Amplify-and-Modulo for Gaussian Two-Way Relay Channel

Silas L. Fong
Department of Electronic Engineering
City University of Hong Kong
Kowloon, Hong Kong
Email: lhfong5@ie.cuhk.edu.hk

Li Ping
Department of Electronic Engineering
City University of Hong Kong
Kowloon, Hong Kong
Email: eeliping@cityu.edu.hk

Chi Wan Sung
Department of Electronic Engineering
City University of Hong Kong
Kowloon, Hong Kong
Email: albert.sung@cityu.edu.hk

***Abstract*—We consider a two-way relay channel (TWRC) in which two terminals exchange messages with the help of a relay between them. The two terminals transmit messages to the relay through the Multiple Access Channel (MAC) and the relay transmits messages to the two terminals through the Broadcast Channel (BC). We assume that the MAC and the BC do not interfere with each other, and each terminal receives signals only from the relay but not the other terminal. All the nodes are assumed to be full-duplex, which means that they can transmit and receive information at the same time. A transmission scheme for the Gaussian TWRC is said to be *analog-relaying* if the relay does not need any codebook for encoding. The simplest analog-relaying scheme is amplify-and-forward (AF), under which the relay amplifies the received codeword and forwards the resultant codeword to the two terminals. In this paper, we propose a new analog-relaying scheme called *amplify-and-modulo* (AM) based on lattice operations. AM is a slight modification of AF. Under AM, the relay first amplifies the received codeword followed by reducing the power of the amplified codeword using the modulo-lattice operation, and then forwards the resultant codeword to the two terminals. After receiving the codeword transmitted by the relay, each terminal subtracts its own information before decoding. We prove an achievable rate region for AM, and obtain a necessary and sufficient condition under which AM outperforms AF. In addition, we show by graph that AM can achieve a strictly higher equal-rate than AF and another existing analog-relaying scheme together under some scenario.**

I. Introduction

We consider a two-way relay channel (TWRC) [1], in which two terminals exchange messages with the help of a relay between them. The two terminals transmit messages to the relay through the Multiple Access Channel (MAC) and the relay transmits messages to the two terminals through the Broadcast Channel (BC). We assume that the MAC and the BC do not interfere with each other, and each terminal receives signals only from the relay but not the other terminal. All the nodes are assumed to be full-duplex, which means that they can transmit and receive information at the same time.

For the TWRC described above, although several outer bounds on the capacity region have been proved in [2]–[4] and several achievable rate regions have been obtained in [5]–[11], the capacity region of the TWRC is unknown. *Amplify-and-forward* (AF) for the Gaussian TWRC has been studied in [5]. Under AF, the relay first amplifies the received codeword and then forwards the amplified codeword to the two terminals. One drawback of AF is that the noise received at the relay is also amplified. *Decode-and-forward-by-binning* (DFB) for a general TWRC was first proposed in [6]. Under DFB, the relay decodes the messages transmitted by the terminals before encoding the messages by random binning. It is shown in [6] that DFB is optimal among the transmission schemes under which the relay decodes the messages transmitted by the terminals. *Noisy network coding* (NNC) has been proposed for a general multi-terminal network in [11], which shows that NNC can achieve the capacity region of the Gaussian TWRC within 1 bit. *Nested lattice code* (NLC) was proposed for the Gaussian TWRC in [8] to achieve the capacity region within 1/2 bit. Although NLC is almost optimal when the signal-to-noise ratio (SNR) of the MAC is high, it is outperformed by DFB when the SNR of the MAC is low. *Modulo-and-forward* (MF), which performs better than AF and DFB under some scenario, has been proposed for the Gaussian TWRC in [10].

A transmission scheme for the Gaussian TWRC is said to be *analog-relaying* if the relay does not need any codebook for encoding. Note that MF is not analog-relaying unless the two terminals use the same power to transmit [10]. To facilitate understanding, MF is also called *analog-MF* if the two terminals use the same power to transmit. Among AF, DFB, NNC and NLC, only AF is analog-relaying, which means the encoding complexity at the relay for AF is lower than that for the other schemes. In this paper, we propose a new analog-relaying scheme called *amplify-and-modulo* (AM), which is based on the modulo-lattice operation for analog signals [12,13]. AM is a slight modification of AF. Under AM, the relay first amplifies the received codeword followed by reducing the power of the amplified codeword using the modulo-lattice operation, and then forwards the resultant codeword to the two terminals. After receiving the codeword transmitted by the relay, each terminal subtracts its own information before decoding.

This paper is organized as follows. Section II presents the notation of this paper. Sections III and IV review some well-known results on weak typicality and lattice properties respectively. Section V proposes AM, which is based on the modulo-lattice operation, for the Gaussian TWRC. An achievable rate region for AM is evaluated using weak typicality and some lattice properties. In Section VI, we compare AM with AF and obtain a necessary and sufficient condition under which AM

978-1-4673-2569-1/12/$31.00 ©2012 IEEE

outperforms AF. In addition, we show by graph that AM can achieve a strictly higher equal-rate than AF and analog-MF, the other two analog-relaying schemes, as well as DFB together under some scenario. Section VII concludes this paper.

II. Notation

We use $Pr\{E\}$ to represent the probability of an event E. We use a capital letter X to denote a random variable with alphabet $\mathcal{X}$, and use the small letter x to denote the realization of X. We use $E[X]$ to represent the expectation of a random variable X. We use X^n to denote a random column vector $[X_1\ X_2\ \ldots\ X_n]^T$, where the components X_k have the same alphabet $\mathcal{X}$. We let $p_X(x)$ and $p_{X^n}(x^n)$ denote the probability mass functions of the discrete random variables X and X^n respectively. We let $f_X(x)$ and $f_{X^n}(x^n)$ denote the probability density functions of the random variables X and X^n respectively. We let $f_{Y|X}(y|x)$ denote the probability density function of Y conditioned on the event $\{X = x\}$ for any continuous random variable Y and any general random variable X. For simplicity, we drop the subscript of a notation if there is no ambiguity. We let $\mathcal{N}(0, N)$ denote a Gaussian random variable whose mean and variance are 0 and N respectively. We let $\mathcal{N}(0, N)^n$ denote n independent copies of $\mathcal{N}(0, N)$. To facilitate discussion, we let $(a)^+$ and $\mathbb{R}^2_+$ denote $\max\{0, a\}$ and the set of all pairs of non-negative real numbers respectively.

