Nested Lattice Codes for Gaussian Relay Networks with Interference

Wooseok Nam, *Student Member, IEEE,* Sae-Young Chung, *Senior Member, IEEE,* and Yong H. Lee, *Senior Member, IEEE*
School of EECS, KAIST,
Daejeon, Republic of Korea
E-mail: wsnam@stein.kaist.ac.kr, sychung@ee.kaist.ac.kr, yohlee@ee.kaist.ac.kr

Abstract

In this paper, a class of relay networks is considered. We assume that, at a node, outgoing channels to its neighbors are orthogonal, while incoming signals from neighbors can interfere with each other. We are interested in the multicast capacity of these networks. As a subclass, we first focus on Gaussian relay networks with interference and find an achievable rate using a lattice coding scheme. It is shown that there is a constant gap between our achievable rate and the information theoretic cut-set bound. This is similar to the recent result by Avestimehr, Diggavi, and Tse, who showed such an approximate characterization of the capacity of general Gaussian relay networks. However, our achievability uses a structured code instead of a random one. Using the same idea used in the Gaussian case, we also consider linear finite-field symmetric networks with interference and characterize the capacity using a linear coding scheme.

Index Terms

Wireless networks, multicast capacity, lattice codes, structured codes, multiple-access networks, relay networks

I. Introduction

Characterizing the capacity of general relay networks has been of great interest for many years. In this paper, we confine our interest to the capacity of single source multicast relay networks, which is still an open problem. For instance, the capacity of single relay channels is still unknown except for some special cases [1]. However, if we confine the class of networks further, there are several cases in which the capacity is characterized.

Recently, in [2], the multicast capacity of wireline networks was characterized. The capacity is given by the max-flow min-cut bound, and the key ingredient to achieve the bound is a new coding technique called network coding. Starting from this seminal work, many efforts have been made to incorporate wireless effects in the network model, such as broadcast, interference, and noise. In [3], the broadcast nature was incorporated into the network model by requiring each relay node to send the same signal on all outgoing channels, and the unicast capacity was determined. However, the model assumed that the network is deterministic (noiseless) and has no interference in reception at each node. In [4], the work was extended to multicast capacity. In [5], the interference nature was also incorporated, and an achievable multicast rate was computed. This achievable rate has a cut-set-like representation and meets the information theoretic cut-set bound [27] in some special cases. To incorporate the noise, erasure networks with broadcast or interference only were considered in [7], [8]. However, the network models in [7], [8] assumed that the side information on the location of all erasures in the network is provided to destination nodes. Noisy networks without side information at destination nodes were considered in [12] and [13] for finite-field additive noise and erasure cases, respectively.

This work was supported 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]

Along the same lines of the previous work on wireless networks mentioned above, we consider the multicast problem in a special class of networks called relay networks with interference. More specifically, we assume that all outgoing channels at each node are orthogonal, e.g., using frequency or time division multiplexing, but signals incoming from multiple neighbor nodes to a node can interfere with each other. Since wireless networks are often interference limited, our setup focuses on the more important aspect of them. This model covers those networks considered in [8], [9], [10], [12]. Our interest in the relay networks with interference was inspired by [14], in which the capacity of single relay channels with interference was established. In this paper, we focus on two special subclasses of general networks with interference; Gaussian relay networks with interference and linear finite-field symmetric networks with interference.

For the Gaussian relay networks with interference, we propose a scheme based on nested lattice codes [19] which are formed from a lattice chain and compute an achievable multicast rate. The basic idea of using lattice codes is to exploit the structural gain of *computation coding* [11], which corresponds to a kind of combined channel and network coding. Previously, lattices were used in Gaussian networks in [10], and an achievability was shown. However, our network model differs from the one in [10] in that we assume general unequal power constraints for all incoming signals at each node, while an equal power constraint was mainly considered in [10]. In addition, our lattice scheme is different from that in [10] in that we use lattices to produce nested lattice codes, while lattices were used as a source code in [10].

We also show that our achievable rate is within a constant number of bits from the information theoretic cut-set bound of the network. This constant depends only on the network topology and not on other parameters, e.g., transmit powers and noise variances. This is similar to the recent result in [6], which showed an approximate capacity characterization for general Gaussian relay networks using a random coding scheme. However, our achievability uses a structured code instead of a random one. Thus, our scheme has a practical interest because structured codes may reduce the complexity of encoding and decoding.

Finally, we introduce a model of linear finite-field symmetric networks with interference, which generalizes those in [12], [13]. In the finite-field case, we use a linear coding scheme, which corresponds to the finite-field counterpart of the lattice coding scheme. The techniques for deriving an achievable rate for the finite-field network are basically the same as those for the Gaussian case. However, in this case, the achievable rate always meets the information theoretic cut-set bound, and, thus, the capacity is fully established.

This paper is organized as follows. Section II defines notations and parameters used in this paper and introduces the network model and the problem of interest. In Section III, we analyze Gaussian relay networks with interference and give the upper and lower bounds for the multicast capacity. In Section IV, we define a model of linear finite-field symmetric networks with interference and present the multicast capacity. Section V concludes the paper.

II. Relay networks with interference

A. System model and notations

We begin with a description of the class of networks that will be dealt with in this paper. The memoryless relay networks with interference are characterized such that all outgoing channels from a node to its neighbors are orthogonal to each other. We still assume that incoming signals at a node can interfere with each other through a memoryless multiple-access channel (MAC). An example of this class of networks is shown in Fig. 1. Some special cases and subclasses of these networks have been studied in many previous works [8], [9], [10], [13], [14].

We will begin by giving a detailed description of the network and some definitions of the parameters. The network is represented by a directed graph $\mathcal{G} = (V, E)$, where $V = \{1, \ldots, |V|\}$ is a vertex set and $E \subseteq V \times V$ is an edge set. Each vertex and edge correspond to a communication node and a channel in the network, respectively. In this paper, we focus on a multicast network: vertex 1 represents the source

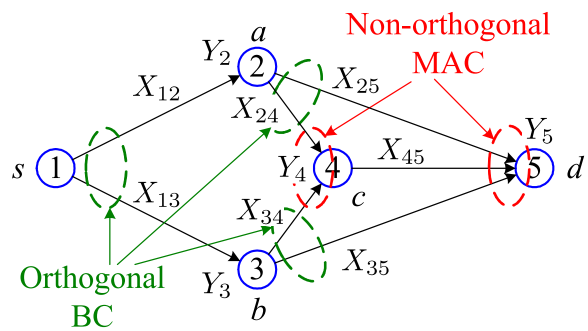

Fig. 1. Example of general memoryless relay network with interference.

node and is denoted by s, and the set of destination nodes is denoted by D, where $s \notin D$. It will be assumed that the source node has no incoming edge, and the destination nodes have no outgoing edge. All the other nodes, which are neither the source nor the destination, are called the relay nodes. Since all broadcast channels in the network are orthogonal, we associate a discrete or continuous random variable $X_{u,v}^{(t)}$ at time t with edge $(u,v) \in E$ as a channel input (output of a node). As a channel output (input of a node), we associate a discrete or continuous random variable $Y_v^{(t)}$ at time t with node $v \in V \setminus \{1\}$. From now on, we sometimes drop the superscript '$^{(t)}$' when doing so causes no confusion.

