Capacity of the Gaussian Two-Way Relay Channel to Within $\frac{1}{2}$ Bit

Wooseok Nam, *Member, IEEE*, Sae-Young Chung, *Senior Member, IEEE*, and Yong H. Lee, *Senior Member, IEEE*

***Abstract*—In this paper, a Gaussian two-way relay channel, where two source nodes exchange messages with each other through a relay, is considered. We assume that all nodes operate in full-duplex mode and there is no direct channel between the source nodes. We propose an achievable scheme composed of nested lattice codes for the uplink and structured binning for the downlink. Unlike conventional nested lattice codes, our codes utilize two different shaping lattices for source nodes based on a three-stage lattice partition chain, which is a key ingredient for producing the best gap-to-capacity results to date. Specifically, for all channel parameters, the achievable rate region of our scheme is within $\frac{1}{2}$ bit from the capacity region for each user and its sum rate is within $\log \frac{3}{2}$ bit from the sum capacity.**

***Index Terms*—Two-way relay channel, wireless networks, network coding, lattice codes.**

I. Introduction

We consider a two-way relay channel (TRC), as shown in Fig. 1(a). Nodes 1 and 2 want to exchange messages with each other, and a relay node facilitates the communication between them. This TRC can be thought of as a basic building block of general wireless networks, along with the relay channel [1], the interference channel [2], etc. Recently, there has been a great deal of interest in characterizing the capacity of wireless networks. Inspired by network coding [3], TRC has been studied in the context of network coding for wireless networks due to its simple structure. However, the capacity region of the general TRC is still unknown.

In [4], several classical relaying strategies for the one-way relay channel [1], such as amplify-and-forward (AF), decode-and-forward (DF), and compress-and-forward (CF), were extended and applied to the TRC. AF relaying is a very simple and practical strategy, but due to the noise amplification, its performance suffers at low signal-to-noise ratios (SNRs). DF relaying requires the relay to decode all the source messages and, thus, does not suffer from the noise amplification. In [5], it was shown that the achievable rate region of DF relaying can be improved by applying network coding to the decoded messages at the relay. This scheme is optimal in some cases [6], but it is generally subject to *multiplexing loss* [7].

In general, the relay does not need to reconstruct all the messages, but only needs to pass sufficient information to the destination nodes to do so. CF or partial DF relaying strategies for the TRC, in which the relay does not fully decode the source messages, were studied in [8]. In [9], a deterministic approach was used to achieve the information theoretic cut-set bound [26] within $\frac{3}{2}$ bits for each user when applied to the Gaussian TRC.

In this paper, we focus on the Gaussian TRC with full-duplex nodes and no direct communication links between the source nodes. Such a Gaussian TRC is shown in Fig. 1(b), and it is essentially the same as those considered in [8]–[11]. For the uplink, i.e., the channel from the source nodes to the relay, we propose a scheme based on nested lattice codes [12] formed from a lattice partition chain [13]. This scheme is borrowed from our previous work on the relay networks with interference in [14], [15]. By using nested lattice codes for the uplink, we can exploit the structural gain of *computation coding* [16], which corresponds to a kind of combined channel and network coding. For the downlink, i.e., the channel from the relay to the destination nodes, we see the channel as a broadcast channel with receiver side information [6], [17]–[19], since the receiver nodes know their own transmitted messages. In such a channel, the capacity region can be achieved by random binning of messages [17]. In our strategy, a structural binning of messages, rather than the random one, is naturally introduced by the lattice codes used in the uplink. Thus, at each destination node, together with the side information, i.e., its own message, this binning information can be exploited for decoding.

In fact, our work is not the first to apply lattice codes to the Gaussian TRC. In [11], nested lattice codes were already proposed for the Gaussian TRC. In [11], however, it was assumed that the channel is symmetric, i.e., all source and relay nodes have the same transmit powers and noise variances, which makes it easier to use the same nested lattice codes at both sources.

We propose a new nested lattice coding scheme based on a three-stage lattice partition chain. In this scheme, the source nodes still share the same coding lattice but can choose different shaping lattices according to their transmit powers. Using this scheme, we show that we can in fact achieve the cut-set bound within $\frac{1}{2}$ bit for each user[1] for any channel parameters, e.g.,

[1]This is $\frac{1}{2}$ bit per real dimension or 1 bit per complex dimension.

Manuscript received February 15, 2009; revised January 22, 2010. Date of current version October 20, 2010. This work was supported in part by the IT R&D program of MKE/IITA [2008-F-004-01, 5G mobile communication systems based on beam division multiple access and relays with group cooperation] and by the MKE under ITRC supervised by NIPA [NIPA-2010-(C1090-1011-0005)]. The material in this paper was presented in part at the International Zurich Seminar (IZS), Zurich, Switzerland, March 2008.

The authors are with the Department of Electrical Engineering, KAIST, Daejeon, Korea (e-mail: wsnam@stein.kaist.ac.kr; sychung@ee.kaist.ac.kr; yohlee@ee.kaist.ac.kr).

Communicated by E. Erkip, Associate Editor for Shannon Theory.