III. Weak Typicality

Let $(X_1, X_2, \ldots, X_k)$ denote a finite collection of continuous random variables with some fixed joint distribution $f(x_1, x_2, \ldots, x_k)$. Let S denote an ordered subset of these random variables and consider n independent copies of S. Then, $f_{S^n}(s^n) = \prod_{i=1}^n f_{S_i}(s_i)$ for all $s^n \in \mathcal{S}^n$. We use weak typicality [2,14] to prove our results in this paper, and the results of weak typicality are summarized in the following definition and lemma.

Definition 1 (Typical sequence [2,14]):
The set $A_\epsilon^{(n)}(X_1, X_2, \ldots, X_k)$ of ϵ-typical n-sequence $(x_1^n, x_2^n, \ldots, x_k^n)$ is

$$\left\{ \begin{array}{l} (x_1^n, x_2^n, \ldots, x_k^n) \in \\ \mathcal{X}_1^n \times \mathcal{X}_2^n \times \ldots \times \mathcal{X}_k^n \end{array} \middle| \begin{array}{l} \left| -\frac{1}{n} \log_2 f(s^n) - h(S) \right| \le \epsilon \\ \text{for all } S \subseteq \{X_1, X_2, \ldots, X_k\} \end{array} \right\}.$$

Lemma 1 (Typical set [2,14]): Fix $\epsilon > 0$.

(i) For all $S \subseteq \{X_1, X_2, \ldots, X_k\}$, $\int_{s^n \in A_\epsilon^{(n)}(S)} f_{S^n}(s^n) > 1 - \epsilon$ for sufficiently large n.

(ii) For any $S_1, S_2 \subseteq \{X_1, X_2, \ldots, X_k\}$ with the probability density function $f_{S_1,S_2}(s_1, s_2)$, let $\tilde{S}_1, \tilde{S}_2 \subseteq \{X_1, X_2, \ldots, X_k\}$ such that $f_{\tilde{S}_1^n, \tilde{S}_2^n}(s_1^n, s_2^n) = \prod_{i=1}^n f_{S_1}(s_{1,i}) f_{S_2}(s_{2,i})$. Then,

$$\int_{(s_1^n, s_2^n) \in A_\epsilon^{(n)}(S_1,S_2)} f_{\tilde{S}_1^n, \tilde{S}_2^n}(s_1^n, s_2^n) \le 2^{-n(I(S_1;S_2)-3\epsilon)}$$

for sufficiently large n.

IV. Lattice Properties

An n-dimensional lattice Λ is a set which can be written as $\Lambda = \{\lambda \in \mathbb{R}^n : \lambda = G\boldsymbol{x} \text{ for some } \boldsymbol{x} \in \mathbb{Z}^n\}$ for some $n \times n$ real-valued non-singular matrix G. In other words, an n-dimensional lattice Λ is an n-dimensional discrete subgroup of the Euclidean space $\mathbb{R}^n$ with the ordinary vector addition operation. For each $\boldsymbol{x} \in \mathbb{R}^n$, let $\boldsymbol{x} + \Lambda$ denote the coset of Λ in $\mathbb{R}^n$ that contains $\boldsymbol{x}$. For every $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$, let $||\boldsymbol{x}|| = \sqrt{(\sum_{i=1}^n x_i^2)}$ be the Euclidean norm of $\boldsymbol{x}$. A fundamental Voronoi region of $\Lambda \subset \mathbb{R}^n$, denoted by $\mathcal{V}(\Lambda)$, is a set of coset representatives of the cosets of Λ such that each coset representative $\boldsymbol{x} \in \mathcal{V}(\Lambda)$ is an element with minimum Euclidean norm in the coset $\boldsymbol{x} + \Lambda$. Note that there are more than one fundamental Voronoi regions because there always exists some coset of Λ in $\mathbb{R}^n$ which has two elements with minimum Euclidean norm. In this paper, we let $\mathcal{V}(\Lambda)$ denote an arbitrary fundamental Voronoi region of Λ and let $\text{Vol}(\Lambda)$ denote the volume of $\mathcal{V}(\Lambda)$. Every $\boldsymbol{x} \in \mathbb{R}^n$ can be uniquely written as $\boldsymbol{x} = \lambda + \boldsymbol{r}$ for some $\lambda \in \Lambda$ and some $\boldsymbol{r} \in \mathcal{V}(\Lambda)$. We let $\boldsymbol{x}$ (mod Λ) denote $\boldsymbol{r}$. For a comprehensive introduction to lattices, we refer the reader to [15].

We let $\mathcal{B}(r) = \{\boldsymbol{x} \in \mathbb{R}^n : ||\boldsymbol{x}|| \le r\}$ be the closed ball centered at the origin in $\mathbb{R}^n$ with radius r, and let $V_{\mathcal{B}}(r)$ denote the volume of $\mathcal{B}(r)$. Let r_Λ^{effec} denote the "effective radius" of $\mathcal{V}(\Lambda)$ that satisfies $V_{\mathcal{B}}(r_\Lambda^{\text{effec}}) = \text{Vol}(\Lambda)$. For any sets A and B in $\mathbb{R}^n$, let $A+B$ denote $\{x+y : x \in A, y \in B\}$. Given a lattice Λ, the set $\Lambda + \mathcal{B}(r)$ is a *covering of* $\mathbb{R}^n$ if $\mathbb{R}^n \subseteq \Lambda + \mathcal{B}(r)$, i.e., each point in $\mathbb{R}^n$ is covered by at least one sphere centered at a lattice point with radius r. Define the covering radius of the lattice r_Λ^{cov} by $r_\Lambda^{\text{cov}} = \min\{r : \Lambda + \mathcal{B}(r) \text{ is a covering}\}$. For any lattice Λ, $\mathcal{V}(\Lambda) \subsetneq \mathcal{B}(r_\Lambda^{\text{cov}})$ and $r_\Lambda^{\text{effec}} < r_\Lambda^{\text{cov}}$ by the definitions of $\mathcal{V}(\Lambda)$, r_Λ^{cov} and r_Λ^{effec}.

