\bar{\mathbf{z}}_{\mathrm{R},M} \notin \mathcal{E}) \text{ or } \mathcal{A}\} \leq \Pr\{\bar{\mathbf{z}}_{\mathrm{R},M} \notin \mathcal{E}\} + \Pr\{\mathcal{A}\}. \quad (41)$$

From [38, Theorem 4], for arbitrary $\delta > 0$, we get

$$\bar{p}_e \triangleq E\{p_e(\mathcal{E}|\Lambda_{C,M})\} \leq \Pr(\bar{\mathbf{z}}_{\mathrm{R},M} \notin \mathcal{E}) + (1+\delta)\frac{V(\mathcal{E})}{V(\Lambda_C)}. \quad (42)$$

Noting that each element of $\bar{\mathbf{z}}_{\mathrm{R},M}$ is an i.i.d. white Gaussian random variable with zero mean and variance $1/{r_{M,M}}^2$, we consider the decision region $\mathcal{E}$, for some $\delta' > 0$, defined by

$$\mathcal{E} \triangleq \left\{\mathbf{x} \in \mathbb{R}^T : \|\mathbf{x}\|^2 \leq \frac{T(1+\delta')}{{r_{M,M}}^2}\right\} \quad (43)$$

From the standard typicality arguments, it can be shown that, for any $\delta' > 0$ and $\epsilon > 0$, there exist T^* such that for all $T > T^*$

$$\Pr(\bar{\mathbf{z}}_{\mathrm{R},M} \notin \mathcal{E}) < \epsilon/2. \quad (44)$$

Therefore, for sufficiently large $\mathbf{T}$, the probability of error can be written as

$$\bar{p}_e^{(M)} \leq \epsilon/2 + (1+\delta)\frac{V(\mathcal{E})}{V(\Lambda_C)}. \quad (45)$$

Because Λ_C and $\Lambda_{i,M}$ are designed such that $V(\Lambda_{i,M}) = |\mathcal{C}_{i,M}| \cdot V(\Lambda_C)$ and $|\mathcal{C}_{1,M}| = 2^{TR_{i,M}}$, for $i = 1,2$, we have

$$\frac{V(\mathcal{E})}{V(\Lambda_C)} = \frac{V(\mathcal{E})2^{TR_{i,M}}}{V(\Lambda_{i,M})} = 2^{TR_{i,M}} \cdot (1+\delta')^{\mathrm{T}} \times(\gamma_{i,M} \cdot {r_{M,M}}^2)^{-T/2}, \quad (46)$$

and thus, we have

$$\begin{aligned}\bar{p}_e^{(M)} &\leq \epsilon/2 + (1+\delta)(1+\delta') \times \left(\gamma_{i,M} \cdot {r_{M,M}}^2\right)^{-T/2} \cdot 2^{TR_{i,M}} \\ &= \epsilon/2 + 2^{\log_2\left((1+\delta)(1+\delta')\left(\gamma_{i,M} \cdot {r_{M,M}}^2\right)^{-T/2}\right)} \times 2^{TR_{i,M}} \\ &= \epsilon/2 + 2^{-\frac{T}{2}\log_2\left(\gamma_{i,M} \cdot {r_{M,M}}^2\right) + TR_{i,M}} \times 2^{\log_2\left((1+\delta)(1+\delta')^{\mathrm{T}}\right)}. \end{aligned} \quad (47)$$

References

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. Prob.*, vol. 1, Berkely, CA, 1961, pp. 611–644.

[2] S. Zhang, S.-C. Liew, and P. P. Lam, "Hot topic: Physical-layer network coding," in *Proc. ACM MobiCom 2006*, Los Angeles, CA, Sept. 2006.

[3] S. Katti, R. Hariharan, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," in *Proc. ACM SIGCOMM*, Pista, Italy, Sept. 2006.

[4] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[5] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," in *Proc. IEEE Int'l Conf. Commun. (ICC)*, Glasgow, Scotland, June 2007.

[6] R. Zhang, C. C. Chai, Y.-C. Liang, and S. Cui, "On capacity region of two-way multi-antenna relay channel with analogue network coding," in *Proc. IEEE Int'l Conf. Commun. (ICC)*, Dresden, Germany, June 2009.

[7] R. Vaze and R. W. Heath Jr., "Optimal amplify and forward strategy for two-way relay channel with multiple relays," in *IEEE Inform. Theory Workshop (ITW)*, Volos, Greece, June 2009.

[8] Y. Zhao, R. Adve, and T. J. Lim, "Improving amplify-and-forward relay networks : optimal power allocation versus selection," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3114–3123, Aug. 2007.

[9] 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.

[10] N. Lee, H. J. Yang, and J. Chun, "Achievable sum-rate maximizing AF relay beamforming scheme in two-way relay channels," in *IEEE Int'l Conf. Commun. (ICC) Workshop*, Beijing, China, May 2008.

[11] T. Unger and A. Klein, "On the performance of two-way relaying with multiple-antenna relay stations," in *IEEE Mobile and Wireless*, Budapest, Hungary, July 2007.

[12] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with anlaogue network coding," *IEEE J. Select. Areas Commun.*, vol. 27, no. 5, pp. 699–712, June 2009.