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.2010.2069150

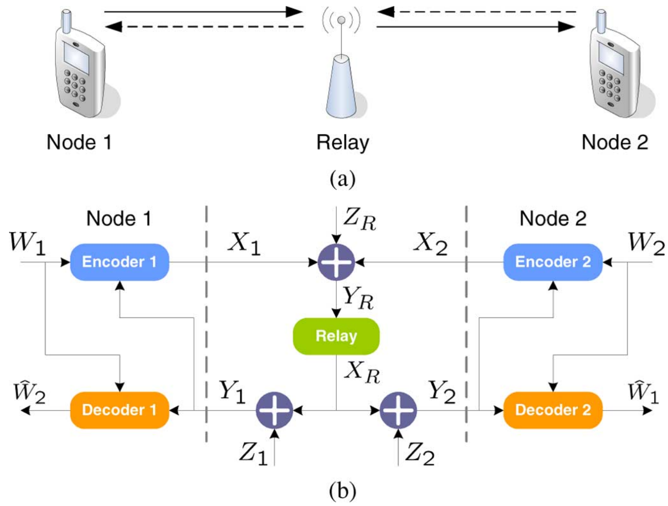

Fig. 1. Gaussian two-way relay channel.

transmit powers and noise variances.[2] We also show that the sum-rate achieved by our scheme is within $\log\frac{3}{2} \simeq 0.58$ bits from the cut-set bound for all channel parameters. Moreover, these gaps vanish as the uplink SNRs increase. Thus, compared to the previous works [8]–[11], our result gives the tightest capacity characterization of the Gaussian TRC.

This paper is organized as follows. In Section II, we present the channel model and define related parameters. The cut-set bound on the capacity region is given in Section III. Section IV illustrates our achievable scheme and computes its achievable rate region. Section V concludes the paper.

II. SYSTEM MODEL

We consider a Gaussian two-way relay channel, as shown in Fig. 1(b). We assume that the source and relay nodes operate in full-duplex mode and there is no direct path between the two source nodes. The variables of the channel are as follows:

- $W_i \in \{1,\ldots,2^{nR_i}\}$: message of node i;
- $\mathbf{X}_i = [X_i^{(1)},\ldots,X_i^{(n)}]^T$: channel input of node i;
- $\mathbf{Y}_R = [Y_R^{(1)},\ldots,Y_R^{(n)}]^T$: channel output at the relay;
- $\mathbf{X}_R = [X_R^{(1)},\ldots,X_R^{(n)}]^T$: channel input of the relay;
- $\mathbf{Y}_i = [Y_i^{(1)},\ldots,Y_i^{(n)}]^T$: channel output at node i;
- $\hat{W}_i \in \{1,\ldots,2^{nR_i}\}$: estimated message of node i;

where $i \in \{1,2\}$, n is the number of channel uses, and R_i denotes the rate of the message of node i. We assume that the messages W_1 and W_2 are uniformly distributed and independent of each other. Node i transmits $X_i^{(t)}$ at time t to the relay through the uplink channel specified by

$$Y_R^{(t)} = X_1^{(t)} + X_2^{(t)} + Z_R^{(t)}$$

where $Z_R^{(t)}$ is an independent identically distributed (i.i.d.) Gaussian random variable with zero mean and variance of σ_R^2. The transmit signal $X_i^{(t)}$ is determined as a function of message W_i and past channel outputs $\mathbf{Y}_i^{t-1} = [Y_i^{(1)},\ldots,Y_i^{(t-1)}]^T$, i.e., $X_i^{(t)} = f_i^{(t)}(W_i,\mathbf{Y}_i^{t-1})$. There are power constraints P_i, $i \in \{1,2\}$ on the transmitted signals

$$\frac{1}{n}\sum_{t=1}^{n}\left(X_i^{(t)}\right)^2 \le P_i, i = 1,2.$$

At the same time, the relay transmits $X_R^{(t)}$ to nodes 1 and 2 through the downlink channel specified by

$$Y_i^{(t)} = X_R^{(t)} + Z_i^{(t)}, i \in \{1,2\}$$

where $Z_i^{(t)}$ is an i.i.d. Gaussian random variable with zero mean and variance of σ_i^2. The power constraint at the relay is given by

$$\frac{1}{n}\sum_{t=1}^{n}\left(X_R^{(t)}\right)^2 \le P_R.$$

Since the relay has no messages of its own, $X_R^{(t)}$ is formed as a function of past channel outputs $\mathbf{Y}_R^{t-1} = [Y_R^{(1)},\ldots,Y_R^{(t-1)}]^T$, i.e., $X_R^{(t)} = f_R^{(t)}(\mathbf{Y}_R^{t-1})$. At node 1, the message estimate $\hat{W}_2 = g_1(W_1,\mathbf{Y}_1)$ is computed from the received signal $\mathbf{Y}_1$ and its message W_1. Decoding at node 2 is performed similarly. Now, the average probability of error is defined as

$$P_e = \Pr\{\hat{W}_1 \ne W_1 \text{ or } \hat{W}_2 \ne W_2\}.$$

For the aforementioned TRC, we say that a rate pair (R_1,R_2) is achievable if a sequence of encoding and decoding functions exists such that the error probability vanishes as n tends to infinity. The capacity region of the TRC is defined as the closure of all achievable rate pairs.