Definition 2: A sequence of lattices $\{\Lambda_n\}_{n=1,2,\ldots}$, where each Λ_n is an n-dimensional lattice, is said to be *good for covering* if $\lim_{n\to\infty} \frac{r_{\Lambda_n}^{\text{cov}}}{r_{\Lambda_n}^{\text{effec}}} = 1$.

Definition 3: Let $Z^n = \mathcal{N}(0, N)^n$. A sequence of lattices $\{\Lambda_n\}_{n=1,2,\ldots}$, where each Λ_n is an n-dimensional lattice, is said to be *good for AWGN channel coding with noise variance* N if for any $\epsilon > 0$, $r_{\Lambda_n}^{\text{effec}} < \sqrt{n(N+\epsilon)}$ and $Pr\{Z^n \notin \mathcal{V}(\Lambda_n)\} < \epsilon$ for sufficiently large n.

V. Gaussian Two-Way Relay Channel

The TWRC consists of three nodes – two terminals indexed by 1 and 2, and a relay r between them. Node 1 and node 2 do not communicate directly, but communicate through node r using two different channels. Node 1 and node 2 choose messages W_1 and W_2 independently and exchange the messages through node r in n time slots as follows. For each $i \in \{1, 2\}$, node i transmits $X_{i,k}$ through the multiple access channel (MAC) in the k^{th} time slot. The received symbol in the same time slot at r is $Y_{r,k} = X_{1,k} + X_{2,k} + Z_{r,k}$ where $Z_{r,k}$ is a Gaussian random variable independent of $(X_{1,k}, X_{2,k})$. In addition, r transmits $X_{r,k}$ through the broadcast channel (BC) in the k^{th} time slot. The received symbol in the same time slot at node i is $Y_{i,k} = X_{r,k} + Z_{i,k}$ for $i = 1, 2$, where $Z_{i,k}$ is a

Gaussian random variable independent of $X_{r,k}$. After n time slots, node 1 declares $\hat{W}_2$ to be the transmitted W_2 based on Y_1^n and W_1, and node 2 declares $\hat{W}_1$ to be the transmitted W_1 based on Y_2^n and W_2. For each $i \in \{1,2,r\}$, any codeword $[x_{i,1}\ x_{i,2}\ \ldots\ x_{i,n}]^T$ that is transmitted over the channel should satisfy $\frac{1}{n}\sum_{k=1}^{n} x_{i,k}^2 \leq P_i$, where P_i represents the maximum power constraint for node i.

Definition 4: The Gaussian TWRC consists of $Z_i = \mathcal{N}(0, N_i)$ for each $i \in \{1,2,r\}$, a probability density function $f_1(y_r|x_1,x_2)$ representing the MAC and a probability density function $f_2(y_1,y_2|x_r)$ representing the BC such that for all real numbers x_1, x_2, x_r, y_1, y_2 and y_r,

$$\begin{aligned} f_1(y_r|x_1,x_2) &= f_{Z_r|X_1,X_2}(y_r - x_1 - x_2|x_1,x_2) \\ &= f_{Z_r}(y_r - x_1 - x_2) \end{aligned}$$

and

$$\begin{aligned} f_2(y_1,y_2|x_r) &= f_{Z_1,Z_2|X_r}(y_1 - x_r, y_2 - x_r|x_r) \\ &= f_{Z_1,Z_2}(y_1 - x_r, y_2 - x_r). \end{aligned}$$

In addition, we require that for any two finite discrete random variables X_1 and X_2 with an input distribution $p(x_1,x_2)$ for the MAC and any finite discrete random variable X_r with an input distribution $p(x_r)$ for the BC, the relationship among X_1, X_2, X_r, the output Y_r of the MAC and the outputs Y_1 and Y_2 of the BC satisfy

$$Pr\{X_1 \leq x_1, X_2 \leq x_2, X_r \leq x_r, Y_1 \leq y_1, Y_2 \leq y_2, Y_r \leq y_r\} = \int_{-\infty}^{y_r} f_1(v_r|x_1,x_2)p(x_1,x_2)dv_r \int_{-\infty}^{y_1}\int_{-\infty}^{y_2} f_2(v_1,v_2|x_r)p(x_r)dv_2dv_1$$

for all real numbers x_1, x_2, x_r, y_1, y_2 and y_r. The Gaussian TWRC is denoted by (N_1,N_2,N_r).

Definition 5: An (n,M_1,M_2)-code on the channel (N_1,N_2,N_r) subject to maximum power constraints P_1, P_2 and P_r consists of the following:

1) A message set $\mathcal{W}_i = \{1,2,\ldots,M_i\}$ at node i for each $i \in \{1,2,r\}$, where $\mathcal{W}_r = \emptyset$.
2) An encoding function $\alpha_{i,k} : \mathcal{W}_i \times \mathcal{Y}_i^{k-1} \to \mathcal{X}_i$ at node i for each $i \in \{1,2,r\}$ and each $k \in \{1,2,\ldots,n\}$, where $\alpha_{i,k}$ is the encoding function in the k^{th} time slot such that $X_{i,k} = \alpha_{i,k}(W_i, Y_i^{k-1})$. In addition, every codeword x_i^n must satisfy the power constraint $\sum_{k=1}^{n} x_{i,k}^2 \leq nP_i$.
3) A decoding function $g_i : \mathcal{W}_i \times \mathcal{Y}_i^n \to \mathcal{W}_j$ at node i for each $(i,j) \in \{(1,2),(2,1)\}$ such that $g_i(W_i, Y_i^n) = \hat{W}_j$.

