Lattice strategies for a multi-pair bi-directional relay network

Sang Joon Kim, Besma Smida and Natasha Devroye

***Abstract*—We consider a cellular network inspired channel model which consists of the bi-directional exchange of information between a base-station and M terminal nodes with the help of a relay. The base-station has a message for each of M terminals, and conversely, each terminal node has a message for the base-station. A single relay assists the bi-directional communication endeavor. We assume an AWGN channel model with direct links (omitted in previous studies) between the base-station, relay, and half-duplex nodes. In this scenario, we derive achievable rate regions for two temporal protocols – needed in half-duplex networks – which indicate which users transmit when. These achievable rate regions are based on a novel lattice encoding and decoding strategy which outperforms previously derived regions using random-coding based decode-and-forward strategies under certain channel conditions. The terminal nodes employ nested lattice codes, and the relay decodes a series of codeword combinations – one of the main novelties of our scheme – from which it deduce the sum-codewords of the base-station to terminal node i, which it then broadcasts. This scheme differs markedly from previously considered successive-decoding based lattice strategies and provides a more general framework for the joint decoding of lattice codewords in a MAC. Numerical evaluations of our lattice-based inner bounds are shown to improve upon previous random-coding based schemes under certain channel conditions, and are compared to half-duplex cut-set outer bounds. We further demonstrate a constant sum-rate gap result using this lattice-based scheme for symmetric channels for one of the two protocols.**

***Index Terms*—bi-directional relay channel, lattice codes**

I. Introduction

We consider bi-directional communication over additive white Gaussian noise (AWGN) channels between a base-station and M terminal nodes with the help of a relay. This models cellular communication systems, where the uplink and downlink streams of data are jointly considered rather than treated as two uni-directional links, as is currently the norm. We further attempt to model realistic systems by assuming 1) the base-station, relay and terminals are half-duplex, and 2) direct links exist between the base-station and terminal nodes. The physical channel model is given in Fig. 1.

Contributions. We extend prior work on this channel model [7] in which Decode-and-Forward (DF) strategies at the relay with and without Compress-and-Forward (CF) based cooperation strategies at the terminal nodes were used to derive achievable rate regions for three temporal protocols. Three temporal

Sang Joon Kim is with the Samsung Advanced Institute of Technology, Korea. Email: sangjoon0919.kim@samsung.com. Besma Smida is with the Department of Electrical and Computer Engineering, Purdue University Calumet, Hammond, IN 46323. E-mail: bsmida@purdue.edu. Natasha Devroye is with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL 60607. Email: devroye@ece.uic.edu. The work of N. Devroye was partially supported by NSF under award 1053933.

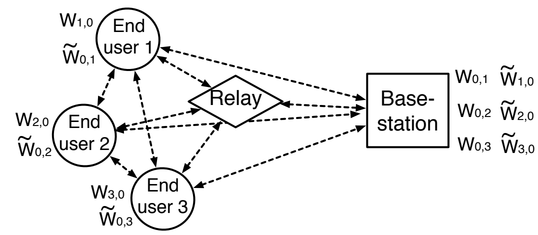

Fig. 1. Our physical channel model consists of multiple independent bi-directional desired communication flows (indicated by arrows) between multiple terminal nodes and a single base-station.

protocols, which indicate which user(s) transmit when and are needed to ensure the half-duplex constraints, were considered: 1) Full Multiple Access and Broadcast Channel (FMABC), 2) Partial Multiple Access and Broadcast Channel (PMABC), and 3) Full Time Division Broadcast Channel (FTDBC). In all these protocols, the relay utilized a DF strategy together with random coding and jointly typical decoding at all nodes; in this paper we derive achievable rate regions by allowing the terminals and base-station to employ structured lattice codes, which in turn permits the relay to decode message combinations rather than individual messages. In particular, we propose a novel and markedly different decoding strategy at the relay in which it decodes combinations of "sums of messages" and uses these to obtain the individual "sum" of the messages from the base-station to user i and vice-versa (for all M users), as opposed to the previously proposed successive decoding strategy [6], in which the relay decodes the modulo "sums of messages" individually. Hence, for the channel illustrated in Fig. 1, we proposed a scheme where the relay successively decodes the lattice points corresponding to $W_{10} + W_{01}$, $W_{20} + W_{02}$ and $W_{30} + W_{30}$ not by decoding these directly, but rather by exploiting the lattice structure to decode the combinations of the sums, e.g. the lattice points corresponding to $W_{10} + W_{01} + W_{20} + W_{02} + W_{30} + W_{30}$, $W_{10} + W_{01} + W_{20} + W_{02}$, and $W_{10} + W_{01}$. The "sum of messages" combinations are chosen carefully in order to guarantee both the recuperation of individual sums and the reduction the interference level at each decoding step. We present lattice based achievable rate regions for the FMABC and PMABC protocols; lattice-based strategies are unlikely to outperform random codes in FTDBC protocols where only a single node transmits at a time.

Related work. Lattice coding strategies were demonstrated to achieve within 1/2 bit from the cut-set outer bound for the two-way relay channel in [9] (and related [1], [10], [13]) for full-duplex nodes in the absence of direct links between the terminal nodes. A lattice-based scheme employing a lattice list decoder was proposed in [12] which extends results to two-way relays channels with direct links. When multiple bi-directional data streams wish to be exchanged between pairs of nodes (rather than between a base-station and multiple terminal nodes as considered here) through the help of a relay, a superposition of lattice and random codes have been exploited in [11] to derive achievable rate regions. A more general channel model is considered in [6], where sequential decoding of pairs of lattice points is used at the relay. In all these schemes, the central idea behind the use of lattices is to exploit their structure to decode a sum of lattice points at the relay rather than having them decode all individual messages as would be needed in random-coding based strategies. This is particularly useful in two-way strategies where this sum may be forwarded by the relay, allowing terminal nodes to exploit their own message side-information to cancel out the other user's message from the sum and obtain the desired message.