III. AN OUTER BOUND FOR THE CAPACITY REGION

By the cut-set bound [26], if a rate pair (R_1,R_2) is achievable for a general TRC, a joint probability distribution $p(x_1,x_2,x_R)$ exists such that

$$\begin{aligned} R_1 &\le \min\{I(X_1;Y_R,Y_2|X_R,X_2) \\ &\qquad I(X_1,X_R;Y_2|X_2)\} \quad (1a) \\ R_2 &\le \min\{I(X_2;Y_R,Y_1|X_R,X_1) \\ &\qquad I(X_2,X_R;Y_1|X_1)\}. \quad (1b) \end{aligned}$$

In particular, for the Gaussian TRC under consideration, we can use the fact that there is no direct path between nodes 1 and 2, and all terms under the minimization are maximized by the product distribution $p(x_1,x_2,x_R) = p(x_1)p(x_2)p(x_R)$, where $p(x_1)$, $p(x_2)$, and $p(x_R)$ are Gaussian probability density functions with zero means and variances of P_1, P_2, and P_R, respectively. In [7], it was shown that (1) reduces to

$$R_1 \le \min\left\{\frac{1}{2}\log\left(1+\frac{P_1}{\sigma_R^2}\right), \frac{1}{2}\log\left(1+\frac{P_R}{\sigma_2^2}\right)\right\} \quad (2a)$$

$$R_2 \le \min\left\{\frac{1}{2}\log\left(1+\frac{P_2}{\sigma_R^2}\right), \frac{1}{2}\log\left(1+\frac{P_R}{\sigma_1^2}\right)\right\}. \quad (2b)$$

[2]The same result was shown in our previous work [10]. However, the result in [10] relied on the existence of lattices satisfying a certain condition. In this paper, we further refine the condition and show the existence of such lattices for all channel parameters, thus making our results fully general.

IV. An Achievable Rate Region for the Gaussian TRC

In this section, we compute an achievable rate region for the Gaussian TRC. For the uplink, we consider using nested lattice codes, which are formed from a lattice partition chain. For the downlink, we use a structured binning of messages at the relay, which is naturally introduced by the nested lattice codes. The destination nodes decode each other's message using this binning information and their own messages as side information.

The main result of this section is as follows.

Theorem 1: For a Gaussian TRC, as shown in Fig. 1(b), we can achieve the following region:

$$R_1 \leq \min\left\{\left[\frac{1}{2}\log\left(\frac{P_1}{P_1+P_2}+\frac{P_1}{\sigma_R^2}\right)\right]^+ \right.$$
$$\left. \frac{1}{2}\log\left(1+\frac{P_R}{\sigma_2^2}\right)\right\} \tag{3a}$$
$$R_2 \leq \min\left\{\left[\frac{1}{2}\log\left(\frac{P_2}{P_1+P_2}+\frac{P_2}{\sigma_R^2}\right)\right]^+ \right.$$
$$\left. \frac{1}{2}\log\left(1+\frac{P_R}{\sigma_1^2}\right)\right\} \tag{3b}$$

where $[x]^+ \triangleq \max\{x,0\}$.

If we compare the right-hand sides (RHS) of (2a) and (3a), the second terms in the minimizations are the same and the first terms differ by at most $\frac{1}{2}$ bit. Then from a simple inequality $\min\{a_1,a_2\}-\min\{b_1,b_2\} \leq \max\{a_1-b_1,a_2-b_2\}$, the RHSs of (2a) and (3a) differ by at most $\frac{1}{2}$ bit. The same holds for (2b) and (3b), and thus the achievable rate region (3) is within $\frac{1}{2}$ bit of the outer bound (2) for each user regardless of channel parameters such as the transmit powers and noise variances. Moreover, as the uplink SNRs $\frac{P_1}{\sigma_R^2}$ and $\frac{P_2}{\sigma_R^2}$ increase, the gap vanishes and our achievable region asymptotically approaches the capacity region of the Gaussian TRC.

We prove Theorem 1 in the following subsections.

Remark 1: After some algebra, it can be easily shown that the sum rate achieved by our scheme is within $\log\frac{3}{2} \simeq 0.58$ bits from the cut-set bound for all channel parameters. See the Appendix for the proof. Note that [20], [21] recently showed that the same $\frac{1}{2}$ bit gap can be achieved for each user for the Gaussian TRC using noisy network coding and layered noisy network coding, respectively. However, the sum rates achieved by [20], [21] were shown to be within 1 bit from the cut-set bound. Therefore, our scheme gives the best gap-to-capacity result to date.

A. Lattice Scheme for the Uplink

For the scheme for the uplink, we consider a lattice coding scheme. We will not cover the full details of lattices and lattice codes. For a comprehensive review, we refer readers to [12], [22], [23], and the references therein.

A *nested lattice code* can be defined in terms of two n-dimensional lattices Λ_C^n and Λ^n, which form a lattice partition Λ_C^n/Λ^n, i.e., $\Lambda^n \subseteq \Lambda_C^n$. The nested lattice code is a lattice code which uses Λ_C^n as codewords and the Voronoi region [13] of Λ^n as a shaping region. For Λ_C^n/Λ^n, we define the *set of coset leaders* as

$$\mathcal{C} = \{\Lambda_C^n \bmod \Lambda^n\} \triangleq \{\Lambda_C^n \cap \mathcal{R}\}$$

where $\mathcal{R}$ is the Voronoi region of Λ^n. Then the coding rate of the nested lattice code is given by

$$R = \frac{1}{n}\log|\mathcal{C}| = \frac{1}{n}\log\frac{\mathrm{Vol}(\Lambda^n)}{\mathrm{Vol}\,(\Lambda_C^n)}$$

where $\mathrm{Vol}(\,\cdot\,)$ denotes the volume of the Voronoi region of a lattice, which is given by

$$\mathrm{Vol}(\Lambda^n) = \int_{\mathcal{R}} d\mathbf{x}.$$

As another important parameter of a lattice, the second moment per dimension associated with $\mathcal{R}$ is defined as

$$\sigma^2(\Lambda^n) = \frac{1}{\mathrm{Vol}(\Lambda^n)}\cdot\frac{1}{n}\int_{\mathcal{R}}\|\mathbf{x}\|^2 d\mathbf{x}.$$

To quantify certain optimal properties of lattices, terms like *Rogers-good* and *Poltyrev-good* are frequently used. Roughly speaking, a sequence of lattices is called Rogers-good if it is asymptotically efficient for sphere covering and Poltyrev-good if it is good for additive white Gaussian noise (AWGN) channel coding. For more detail notion of goodness of lattices, see [22]. It is shown in [12] that nested lattice codes can achieve the capacity of the AWGN channel with *minimum Euclidean distance lattice decoding* [24] if Λ^n is simultaneously Rogers-good and Poltyrev-good and Λ_C^n is Poltyrev-good.