At node $v \in V$, the set of incoming and outgoing nodes are denoted by

$$\Delta(v) = \{u : (u,v) \in E\},$$
$$\Theta(v) = \{w : (v,w) \in E\}.$$

Set $S \subset V$ is called a cut if it contains node s and its complement S^c contains at least one destination node $d \in D$, i.e., $S^c \cap D \neq \emptyset$. Let Γ denote the set of all cuts. The boundaries of S and S^c are defined as

$$\bar{S} = \{u : \exists v \text{ s.t. } (u,v) \in E, u \in S, v \in S^c\},$$
$$\bar{S}^c = \{v : \exists u \text{ s.t. } (u,v) \in E, u \in S, v \in S^c\}.$$

For node $v \in S^c$, the set of incoming nodes across S is defined as

$$\Delta_S(v) = \Delta(v) \cap S = \Delta(v) \cap \bar{S}.$$

For any sets $S_1 \subseteq V$ and $S_2 \subseteq V$, we define

$$X_{S_1,S_2} = \{X_{u,v} : (u,v) \in E, u \in S_1, v \in S_2\},$$
$$Y_{S_1} = \{Y_v : v \in S_1\},$$

and

$$X_{\Delta(v)} = \{X_{u,v} : u \in \Delta(v)\}.$$

Using the aforementioned notations, we can formally define the class of networks of interest. The memoryless relay network with interference is characterized by the channel distribution function

$$p\left(y_V | x_{V,V}\right) = p\left(y_2 | x_{\Delta(2)}\right) p\left(y_3 | x_{\Delta(3)}\right) \cdots p\left(y_M | x_{\Delta(M)}\right)$$

over all input and output alphabets.

B. Coding for the relay network with interference

The multicast over the relay network consists of encoding functions $f_{u,v}^{(t)}(\cdot)$, $(u,v) \in E$, $t = 1, \ldots, N$, and decoding functions $g_d(\cdot)$, $d \in D$. The source node s has a random message W that is uniform over $\{1, \ldots, M\}$ and transmits

$$X_{s,w}^{(t)} = f_{s,w}^{(t)}(W)$$

at time t on the outgoing channels (s,w), $w \in \Theta(s)$. The relay node v transmits

$$X_{v,w}^{(t)} = f_{v,w}^{(t)}(Y_v^{t-1})$$

at time t on the outgoing channels (v,w), $w \in \Theta(v)$, where $Y_v^{t-1} = \left(Y_v^{(1)}, \ldots, Y_v^{(t-1)}\right)$. At destination node $d \in D$, after time N, an estimate of the source message is computed as

$$\hat{W} = g_d\left(Y_d^N\right).$$

Then, the probability of error is

$$P_e = \Pr\left\{\bigcup_{d \in D}\left\{g_d(Y_d^N) \neq W\right\}\right\}. \tag{1}$$

We say that the multicast rate R is *achievable* if, for any $\epsilon > 0$ and for all sufficiently large N, encoders and decoders with $M \geq 2^{NR}$ exist such that $P_e \leq \epsilon$. The *multicast capacity* is the supremum of the achievable multicast rates.

As stated in Section I, we are interested in characterizing the multicast capacity of the memoryless relay networks with interference. However, as shown in [13], even for a relatively simple parallel relay channel, finding the capacity is not easy. Thus, we further restrict our interest to the Gaussian networks in Section III and the linear finite-field symmetric networks in Section IV.

III. Gaussian relay networks with interference

In this section, we consider Gaussian relay networks with interference. At node v at time t, the received signal is given by

$$Y_v^{(t)} = \sum_{u \in \Delta(v)} X_{u,v}^{(t)} + Z_v^{(t)},$$

where $Z_v^{(t)}$ is an independent identically distributed (i.i.d.) Gaussian random variable with zero mean and unit variance. For each block of channel input $\left(X_{u,v}^{(1)}, \ldots, X_{u,v}^{(n)}\right)$, we have the average power constraint given by

$$\frac{1}{n}\sum_{t=1}^{n}\left(X_{u,v}^{(t)}\right)^2 \leq P_{u,v}.$$

In [10], Nazer et al. studied the achievable rate of the Gaussian relay networks with interference for the equal power constraint case, where $P_{u,v} = P_v$ for all $u \in \Delta(v)$. In our work, we generalize it such that $P_{u,v}$'s can be different. The main result of this section is as follows.

Theorem 1: For a Gaussian relay network with interference, an upper bound for the multicast capacity is given by

$$\min_{S \in \Gamma} \sum_{v \in \bar{S}^c} C\left(\left(\sum_{\substack{u \in \\ \Delta_S(v)}} \sqrt{P_{u,v}}\right)^2\right), \tag{2}$$

where $C(x) = \frac{1}{2}\log(1+x)$. For the same network, we can achieve all rates up to

$$\min_{S\in\Gamma}\sum_{v\in\bar{S}^c}\left[\frac{1}{2}\log\left(\left(\frac{1}{\sum_{\substack{u\in\\ \Delta(v)}} P_{u,v}}+1\right)\cdot\max_{\substack{u\in\\ \Delta_S(v)}} P_{u,v}\right)\right]^+, \tag{3}$$

where $[x]^+ \triangleq \max\{x, 0\}$. Furthermore, the gap between the upper bound and the achievable rate is bounded by

$$\sum_{v\in V\setminus\{1\}} \log\left(|\Delta(v)|\right). \tag{4}$$

Remark 1: Note that, in the equal power case, i.e., $P_{u,v} = P$, the achievable multicast rate (3) has terms in the form of $\log\left(\frac{1}{K}+P\right)$ for some integer $K \geq 1$. Similar forms of achievable rate were observed in [10], [15], [16], [25] for some equal power Gaussian networks.

The following subsections are devoted to proving Theorem 1.

A. Upper bound

The cut-set bound [27] of the network is given by

$$R \leq \max_{p(x_{V,V})}\min_{S\in\Gamma} I\left(X_{S,V}; Y_{S^c}|X_{S^c,V}\right). \tag{5}$$

Though the cut-set bound is a general and convenient upper bound for the capacity, it is sometimes challenging to compute the exact cut-set bound in a closed form. This is due to the optimization by the joint probability density function (pdf) $p(x_{V,V})$. In some cases, such as the finite-field networks in [5], [8], [12], [13], it is easy to compute the cut-set bound because a product distribution maximizes it. For the Gaussian case, however, the optimizing distribution for the cut-set bound is generally not a product distribution.

Thus, we consider another upper bound which is easier to compute than the cut-set bound. This bound is referred to as the *relaxed cut-set bound* and given by

$$R \leq \min_{S\in\Gamma}\max_{p(x_{V,V})} I\left(X_{S,V}; Y_{S^c}|X_{S^c,V}\right). \tag{6}$$

Due to the max-min inequality, the relaxed cut-set bound is looser than the original cut-set bound (5). For the relay network with interference, we can further simplify (6) as

$$\begin{aligned} I(X_{S,V}; Y_{S^c}|X_{S^c,V}) &= I(X_{S,S}, X_{S,S^c}; Y_{S^c}|X_{S^c,V}) \\ &= I(X_{S,S^c}; Y_{S^c}|X_{S^c,V}) \\ &= I(X_{\bar{S},\bar{S}^c}; Y_{\bar{S}^c}|X_{S^c,V}), \end{aligned}$$

where the second and the third equalities follow by the structure of the network, i.e.,

- $X_{S,S} \to (X_{S,S^c}, X_{S^c,V}) \to Y_{S^c}$,
- $(X_{S,S^c}, Y_{\bar{S}^c}) \to X_{S^c,V} \to Y_{S^c\setminus\bar{S}^c}$,
- $X_{S,S^c} = X_{\bar{S},\bar{S}^c}$.