II. Preliminaries

A. Notation and definitions

There exist a set of terminal nodes $\mathcal{M} := \mathcal{B} \cup \{0\} = \{0, 1, 2, \cdots, \mathsf{m}\}$ - a base station (node 0) and a subset of terminal nodes $\mathcal{B} := \{1, 2, \cdots, \mathsf{m}\}$ - and a relay r which aids in the communication between the terminal nodes and the base station. We use $R_{i,j}$ to denote the rate of communication of message $W_{i,j}$ from node i to node j, and let $R_{S,T}$ be the vector of rates of messages from set S to set T where $S, T \subseteq \mathcal{M}$. Each end user communicates with the base station bi-directionally and no information is exchanged between end users, i.e. $R_{i,j} = 0$ (or is undefined) for all $i, j \in \mathcal{B}$, $R_{0,i}$ and $R_{i,0}$ may be non-zero for $i \in \mathcal{B}$. At channel use k, we use X_i^k and Y_i^k to denote the input and received signal distributions of node i. Because of the half-duplex constraint, not all nodes transmit/receive during all phases and we use the dummy symbol $\varnothing$ to denote that there is no input or no output at a node during a phase. $\Delta_{i,n}$ is the phase duration of phase i with block size n and Δ_i is the phase duration of phase i when $n \to \infty$. Each node i has channel input and output alphabets $\mathcal{X}_i \in \mathbb{C} \cup \varnothing$ and $\mathcal{Y}_i \in \mathbb{C} \cup \varnothing$, which are related through an AWGN memoryless channel. Lower case letters x_i denote instances of the upper case X_i which lie in the calligraphic alphabets $\mathcal{X}$. Boldface $\mathbf{x}_i$ represents a vector indexed by time at node i. It is convenient to denote by $\mathbf{x}_S := \{\mathbf{x}_i | i \in S\}$, a set of vectors indexed by time. $[x]^+ := \max\{x, 0\}$. A^t is transpose of matrix A. Finally, let $S(j) := \{i | i < j, i \in S\}$.

B. Lattice codes

An n-dimensional lattice Λ is defined as a discrete subgroup in Euclidean space $\mathbb{C}^n$ associated with vector addition, i.e., for any $\lambda_1, \lambda_2 \in \Lambda$, $\lambda_1 + \lambda_2 \in \Lambda$. $\mathcal{V}$ is the Voronoi region of Λ and Vol($\mathcal{V}$) is the volume of $\mathcal{V}$. $\sigma^2(\Lambda)$ is the second moment per dimension of $\mathcal{V}$. The nearest lattice quantization function $Q(\mathbf{x})$ and modulo function $\bmod \Lambda$ are defined as

$$Q(\mathbf{x}) := \min_{\lambda \in \Lambda} ||\mathbf{x} - \lambda|| \quad (1)$$

$$\mathbf{x} \bmod \Lambda := \mathbf{x} - Q(\mathbf{x}), \quad (2)$$

respectively. Suppose that Λ and Λ_C form a lattice partition Λ_C/Λ, i.e. $\Lambda \subseteq \Lambda_C$. The nested lattice code C is defined as the set of coset leaders as $C = \{\Lambda_C \bmod \Lambda\}$.

Construction needed for relay decoding. We define a vector of lattice codewords $\mathbf{c} := [c_1, c_2, \cdots, c_m] \in C^{1\times m}$ and a *combination matrix* $\mathbf{G} \in \{0,1\}^{m\times l}$ which satisfies:

(a) There exists a row of $\mathbf{G}$ in which all elements are 1.
(b) There exists $\mathbf{D} \in \mathbb{Z}^{l\times m}$ such that $\mathbf{GD} = \mathbf{I}$.

Then we also define the corresponding lattice codeword vector $\mathbf{t} = [t_1, t_2, \cdots, t_l] \in C^{1\times l}$, as $\mathbf{t} = \mathbf{cG} \bmod \Lambda$. From the second property of $\mathbf{G}$, $\mathbf{c} = \mathbf{tD} \bmod \Lambda$ as well. Intuitively, we will use this combination matrix $\mathbf{G}$ at the relay to decode different combinations $\mathbf{t}$ of lattice codewords at the relay. As we will see, the first property of $\mathbf{G}$ will ensure that all decoded sums are uniformly distributed in the coarsest lattice, while the second property ensures that the original codewords $\mathbf{c}$ may be recovered from the combinations $\mathbf{t}$.

Lemma 1: A combination matrix $\mathbf{G} \in \{0,1\}^{m\times m}$, that satisfies the two properties (a) and (b), always exists.

Proof: Let $\mathbf{G}$ be a upper triangular matrix in $\{0,1\}^{m\times m}$, and let $\mathbf{D}$ be the matrix in $\mathbb{Z}^{m\times m}$ such that $d_{ii} = 1$, $d_{i(i+1)} = -1$, otherwise $d_{ij} = 0$. Then the first row of $\mathbf{G}$ is $\mathbf{1}$ and $\mathbf{GD} = \mathbf{I}$. ■

C. Protocols for multiple terminal-node pairs

The total transmission time is divided into two time *divisions*, each of which may consist of one or more *phases*. During the first time division - called the *multiple access* division, the terminal nodes transmit to the relay. During the second time division - called the *broadcast* division - the relay transmits to the terminal nodes. In the multiple access division, we consider four cases: 1) all terminal nodes transmit for the whole duration, 2) 0 uses the whole duration and the other terminal nodes $1, \cdots, \mathsf{m}$ transmit sequentially. We denote the first protocol as the *Full Multiple Access Broadcast* (FMABC) protocol, and the second protocol as the *Partial Multiple Access Broadcast* (PMABC) protocol, as in [7].

III. Gaussian noise channel

We consider the AWGN channel, where at channel use k :

$$\mathbf{Y}[k] = \mathbf{H}\mathbf{X}[k] + \mathbf{Z}[k] \quad (3)$$

where $\mathbf{Y}[k]$, $\mathbf{X}[k]$ and $\mathbf{Z}[k]$ are in $(\mathbb{C}^*)^{(m+2)\times 1} = (\mathbb{C} \cup \{\varnothing\})^{(m+2)\times 1}$, and $\mathbf{H} \in \mathbb{C}^{(m+2)\times(m+2)}$. $\mathbf{Y}[k] = [Y_0[k], Y_1[k], \cdots, Y_{\mathsf{m}}[k], Y_{\mathsf{r}}[k]]^t$ and $\mathbf{X}[k]$, $\mathbf{Z}[k]$ and $\mathbf{H}$ are similarly defined. If node i is not in transmission mode at time k, $X_i[k] = \varnothing$, which means that the input symbol does not exist in the above mathematical channel model. $h_{i,j}$ is the effective channel gain between transmitter i and receiver j. For each node i and transmission phase ℓ, we impose average power constraints $\frac{1}{\Delta_{\ell,n}\cdot n} \sum_{k=\Delta_{\overline{\ell-1},n}n+1}^{\Delta_{\bar{\ell},n}n} ||X_i^k||^2 \le P_i$, where $\Delta_{\bar{\ell},n} = \Delta_{1,n} + \cdots + \Delta_{\ell,n}$. We assume that each node is aware of all channel gains, i.e., full CSI. The noise at all receivers is independent, of unit power, additive, white Gaussian, complex and circularly symmetric.