In this subsection, we design nested lattice codes based on a three-way lattice partition chain for the Gaussian TRC. In the following argument, we assume that $P_1 \geq P_2$ without loss of generality. Now, let us first consider a theorem that is a key to our code construction.

Theorem 2: For any $P_1 \geq P_2 \geq 0$, a sequence of n-dimensional lattice partition chains [13] $\Lambda_C^n/\Lambda_2^n/\Lambda_1^n$, i.e., $\Lambda_1^n \subseteq \Lambda_2^n \subseteq \Lambda_C^n$, exists that satisfies the following properties.

a) Λ_1^n and Λ_2^n are simultaneously Rogers-good and Poltyrev-good while Λ_C^n is Poltyrev-good.

b) For any $\epsilon > 0$, $P_i - \epsilon \leq \sigma^2(\Lambda_i^n) \leq P_i$, $i \in \{1,2\}$, for sufficiently large n.

c) The coding rate of the nested lattice code associated with the lattice partition Λ_C^n/Λ_2^n can approach any value $\gamma \geq 0$ as n tends to infinity, i.e.

$$R_2 = \frac{1}{n}\log|\mathcal{C}_2| = \frac{1}{n}\log\left(\frac{\mathrm{Vol}\,(\Lambda_2^n)}{\mathrm{Vol}\,(\Lambda_C^n)}\right) = \gamma + o_n(1) \tag{4}$$

where $\mathcal{C}_2 = \{\Lambda_C^n \bmod \Lambda_2^n\}$ and $o_n(1) \to 0$ as $n \to \infty$. Furthermore, the coding rate of the nested lattice code associated with Λ_C^n/Λ_1^n is given by

$$\begin{aligned} R_1 &= \frac{1}{n}\log|\mathcal{C}_1| \\ &= \frac{1}{n}\log\left(\frac{\mathrm{Vol}\,(\Lambda_1^n)}{\mathrm{Vol}\,(\Lambda_C^n)}\right) \\ &= R_2 + \frac{1}{2}\log\left(\frac{P_1}{P_2}\right) + o_n(1) \end{aligned}$$

where $\mathcal{C}_1 = \{\Lambda_C^n \bmod \Lambda_1^n\}$.

Proof: See [15, Proof of Theorem 2]. ■

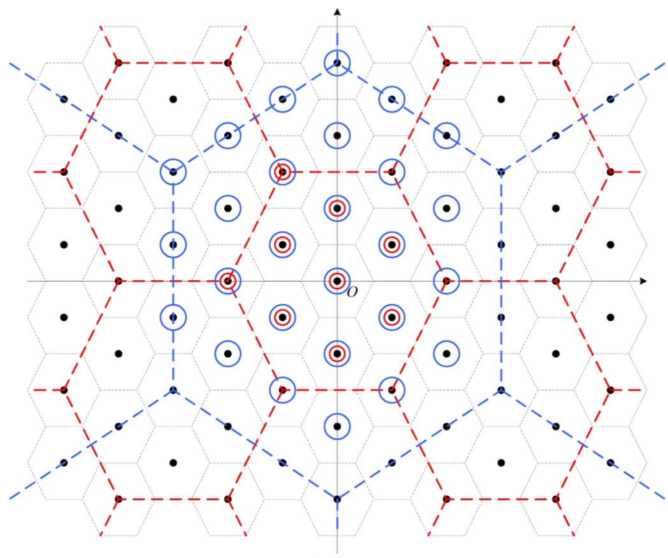

Fig. 2. Example of a lattice partition chain and sets of coset leaders. $\mathcal{C}_2 \subseteq \mathcal{C}_1 \subseteq \Lambda_C$.

For instance, a lattice partition chain and the corresponding sets of coset leaders are visualized in Fig. 2 for the two-dimensional case.

Encoding: Let us think of a lattice partition chain (more precisely, a sequence of lattice partition chains) and sets of coset leaders as described in Theorem 2. We use $\mathcal{C}_1$ and $\mathcal{C}_2$ for nodes 1 and 2 respectively. For node i, the message set $\{1,\ldots,2^{nR_i}\}$ is one-to-one mapped to $\mathcal{C}_i$. Thus, to transmit a message, node i chooses $\mathbf{W}_i \in \mathcal{C}_i$ associated with the message and sends

$$\mathbf{X}_i = (\mathbf{W}_i + \mathbf{U}_i) \bmod \Lambda_i$$

where $\mathbf{U}_i$ is a random dither vector with $\mathbf{U}_i \sim \mathrm{Unif}(\mathcal{R}_i)$ and $\mathcal{R}_i$ denotes the Voronoi region of Λ_i (we suppressed the superscript 'n' for simplicity). The dither vectors $\mathbf{U}_i, i \in \{1,2\}$, are independent of each other and also independent of the messages and the noise. We assume that each $\mathbf{U}_i$ is known to the source nodes and the relay. Note that, due to the *crypto-lemma* [23], $\mathbf{X}_i$ is uniformly distributed over $\mathcal{R}_i$ and independent of $\mathbf{W}_i$. Thus, the average transmit power of node i is equal to $\sigma^2(\Lambda_i)$, which approaches P_i as n tends to infinity, and the power constraint is met.