[13] S. Kim, P. Mitran, and V. Tarokh, "Performance bounds for bidirectional coded cooperation protocols," *IEEE Trans. Inform. Theory*, vol. 54, no. 11, pp. 5235–5341, Nov. 2008.

[14] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, "Broadcast capacity region of two-phase bidirectional relaying," *IEEE Trans. Inform. Theory*, vol. 54, no. 1, pp. 454–458, Jan. 2008.

[15] R. Knopp, "Two-way wireless communication via a relay station," in *GDR-ISIS meeting*, Paris, France, Mar. 2007.

[16] D. Gündüz, E. Tuncel, and J. Nayak, "Rate regions for the separated two-way relay channel," in *Proc. 46th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2008.

[17] K. Narayanan, M. P. Wilson, and A. Sprintson, “Joint physical layer coding and network coding for bi-directional relaying,” in *Proc. 45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2007.

[18] C. Schnurr, S. Stańczak, and T. J. Oechtering, “Achievable rates for the restricted half-duplex two-way relay channel under a partial-decode-and-forward protocol,” in *IEEE Inform. Theory Workshop (ITW)*, Porto, Portugal, May 2008.

[19] S. Zhang and S. C. Liew, “Capacity of two-way relay channel,” available at http://arxiv.org/abs/0804.3120.

[20] W. Nam, S. Chung, and Y. Lee, “Capacity bounds for two-way relay channels,” in *Proc. Int'l Zurich Seminar on Communications (IZS)*, Zurich, Switzerland, Mar. 2008.

[21] ——, “Capacity of the Gaussian two-way relay channel to within $\frac{1}{2}$ bit,” *IEEE Trans. Inform. Theory*, to appear.

[22] B. Rankov and A. Wittneben, “Achievable rate regions for the two-way relay channel,” in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT)*, July 2006, pp. 1668–1672.

[23] A. S. Avestimehr, A. Sezgin, and D. N. C. Tse, “Approximate capacity of the two-way relay channel: A deterministic approach,” in *Proc. 46th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2008.

[24] ——, “Capacity of the two-way relay channel within a constant gap,” *European Transactions on Telecommunications*, to appear.

[25] A. Sezgin, M. A. Khajehnejad, A. S. Avestimehr, and B. Hassibi, “Approximate capacity region of the two-pair bidirectional Gaussian relay network,” in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT)*, Seoul, Korea, July 2009.

[26] A. S. Avestimehr, M. A. Khajehnejad, A. Sezgin, and B. Hassibi, “Capacity region of the deterministic multi-pair bi-directional relay network,” in *IEEE Inform. Theory Workshop (ITW)*, Volos, Greece, June 2009.

[27] D. Gündüz, A. Yener, A. Goldsmith, and H. V. Poor, “The multi-way relay channel,” in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT)*, Seoul, Korea, June 2009.

[28] T. K.-Akino, P. Popovski, and V. Tarokh, “Optimized constellations for two-way wireless relaying with physical network coding,” *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 773–787, June 2009.

[29] D. Gündüz, A. Goldsmith, and H. V. Poor, “MIMO two-way relay channel: Diversity-multiplexing trade-off analysis,” in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Oct. 2008.

[30] R. Vaze and R. W. H. Jr., “On the capacity and diversity-multiplexing tradeoff of the two-way relay channel,” *IEEE Trans. Inform. Theory*, submitted for publication, (available at http://arxiv.org/abs/0810.3900).

[31] T. J. Oechtering and H. Boche, “Optimal transmit strategies in multi-antenna bidirectional relaying,” in *Proc. IEEE Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, HI, Apr. 2007.

[32] I. Hammerström, M. Kuhn, C. Eşli, J. Zhao, A. Wittneben, and G. Bauch, “MIMO two-way relaying with transmit CSI at the relay,” in *Proc. IEEE Signal Process. Advances in Wireless Commun. (SPAWC)*, Helsinki, Finland, June 2007.

[33] R. Vaze and R. W. Heath Jr., “Capacity scaling for mimo two-way relaying,” in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT)*, Nice, France, June 2007.

[34] H. J. Yang, K. Lee, and J. Chun, “Zero-forcing based two-phase relaying,” in *Proc. IEEE Int'l Conf. Commun. (ICC)*, Glasgow, Scotland, June 2007.

[35] H. J. Yang and J. Chun, “Generalized schur decomposition-based two-way relaying for wireless MIMO systems,” in *Proc. IEEE Globecom*, New Orleans, LA, Nov.-Dec. 2008.

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

[37] U. Erez and R. Zamir, “Achieving 1/2(1+SNR) on the AWGN channel with lattice encdoing and decoding,” *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[38] H.-A. Loeliger, “Averaging bounds for lattices and linear codes,” *IEEE Trans. Inform. Theory*, vol. 43, no. 6, Nov. 1997.

[39] H. E. Gamal, G. Caire, and M. O. Damen, “Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels,” *IEEE Trans. Inform. Theory*, vol. 50, no. 6, June 2004.

[40] T. J. Oechtering, R. F. Wyrembelski, and H. Boche, “Multi-antenna bidirectional broadcast channels-optimal transmit strategies,” *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1948–1958, May 2009.

[41] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiely, 1991.