For cut S, the mutual information $I(X_{\bar{S},\bar{S}^c}; Y_{\bar{S}^c}|X_{S^c,V})$ is maximized when there is a perfect coherence between all inputs to a Gaussian MAC across the cut. Thus, we have

$$\max_{p(x_{V,V})} I(X_{\bar{S},\bar{S}^c}; Y_{\bar{S}^c}|X_{S^c,V}) = \sum_{v\in\bar{S}^c} C\left(\left(\sum_{\substack{u\in\\ \Delta_S(v)}}\sqrt{P_{u,v}}\right)^2\right). \tag{7}$$

Then by (6) and (7), the upper bound (2) follows.

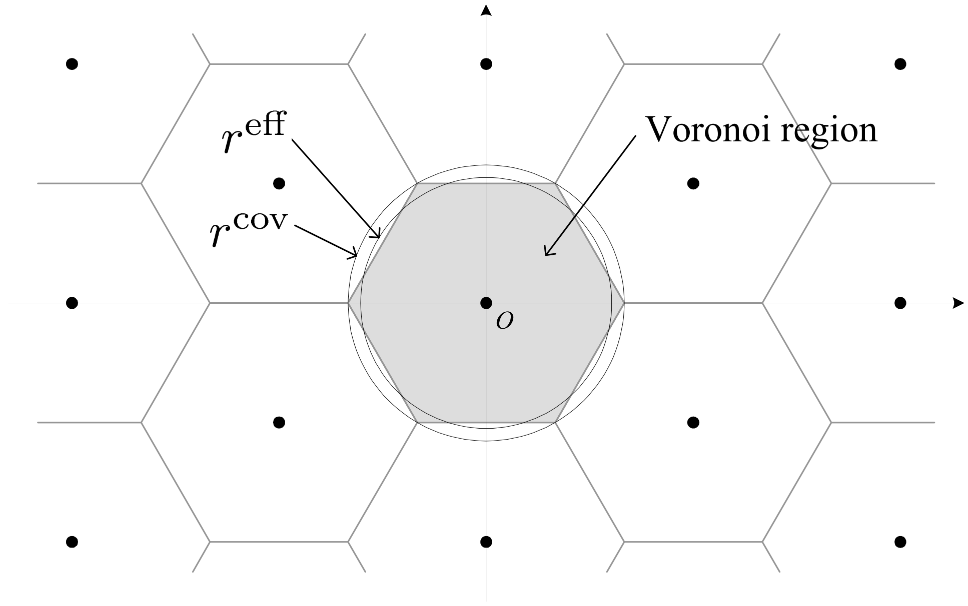

Fig. 2. Example: two-dimensional lattice constellation.

B. Lattices and nested lattice codes

Before proving the achievable part of Theorem 1, let us establish some preliminaries for the lattices and nested lattice codes, which are key ingredients of our achievability proof. For a more comprehensive review on lattices and nested lattice codes, see [19], [20], [23]. An n-dimensional lattice Λ is defined as a discrete subgroup of Euclidean space $\mathbb{R}^n$ with ordinary vector addition. This implies that for any lattice points $\lambda, \lambda' \in \Lambda$, we have $\lambda + \lambda' \in \Lambda$, $\lambda - \lambda' \in \Lambda$, and $\mathbf{0} \in \Lambda$. The nearest neighbor lattice quantizer associated with Λ is defined as

$$Q(\mathbf{x}) = \arg\min_{\lambda \in \Lambda} \|\mathbf{x} - \lambda\|,$$

and the $\bmod \Lambda$ operation is

$$\mathbf{x} \bmod \Lambda = \mathbf{x} - Q(\mathbf{x}).$$

The (fundamental) Voronoi region of Λ, denoted by $\mathcal{R}$, is defined as the set of points in $\mathbb{R}^n$ closer to the origin than to any other lattice points, i.e.,

$$\mathcal{R} = \{\mathbf{x} : Q(\mathbf{x}) = \mathbf{0}\},$$

where ties are broken arbitrarily. In Fig. 2, an example of a two-dimensional lattice, and its Voronoi region are depicted.

We now define some important parameters that characterize the lattice. The covering radius of the lattice r^{cov} is defined as the radius of a sphere circumscribing around $\mathcal{R}$, i.e.,

$$r^{\mathrm{cov}} = \min\{r : \mathcal{R} \subseteq r\mathcal{B}\},$$

where $\mathcal{B}$ is an n-dimensional unit sphere centered at the origin, and, thus, $r\mathcal{B}$ is a sphere of radius r. In addition, the effective radius of Λ, denoted by r^{eff}, is the radius of a sphere with the same volume as $\mathcal{R}$, i.e.,

$$r^{\mathrm{eff}} = \left(\frac{\mathrm{Vol}(\mathcal{R})}{\mathrm{Vol}(\mathcal{B})}\right)^{\frac{1}{n}},$$

where $\mathrm{Vol}(\cdot)$ denotes the volume of a region. The second moment per dimension of Λ is defined as the second moment per dimension associated with $\mathcal{R}$, which is given by

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

In the rest of this paper, we also use $\mathrm{Vol}(\Lambda)$ and $\sigma^2(\Lambda)$, which have the same meaning as $\mathrm{Vol}(\mathcal{R})$ and $\sigma^2(\mathcal{R})$, respectively. Finally, we define the normalized second moment of Λ as

$$G(\Lambda) = \frac{\sigma^2(\mathcal{R})}{(\mathrm{Vol}(\mathcal{R}))^{2/n}}.$$

For any Λ, $G(\Lambda)$ is greater than $\frac{1}{2\pi e}$, which is the normalized second moment of a sphere whose dimension tends to infinity.

Goodness of lattices

We consider a sequence of lattices Λ^n. The sequence of lattices is said to be *Rogers-good* if

$$\lim_{n\to\infty} \frac{r^{\mathrm{cov}}}{r^{\mathrm{eff}}} = 1,$$

which implies that Λ^n is asymptotically efficient for sphere covering [20]. This also implies the goodness of Λ^n for mean-square error quantization, i.e.,

$$\lim_{n\to\infty} G(\Lambda^n) = \frac{1}{2\pi e}.$$

We now define the goodness of lattices related to the channel coding for the additive white Gaussian noise (AWGN) channel. A sequence of lattices is said to be *Poltyrev-good* if, for $\bar{\mathbf{Z}} \sim \mathcal{N}(\mathbf{0}, \bar{\sigma}^2\mathbf{I})$,

$$\Pr\{\bar{\mathbf{Z}} \notin \mathcal{R}\} \le e^{-nE_P(\mu)}, \tag{8}$$

where $E_P(\cdot)$ is the Poltyrev exponent [22] and μ is the volume-to-noise ratio (VNR) defined as

$$\mu = \frac{(\mathrm{Vol}(\mathcal{R}))^{2/n}}{2\pi e\bar{\sigma}^2}.$$

Note that (8) upper bounds the error probability of the nearest lattice point decoding (or equivalently, Euclidean lattice decoding) when we use lattice points as codewords for the AWGN channel. Since $E_P(\mu) > 0$ for $\mu > 1$, a necessary condition for reliable decoding is $\mu > 1$.