Decoding: The received vector at the relay is given by

$$\mathbf{Y}_R = \mathbf{X}_1 + \mathbf{X}_2 + \mathbf{Z}_R$$

where $\mathbf{Z}_R = [Z_R^{(1)}, \ldots, Z_R^{(n)}]^T$. Upon receiving $\mathbf{Y}_R$, the relay computes

$$\begin{aligned}\tilde{\mathbf{Y}}_R &= \left(\alpha \mathbf{Y}_R - \sum_{j=1}^{2} \mathbf{U}_j\right) \bmod \Lambda_1\\ &= \Bigg[\sum_{j=1}^{2}(\mathbf{W}_j + \mathbf{U}_j) \bmod \Lambda_j - \sum_{j=1}^{2}\mathbf{X}_j\\ &\quad + \alpha\sum_{j=1}^{2}\mathbf{X}_j + \alpha\mathbf{Z}_R - \sum_{j=1}^{2}\mathbf{U}_j\Bigg] \bmod \Lambda_1\\ &= \Bigg[\sum_{j=1}^{2}(\mathbf{W}_j - Q_j(\mathbf{W}_j + \mathbf{U}_j))\\ &\quad - (1-\alpha)(\mathbf{X}_1 + \mathbf{X}_2) + \alpha\mathbf{Z}_R\Bigg] \bmod \Lambda_1\\ &= \left(\mathbf{T} + \tilde{\mathbf{Z}}_R\right) \bmod \Lambda_1 \end{aligned} \quad (5)$$

where

$$\begin{aligned}\mathbf{T} &= \left[\sum_{j=1}^{2}(\mathbf{W}_j - Q_j(\mathbf{W}_j + \mathbf{U}_j))\right] \bmod \Lambda_1\\ &= [\mathbf{W}_1 + \mathbf{W}_2 - Q_2(\mathbf{W}_2 + \mathbf{U}_2)] \bmod \Lambda_1 \end{aligned} \quad (6)$$

$$\tilde{\mathbf{Z}}_R = -(1-\alpha)(\mathbf{X}_1 + \mathbf{X}_2) + \alpha\mathbf{Z}_R \quad (7)$$

$\alpha \in [0,1]$ is a scaling factor, and $Q_j(\cdot)$ denotes the nearest neighbor lattice quantizer associated with Λ_j. In the third equality of (5), we used the property that $Q_j(\mathbf{x}) = \mathbf{x} - \mathbf{x} \bmod \Lambda_j$. If we let α be the minimum mean-square error (MMSE) coefficient

$$\alpha = \frac{P_1 + P_2}{P_1 + P_2 + \sigma_R^2}$$

the variance of the effective noise (7) satisfies

$$\frac{1}{n}E\left\{\left\|\tilde{\mathbf{Z}}_R\right\|^2\right\} \le \frac{(P_1+P_2)\sigma_R^2}{P_1 + P_2 + \sigma_R^2}.$$

From the chain relation of lattices in Theorem 2, it follows that $\mathbf{T} \in \mathcal{C}_1$. Moreover, using the crypto-lemma, it is obvious that $\mathbf{T}$ is uniformly distributed over $\mathcal{C}_1$ and independent of $\tilde{\mathbf{Z}}_R$ [15, Lemma 2].

The relay attempts to recover $\mathbf{T}$ from $\tilde{\mathbf{Y}}_R$ instead of recovering $\mathbf{W}_1$ and $\mathbf{W}_2$ separately. Thus, the lattice scheme inherits the idea of computation coding [16] and physical-layer network coding [25]. Also, by not requiring the relay to decode both messages, $\mathbf{W}_1$ and $\mathbf{W}_2$, we can avoid multiplexing loss [5] at the relay. The method of decoding we consider is minimum Euclidean distance lattice decoding [12], [24], which finds the closest point to $\tilde{\mathbf{Y}}_R$ in Λ_C. Thus, the estimate of $\mathbf{T}$ is given by $\hat{\mathbf{T}} = Q_C(\tilde{\mathbf{Y}})$, where $Q_C(\cdot)$ denotes the nearest neighbor lattice quantizer associated with Λ_C. Then, from the lattice symmetry and the independence between $\mathbf{T}$ and $\tilde{\mathbf{Z}}_R$, the probability of decoding error is given by

$$\begin{aligned}p_e &= \Pr\{\hat{\mathbf{T}} \ne \mathbf{T}\}\\ &= \Pr\{\tilde{\mathbf{Z}}_R \bmod \Lambda_1 \notin \mathcal{R}_C\}\end{aligned} \quad (8)$$

where $\mathcal{R}_C$ denotes the Voronoi region of Λ_C. We then have the following theorem.