Definition 6: For an (n,M_1,M_2)-code on the Gaussian TWRC (N_1,N_2,N_r) subject to maximum power constraints P_1, P_2 and P_r, the average probabilities of decoding error of W_1 and W_2 are defined as $P_{e,1}^n = Pr\{g_2(W_2,Y_2^n) \neq W_1\}$ and $P_{e,2}^n = Pr\{g_1(W_1,Y_1^n) \neq W_2\}$ respectively.

Definition 7: A rate pair (R_1,R_2) is achievable if there exists a sequence of (n,M_1,M_2)-codes for the Gaussian TWRC (N_1,N_2,N_r) subject to maximum power constraints P_1, P_2 and P_r such that $\lim_{n\to\infty} \frac{\log_2 M_1}{n} \geq R_1$, $\lim_{n\to\infty} \frac{\log_2 M_2}{n} \geq R_2$, $\lim_{n\to\infty} P_{e,1}^n = 0$ and $\lim_{n\to\infty} P_{e,2}^n = 0$.

The following lemma is due to [16].

Lemma 2: For each $N > 0$, there exists a sequence of lattices $\{\Lambda_n\}_{n=1,2,\ldots}$ which is good for both covering and AWGN channel coding with noise variance N.

Lemma 3: Let $Z = \mathcal{N}(0,N)$. Let Z^n denote n independent copies of Z. For any $\epsilon > 0$, there exists for sufficiently large n an n-dimensional lattice Λ with $\mathcal{V}(\Lambda) \subsetneq \mathcal{B}(\sqrt{n(N+2\epsilon)})$ such that $Pr\{Z^n \notin \mathcal{V}(\Lambda)\} < \epsilon$.

Proof: For each $N > 0$, there exists by Lemma 2 a sequence of lattices $\{\Lambda_n\}_{n=1,2,\ldots}$ which is good for both covering and AWGN channel coding with noise variance N. Fix an $\epsilon > 0$. Then, it follows from Definitions 2 and 3 that for sufficiently large n, $r_{\Lambda_n}^{\text{cov}} < (\sqrt{1+\epsilon/(N+\epsilon)})r_{\Lambda_n}^{\text{effec}}$ and $r_{\Lambda_n}^{\text{effec}} < \sqrt{n(N+\epsilon)}$, which imply that $r_{\Lambda_n}^{\text{cov}} < \sqrt{n(N+2\epsilon)}$. In addition, it follows from Definition 3 that $Pr\{Z^n \notin \mathcal{V}(\Lambda_n)\} < \epsilon$ for sufficiently large n. ■

Let $\mathcal{R}_{\text{AM}}$ denote the set

$$\left\{ (R_1,R_2) \in \mathbb{R}_+^2 \;\middle|\; \begin{array}{l} R_1 \leq \frac{1}{2}\log_2\Big(1+\frac{P_1(P_r-\max\{N_1,N_2\})^+}{N_r(P_r-\max\{N_1,N_2\})^+ + N_2(\max\{P_1,P_2\}+N_r)}\Big), \\ R_2 \leq \frac{1}{2}\log_2\Big(1+\frac{P_2(P_r-\max\{N_1,N_2\})^+}{N_r(P_r-\max\{N_1,N_2\})^+ + N_1(\max\{P_1,P_2\}+N_r)}\Big) \end{array} \right\}. \quad (1)$$

Lemma 4: Let (R_1,R_2) be a point in the interior of $\mathcal{R}_{\text{AM}}$. For sufficiently small $\delta > 0$, there exists for sufficiently large n a $(2n, \lceil 2^{nR_1}\rceil, \lceil 2^{nR_2}\rceil)$-code for the Gaussian TWRC (N_1,N_2,N_r) subject to maximum power constraints P_1, P_2 and P_r such that $P_{e,1}^{2n} < 2(\beta^2+2)\delta$ and $P_{e,2}^{2n} < 2(\beta^2+2)\delta$, where

$$\beta = \sqrt{\frac{(P_r - \max\{N_1,N_2\})^+}{\max\{P_1,P_2\}+N_r}}. \quad (2)$$

Proof: Since the theorem is trivial when $P_r - \max\{N_1,N_2\} \leq 0$ (cf. (1)), we assume that $P_r - \max\{N_1,N_2\} > 0$. Then, β in (2) is a positive real number. Since (R_1,R_2) is in the interior of $\mathcal{R}_{\text{AM}}$, it follows that for sufficiently small $\delta > 0$,

$$\begin{aligned} R_1 < \frac{1}{2}\log_2(1+(P_1-\delta)(P_r-\max\{N_1,N_2\})/(N_r(P_r- \\ \max\{N_1,N_2\}) + N_2(\max\{P_1,P_2\}+N_r)) - 3\delta. \end{aligned} \quad (3)$$

By Lemma 3, there exists for sufficiently large n an n-dimensional lattice Λ with

$$\mathcal{V}(\Lambda) \subsetneq \mathcal{B}(\sqrt{nP_r}) \quad (4)$$

such that

$$Pr\{\mathcal{N}(0, P_r - \beta^2\delta)^n \notin \mathcal{V}(\Lambda)\} < \beta^2\delta/2. \quad (5)$$

Consider the following $(2n, \lceil 2^{nR_1}\rceil, \lceil 2^{nR_2}\rceil)$-code. The first n time slots are allocated to the MAC and the last n time slots are allocated to the BC. Therefore, only node 1 and node 2 transmit during the first n time slots and only r transmits

during the last n time slots. To facilitate discussion, let Y_i^n denote the output at node i for each $i \in \{1, 2, r\}$.