A. Channel inputs for the FMABC protocol

We will use pairwise (base-station to user i) lattice coding, where codebook generation is fully explained in the proofs. We construct channel input X_0 with m independent inputs $X_{01}, \cdots, X_{0m}$ as $X_0^k = X_{01}^k + \cdots X_{0m}^k$, with average power constraint for each X_{0i}^k in phase 1 as $\frac{1}{\Delta_{1,n}\cdot n}\sum_{k=1}^{\Delta_{1,n}n}||X_{0i}^k||^2 \le P_{0i}$, with $P_0 = \sum_{i=1}^m P_{0i}$ and independent X_is.

In the relay broadcasting phase we apply an extension of Marton's broadcast channel region in Theorem 1 of [7] to the Gaussian channel. We use the following relationship between input signals and auxiliary variables of transmitter r, which is similar to Costa's setup in [2]: $\mathbf{U}_\mathsf{r} = \mathbf{\Gamma}\mathbf{V}_\mathsf{r}$, with $X_\mathsf{r} = V_{\mathsf{r}1} + \cdots V_{\mathsf{r}m}$. where $\mathbf{U}_\mathsf{r}$ and $\mathbf{V}_\mathsf{r}$ are vectors of length m and $\mathbf{\Gamma} \in \mathbb{R}^{m\times m}$. Also the $V_{\mathsf{r}i}$ follow the distributions $V_{\mathsf{r}i} \sim \mathcal{CN}(0, P_{\mathsf{r}i})$, where $\sum_{i=1}^m P_{\mathsf{r}i} = P_\mathsf{r}$ and $V_{\mathsf{r}i}$ are independent. Let $U_{\mathsf{r}S} := \{U_{\mathsf{r}i_1}, \cdots U_{\mathsf{r}i_s}\}$, where $S = \{i_1, \cdots, i_s\}$. Then $|U_{\mathsf{r}S}| := \det(\mu(U_{\mathsf{r}S}))$, where $\mu(U_{\mathsf{r}S})$ is the covariance matrix of $U_{\mathsf{r}S}$. Finally, we define:

$$I(U_S; V_T) = \log_2\left(\frac{|U_S||V_T|}{|U_S \cup V_T|}\right) =: C_B(U_S, V_T). \quad (4)$$

B. Channel inputs for the PMABC protocol

In phase ℓ node 0 constructs channel input X_0^k as $X_0^k = X_{01}^k + \cdots + X_{0m}^k$, for $\Delta_{\overline{\ell-1},n}n + 1 \le k \le \Delta_{\bar{\ell},n}n$ with average power constraints $\frac{1}{\Delta_{\ell,n}\cdot n}\sum_{k=\Delta_{\overline{\ell-1},n}n+1}^{\Delta_{\bar{\ell},n}n}||X_{0\ell}^k||^2 \le P_{0\ell}^{(\ell)}$, and $X_{0i} \sim \mathcal{CN}(0, P_{0i}^{(\ell)})$ $(i \ne l)$, where $\sum_{j=1}^m P_{0j}^{(\ell)} = P_0$ and X_{0j} are all independent. Similar to the relay broadcasting, we use Costa's setup to broadcast X_{0i} $(i \ne \ell)$ as $\mathbf{U}_0^{(\ell)} = \mathbf{\Gamma}^{(\ell)}\mathbf{X}_0^{(\ell)}$, where $\mathbf{U}_0^{(\ell)} = (U_{01}, \cdots, U_{0(\ell-1)}, U_{0(\ell+1)}, \cdots, U_{0m})$ and $\mathbf{X}_0^{(\ell)} = (X_{01}, \cdots, X_{0(\ell-1)}, X_{0(\ell+1)}, \cdots, X_{0m})$ are vectors of length $m-1$ and $\mathbf{\Gamma}^{(\ell)} \in \mathbb{R}^{(m-1)\times(m-1)}$. In phase $m+1$, the relay uses the FMABC broadcasting scheme.

IV. Achievable rate regions

We now derive achievable rate regions for the FMABC and PMABC protocols with nested lattice codes.

A. FMABC protocol

We assume that the powers $\{P_{0i}\}$, $\{P_{\mathsf{r}i}\}$ and $\mathbf{\Gamma}$, are all predetermined. We restrict the i^{th} column of $\mathbf{G}$ to be $\mathbf{1}$ if $|h_{i,\mathsf{r}}|^2 P_i$ or $|h_{0,\mathsf{r}}|^2 P_{0i}$ is the maximum over all i.

Theorem 2: An achievable rate region of the half-duplex bi-directional Gaussian relay channel under the FMABC protocol is the convex hull of the set of all points $(R_{0,b}, R_{b,0})$ for all $b \in \mathcal{B}$ satisfying

$$R_{0,i} < \Delta_1 \min_{j\in g(i)}\left\{\left[\log_2\left(\frac{|h_{0,\mathsf{r}}|^2 P_{0i}}{\sum_{k=1}^m g_{jk}(|h_{0,\mathsf{r}}|^2 P_{0k} + |h_{k,\mathsf{r}}|^2 P_k)} + \frac{|h_{0,\mathsf{r}}|^2 P_{0i}}{\sum_{k=1}^m (1-g_{jk})(|h_{0,\mathsf{r}}|^2 P_{0k} + |h_{k,\mathsf{r}}|^2 P_k) + 1}\right)\right]^+\right\} \quad (5)$$