Theorem 3: Let

$$R_1^* = \left[\frac{1}{2}\log\left(\frac{P_1}{P_1+P_2} + \frac{P_1}{\sigma_R^2}\right)\right]^+.$$

For any $\bar{R}_1 < R_1^*$ and a lattice partition chain as described in Theorem 2 with R_1 approaching $\bar{R}_1$, i.e., $R_1 = \bar{R}_1 + o_n(1)$, the error probability under minimum Euclidean distance lattice decoding (8) is bounded by

$$p_e \le e^{-n(E_P(2^{2(R_1^* - \bar{R}_1)}) - o_n(1))}$$

where $E_P(\cdot)$ is the Poltyrev exponent [24] given by

$$E_P(x) = \begin{cases} \frac{x}{2} = \frac{1}{2}(x - 1 - \ln x), & 1 \le x < 2 \\ \frac{1}{2}\left(1 + \ln \frac{x}{4}\right), & 2 \le x < 4 \\ \frac{x}{8}, & x \ge 4. \end{cases}$$

Proof: See the proof of Theorem 3 in [15]. ■

According to Theorem 3, the error probability vanishes as $n \to \infty$ if $\bar{R}_1 < R_1^*$ since $E_p(x) > 0$ for $x > 1$. This implies that the nested lattice code can have any rate below R_1^* for the reliable decoding of $\mathbf{T}$. Thus, by c) of Theorem 2 and Theorem 3, the error probability at the relay vanishes as $n \to \infty$ if

$$R_i < \left[\frac{1}{2}\log\left(\frac{P_i}{P_1 + P_2} + \frac{P_i}{\sigma_R^2}\right)\right]^+, \quad i = 1, 2. \tag{9}$$

B. Downlink Phase

Recall that, in the previous subsection, we have assumed that $P_1 \ge P_2$ without loss of generality and thus $R_1 \ge R_2$ by uplink code construction. We now generate 2^{nR_1} n-sequences with each element i.i.d. according to $\mathcal{N}(0, P_R)$. These sequences form a codebook $\mathcal{C}_R$. We assume one-to-one correspondence between each $\mathbf{t} \in \mathcal{C}_1$ and a codeword $\mathbf{X}_R \in \mathcal{C}_R$. To make this correspondence explicit, we use the notation $\mathbf{X}_R(\mathbf{t})$. After the relay decodes $\hat{\mathbf{T}}$, it transmits $\mathbf{X}_R(\hat{\mathbf{T}})$ at the next block to nodes 1 and 2. We assume that there is no error in the uplink, i.e., $\hat{\mathbf{T}} = \mathbf{T}$. Under this condition, $\hat{\mathbf{T}}$ is uniform over $\mathcal{C}_1$, and, thus, $\mathbf{X}_R(\hat{\mathbf{T}})$ is also uniformly chosen from $\mathcal{C}_R$.

Upon receiving $\mathbf{Y}_1 = \mathbf{X}_R + \mathbf{Z}_1$, where $\mathbf{Z}_1 = [Z_1^{(1)}, \ldots, Z_1^{(n)}]^T$, node 1 estimates the relay message $\hat{\mathbf{T}}$ as $\hat{\mathbf{T}}_1$ if a unique codeword $\mathbf{X}_R(\hat{\mathbf{T}}_1) \in \mathcal{C}_{R,1}$ exists such that $(\mathbf{X}_R(\hat{\mathbf{T}}_1), \mathbf{Y}_1)$ are jointly typical, where

$$\mathcal{C}_{R,1} = \{\mathbf{X}_R(\mathbf{t}) : \mathbf{t} = [\mathbf{W}_1 + \mathbf{w}_2 - Q_2(\mathbf{w}_2 + \mathbf{U}_2)] \bmod \Lambda_1, \mathbf{w}_2 \in \mathcal{C}_2\}.$$

Then, from the knowledge of $\mathbf{W}_1$ and $\hat{\mathbf{T}}_1$, node 1 estimates the message of node 2 as

$$\hat{\mathbf{W}}_2 = (\hat{\mathbf{T}}_1 - \mathbf{W}_1) \bmod \Lambda_2. \tag{10}$$

Given $\hat{\mathbf{T}} = \mathbf{T}$, we have $\hat{\mathbf{W}}_2 = \mathbf{W}_2$ if and only if $\hat{\mathbf{T}}_1 = \hat{\mathbf{T}}$. Note that $|\mathcal{C}_{R,1}| = 2^{nR_2}$. Thus, from the argument of random coding and jointly typical decoding [26], we have

$$\Pr\{\hat{\mathbf{T}}_1 \ne \hat{\mathbf{T}} | \hat{\mathbf{T}} = \mathbf{T}\} \to 0 \tag{11}$$

as $n \to \infty$ if

$$R_2 < \frac{1}{2}\log\left(1 + \frac{P_R}{\sigma_1^2}\right). \tag{12}$$

Similarly, at node 2, the relay message is estimated to be $\hat{\mathbf{T}}_2$ if there exists a unique codeword $\mathbf{X}_R(\hat{\mathbf{T}}_2) \in \mathcal{C}_{R,2}$ such that $(\mathbf{X}_R(\hat{\mathbf{T}}_2), \mathbf{Y}_2)$ are jointly typical, where

$$\mathcal{C}_{R,2} = \{\mathbf{X}_R(\mathbf{t}) : \mathbf{t} = [\mathbf{w}_1 + \mathbf{W}_2 - Q_2(\mathbf{W}_2 + \mathbf{U}_2)] \bmod \Lambda_1, \mathbf{w}_1 \in \mathcal{C}_1\}.$$

Then the message of node 1 is estimated as

$$\hat{\mathbf{W}}_1 = [\hat{\mathbf{T}}_2 - \mathbf{W}_2 + Q_2(\mathbf{W}_2 + \mathbf{U}_2)] \bmod \Lambda_1. \tag{13}$$

Since $|\mathcal{C}_{R,1}| = 2^{nR_1}$, we have

$$\Pr\{\hat{\mathbf{T}}_2 \ne \hat{\mathbf{T}} | \hat{\mathbf{T}} = \mathbf{T}\} \to 0 \tag{14}$$

as $n \to \infty$ if

$$R_1 < \frac{1}{2}\log\left(1 + \frac{P_R}{\sigma_2^2}\right). \tag{15}$$

Note that, in the downlink, although the channel setting is broadcast, nodes 1 and 2 achieve their point-to-point channel capacities (12) and (15) without being affected by each other. This is because of the side information on the transmitted message at each node and binning of messages [6], [17]–[19]. In our scheme, the relation in (6) represents how the message pair $(\mathbf{W}_1, \mathbf{W}_2)$ is binned to $\mathbf{T}$.