Codebook generation for W_1 and W_2: Let

$$X_i = \mathcal{N}(0, P_i - \delta) \tag{6}$$

for each $i \in \{1, 2\}$. For each $i \in \{1, 2\}$, generate $\lceil 2^{nR_i} \rceil$ independent codewords $x_i^n(j)$, $j = 1, 2, \ldots, \lceil 2^{nR_i} \rceil$, of length n, generating each codeword i.i.d.$\sim X_i^n$. The codebooks are revealed to both node 1 and node 2. If node 1 and node 2 choose W_1 and W_2 respectively, they transmit $x_1^n(W_1)$ and $x_2^n(W_2)$ respectively. An encoding error is declared if $||x_1^n(W_1)||^2 > nP_1$ or $||x_2^n(W_2)||^2 > nP_2$.

Amplify-and-modulo for Y_r^n: After receiving Y_r^n during the first n time slots, r transmits

$$X_r^n = (\beta Y_r^n) \text{ (mod } \Lambda) \tag{7}$$

during the last n time slots. An encoding error is declared if $||X_r^n||^2 > nP_r$.

Decoding of W_1: Fix the codebooks for W_1 and W_2. Suppose node 2 transmits $x_2^n(j)$. After receiving Y_2^n during the last n time slots, node 2 constructs

$$\hat{Y}_2^n = (Y_2^n - \beta x_2^n(j)) \text{ (mod } \Lambda). \tag{8}$$

Then, node 2 computes the set

$$\left\{ x_1^n(e) \in \mathcal{X}_1^n \,\middle|\, (x_1^n(e), \hat{Y}_2^n) \in A_\delta^{(n)}(X_1, \beta X_1 + \beta Z_r + Z_2) \right\},$$

denoted by $V_2(j)$. If $V_2(j)$ contains exactly one element $x_1^n(e)$, node 2 declares that $\hat{W}_1 = e$. Otherwise, node 2 declares a decoding error.

Decoding of W_2: It is symmetric to the decoding of W_1.

Average values of $P_{e,1}^{2n}$ and $P_{e,2}^{2n}$ over the random codebooks ignoring any encoding error: We first calculate $E[P_{e,1}^{2n}] = Pr\{\hat{W}_1 \neq W_1\}$ and $E[P_{e,2}^{2n}] = Pr\{\hat{W}_2 \neq W_2\}$ ignoring any encoding error, where each expectation is taken with respect to the random codebooks for W_1 and W_2. The calculation of $E[P_{e,1}^{2n}]$ is as follows. Since $E[P_{e,1}^{2n}]$ does not depend on a particular choice of (W_1, W_2) due to the symmetry of random codebooks for W_1 and W_2, we assume $(W_1, W_2) = (1, 1)$ has been sent without loss of generality. Let $\mathcal{E}_1 = \{V_2(1) \text{ does not contain } X_1^n(1)\}$ and $\mathcal{E}_2 = \{V_2(1) \text{ contains an element other than } X_1^n(1)\}$ be two events such that $\{\hat{W}_1 \neq W_1\} = \mathcal{E}_1 \cup \mathcal{E}_2$ and

$$E[P_{e,1}^{2n}] = Pr\{\hat{W}_1 \neq W_1\} \leq Pr\{\mathcal{E}_1\} + Pr\{\mathcal{E}_2\}. \tag{9}$$

Consider the following chain of equalities:

$$\begin{aligned}
\hat{Y}_2^n &\overset{\text{(a)}}{=} (Y_2^n - \beta X_2^n(1)) \text{ (mod } \Lambda) \\
&\overset{\text{(b)}}{=} (X_r^n + Z_2^n - \beta X_2^n(1)) \text{ (mod } \Lambda) \\
&\overset{\text{(c)}}{=} (\beta Y_r^n \text{ (mod } \Lambda) + Z_2^n - \beta X_2^n(1)) \text{ (mod } \Lambda) \\
&\overset{\text{(d)}}{=} (\beta X_1^n(1) + \beta Z_r^n + Z_2^n) \text{ (mod } \Lambda)
\end{aligned} \tag{10}$$

where

(a) follows from (8).
(b) follows from Definition 4 that $Y_2^n = X_r^n + Z_2^n$.
(c) follows from (7).
(d) follows from Definition 4 that $Y_r^n = X_1^n(1) + X_2^n(1) + Z_r^n$.

Since

$$\beta X_1^n(1) + \beta Z_r^n + Z_2^n = \mathcal{N}(0, \beta^2 P_1 + \beta^2 N_r + N_2 - \beta^2 \delta)^n$$

by construction and

$$\beta^2 P_1 + \beta^2 N_r + N_2 - \beta^2 \delta \leq P_r - \beta^2 \delta$$

by (2), it then follows from (5) that

$$Pr\left\{ \begin{matrix} \beta X_1^n(1) + \beta Z_r^n + Z_2^n \\ = (\beta X_1^n(1) + \beta Z_r^n + Z_2^n) \text{ (mod } \Lambda) \end{matrix} \right\} > 1 - \beta^2 \delta / 2,$$

which implies from (10) that

$$Pr\{\hat{Y}_2^n = \beta X_1^n(1) + \beta Z_r^n + Z_2^n\} > 1 - \beta^2 \delta. \tag{11}$$

By Lemma 1(i), for sufficiently large n,

$$Pr\left\{ \begin{matrix} (X_1^n(1), \beta X_1^n(1) + \beta Z_r^n + Z_2^n) \\ \in A_\delta^{(n)}(X_1, \beta X_1 + \beta Z_r + Z_2) \end{matrix} \right\} > 1 - \delta. \tag{12}$$