$$R_{i,0} < \Delta_1 \min_{j\in g(i)}\left\{\left[\log_2\left(\frac{|h_{i,\mathsf{r}}|^2 P_i}{\sum_{k=1}^m g_{jk}(|h_{0,\mathsf{r}}|^2 P_{0k} + |h_{k,\mathsf{r}}|^2 P_k)} + \frac{|h_{i,\mathsf{r}}|^2 P_i}{\sum_{k=1}^m (1-g_{jk})(|h_{0,\mathsf{r}}|^2 P_{0k} + |h_{k,\mathsf{r}}|^2 P_k) + 1}\right)\right]^+\right\} \quad (6)$$

$$R_{\{0\},S} < \sum_{i\in S}\Delta_2 C_B(U_{\mathsf{r}i}, Y_i) - \Delta_2 C_B(U_{\mathsf{r}i}, U_{\mathsf{r}S(i)}) \quad (7)$$

$$R_{S,\{0\}} < \Delta_2 C_B(U_{\mathsf{r}S}, Y_0 \cup U_{\mathsf{r}\bar{S}}) \quad (8)$$

where $g(i) := \{j | g_{ij} = 1\}$ for $i \in \mathcal{B}$, $S \subseteq \mathcal{B}$ over all possible combination matrices $\mathbf{G}$. □

Remark 3: The bounds in (5) and (6) result from the multiple access period and are based on nested lattice codes. The bounds in (7) and (8) are the result of the relay broadcast period and are derived from Theorem 4 in [8]. Thus, we focus on the proof of the bounds (5) and (6).

Proof: First, for convenience of analysis we define $p(i)$ as the subscript of P of i^{th} largest value in $\{|h_{1,\mathsf{r}}|^2 P_1, |h_{2,\mathsf{r}}|^2 P_2, \cdots, |h_{m,\mathsf{r}}|^2 P_m, |h_{0,\mathsf{r}}|^2 P_{01}, \cdots, |h_{0,\mathsf{r}}|^2 P_{0m}\}$.

Code generation: Generate $\Delta_{1,n}\cdot n$ dimensional lattice chains $\Lambda_{p(1)} \subseteq \Lambda_{p(2)} \subseteq \cdots \subseteq \Lambda_C$ such that Λ_i and Λ_{0i} $(1 \le i \le m)$ are all both Rogers-good and Poltyrev-good, while Λ_C is Poltyrev-good and for any $\delta > 0$, $|h_{0,\mathsf{r}}|^2 P_{0i} - \delta < \sigma^2(\Lambda_{0i}) < |h_{0,\mathsf{r}}|^2 P_{0i}$ and $|h_{i,\mathsf{r}}|^2 P_i - \delta < \sigma^2(\Lambda_i) < |h_{i,\mathsf{r}}|^2 P_i$ for sufficiently large n. The codeword $\mathcal{C}_{0i}$ for message $W_{0,i}$ is the set of coset leaders of the lattice partition Λ_C/Λ_{0i} as $\mathcal{C}_{0i} = \{\Lambda_C \bmod \Lambda_{0i}\}$. Similarly, we define the codeword $\mathcal{C}_i$ for message $W_{i,0}$. Then each code rate is

$$R_{0,i} = \frac{1}{\Delta_{1,n}\cdot n}\log_2|C_{0,i}| = \frac{1}{\Delta_{1,n}\cdot n}\log_2\frac{|\Lambda_{0,i}|}{|\Lambda_C|} + o_n(1) \quad (9)$$

$$R_{i,0} = \frac{1}{\Delta_{1,n}\cdot n}\log_2|C_i| = \frac{1}{\Delta_{1,n}\cdot n}\log_2\frac{|\Lambda_i|}{|\Lambda_C|} + o_n(1). \quad (10)$$

Such a lattice code chain always exists by Theorem 2 in [9].

Encoding: During phase 1, node 0 encodes $\mathbf{c}_{0i}(W_{0,i}) \in \mathcal{C}_{0i}$ and constructs $\mathbf{x}_{0i} = \frac{1}{h_{0,\mathsf{r}}}((\mathbf{c}_{0i} - \mathbf{u}_{0i}) \bmod \Lambda_{0i})$, where $\mathbf{u}_{0i}$ is a dither known to every node with uniformly distributed in Λ_{0i}. Similarly, node i encodes $\mathbf{c}_i(W_{i,0}) \in \mathcal{C}_i$ and constructs $\mathbf{x}_i = \frac{1}{h_{i,\mathsf{r}}}((\mathbf{c}_i - \mathbf{u}_i) \bmod \Lambda_i)$.

Decoding: The key novelty of our scheme lies in the decoding of various combinations of lattice codeword sums, which are done in a way to ensure that all the individual "sums of codewords" corresponding to $W_{0,i} + W_{i,0}$ may be obtained.

We first define $\tilde{\mathbf{c}}_i^{sum} = \mathbf{c}_{0i} - Q_{0i}(\mathbf{c}_{0i} - \mathbf{u}_{0i}) + \mathbf{c}_i - Q_i(\mathbf{c}_i - \mathbf{u}_i)) \bmod \Lambda_{p(1)}$ $(1 \le i \le m)$, where $Q_{0i}(x)$ and $Q_i(x)$ are the lattice quantizers of Λ_{0i} and Λ_i, respectively. Then we define the length-m vector $\tilde{\mathbf{c}}^{sum} = [\tilde{\mathbf{c}}_1^{sum}, \cdots, \tilde{\mathbf{c}}_m^{sum}]$ and construct the length-l vector $\tilde{\mathbf{t}} = (\tilde{\mathbf{c}}^{sum}\mathbf{G}) \bmod \Lambda_{p(1)}$.