Nested lattices codes

Now we consider two lattices Λ and Λ_C. Assume that Λ is coarse compared to Λ_C in the sense that $\mathrm{Vol}(\Lambda) \ge \mathrm{Vol}(\Lambda_C)$. We say that the coarse lattice Λ is a sublattice of the fine lattice Λ_C if $\Lambda \subseteq \Lambda_C$ and call the quotient group (equivalently, the set of cosets of Λ relative to Λ_C) Λ_C/Λ a *lattice partition*. For the lattice partition, the *set of coset leaders* is defined as

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

and the *partitioning ratio* is

$$\rho = |\mathcal{C}|^{\frac{1}{n}} = \left(\frac{\mathrm{Vol}(\Lambda)}{\mathrm{Vol}(\Lambda_C)}\right)^{\frac{1}{n}}.$$

Formally, a lattice code is defined as an intersection of a lattice (possibly translated) and a bounding (shaping) region, which is sometimes a sphere. A *nested lattice code* is a special class of lattice codes, whose bounding region is the Voronoi region of a sublattice. That is, the nested lattice code is defined in terms of lattice partition Λ_C/Λ, in which Λ_C is used as codewords and Λ is used for shaping. The coding rate of the nested lattice code is given by

$$\frac{1}{n}\log|\mathcal{C}| = \log\rho.$$

Nested lattice codes have been studied in many previous articles [18], [19], [23], [24], and proved to have many useful properties, such as achieving the capacity of the AWGN channel. In the next subsection, we deal with the nested lattice codes for the achievability proof of Theorem 1.

C. Nested lattice codes for a Gaussian MAC

As an achievable scheme, we use a lattice coding scheme. In [10], lattices were also used to prove an achievable rate of Gaussian relay networks with interference (called Gaussian MAC networks). However, they used the lattice as a source code with a distortion and then related the achievable distortion to the information flow through the network. Our approach is different from [10] in that we use lattices to produce coding and shaping lattices, and form nested lattice codes. As a result, our approach can handle unequal power constraints where incoming links have different power at a MAC. Our scheme is a generalization of the nested lattice codes used for the Gaussian two-way relay channel in [15], [16].

Let us consider a standard model of a Gaussian MAC with K input nodes:

$$Y = \sum_{j=1}^{K} X_j + Z, \tag{9}$$

where Z denotes the AWGN process with zero mean and unit variance. Each channel input X_i is subject to the average power constraint P_i, i.e., $\frac{1}{n}\sum_{t=1}^{n}(X_i^{(t)})^2 \le P_i$. Without loss of generality, we assume that $P_1 \ge P_2 \ge \cdots \ge P_K$.

The standard MAC in (9) is a representative of MACs in the Gaussian relay network with interference. Now, we introduce encoding and decoding schemes for the standard MAC. Let us first consider the following theorem which is a key for our code construction.

Theorem 2: For any $P_1 \ge P_2 \ge \cdots \ge P_K \ge 0$ and $\gamma \ge 0$, a sequence of n-dimensional lattice chains $\Lambda_1^n \subseteq \Lambda_2^n \subseteq \cdots \subseteq \Lambda_K^n \subseteq \Lambda_C^n$ exists that satisfies the following properties.

a) Λ_i^n, $1 \le i \le K$, are simultaneously Rogers-good and Poltyrev-good while Λ_C^n is Poltyrev-good.

b) For any $\delta > 0$, $P_i - \delta \le \sigma^2(\Lambda_i^n) \le P_i$, $1 \le i \le K$, for sufficiently large n.

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

$$R_K \triangleq \frac{1}{n}\log|\mathcal{C}_K| = \gamma + o_n(1),$$

where $\mathcal{C}_K = \{\Lambda_C^n \bmod \Lambda_K^n\}$ and $o_n(1) \to 0$ as $n \to \infty$. Furthermore, for $1 \le i \le K-1$, the coding rate of the nested lattice code associated with Λ_C^n/Λ_i^n is given by

$$R_i \triangleq \frac{1}{n}\log|\mathcal{C}_i| = R_K + \frac{1}{2}\log\left(\frac{P_i}{P_K}\right) + o_n(1),$$

where $\mathcal{C}_i = \{\Lambda_C^n \bmod \Lambda_i^n\}$.

Proof: See Appendix A. ∎

A conceptual representation of the lattice chain and the corresponding sets of coset leaders are given in Fig. 3 for a two-dimensional case.

Encoding

We consider a lattice chain as described in Theorem 2. We assign the i-th input node to the MAC with the set of coset leaders $\mathcal{C}_i$. For each input node, the message set $\{1,\dots,2^{nR_i}\}$ is arbitrarily mapped onto

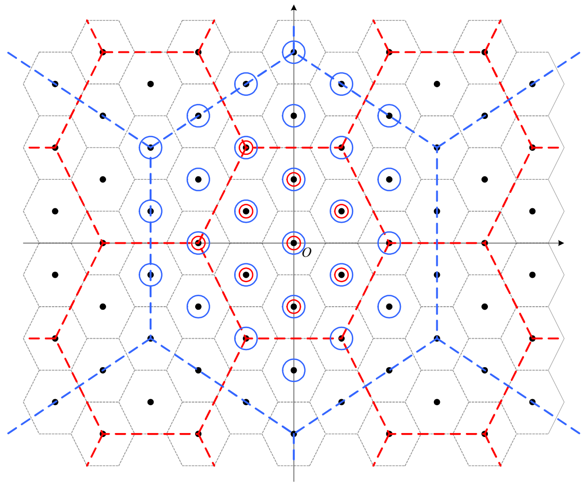

Fig. 3. Example of lattice chain and sets of coset leaders.

$\mathcal{C}_i$. We also define random dither vectors $\mathbf{U}_i \sim \mathrm{Unif}(\mathcal{R}_i)$, $1 \le i \le K$, where $\mathcal{R}_i$ denotes the Voronoi region of Λ_i (we dropped the superscript 'n' for simplicity). These dither vectors are independent of each other and also independent of the message of each node and the noise. We assume that each $\mathbf{U}_i$ is known to both the i-th input node and the receiver. To transmit a message that is uniform over $\{1, \ldots, 2^{nR_i}\}$, node i chooses $\mathbf{W}_i \in \mathcal{C}_i$ associated with the message and sends

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

Let us introduce a useful lemma, which is known as the *crypto-lemma* and frequently used in the rest of this paper. The lemma is given in [23], and we repeat it here for completeness.

Lemma 1 (Crypto-lemma [23]): Let $\mathcal{C}$ be a finite or compact group with group operation $+$. For independent random variables a and b over $\mathcal{C}$, let $c = a + b$. If a is uniform over $\mathcal{C}$, then c is independent of b and uniform over $\mathcal{C}$.

By Lemma 1, $\mathbf{X}_i$ is uniformly distributed over $\mathcal{R}_i$ and independent of $\mathbf{W}_i$. Thus, regardless of $\mathbf{W}_i$, the average transmit power of node i is equal to $\sigma^2(\Lambda_i)$, which approaches P_i as n tends to infinity. Thus, the power constraint is met.