Let $\hat{\mathcal{E}}$ denote the event $\{\hat{Y}_2^n = \beta X_1^n(1) + \beta Z_r^n + Z_2^n\}$. Then,

$$\begin{aligned}
&Pr\{\mathcal{E}_1\} \\
&= Pr\{\mathcal{E}_1 \cap \hat{\mathcal{E}}\} + Pr\{\mathcal{E}_1 \cap \hat{\mathcal{E}}^c\} \\
&\overset{\text{(a)}}{<} Pr\{\{X_1^n(1) \notin V_2(1)\} \cap \hat{\mathcal{E}}\} + \beta^2 \delta \\
&= Pr\{\{(X_1^n(1), \hat{Y}_2^n) \notin A_\delta^{(n)}(X_1, \beta X_1 + \beta Z_r + Z_2)\} \cap \hat{\mathcal{E}}\} + \beta^2 \delta \\
&\leq Pr\left\{ \begin{matrix} (X_1^n(1), \beta X_1^n(1) + \beta Z_r^n + Z_2^n) \\ \notin A_\delta^{(n)}(X_1, \beta X_1 + \beta Z_r + Z_2/\beta) \end{matrix} \right\} + \beta^2 \delta \\
&\overset{\text{(b)}}{<} (\beta^2 + 1)\delta,
\end{aligned} \tag{13}$$

where (a) follows from (11) and (b) follows from (12). For each $k \in \{2, 3, \ldots, \lceil 2^{nR_1} \rceil\}$, let $\mathcal{E}_{2,k}$ denote the event $\{X_1^n(k) \in V_2(1)\} \cap \hat{\mathcal{E}}$. Then,

$$\begin{aligned}
Pr\{\mathcal{E}_2\} &= Pr\{\mathcal{E}_2 \cap \hat{\mathcal{E}}\} + Pr\{\mathcal{E}_2 \cap \hat{\mathcal{E}}^c\} \\
&\overset{\text{(a)}}{<} Pr\{\mathcal{E}_2 \cap \hat{\mathcal{E}}\} + \beta^2 \delta \\
&= Pr\left\{ \bigcup_{k=2}^{\lceil 2^{nR_1} \rceil} \mathcal{E}_{2,k} \right\} + \beta^2 \delta \\
&\leq \sum_{k=2}^{\lceil 2^{nR_1} \rceil} Pr\{\mathcal{E}_{2,k}\} + \beta^2 \delta,
\end{aligned} \tag{14}$$

where (a) follows from (11). Since

$$\begin{aligned}
\mathcal{E}_{2,k} &= \{\{(X_1^n(k), \hat{Y}_2^n) \in A_\delta^{(n)}(X_1, \beta X_1 + \beta Z_r + Z_2)\} \cap \hat{\mathcal{E}}\} \\
&\subseteq \left\{ \begin{matrix} (X_1^n(k), \beta X_1^n(1) + \beta Z_r^n + Z_2^n) \\ \in A_\delta^{(n)}(X_1, \beta X_1 + \beta Z_r + Z_2) \end{matrix} \right\}
\end{aligned}$$

and

$$\begin{aligned}
&f_{X_1^n(k), \beta X_1^n(1) + \beta Z_r^n + Z_2^n}(x^n, y^n) \\
&\quad = \prod_{i=1}^{n} f_{X_1}(x_i) f_{\beta X_1 + \beta Z_r + Z_2}(y_i),
\end{aligned}$$

it follows from Lemma 1(ii) that

$$Pr\{\mathcal{E}_{2,k}\} \le 2^{-n(I(X_1;\beta X_1+\beta Z_r+Z_2)-3\delta)}. \quad (15)$$

Following from (14), we obtain

$$\begin{aligned} Pr\{\mathcal{E}_2\} &\le \sum_{k=2}^{\lceil 2^{nR_1}\rceil} Pr\{\mathcal{E}_{2,k}\} + \beta^2\delta \\ &\overset{(a)}{\le} \sum_{k=2}^{\lceil 2^{nR_1}\rceil} (2^{-n(I(X_1;\beta X_1+\beta Z_r+Z_2)-3\delta)} + \beta^2\delta \\ &< 2^{nR_1}2^{-n(I(X_1;\beta X_1+\beta Z_r+Z_2)-3\delta)} + \beta^2\delta \quad (16) \end{aligned}$$

where (a) follows from (15). In addition,

$$\begin{aligned} &I(X_1;\beta X_1+\beta Z_r+Z_2) \\ &\ge h(\beta X_1+\beta Z_r+Z_2) - h(\beta Z_r+Z_2) \\ &\overset{(a)}{=} \frac{1}{2}\log_2\left(1+\frac{\beta^2(P_1-\delta)}{\beta^2 N_r+N_2}\right) \\ &\overset{(b)}{=} \frac{1}{2}\log_2(1+(P_1-\delta)(P_r-\max\{N_1,N_2\})/((P_r- \\ &\quad \max\{N_1,N_2\})N_r+(\max\{P_1,P_2\}+N_r)N_2), \quad (17) \end{aligned}$$

where

(a) follows from (6) and the fact that the differential entropy of $\mathcal{N}(0,N)$ is $\log_2(\sqrt{2\pi eN})$.
(b) follows from (2).

It then follows from (3) and (17) that

$$R_1 < I(X_1;\beta X_1+\beta Z_r+Z_2) - 3\delta,$$

which implies from (16) that $Pr\{\mathcal{E}_2\} < (\beta^2+1)\delta$ for sufficiently large n, which then implies from (9) and (13) that

$$E[P_{e,1}^{2n}] = Pr\{\hat{W}_1 \ne W_1\} < 2(\beta^2+1)\delta \quad (18)$$

for sufficiently large n. By symmetry,

$$E[P_{e,2}^{2n}] = Pr\{\hat{W}_2 \ne W_2\} < 2(\beta^2+1)\delta.$$

Bounding encoding errors: An encoding error occurs if any one of the following events occurs: $\{||X_1^n(W_1)||^2 > nP_1\}$, $\{||X_2^n(W_2)||^2 > nP_2\}$ and $\{||X_r^n||^2 > nP_r\}$, denoted by $\mathcal{E}_{P_1}$, $\mathcal{E}_{P_2}$ and $\mathcal{E}_{P_r}$ respectively. By the weak law of large numbers, it follows from (6) that for each $i \in \{1,2\}$,

$$Pr\{||X_i^n||^2 \le nP_i\} \ge 1-\delta \quad (19)$$

for sufficiently large n. In addition,

$$\begin{aligned} Pr\{||X_r^n||^2 \le nP_r\} &\overset{(a)}{=} Pr\{||\beta Y_r^n \pmod\Lambda||^2 \le nP_r\} \\ &\overset{(b)}{=} 1, \quad (20) \end{aligned}$$

where (a) follows from (7) and (b) follows from (4).