At the end of phase 1, node r decodes $\tilde{\mathbf{t}}_i$ $(1 \le i \le l)$. Node r constructs the $\Delta_{1,n}\cdot n$-length vector

$$\tilde{\mathbf{y}}_{\mathsf{r},i} = (\alpha_i \mathbf{y}_\mathsf{r} + \sum_{j=1}^m g_{ji}(\mathbf{u}_{0j} + \mathbf{u}_j)) \bmod \Lambda_{p(1)} \quad (11)$$

$$= (\tilde{\mathbf{t}}_i + \tilde{\mathbf{z}}_i) \bmod \Lambda_{p(1)}, \quad (12)$$

where $\tilde{\mathbf{z}}_i = -(1-\alpha_i)\left(\sum_{j=1}^m g_{ji}(h_{0,j}\mathbf{x}_{0j} + h_{j,0}\mathbf{x}_j)\right) + \alpha_i\left(\sum_{j=i}^m (1-g_{ji})(h_{0,j}\mathbf{x}_{0j} + h_{j,0}\mathbf{x}_j) + \mathbf{z}_\mathsf{r}\right)$. Since $\mathbf{c}_{p(1)} -$

$Q_{p(1)}(\mathbf{c}_{p(1)} - \mathbf{u}_{p(1)}) \bmod \Lambda_{p(1)} = \mathbf{c}_{p(1)}$ is uniformly distributed in $C_{p(1)}$ and since the corresponding column of $\mathbf{G}$ is $\mathbf{1}$, by the crypto-lemma [4], $\tilde{\mathbf{t}}_i$ is also uniformly distributed in $\Lambda_{p(1)}$ and is independent of $\tilde{\mathbf{z}}_i$ $\forall i$. To minimize the mean square error (MMSE) we take

$$\alpha_i = \frac{\sum_{j=1}^{m} g_{ji}(|h_{0,j}|^2 P_{0j} + |h_{j,0}|^2 P_j)}{\sum_{j=1}^{m}(|h_{0,j}|^2 P_{0j} + |h_{j,0}|^2 P_j) + 1}. \quad (13)$$

Then the relay decodes $\tilde{\mathbf{t}}_i$ as the nearest lattice point in Λ_C, i.e., $\tilde{\mathbf{t}}_i = Q_C(\tilde{\mathbf{y}}_{\mathsf{r},i})$ by using Euclidian lattice decoding [3]. After decoding all $\tilde{\mathbf{t}}_i$s, relay r constructs $\tilde{\mathbf{c}}^{sum} = \tilde{\mathbf{t}}\mathbf{D} \bmod \Lambda_{p(1)}$, and the relay r has, $\forall i \in [1, m]$:

$$\tilde{\mathbf{c}}_i^{sum} = (\mathbf{c}_{0i} - Q_{0i}(\mathbf{c}_{0i} - \mathbf{u}_{0i}) + \mathbf{c}_i - Q_i(\mathbf{c}_i - \mathbf{u}_i)) \bmod \Lambda_{p(1)}.$$

The error analysis is provided on Natasha Devroye's web-site. ■

B. PMABC protocol

We assume that the powers $\{P_{0i}^{(\ell)}\}$'s, $\mathbf{\Gamma}^{(\ell)}$ $(1 \le \ell \le m)$, $\{Pri\}'s$, and $\mathbf{\Gamma}$, are all predetermined.

Theorem 4: An achievable rate region of the half-duplex bi-directional Gaussian relay channel under the PMABC protocol is the convex hull of the set of all points $(R_{0,b}, R_{b,0})$ for all $b \in \mathcal{B}$ satisfying (14) and (15) for $S \subseteq \mathcal{B}$. □

Remark 5: In the multiple access period, node 0 constructs a lattice code to transmit a portion of $W_{0,i}$ during phase i to relay r and uses Marton's broadcasting to transmit other messages through direct links. These two sets of codewords are treated as interference at the appropriate receivers.

Proof: First, node 0 divides each $W_{0,i}$ into m sub-messages as $W_{0,i} = \{W_{0,i|1}, \cdots W_{0,i|m}\}$.

During phase i $(1 \le i \le m)$: Node 0 and i generate $\Delta_{i,n} \cdot n$-length lattice codes $\mathbf{x}_{0i}^{(i)}(w_{0,i|i})$ and $\mathbf{x}_i^{(i)}(w_{i,0})$ by the same code construction of the FMABC protocol, respectively. Additionally node 0 randomly generates $\mathbf{x}_{0j}^{(i)}(w_{0,j|i})$ with power constraint P_{0j} for $j \neq i$. Such a code chain always exists by Theorem 2 in [9]. This is the same code construction as in [9] except for additional interference terms such that $\mathbf{x}_{0j}^{(i)}$ $(j \neq i)$ when the relay decodes the lattice sum. Therefore, the bounds below are sufficient for the reliable decoding at relay:

$$R_{0,i|i} < \Delta_i \left[\log_2\left(\frac{|h_{0,\mathsf{r}}|^2 P_{0i}^{(i)}}{|h_{0,\mathsf{r}}|^2 P_{0i}^{(i)} + |h_{i,\mathsf{r}}|^2 P_i} + \frac{|h_{0,\mathsf{r}}|^2 P_{0i}^{(i)}}{\sum_{j\neq i}(|h_{0,\mathsf{r}}|^2 P_{0j}^{(i)}) + 1}\right)\right]^+ \quad (16)$$

$$R_{i,0} < \Delta_i \left[\log_2\left(\frac{|h_{i,\mathsf{r}}|^2 P_i}{|h_{0,\mathsf{r}}|^2 P_{0i}^{(i)} + |h_{i,\mathsf{r}}|^2 P_i} + \frac{|h_{i,\mathsf{r}}|^2 P_i}{\sum_{j\neq i}(|h_{0,\mathsf{r}}|^2 P_{0j}^{(i)}) + 1}\right)\right]^+ \quad (17)$$

To broadcast $W_{0,j|i}$ $(j \neq i)$ node 0 generates random codes and use Marton's broadcasting. Thus we have,

$$\sum_{j\in S} R_{0,j|i} < \sum_{j \in S} \Delta_i C_B(U_{0j}^{(i)}, Y_j) - \Delta_i C_B(U_{0j}^{(i)}, U_{0S(j)}^{(i)}) \quad (18)$$

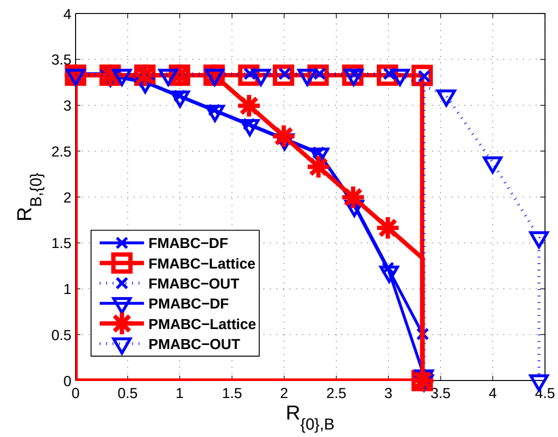

Fig. 2. Comparison of protocol and coding gains with $P_0 = P_1 = P_2 = P_\mathsf{r} = 20$ dB.

for any $S \subseteq \mathcal{B} \setminus \{i\}$.