Decoding

Upon receiving $\mathbf{Y} = \sum_{j=1}^{K} \mathbf{X}_j + \mathbf{Z}$, where $\mathbf{Z}$ is a vector of i.i.d. Gaussian noise with zero mean and unit variance, the receiver computes

$$\begin{aligned}
\tilde{\mathbf{Y}} &= \left(\alpha \mathbf{Y} - \sum_{j=1}^{K} \mathbf{U}_j\right) \bmod \Lambda_1 \\
&= \left[\sum_{j=1}^{K} (\mathbf{W}_j + \mathbf{U}_j) \bmod \Lambda_j - \sum_{j=1}^{K} \mathbf{X}_j \right. \\
&\qquad \left. + \alpha \sum_{j=1}^{K} \mathbf{X}_j + \alpha \mathbf{Z} - \sum_{j=1}^{K} \mathbf{U}_j \right] \bmod \Lambda_1 \\
&= \left(\mathbf{T} + \tilde{\mathbf{Z}}\right) \bmod \Lambda_1,
\end{aligned}$$

where

$$\mathbf{T} = \left[\sum_{j=1}^{K} (\mathbf{W}_j - Q_j(\mathbf{W}_j + \mathbf{U}_j))\right] \bmod \Lambda_1$$
$$= \left[\mathbf{W}_1 + \sum_{j=2}^{K} (\mathbf{W}_j - Q_j(\mathbf{W}_j + \mathbf{U}_j))\right] \bmod \Lambda_1, \tag{10}$$
$$\tilde{\mathbf{Z}} = -(1-\alpha)\sum_{j=1}^{K} \mathbf{X}_j + \alpha \mathbf{Z},$$

$0 \le \alpha \le 1$ is a scaling factor, and $Q_j(\cdot)$ denotes the nearest neighbor lattice quantizer associated with Λ_j. We choose α as the minimum mean-square error (MMSE) coefficient to minimize the variance of the effective noise $\tilde{\mathbf{Z}}$. Thus,

$$\alpha = \frac{\sum_{j=1}^{K} P_j}{\sum_{j=1}^{K} P_j + 1},$$

and the resulting noise variance satisfies

$$\frac{1}{n} E\left\{ \left\| \tilde{\mathbf{Z}} \right\|^2 \right\} \le \frac{\sum_{j=1}^{K} P_j}{\sum_{j=1}^{K} P_j + 1}. \tag{11}$$

Note that, though the relation in (11) is given by an inequality, it becomes tight as $n \to \infty$ by Theorem 2. By the chain relation of the lattices in Theorem 2, it is easy to show that $\mathbf{T} \in \mathcal{C}_1$. Regarding $\mathbf{T}$, we have the following lemma.

Lemma 2: $\mathbf{T}$ is uniform over $\mathcal{C}_1$ and independent of $\tilde{\mathbf{Z}}$.

Proof: Define $\tilde{\mathbf{W}} \triangleq \sum_{j=2}^{K} (\mathbf{W}_j - Q_j(\mathbf{W}_j + \mathbf{U}_j)) \bmod \Lambda_1$, and, thus, $\mathbf{T} = \left(\mathbf{W}_1 + \tilde{\mathbf{W}}\right) \bmod \Lambda_1$. Note that $\tilde{\mathbf{W}}$ is correlated with $\mathbf{X}_i$, $2 \le i \le K$, and $\tilde{\mathbf{Z}}$. Since $\mathbf{W}_1$ is uniform over $\mathcal{C}_1$ and independent of $\tilde{\mathbf{W}}$, $\mathbf{T}$ is independent of $\tilde{\mathbf{W}}$ and uniformly distributed over $\mathcal{C}_1$ (crypto-lemma). Hence, if $\mathbf{T}$ and $\tilde{\mathbf{Z}}$ are correlated, it is only through $\mathbf{W}_1$. However, $\mathbf{W}_1$ and $\tilde{\mathbf{Z}}$ are independent of each other, and, consequently, $\mathbf{T}$ is also independent of $\tilde{\mathbf{Z}}$. ■

The receiver tries to retrieve $\mathbf{T}$ from $\tilde{\mathbf{Y}}$ instead of recovering $\mathbf{W}_i$, $1 \le i \le K$, separately. For the decoding method, we consider *Euclidean lattice decoding* [19]-[23], which finds the closest point to $\tilde{\mathbf{Y}}$ in Λ_C. From the symmetry of the lattice structure and the independence between $\mathbf{T}$ and $\tilde{\mathbf{Z}}$ (Lemma 2), the probability of decoding error is given by

$$p_e = \Pr\left\{\mathbf{T} \ne Q_C\left(\tilde{\mathbf{Y}}\right)\right\}$$
$$= \Pr\left\{\tilde{\mathbf{Z}} \bmod \Lambda_1 \notin \mathcal{R}_C\right\}, \tag{12}$$

where $Q_C(\cdot)$ denotes the nearest neighbor lattice quantizer associated with Λ_C and $\mathcal{R}_C$ denotes the Voronoi region of Λ_C. Then, we have the following theorem.

Theorem 3: Let

$$R_1^* = \left[\frac{1}{2}\log\left(\frac{P_1}{\sum_{j=1}^{K} P_j} + P_1\right)\right]^+.$$

For any $\bar{R}_1 < R_1^*$ and a lattice 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 Euclidean lattice decoding (12) is bounded by

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

Proof: See Appendix B. ■

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 achieve any rate below R_1^*. Thus, by c) of Theorem 2 and Theorem 3, the coding rate R_i, $1 \leq i \leq K$, can approach R_i^* arbitrarily closely while keeping p_e arbitrarily small for sufficiently large n, where

$$R_i^* = \left[\frac{1}{2}\log\left(\frac{P_i}{\sum_{j=1}^{K} P_j} + P_i\right)\right]^+. \tag{13}$$

Remark 2: In theorem 3, we showed the error exponent of lattice decoding and the achievability of R_1 directly followed. However, if we are only interested in finding the achievability of R_1, not in the error exponent, we can use the argument on the bounding behavior of lattice decoding in [21], which gives the same result in a much simpler way.

Remark 3: Since $P_1 \geq \cdots \geq P_K$, we have $R_1^* \geq \cdots \geq R_K^*$. Now, consider the case that, for some $\hat{i} < K$, the rates R_i^*, $\hat{i}+1 \leq i \leq K$, are zero while R_i^*, $1 \leq i \leq \hat{i}$, are nonzero. In this situation, nodes $\hat{i}+1, \ldots, K$ cannot transmit any useful information to the receiver, and, thus, we can turn them off so as not to hinder the transmissions of nodes $1, \ldots, \hat{i}$. Then, the variance of $\tilde{\mathbf{Z}}$ decreases and we have extended rates given by

$$R_i^* = \left[\frac{1}{2}\log\left(\frac{P_i}{\sum_{j=1}^{\hat{i}} P_j} + P_i\right)\right]^+, \ 1 \leq i \leq \hat{i}.$$

However, for the ease of exposition, we do not consider the transmitter turning-off technique and assume that nodes $\hat{i}+1, \ldots, K$ just transmit $\mathbf{X}_i = \mathbf{U}_i$ when their coding rates are zero.

D. Achievable multicast rate

We consider B blocks of transmissions from the source to destinations. Each block consists of n channel uses. In block $k \in \{1, \ldots, B\}$, an independent and uniform message $W[k] \in \{1, \ldots, 2^{nR}\}$ is sent from the source node s. It takes at most $L \triangleq B + |V| - 2$ blocks for all the B messages to be received by destination nodes. After receiving L blocks, destination nodes decode the source message $W \triangleq (W[1], \ldots, W[B])$. Thus, the overall rate is $\frac{B}{L}R$, which can be arbitrarily close to R by choosing B sufficiently large.

Time-expanded network

For ease of analysis, we consider the B blocks of transmissions over the time-expanded network [2], [5], $\mathcal{G}_{\mathrm{TE}}$, obtained by unfolding the original network $\mathcal{G}$ over $L+1$ time stages. In $\mathcal{G}_{\mathrm{TE}}$, node $v \in V$ at block k appears as $v[k]$, and $v[k]$ and $v[k']$ are treated as different nodes if $k \neq k'$. There are a virtual source and destination nodes, denoted by s_{TE} and d_{TE}, respectively. We assume that s_{TE} and $s[k]$'s are connected through virtual error-free infinite-capacity links, and, similarly, d_{TE} and $d[k]$'s are. For instance, the network in Fig. 1 is expanded to the network in Fig. 4. Dealing with the time-expanded network does not impose any constraints on the network. Any scheme for the original network can be interpreted to a scheme for the time-expanded network and vice-versa. In our case, the transmissions of B messages $W[k]$, $k = 1, \ldots, B$, from s to $d \in D$ over $\mathcal{G}$ correspond to the transmission of a single message W from s_{TE} to $d_{\mathrm{TE}} \in D_{\mathrm{TE}}$ over $\mathcal{G}_{\mathrm{TE}}$, where D_{TE} denotes the set of virtual destination nodes.