C. Achievable Rate Region

Clearly, the message estimates (10) and (13) are exact if and only if $\hat{\mathbf{T}}_1 = \hat{\mathbf{T}}_2 = \mathbf{T}$. Thus, the error probability is given by

$$\begin{aligned} P_e &= \Pr\{\hat{\mathbf{T}}_1 \ne \mathbf{T} \text{ or} \hat{\mathbf{T}}_2 \ne \mathbf{T}\} \\ &\le \Pr\{\hat{\mathbf{T}}_1 \ne \hat{\mathbf{T}} \text{or} \hat{\mathbf{T}}_2 \ne \hat{\mathbf{T}} \text{ or } \hat{\mathbf{T}} \ne \mathbf{T}\} \\ &\le \Pr\{\hat{\mathbf{T}} \ne \mathbf{T}\} + \Pr\{\hat{\mathbf{T}}_1 \ne \hat{\mathbf{T}} | \hat{\mathbf{T}} = \mathbf{T}\} \\ &\quad + \Pr\{\hat{\mathbf{T}}_2 \ne \hat{\mathbf{T}} | \hat{\mathbf{T}} = \mathbf{T}\} \end{aligned} \tag{16}$$

By Theorem 3, the first term of (16) vanishes as $n \to \infty$ if $R_i < R_i^*$, $i \in \{1, 2\}$. Also, by (11) and (14), the second and third terms also vanish as $n \to \infty$ if (12) and (15) hold. Thus, the achievable rate region (3) follows from (9), (12), and (15).

V. Conclusion

In this paper, we considered the Gaussian TRC. An achievable scheme was presented based on nested lattice codes for the uplink and structured binning for the downlink. In particular, our nested lattice codes are distinguished from the previous ones in that we use different shaping lattices for each source node to satisfy their different transmit power constraints. Our scheme achieves rates within $\frac{1}{2}$ bit from the capacity region for each user and within $\log \frac{3}{2}$ bit from the sum capacity, which gives the best gap-to-capacity result to date. Though the capacity region is very nearly reached, the exact capacity region of the Gaussian TRC is still an open problem.

APPENDIX
PROOF FOR THE $\log\frac{3}{2}$ BITS GAP TO THE SUM CAPACITY

To show the $\log\frac{3}{2}$ bits gap to the sum capacity, it is sufficient to prove the following inequality:

$$\begin{aligned}\frac{1}{2}\log(1+x)+\frac{1}{2}\log(1+y)\\ -\left[\frac{1}{2}\log\left(\frac{x}{x+y}+x\right)\right]^{+}\\ -\left[\frac{1}{2}\log\left(\frac{y}{x+y}+y\right)\right]^{+}\\ \le\log\frac{3}{2}\end{aligned}\tag{17}$$

for $x>0$ and $y>0$. We denote the left-hand side (LHS) of (17) as $g(x,y)$. We prove the inequality by dividing the region of x and y, and showing $g(x,y)\le\log\frac{3}{2}$ in each region.

A. $\frac{x}{x+y}+x>1$, $\frac{y}{x+y}+y>1$: In this case, $g(x,y)$ reduces to

$$\begin{aligned}g(x,y)&=\frac{1}{2}\log(1+x)+\frac{1}{2}\log(1+y)\\&\quad-\frac{1}{2}\log\left(\frac{x}{x+y}+x\right)\\&\quad-\frac{1}{2}\log\left(\frac{y}{x+y}+y\right)\\&=\frac{1}{2}\log\left(1+\frac{y}{x(1+x+y)}\right)\\&\quad+\frac{1}{2}\log\left(1+\frac{x}{y(1+x+y)}\right).\end{aligned}$$

From $\frac{x}{x+y}+x>1$ and $\frac{y}{x+y}+y>1$, we have $\frac{y}{x(1+x+y)}<\frac{y}{x+y}$ and $\frac{x}{y(1+x+y)}<\frac{x}{x+y}$, and, thus

$$\begin{aligned}g(x,y)&<\frac{1}{2}\log\left(1+\frac{y}{x+y}\right)+\frac{1}{2}\log\left(1+\frac{x}{x+y}\right)\\&\le\log\frac{3}{2}\end{aligned}$$

where the second inequality is by Jensen's inequality.

B. $\frac{x}{x+y}+x>1$, $\frac{y}{x+y}+y\le 1$: We have

$$\begin{aligned}g(x,y)=\frac{1}{2}\log(1+x)+\frac{1}{2}\log(1+y)\\-\frac{1}{2}\log\left(\frac{x}{x+y}+x\right).\end{aligned}$$

From $\frac{x}{x+y}+x>1$, $\frac{y}{x(1+x+y)}<\frac{y}{x+y}$ holds, and thus

$$\begin{aligned}g(x,y)&<\frac{1}{2}\log\left(1+\frac{y}{x+y}\right)+\frac{1}{2}\log(1+y)\\&\le\log\left(1+\frac{1}{2}\left(\frac{y}{x+y}+y\right)\right)\\&\le\log\frac{3}{2}\end{aligned}$$

where the second inequality is by Jensen's inequality, and the third by $\frac{y}{x+y}+y\le 1$.