During phase $m+1$: The relay r broadcasts the lattice sums as in Theorem 4 in [8]. Thus we have

$$\sum_{i\in S} R_{0,i|i} < \sum_{i \in S} \Delta_{m+1} C_B(U_{\mathsf{r}i}, Y_i) - \Delta_{m+1} C_B(U_{\mathsf{r}i}, U_{\mathsf{r}S(i)}) \quad (19)$$

$$R_{S,\{0\}} < \Delta_{m+1} C_B(U_{\mathsf{r}S}, Y_0 \cup U_{\mathsf{r}\bar{S}}) \quad (20)$$

Combining (16) – (20) and [8, Lemma 23] yields (14)–(15). ■

V. Performance Comparison and Constant Gap

A. Numerical analysis with $m = 2$

In this section, we numerically evaluate the rate regions obtained in the previous section for the case of two terminal nodes $m = 2$, where $Z_i \sim \mathcal{CN}(0, 1)$, $\forall i \in \{0, 1, 2, \mathsf{r}\}$. We use the following channel coefficients for the comparison, and note that the relative performance of the different schemes may change under different channel conditions.

$$\mathbf{H} = \begin{bmatrix} 0 & 0.3 & 0.05 & 1 \\ 0.3 & 0 & 1.5 & 1 \\ 0.05 & 1.5 & 0 & 0.2 \\ 1 & 1 & 0.2 & 0 \end{bmatrix} \quad (21)$$

We first compare the achievable rate regions and outer bounds of the FMABC and PMABC protocols. For achievability, we use the DF scheme in [8] and the Lattice code scheme in the Section IV. Cut-set outer bounds are taken directly from [8]. We project the regions in $(R_{\{0\},\mathcal{B}}, R_{\mathcal{B},\{0\}})$ 2-dimensional space. We then proceed to examine the maximal sum-rate $R_{0,1} + R_{0,2} + R_{1,0} + R_{2,0}$.

Fig. 2 illustrates the achievable regions and outer bounds with transmit powers $P_0 = P_1 = P_2 = P_\mathsf{r} = 20$ dB. In the FMABC protocol the lattice forwarding outperforms DF and closely approaches the outer bounds. This is because 1) the sum-rate bound on $R_{0,1} + R_{0,2} + R_{1,0} + R_{2,0}$ does not exist in the lattice region, and 2) the penalty (loss of 1+ term inside $\log$) in the lattice forwarding becomes negligible as the

$$R_{\{0\},S} < \sum_{j=1}^{m} \sum_{i\in S\setminus\{j\}} \Delta_j C_B(U_{0i}^{(j)}, Y_i) - \Delta_j C_B(U_{0i}^{(j)}, U_{0S(i)}^{(j)}) +$$

$$\min\left\{ \sum_{i\in S} \Delta_i \left[\log_2\left(\frac{|h_{0,\mathrm{r}}|^2 P_{0i}^{(i)}}{|h_{0,\mathrm{r}}|^2 P_{0i}^{(i)} + |h_{i,\mathrm{r}}|^2 P_i} + \frac{|h_{0,\mathrm{r}}|^2 P_{0i}^{(i)}}{\sum_{j\neq i}(|h_{0,\mathrm{r}}|^2 P_{0j}^{(i)}) + 1}\right)\right]^+ , \sum_{i\in S} \Delta_{m+1} C_B(U_{\mathrm{r}i}, Y_i) - \Delta_{m+1} C_B(U_{\mathrm{r}i}, U_{\mathrm{r}S(i)}) \right\} \quad (14)$$

$$R_{S,\{0\}} < \min\left\{ \sum_{i\in S} \Delta_i \left[\log_2\left(\frac{|h_{i,\mathrm{r}}|^2 P_i}{|h_{0,\mathrm{r}}|^2 P_{0i}^{(i)} + |h_{i,\mathrm{r}}|^2 P_i} + \frac{|h_{i,\mathrm{r}}|^2 P_i}{\sum_{j\neq i}(|h_{0,\mathrm{r}}|^2 P_{0j}^{(i)}) + 1}\right)\right]^+ , \Delta_{m+1} C_B(U_{\mathrm{r}S}, Y_0 \cup U_{\mathrm{r}\bar{S}}) \right\} \quad (15)$$

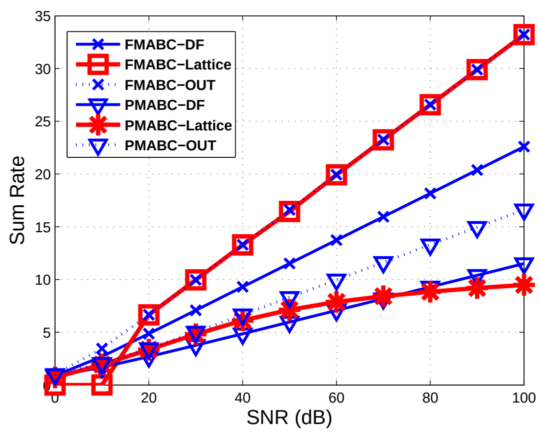

Fig. 3. Sum rate comparison.

SNR increases. The bounds for lattice coding are of the form $\log_2(A+x)$, $(0 < A < 1)$. Since A is less than 1, this results in a gap from the cut-set bound. However, as SNR increases, $\log_2(A+x) \approx \log_2(x)$, i.e., the lattice bound approaches the cut-set. In contrast to the FMABC case, the PMABC lattice scheme achieves a region similar to that of the DF scheme in [8]. Using lattice codes, the sum-rate constraint corresponding to decoding at the relay is removed. However, since the direct link side information is exploited less efficiently than in the DF scheme, there is a trade off between the performance of the random coding and lattice-coding based DF schemes.