A main characteristic of the time-expanded network is that it is always a layered network [5] which has equal length paths from the source to each destination[1]. We define the set of nodes at length k from the virtual source node as

$$V_{\mathrm{TE}}[k] = \{v[k] : v \in V\}$$

[1]Another characteristic is that the time-expanded network is always acyclic [2].

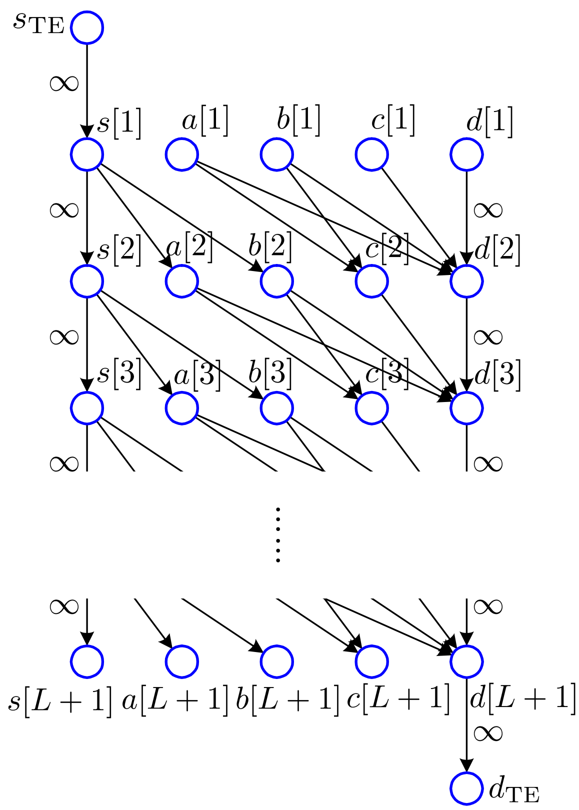

Fig. 4. Time-expansion of the network in Fig. 1.

and call it the k-th layer. We use the subscript '$_{\mathrm{TE}}$' to differentiate parameters of $\mathcal{G}$ and $\mathcal{G}_{\mathrm{TE}}$. The set of nodes and edges of $\mathcal{G}_{\mathrm{TE}}$ are defined as

$$\begin{aligned} V_{\mathrm{TE}} &= \{s_{\mathrm{TE}}\} \cup D_{\mathrm{TE}} \cup \left(\bigcup_{k=1}^{L+1} V_{\mathrm{TE}}[k] \right), \\ E_{\mathrm{TE}} &= \{(u[k], v[k+1]) : (u,v) \in E, k = 1, \ldots, L\} \\ &\quad \cup \{(s[k-1], s[k]) : k = 1, \ldots, L\} \\ &\quad \cup \{(d[k], d[k+1]) : k = 1, \ldots, L\}, \end{aligned}$$

where we define $s[0] = s_{\mathrm{TE}}$ and $d[L+2] = d_{\mathrm{TE}}$. Note that, since $\mathcal{G}_{\mathrm{TE}}$ is layered, edges only appear between adjacent layers. From V_{TE} and E_{TE}, the other parameters, e.g., $\Delta_{\mathrm{TE}}(\cdot)$, $\Theta_{\mathrm{TE}}(\cdot)$, S_{TE}, $\bar{S}_{\mathrm{TE}}$, Γ_{TE}, and $\Delta_{\mathrm{TE},S}(\cdot)$, are similarly defined as $\Delta(\cdot)$, $\Theta(\cdot)$, S, $\bar{S}$, Γ, and $\Delta_S(\cdot)$, respectively.

Encoding

We apply the nested lattice codes in Section III-C over the all Gaussian MACs in the network. Thus, node $v[k]$ is assigned with sets of coset leaders $\mathcal{C}_{v[k],w[k^+]}$, $w[k^+] \in \Theta_{\mathrm{TE}}(v[k])$, where $k^+ \triangleq k+1$. We do not change the lattice scheme over blocks, and, thus, $\mathcal{C}_{v[k],w[k^+]} = \mathcal{C}_{v,w}$

At node $s[k]$, the indices $\left\{1, \ldots, 2^{nR}\right\}$ are uniformly randomly mapped onto vectors in $\mathcal{C}_{s,w}$, $w \in \Theta(s)$. We define the random mapping as $f_{s[k],w[k^+]}(\cdot)$. Then, node $s[k]$ receives $W = (W[1], \ldots, W[B])$ from $s[k^-]$ through the error-free link, where $k^- \triangleq k-1$, and transmits

$$\mathbf{W}_{s[k],w[k^+]} = f_{s[k],w[k^+]}(W[k])$$

on channel $(s[k], w[k^+])$ using a random dither vector $\mathbf{U}_{s[k],w[k^+]}$. At node $v[k]$ that is not $s[k]$ or $d[k]$, the received signal is given by

$$\tilde{\mathbf{Y}}_{v[k]} = \left(\mathbf{T}_{v[k]} + \tilde{\mathbf{Z}}_{v[k]} \right) \bmod \Lambda_v, \tag{14}$$

where

$$\mathbf{T}_{v[k]} = \left[\sum_{\substack{u[k^-] \in \\ \Delta_{\mathrm{TE}}(v[k])}} \left(\mathbf{W}_{u[k^-],v[k]} - Q_{u,v} \left(\mathbf{W}_{u[k^-],v[k]} + \mathbf{U}_{u[k^-],v[k]} \right) \right) \right] \bmod \Lambda_v, \tag{15}$$

and $\tilde{\mathbf{Z}}_{v[k]}$ is an effective noise vector. In (14), Λ_v denotes the lattice associated with the incoming channel to node v with the largest power. Then, $\mathbf{T}_{v[k]}$ is decoded using Euclidean lattice decoding, which yields an estimate $\hat{\mathbf{T}}_{v[k]}$. Next, $\hat{\mathbf{T}}_{v[k]}$ is uniformly and randomly mapped onto vectors in $C_{v,w}$, $w \in \Theta(v)$. This mapping is denoted by $f_{v[k],w[k^+]}(\cdot)$, and node $v[k]$ transmits

$$\mathbf{W}_{v[k],w[k^+]} = f_{v[k],w[k^+]} \left(\hat{\mathbf{T}}_{v[k]} \right)$$

on channel $(v[k], w[k^+])$ using a random dither vector $\mathbf{U}_{v[k],w[k^+]}$. Node $d[k]$, $d \in D$, receives $\tilde{\mathbf{Y}}_{d[k]}$ and computes $\hat{\mathbf{T}}_{d[k]}$. It also receives $\left(\hat{\mathbf{T}}_{d[1]}, \ldots, \hat{\mathbf{T}}_{d[k^-]} \right)$ from $d[k^-]$ through the virtual error-free infinite-capacity link and passes $\left(\hat{\mathbf{T}}_{d[1]}, \ldots, \hat{\mathbf{T}}_{d[k]} \right)$ to node $d[k^+]$.

We assume that all the random mappings $f_{u[k],v[k^+]}$, $(u[k], v[k^+]) \in E_{\mathrm{TE}}$ are done independently.