Existence of a deterministic code without encoding error and with arbitrarily small probabilities of decoding error of W_1 and W_2: By the union bound, it follows from (18), (19) and (20) that

$$Pr\{\{\hat{W}_1 \ne W_1\} \cup \mathcal{E}_{P_1} \cup \mathcal{E}_{P_2} \cup \mathcal{E}_{P_r}\} < 2(\beta^2+2)\delta$$

for sufficiently large n. By symmetry,

$$Pr\{\{\hat{W}_2 \ne W_2\} \cup \mathcal{E}_{P_1} \cup \mathcal{E}_{P_2} \cup \mathcal{E}_{P_r}\} < 2(\beta^2+2)\delta$$

for sufficiently large n. Consequently, there exists for sufficiently large n some deterministic codebooks for W_1 and W_2, denoted by $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively, such that the probabilities of the events

$$\{\hat{W}_1 \ne W_1\} \cup \mathcal{E}_{P_1} \cup \mathcal{E}_{P_2} \cup \mathcal{E}_{P_r}$$

and

$$\{\hat{W}_2 \ne W_2\} \cup \mathcal{E}_{P_1} \cup \mathcal{E}_{P_2} \cup \mathcal{E}_{P_r}$$

are less than $2(\beta^2+2)\delta$. Then, we transform the deterministic codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ into some codebooks $\mathcal{C}_1'$ and $\mathcal{C}_2'$ without encoding error by altering only the codewords which cause encoding errors. The codebooks $\mathcal{C}_1'$ and $\mathcal{C}_2'$ are the desired codebooks without encoding error where $P_{e,1}^{2n} < 2(\beta^2+2)\delta$ and $P_{e,2}^{2n} < 2(\beta^2+2)\delta$. ■

Theorem 1: Any rate pair in $\mathcal{R}_{\text{AM}}$ is achievable.

Proof: For any rate pair (R_1,R_2) in the interior of $\mathcal{R}_{\text{AM}}$, the $(2n,\lceil 2^{nR_1}\rceil,\lceil 2^{nR_2}\rceil)$-code constructed in Lemma 4 can be interleaved in such a way as performed in [17] that for the resultant code, the rate is larger than $(R_1-1/n, R_2-1/n)$ and the probabilities of decoding error of the messages are less than $1/n$. Consequently, it follows from Definition 7 that any rate pair in the interior of $\mathcal{R}_{\text{AM}}$ is achievable. The theorem then follows from Definition 7 that the limit of any convergent sequence of achievable rate pairs is also achievable. ■

VI. Comparison of AF, AM and analog-MF

We compare various analog-relaying schemes in this section. In addition to AF, AM and analog-MF, the recently proposed *noisy analog network coding* (NANC) [18] is analog-relaying. However, no analytical achievable rate region for NANC is provided in [18]. Therefore, we only compare AF, AM and analog-MF.

The rate region achievable by AF [5,10], denoted by $\mathcal{R}_{\text{AF}}$, is

$$\left\{(R_1,R_2) \in \mathbb{R}_+^2 \;\middle|\; \begin{array}{l} R_1 \le \frac{1}{2}\log_2\left(1+\frac{P_1P_r}{P_rN_r+P_1N_2+P_2N_2+N_rN_2}\right), \\ R_2 \le \frac{1}{2}\log_2\left(1+\frac{P_2P_r}{P_rN_r+P_1N_1+P_2N_1+N_rN_1}\right) \end{array}\right\}. \quad (21)$$

The following theorem obtains a necessary and sufficient condition under which AM outperforms AF.

Theorem 2: $\mathcal{R}_{\text{AM}} \supsetneq \mathcal{R}_{\text{AF}}$ if and only if

$$P_r > \frac{\max\{N_1,N_2\}(P_1+P_2+N_r)}{\min\{P_1,P_2\}}. \quad (22)$$

Proof: The theorem follows from the fact that for $P_r - \max\{N_1,N_2\} \ge 0$,

$$\frac{P_1(P_r-\max\{N_1,N_2\})}{N_r(P_r-\max\{N_1,N_2\})+N_2(\max\{P_1,P_2\}+N_r)} > \frac{P_1P_r}{P_rN_r+P_1N_2+P_2N_2+N_rN_2}$$

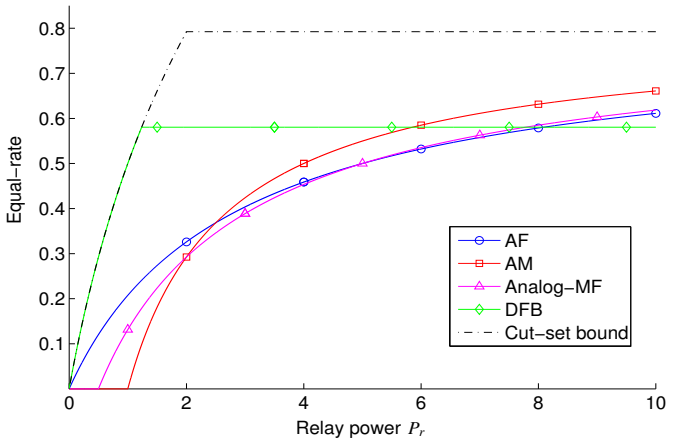

Fig. 1. Equal-rates achievable by AF, AM, analog-MF and DFB.