In Fig. 3 the sum-rate $(R_{0,1}+R_{0,2}+R_{1,0}+R_{2,0})$ is plotted. The FMABC lattice scheme approaches the outer bound when $SNR > 20$. In the low SNR regime, the rates are reduced due to the lattice penalty, i.e, $[\log_2(A+x)]^+ = 0$. In the PMABC protocol, lattice forwarding is better than DF when $SNR < 70$, otherwise DF is better since the channel inputs for $W_{0,j}$ $(j \neq i)$ during the i^{th} phase are treated as interference at the relay using the lattice scheme, while these are exploited to decode the individual messages in the random coding scheme.

B. *Constant gap of the FMABC protocol*

If $|h_{1,\mathrm{r}}| = |h_{2,\mathrm{r}}| = \cdots = |h_{m,\mathrm{r}}|$ and $P_1 = P_2 = \cdots = P_m$, then we achieve within a constant gap from the cut-set outer bound in the multiple access division/phase of the FMABC protocol. For convenience of analysis, we define $R_{S,T}^{in}$ as the maximum rate of the lattice code achievable rate region from set S to set T and $R_{S,T}^{out}$ as the minimum sum-rate of the cut-set outer bound from set S to set T.

Theorem 6: In the multiple access division of the FMABC protocol, if $|h_{1,\mathrm{r}}| = |h_{2,\mathrm{r}}| = \cdots = |h_{m,\mathrm{r}}|$ and $P_1 = P_2 = \cdots = P_m$, $\forall S \subseteq \mathcal{B}$:

$$R_{\{0\},S}^{out} - R_{\{0\},S}^{in} \leq |S|, \quad R_{S,\{0\}}^{out} - R_{S,\{0\}}^{in} \leq |S|. \quad (22)$$

The proof is provided on Natasha Devroye's web-site.

Corollary 7: The achieved sum-rate of Thm 2 is within $2m$ bits from the cut-set outer bound in the FMABC protocol.

Proof: In the relay broadcast period (phase 2), the link $\mathrm{r} \rightarrow 1, 2, \cdots, m$ is a degraded broadcast channel and the capacity region was obtained in [5]. The opposite direction $\mathrm{r} \rightarrow 0$ is a simple point-to-point channel. Thus, there is no gap in the second phase. The gap in the multiple access division/phase is thus $2m$ bits, by Theorem 6. ■

References

[1] A. Avestimehr, A. Sezgin, and D. Tse, "Capacity of the two-way relay channel within a constant gap," *European Transactions in Telecommunications*, 2009.

[2] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. IT-29, no. 3, pp. 439–441, 1983.

[3] U. Erez and R. Zamir, "Achieving $\frac{1}{2}$log(1+snr) on the awgn channel with lattice encoding and decoding," *IEEE Trans. Inform. Theory*, vol. IT-48, no. 6, pp. 1250–1276, 2004.

[4] 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 Annual Allerton Conference*, Oct. 2003.

[5] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Peredac. Inform.*, vol. 10, no. 3, pp. 3–14, 1974.

[6] D. Gunduz, A. Yener, A. Goldsmith, and H. Poor, "The multi-way relay channel," in *Proc. IEEE Int. Symp. Inform. Theory*, Seoul, Jul. 2009.

[7] S. Kim, B. Smida, and N. Devroye, "Capacity bounds on multi-pair two-way communication with a base-station aided by a relay," in *Proc. IEEE Int. Symp. Inform. Theory*, Austin, Jun. 2010.

[8] ——, "Achievable rate regions and outer bounds for a multi-pair bi-directional relay network," 2011, in preparation, available upon request.

[9] W. Nam, S. Y. Chung, and Y. Lee, "Capacity of the Gaussian Two-Way Relay Channel to Within 1/2 Bit," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

[10] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *Arxiv preprint arXiv:0908.2119*, 2009.

[11] A. Sezgin, A. Avestimehr, M. Khajehnejad, and B. Hassibi, "Divide-and-conquer: Approaching the capacity of the two-pair bidirectional Gaussian relay network," *Arxiv preprint arXiv:1001.4271*, 2010.

[12] Y. Song and N. Devroye, "List decoding for nested lattices and applications to relay channels ," in *Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing, Monticello*, 2010.

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

APPENDIX

Error analysis of Theorem 2: $\forall i \in \mathcal{B}$, $j \in g(i)$, let

$$\bar{R}_{0,i}^{(j)} = \left[\log_2\left(\frac{|h_{0,\mathrm{r}}|^2 P_{0i}}{\sum_{k=1}^m g_{kj}(|h_{0,\mathrm{r}}|^2 P_{0k} + |h_{k,\mathrm{r}}|^2 P_k)} + \frac{|h_{0,\mathrm{r}}|^2 P_{0i}}{\sum_{k=1}^m (1-g_{kj})(|h_{0,\mathrm{r}}|^2 P_{0k} + |h_{k,\mathrm{r}}|^2 P_k)+1}\right)\right]^+ \tag{23}$$

$$\bar{R}_{i,0}^{(j)} = \left[\log_2\left(\frac{|h_{i,\mathrm{r}}|^2 P_{i}}{\sum_{k=1}^m g_{kj}(|h_{0,\mathrm{r}}|^2 P_{0k} + |h_{k,\mathrm{r}}|^2 P_k)} + \frac{|h_{i,\mathrm{r}}|^2 P_{i}}{\sum_{k=1}^m (1-g_{kj})(|h_{0,\mathrm{r}}|^2 P_{0k} + |h_{k,\mathrm{r}}|^2 P_k)+1}\right)\right]^+ \tag{24}$$

The decoding error probability is given as:

$$P_e = \Pr\{\cup_{i=1}^l \tilde{\mathbf{t}}_i \neq \mathbf{t}_i\} \tag{25}$$

$$\leq \sum_{i=1}^l \Pr\{\tilde{\mathbf{t}}_i \neq \mathbf{t}_i\} \tag{26}$$

$$= \sum_{i=1}^l \Pr\{\tilde{\mathbf{z}}_i \bmod \Lambda_{p(1)} \notin \mathcal{V}_C\} \tag{27}$$