Decoding

While decoding, a virtual destination node $d_{\mathrm{TE}} \in D_{\mathrm{TE}}$ assumes that there is no error in decoding $\mathbf{T}_{v[k]}$'s in the network and that the network is deterministic. Therefore, with knowledge of all deterministic relations[2] (15) in the network, node d_{TE} decodes W by simulating all 2^{nBR} messages and finding one that yields the received signal $\hat{\mathbf{T}}_{d_{\mathrm{TE}}} \triangleq \left(\hat{\mathbf{T}}_{d[1]}, \ldots, \hat{\mathbf{T}}_{d[L+1]} \right)$.

Calculation of the probability of error

In the above decoding rule, we will declare an error if at least one of the following events occurs.

- $\mathcal{E}_1$: there is an error in decoding $\mathbf{T}_{v[k]}$ at at least one node in the network.
- $\mathcal{E}_2$: a message $W' \neq W$ exists that yields the same received signal $\hat{\mathbf{T}}_{d_{\mathrm{TE}}}$, which is obtained under W, at at least one virtual destination node $d_{\mathrm{TE}} \in D_{\mathrm{TE}}$.

Thus, the error probability is given by

$$\begin{aligned} P_e &= \Pr\{\mathcal{E}_1 \cup \mathcal{E}_2\} \\ &\leq \Pr\{\mathcal{E}_1\} + \Pr\{\mathcal{E}_2 | \mathcal{E}_1^c\}. \end{aligned} \tag{16}$$

Let us consider the first term in (16). Using the union bound, we have

$$\Pr\{\mathcal{E}_1\} \leq \sum_{k=2}^{L+1} \sum_{\substack{v[k] \in V[k] \\ \setminus \{s[k]\}}} p_{e,v[k]},$$

where

$$p_{e,v[k]} \triangleq \Pr\left\{ \hat{\mathbf{T}}_{v[k]} \neq \mathbf{T}_{v[k]} \right\}.$$

Note that the summation is from $k = 2$ since nodes in the first layer do not have any received signal except for node $s[1]$. By Theorem 3, at node $v \in V \setminus \{1\}$ for any $\epsilon > 0$, $p_{e,v[k]}$ is less than $\frac{\epsilon}{2L|V|}$ for sufficiently large n if

$$\begin{aligned} R_{u,v} &= \frac{1}{n} \log |\mathcal{C}_{u,v}| \\ &= \left[\frac{1}{2} \log \left(\left(\frac{1}{\sum_{\substack{u' \in \\ \Delta(v)}} P_{u',v}} + 1 \right) \cdot P_{u,v} \right) - \epsilon \right]^+ \end{aligned} \tag{17}$$

[2]It is assumed that the all random dither vectors are known to destination nodes. Thus, (15) is deterministic.

for all $u \in \Delta(v)$. Therefore, in this case

$$\Pr\{\mathcal{E}_1\} \leq \frac{\epsilon}{2}.$$

Now, we consider the second term in (16). Under the condition $\mathcal{E}_1^c$, we have $\hat{\mathbf{T}}_{v[k]} = \mathbf{T}_{v[k]}$, and, thus, the network is deterministic. Let us use the notation $\mathbf{W}_{u[k^-],v[k]}(W)$ and $\mathbf{T}_{v[k]}(W)$ to explicitly denote the signals under message W. We say that node $v[k]$ can distinguish W and W' if $\mathbf{T}_{v[k]}(W) \neq \mathbf{T}_{v[k]}(W')$. Thus, from the argument of a deterministic network in [5], the error probability is bounded by

$$\begin{aligned}
\Pr\{\mathcal{E}_2|\mathcal{E}_1^c\} &\leq 2^{nBR} \cdot \Pr\left\{ \bigcup_{\substack{d_{\mathrm{TE}}\in \\ D_{\mathrm{TE}}}} \{\mathbf{T}_{d_{\mathrm{TE}}}(W) = \mathbf{T}_{d_{\mathrm{TE}}}(W')\} \right\} \\
&= 2^{nBR} \cdot \sum_{\substack{S_{\mathrm{TE}}\in \\ \Gamma_{\mathrm{TE}}}} \Pr\{\text{Nodes in } S_{\mathrm{TE}} \text{ can distinguish } W,\, W', \text{ and nodes in } S_{\mathrm{TE}}^c \text{ cannot}\}. \qquad (18)
\end{aligned}$$

We briefly denote the probabilities in the summation in (18) as

$$\Pr\left\{\mathcal{D} = S_{\mathrm{TE}}, \bar{\mathcal{D}} = S_{\mathrm{TE}}^c\right\}.$$

Here, we redefine the cut in the time-expanded network $\mathcal{G}_{\mathrm{TE}}$ for convenience sake. From the encoding scheme, since the source message propagates through nodes $s[k]$, $k = 1, \ldots, L+1$, they can clearly distinguish W and W'. Similarly, if a virtual destination node d_{TE} cannot distinguish W and W', nodes $d[k]$, $k = 1, \ldots, L+1$ cannot either. Thus, when we analyze the error probability (18), we can always assume that $s[k] \in S_{\mathrm{TE}}$ and $d[k] \in S_{\mathrm{TE}}^c$, $k = 1, \ldots, L+1$, without loss of generality.

From the fact that $\mathcal{G}_{\mathrm{TE}}$ is layered, we have

$$\begin{aligned}
\Pr\left\{\mathcal{D} = S_{\mathrm{TE}}, \bar{\mathcal{D}} = S_{\mathrm{TE}}^c\right\} &= \Pr\left\{\mathcal{D} = S_{\mathrm{TE}}, \bar{\mathcal{D}} = S_{\mathrm{TE}}^c[1]\right\} \\
&\quad \cdot \prod_{k=2}^{L+1} \Pr\left\{\bar{\mathcal{D}} = S_{\mathrm{TE}}^c[k] | \mathcal{D} = S_{\mathrm{TE}}[k^-], \bar{\mathcal{D}} = S_{\mathrm{TE}}^c[k^-]\right\} \\
&\leq \prod_{k=2}^{L+1} \Pr\left\{\bar{\mathcal{D}} = S_{\mathrm{TE}}^c[k] | \mathcal{D} = S_{\mathrm{TE}}[k^-], \bar{\mathcal{D}} = S_{\mathrm{TE}}^c[k^-]\right\}, \qquad (19)
\end{aligned}$$

where $S_{\mathrm{TE}}[k]$ and $S_{\mathrm{TE}}^c[k]$ denote the sets of nodes in S_{TE} and S_{TE}^c in the k-th layer, i.e.,

$$\begin{aligned}
S_{\mathrm{TE}}[k] &\triangleq S_{\mathrm{TE}} \cap V_{\mathrm{TE}}[k], \\
S_{\mathrm{TE}}^c[k] &\triangleq S_{\mathrm{TE}}^c \cap V_{\mathrm{TE}}[k].
\end{aligned}$$

Also, from the fact that the random mapping for each channel is independent, we have

$$\begin{aligned}
&\Pr\left\{\bar{\mathcal{D}} = S_{\mathrm{TE}}^c[k] | \mathcal{D} = S_{\mathrm{TE}}[k^-], \bar{\mathcal{D}} = S_{\mathrm{TE}}^c[k^-]\right\} \\
&= \prod_{\substack{v[k]\in \\ S_{\mathrm{TE}}^c[k]}} \Pr\left\{\bar{\mathcal{D}} = \{v[k]\} | \mathcal{D} = S_{\mathrm{TE}}[k^-], \bar{\mathcal{D}} = S_{\mathrm{TE}}^c[k^-]\right\}. \qquad (20)
\end{aligned}$$

Then, we have the following lemma.