C. $\frac{x}{x+y}+x\le 1$, $\frac{y}{x+y}+y>1$: Due to the symmetry, we have the same result as in case *B*.

D. $\frac{x}{x+y}+x\le 1$, $\frac{y}{x+y}+y\le 1$: We have

$$\begin{aligned}g(x,y)&=\frac{1}{2}\log(1+x)+\frac{1}{2}\log(1+y)\\&\le\log\left(1+\frac{x+y}{2}\right)\end{aligned}$$

where the inequality is by Jensen's inequality. From $\frac{x}{x+y}+x\le 1$ and $\frac{y}{x+y}+y\le 1$, we have $x+y\le 1$ and, therefore, $g(x,y)$ is bounded by $\log\frac{3}{2}$. □

REFERENCES

[1] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channels," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[2] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probabil.*, vol. 2, pp. 805–814, 1974.

[3] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1204–1216, Oct. 2000.

[4] B. Rankov and A. Wittneben, "Achievable rate region for the two-way relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Seatle, WA, Jul. 2006, pp. 1668–1672.

[5] R. Knopp, "Two-way radio network with a star topology," *Int. Zurich Seminar on Commun.*, Feb. 2006.

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

[7] R. Knopp, "Two-way wireless communication via a relay station," in *Proc. GDR-ISIS Meet.*, Mar. 2007.

[8] D. Gündüz, E. Tuncel, and J. Nayak, "Rate regions for the separated two-way relay channel," in *Proc. 46th Ann. Allerton Conf.*, Sep. 2008.

[9] A. S. Avestimehr, A. Sezgin, and D. N. C. Tse, "Approximate capacity of the two-way relay channel: A deterministic approach," in *Proc. 46th Ann. Allerton Conf.*, Sep. 2008.

[10] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity bounds for two-way relay channels," *Proc. Int. Zurich Seminar Commun.*, Mar. 2008.

[11] K. Narayanan, M. P. Wilson, and A. Sprintson, "Joint physical layer coding and network coding for bi-directional relaying," in *Proc. 45th Ann. Allerton Conf.*, Sep. 2007.

[12] U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1+\mathrm{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2293–2314, Oct. 2004.

[13] G. D. Forney Jr., M. D. Trott, and S.-Y. Chung, "Sphere-bound-achieving coset codes and multilevel coset codes," *IEEE Trans. Inf. Theory*, vol. 46, pp. 820–850, May 2000.

[14] W. Nam and S.-Y. Chung, "Relay networks with orthogonal components," in *Proc. 46th Ann. Allerton Conf.*, Sept. 2008.

[15] W. Nam, S.-Y. Chung, and Y. H. Lee, "Nested lattice codes for Gaussian relay networks with interference," *IEEE Trans. Inf. Theory* [Online]. Available: http://arxiv.org/abs/0902.2436, submitted for publication

[16] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3498–3516, Oct. 2007.

[17] L.-L. Xie, "Network coding and random binning for multi-user channels," in *Proc. 10th Canadian Workshop on Inf. Theory*, Jun. 2007.

[18] Y. Wu, "Broadcasting when receivers know some messages a priori," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007.

[19] G. Kramer and S. Shamai, "Capacity for classes of broadcast channels with receiver side information," in *IEEE Inf. Theory Workshop*, Tahoe , CA, Sep. 2007.

[20] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, Noisy Network Coding [Online]. Available: http://arxiv.org/abs/1002.3188

[21] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Layered noisy network coding," in *Proc. IEEE Wireless Netw. Coding Workshop*, Boston, MA, Jun. 2010.

[22] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3401–3416, Oct. 2005.

[23] G. D. Forney Jr., "On the role of MMSE estimation in approaching the information theoretic limits of linear Gaussian channels: Shannon meets Wiener," in *Proc. 41st Ann. Allerton Conf.*, Oct. 2003.

[24] G. Poltyrev, "On coding without restrictions for the AWGN channel," *IEEE Trans. Inf. Theory*, vol. 40, pp. 409–417, Mar. 1994.
[25] S. Zhang, S. Liew, and P. Lam, "Physical-layer network coding," *ACM Mobicom'06*.
[26] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

Wooseok Nam (S'04–M'09) was born in Busan, Korea, in 1980. He received the B.S., M.S., and Ph.D. degrees from the Department of Electrical Engineering, KAIST, Daejeon, Korea, in 2002, 2004, and 2009, respectively.

His research interests include network information theory, adaptive signal processing, and digital modem design.

Sae-Young Chung (S'89–M'00–SM'07) received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Seoul, Korea, in 1990 and 1992, respectively. He received the Ph.D. degree from the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, in 2000.

From September 2000 to December 2004, he was with Airvana, Inc. Since January 2005, he has been with KAIST, Daejeon, Korea, where he is now an Associate Professor with the Department of Electrical Engineering. His research interests include network information theory, coding theory, and their applications to wireless communications.

Dr. Chung is an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS.

Yong H. Lee (S'81–M'84–SM'98) was born in Seoul, Korea, on July 12, 1955. He received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Seoul, in 1978 and 1980, respectively, and the Ph.D. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1984.

From 1984 to 1988, he was an Assistant Professor with the Department of Electrical and Computer Engineering, State University of New York, Buffalo. Since 1989, he has been with the Department of Electrical Engineering, KAIST, Daejeon, Korea, where he is currently a Professor and the Dean of College of Information Science and Technology. His research activities are in the area of communication signal processing, which includes interference management, resource allocation, synchronization, estimation and detection for CDMA, TDMA, OFDM, and MIMO systems. He is also interested in designing and implementation of transceivers.