By direct extension of Theorem 3 in [9],

$$\Pr\{\tilde{\mathbf{z}}_i \bmod \Lambda_{p(1)} \notin \mathcal{V}_C\} \leq e^{-\Delta_{1,n} n (E_P(2^{R_i^{\min}}) - o_n(1))} \tag{28}$$

where

$$R_i^{\min} = \frac{1}{\Delta_{1,n}} \min_{j \in g_l(i)} \{\Delta_{1,n}\bar{R}_{0,j}^{(i)} - R_{0,j}, \Delta_{1,n}\bar{R}_{j,0}^{(i)} - R_{j,0}\} \tag{29}$$

where $g_l(j) := \{j | g_{ji} = 1\}$ and $E_P(\cdot)$ is the Poltyrev exponent given by

$$E_P(x) = \begin{cases} \frac{1}{2}(x - 1 - \ln x), & 1 < x \leq 2 \\ \frac{1}{2}(\ln \frac{x}{4} + 1), & 2 \leq x \leq 4 \\ \frac{x}{8}, & 4 \leq x. \end{cases} \tag{30}$$

Thus, if $R_i^{\min}$ is greater than 0, i.e., $R_{0,j} < \bar{R}_{0,j}^{(i)}$ and $R_{j,0} < \bar{R}_{j,0}^{(i)}$ $\forall i \in [1, l]$ and $j \in g_l(i)$, $P_e \to 0$ as n goes to infinity, which are in (5) and (6). At the end of phase 1, the relay has $\tilde{\mathbf{c}}_1^{sum}, \cdots \tilde{\mathbf{c}}_m^{sum}$. By one-to-one mapping the relay assign $\tilde{\mathbf{c}}_1^{sum}, \cdots \tilde{\mathbf{c}}_m^{sum}$ to messages $w_1, \cdots, w_m$ and sends them to terminal nodes during phase 2 using the same broadcasting scheme as in [8].

Proof of Theorem 6: For convenience of analysis we note that $|h_{1,\mathrm{r}}|^2 = |h_{2,\mathrm{r}}|^2 = \cdots = |h_{m,\mathrm{r}}|^2 = h_l$, $|h_{0,\mathrm{r}}|^2 = h_d$, $P_1 = P_2 = \cdots = P_m = P$. For a given $S \subseteq \mathcal{B}$ we take $P_{0i} = P_0/|S|, \forall i \in S$ and $P_{0j} = 0, \forall j \notin S$ to maximize $R_{\{0\},S}$. Similarly, node $j \notin S$ sends a known signal which meets the power constraint to make $R_{j,0} = 0$. The relay can then subtract $\mathbf{x}_j$ $(j \notin S)$ before lattice decoding and the corresponding combination matrix $\mathbf{G}_S$ is the upper triangular matrix in $\{0,1\}^{m \times m}$ by setting all zeros in the j^{th} columns and j^{th} rows for $j \notin S$. From Theorem 17 in [8] the cut-set outer bounds are

$$R_{\{0\},S} < \Delta_1 \log_2(1 + h_d P_0) =: R_{\{0\},S}^{out} \tag{31}$$

$$R_{S,\{0\}} < \Delta_1 \log_2(1 + |S| h_l P) =: R_{S,\{0\}}^{out}. \tag{32}$$

We define $p(i) = |\{j | j \leq i, j \in S\}|$ for $i \in S$. For $i \in S, p(i) \neq |S|$ the achievable sum-rate, from (5) and (6) is

$$R_{0,i}^{in} := \max\{R_{0,i}\} \tag{33}$$

$$= \Delta_1 \min_{p(i) \leq k \leq |S|} \left\{ \left[\log_2\left(\frac{\frac{h_d}{|S|}P_0}{k\left(\frac{h_d}{|S|}P_0 + h_l P\right)} + \frac{\frac{h_d}{|S|}P_0}{(|S|-k)\left(\frac{h_d}{|S|}P_0 + h_l P\right)+1}\right)\right]^+ \right\} \tag{34}$$

$$= \Delta_1 \left[\log_2\left(\frac{\frac{h_d}{|S|}P_0}{p(i)\left(\frac{h_d}{|S|}P_0 + h_l P\right)} + \frac{\frac{h_d}{|S|}P_0}{(|S|-p(i))\left(\frac{h_d}{|S|}P_0 + h_l P\right)+1}\right)\right]^+ \tag{35}$$

$$\leq \Delta_1 \log_2\left(1 + \frac{\frac{h_d}{|S|}P_0}{(|S|-p(i))\left(\frac{h_d}{|S|}P_0 + h_l P\right)+1}\right) \tag{36}$$

$$\leq \Delta_1 \log_2\left(1 + \frac{\frac{h_d}{|S|}P_0}{(|S|-p(i))\frac{h_d}{|S|}P_0 + 1}\right) \tag{37}$$

$$\leq \Delta_1 \tag{38}$$

and for $i \in S, p(i) = |S|$

$$R_{0,i}^{in} := \max\{R_{0,i}\} \tag{39}$$

$$= \Delta_1 \left[\log_2\left(\frac{\frac{h_d}{|S|}P_0}{h_d P_0 + |S| h_l P} + \frac{h_d P_0}{|S|}\right)\right]^+ \tag{40}$$

$$\geq \Delta_1 \left[\log_2\left(\frac{h_d P_0}{|S|}\right)\right]^+ \tag{41}$$

$$\geq \Delta_1 \left(\log_2\left(1 + \frac{h_d P_0}{|S|}\right) - 1\right). \tag{42}$$

Therefore,

$$R_{\{0\},S}^{in} \geq \Delta_1 \sum_{k=1}^{|S|} \left(\log_2\left(1 + \frac{\frac{h_d}{|S|}P_0}{(|S|-k)\frac{h_d}{|S|}P_0 + 1}\right) - 1\right) \tag{43}$$

$$\geq \Delta_1 \log_2(1 + h_d P_0) - |S|\Delta_1 \tag{44}$$

$$= R_{\{0\},S}^{out} - |S|\Delta_1. \tag{45}$$

$$\geq R_{\{0\},S}^{out} - |S| \tag{46}$$

Similarly,

$$
\begin{aligned}
R^{in}_{S,\{0\}} &:= \max\{R_{S,\{0\}}\} && (47)\\
&\geq \Delta_1 \log_2(1 + |S| h_l P) - |S|\Delta_1 && (48)\\
&= R^{out}_{S,\{0\}} - |S|\Delta_1 && (49)\\
&\geq R^{out}_{S,\{0\}} - |S|. && (50)
\end{aligned}
$$

From (46) and (50), we obtain (22).