and

$$\frac{P_2(P_r-\max\{N_1,N_2\})}{N_r(P_r-\max\{N_1,N_2\})+N_1(\max\{P_1,P_2\}+N_r)} > \frac{P_2P_r}{P_rN_r+P_1N_1+P_2N_1+N_rN_1}$$

if and only if (22) holds (cf. (21) and (1)). ■

The main difference between AF and AM is the encoding at the relay. The relay under AF simply broadcasts an amplified signal, while the relay under AM truncates an amplified signal (cf. (7)) into a weaker signal that is then broadcast by the relay. The amplifying factor at the relay for AF is $\sqrt{\frac{P_r}{P_1+P_2+N_r}}$ [5,10], denoted by α_{AF}, while the amplifying factor at the relay for AM is $\sqrt{\frac{(P_r-\max\{N_1,N_2\})^+}{\max\{P_1,P_2\}+N_r}}$ (cf. (2)), denoted by α_{AM}. Note that $\alpha_{AM} > \alpha_{AF}$ if and only if (22) holds, which then implies from Theorem 2 that AM outperforms AF if and only if $\alpha_{AM} > \alpha_{AF}$. Under AM, the modulo-lattice operation at the relay allows α_{AM} to be strictly greater than α_{AF} (and hence a strictly better rate for each direction) without violating the relay power constraint (cf. (7) and (20)), which explains why AM can outperform AF under some scenario.

Consider the three analog-relaying schemes AF, AM and analog-MF when $N_1 = N_2 = N_r = 1$ and $P_1 = P_2 = 2$. Then, the equal-rates achievable by AF, AM and analog-MF are $\frac{1}{2}\log_2(1+\frac{2P_r}{P_r+5})$, $\frac{1}{2}\log_2(1+\frac{2P_r-2}{P_r+2})$ and $\frac{1}{2}\log_2(1+\frac{2P_r-1}{P_r+4})$ [10] respectively, and the cut-set outer bound [3] for the equal-rate is $\min\{\frac{1}{2}\log_2(3), \frac{1}{2}\log_2(1+P_r)\}$. Figure 1 displays the equal-rates achievable by AF, AM and analog-MF, and the cut-set outer bound for $0 \le P_r \le 10$. In order to show that AM can outperform some non-analog-relaying scheme, we also display in Figure 1 the equal-rate achievable by DFB, which is $\min\{\frac{1}{4}\log_2(5), \frac{1}{2}\log_2(1+P_r)\}$ [6]. Figure 1 shows that AM achieves a strictly higher equal-rate than AF, analog-MF and DFB together when $6 \le P_r \le 10$.

VII. Conclusion

We propose amplify-and-modulo (AM), a new analog-relaying scheme for the Gaussian TWRC based on the modulo-lattice operation. Under AM, the relay first amplifies the received codeword followed by reducing the power of the amplified codeword using the modulo-lattice operation, and then forwards the resultant codeword to the two terminals. After receiving the codeword transmitted by the relay, each terminal subtracts its own transmitted codeword followed by the modulo-lattice operation before decoding. We prove an achievable rate region for AM, and obtain a necessary and sufficient condition under which AM outperforms AF. In addition, we show by graph that AM can achieve a strictly higher equal-rate than AF and analog-MF, the other two analog-relaying schemes, as well as DFB together under some scenario.

Acknowledgment

The work of the authors was partially supported by a grant from the University Grants Committee (Project No. AoE/E-02/08) of the Hong Kong Special Administrative Region, China.

References

[1] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE ISIT'06*, Jul. 2006, pp. 1668–1672.
[2] T. M. Cover, *Elements of Information Theory*, 2nd ed. Wiley, 2006.
[3] S. L. Fong and R. W. Yeung, "Capacity bounds for full-duplex two-way relay channel with feedback," in *Proc. IEEE Information Theory and Applications Workshop (ITA), UCSD, San Diego, CA*, Feb. 2011.
[4] ——, "Feedback enlarges capacity region of two-way relay channel," in *Proc. IEEE ISIT'11*, Jul. 2011.
[5] R. Knopp, "Two-way radio networks with a star topology," in *Proc. International Zurich Seminar on Communications (IZS)*, 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, pp. 454–458, Jan. 2008.
[7] W. Nam, S.-Y. Chung and Y. H. Lee, "Capacity bounds for two-way relay channels," in *Proc. IEEE International Zurich Seminar on Communications*, Mar. 2008, pp. 144–147.
[8] ——, "Capacity of the Gaussian two-way relay channel to within $\frac{1}{2}$ Bit," *IEEE Trans. Inf. Theory*, vol. 56, pp. 5488–5494, Nov. 2010.
[9] A. S. Avestimehr, A. Sezgin and D. N. C. Tse, "Approximate capacity of the two-way relay channel: A deterministic approach," in *Proc. Allerton Conference on Communication, Control and Computing*, Sep. 2008, pp. 1582–1589.
[10] I.-J. Baik and S.-Y. Chung, "Network coding for two-way relay channels using lattices," in *Proc. IEEE International Conference on Communications (ICC)*, May 2008, pp. 3898–3902.
[11] S. H. Lim, Y.-H. Kim, A. E. Gamal and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, pp. 3132–3152, May 2011.
[12] Y. Kochman and R. Zamir, "Joint WynerZiv/dirty-paper coding by modulo-lattice modulation," *IEEE Trans. Inf. Theory*, vol. 55, pp. 4878–4889, Nov. 2009.
[13] M. P. Wilson, K. Narayanan and G. Caire, "Joint source channel coding with side information using hybrid digital analog codes ," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4922–4940, Oct. 2010.
[14] R. W. Yeung, *Information Theory and Network Coding*. Springer, 2008.
[15] G. D. Forney Jr., "Coset codes–Part I: Introduction and geometrical classification," *IEEE Trans. Inf. Theory*, vol. 34, pp. 1123–1151, Sep. 1988.
[16] 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.
[17] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, pp. 572–584, Sep. 1979.
[18] M. N. Khormuji and M. Skoglund, "Noisy analog network coding for the two-way relay channel," in *Proc. IEEE ISIT'11*, Jul. 2011, pp. 2065–2069.