Lemma 3: Consider the time-expanded network $\mathcal{G}_{\mathrm{TE}}$ with independent uniform random mapping at each node. For any cut[3] S_{TE} in $\mathcal{G}_{\mathrm{TE}}$, we have

$$\Pr\left\{\bar{\mathcal{D}} = \{v[k]\} | \mathcal{D} = S_{\mathrm{TE}}[k^-], \bar{\mathcal{D}} = S_{\mathrm{TE}}^c[k^-]\right\} \leq 2^{-n\left(\max\limits_{\substack{u[k^-]\in \\ \Delta_{\mathrm{TE},S}(v[k])}} R_{u,v}\right)}$$

[3]From the definition, $s[k] \in S_{\mathrm{TE}}$ and $d[k] \in S_{\mathrm{TE}}^c$, $k = 1, \ldots, L+1$.

for node $v[k] \in \bar{S}^c_{\mathrm{TE}}[k]$, where $\bar{S}^c_{\mathrm{TE}}[k] \triangleq \bar{S}^c_{\mathrm{TE}} \cap V_{\mathrm{TE}}[k]$. For node $v[k] \in S^c_{\mathrm{TE}}[k] \setminus \bar{S}^c_{\mathrm{TE}}[k]$, we have

$$\Pr\left\{\bar{\mathcal{D}} = \{v[k]\} | \mathcal{D} = S_{\mathrm{TE}}[k^-], \bar{\mathcal{D}} = S^c_{\mathrm{TE}}[k^-]\right\} = 1.$$

Proof: See Appendix C. ∎

Thus, by (18)-(20) and Lemma 3, it follows that

$$\Pr\{\mathcal{E}_2|\mathcal{E}_1^c\} \leq 2^{nBR} \cdot |\Gamma_{\mathrm{TE}}| \cdot 2^{-n \min\limits_{S_{\mathrm{TE}} \in \Gamma_{\mathrm{TE}}} \sum\limits_{k=2}^{L+1} \sum\limits_{v[k] \in \bar{S}^c_{\mathrm{TE}}[k]} \left(\max\limits_{u[k^-] \in \Delta_{\mathrm{TE},S}(v[k])} R_{u,v} \right)}. \tag{21}$$

We now consider the following lemma.

Lemma 4: In the time-expanded $\mathcal{G}_{\mathrm{TE}}$ with $L+1$ layers, the term in the exponent of (21)

$$\min_{S_{\mathrm{TE}} \in \Gamma_{\mathrm{TE}}} \sum_{k=2}^{L+1} \sum_{\substack{v[k] \in \\ \bar{S}^c_{\mathrm{TE}}[k]}} \left(\max_{\substack{u[k^-] \in \\ \Delta_{\mathrm{TE},S}(v[k])}} R_{u,v} \right)$$

is upper bounded by

$$L \cdot \min_{S \in \Gamma} \sum_{v \in \bar{S}^c} \left(\max_{u \in \Delta_S(v)} R_{u,v} \right),$$

and lower bounded by

$$(L - |\Gamma| + 2) \cdot \min_{S \in \Gamma} \sum_{v \in \bar{S}^c} \left(\max_{u \in \Delta_S(v)} R_{u,v} \right).$$

Proof: See Appendix D. ∎

Therefore, by (17), (21), and Lemma 4, $\Pr\{\mathcal{E}_2|\mathcal{E}_1^c\}$ is less than $\frac{\epsilon}{2}$ for sufficiently large n if

$$R < \frac{L - |\Gamma| + 2}{B} \cdot \min_{S \in \Gamma} \sum_{v \in \bar{S}^c} \left[\frac{1}{2} \log \left(\left(\frac{1}{\sum_{\substack{u \in \\ \Delta(v)}} P_{u,v}} + 1 \right) \cdot \max_{\substack{u \in \\ \Delta_S(v)}} P_{u,v} \right) - \epsilon \right]^+. \tag{22}$$

Thus, the total probability of error (16) is less than ϵ, and the achievability follows from (22).

E. Gap between the upper and lower bounds

To compute the gap between the upper bound (2) and the achievable rate (3), we can rely on the following lemmas.

Lemma 5: Assume that $P_1 \geq \cdots \geq P_K \geq 0$. For any nonempty set $A \subseteq \{1, \ldots, K\}$ and $l = \min A$, we have

$$\begin{aligned} &\frac{1}{2} \log \left(1 + \left(\sum_{j \in A} \sqrt{P_j} \right)^2 \right) \\ &- \left[\frac{1}{2} \log \left(\left(\frac{1}{\sum_{j=1}^K P_j} + 1 \right) P_l \right) \right]^+ \leq \log K. \end{aligned}$$

Lemma 6:

$$\begin{aligned} &\min\{a_1, \ldots, a_k\} - \min\{b_1, \ldots, b_k\} \\ &\leq \max\{(a_1 - b_1), \ldots, (a_k - b_k)\}. \end{aligned}$$

The proof of Lemma 5 is given in Appendix E, and the proof of Lemma 6 is omitted since it is straightforward. Using Lemmas 5 and 6, the gap in (4) directly follows.

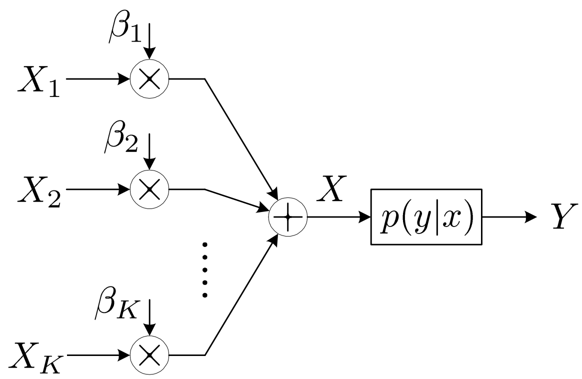

Fig. 5. Linear finite-field symmetric MAC.

IV. Linear finite-field symmetric networks with interference

Let us consider a particular class of discrete memoryless relay networks with interference. The linear finite-field symmetric networks with interference are characterized by a special structure of MACs in them, which is shown in Fig. 5. In more detail, the linear finite-field symmetric network with interference is described as follows:

- Every input alphabet to a MAC at node v is the finite field, $\mathbb{F}_q$.
- The received symbol at node v, $Y_v^{(t)}$, is determined to be the output of a *symmetric discrete memoryless channel* (DMC) $(\mathbb{F}_q, p(y_v|x_v), \mathcal{Y}_v)$ with input

$$X_v^{(t)} = \sum_{u \in \Delta(v)} \beta_{u,v} X_{u,v}^{(t)},$$

  where $\beta_{u,v} \in \mathbb{F}_q \setminus \{0\}$ denotes the channel coefficient. For the definition of the symmetric DMC, see [28, Sec. 4.5].
- The input field size q and channel transition function $p(y_v|x_v)$ associated with node v need not be identical.

A major characteristic of the symmetric DMC is that linear codes can achieve the capacity [28, Sec. 6.2]. Using this, Nazer and Gastpar [11] showed that the *computation capacity* for any linear function of sources can be achieved in the linear finite-field symmetric MAC in Fig. 5. Also, in [12], [13], it was shown that linear codes achieve the multicast capacity of linear finite-field additive noise and erasure networks with interference, which are special cases of the class of networks stated above. Extending this line, we characterize the multicast capacity of the linear finite-field symmetric network with interference.

Theorem 4: The multicast capacity of a linear finite-field symmetric network with interference is given by

$$\min_{S \in \Gamma} \sum_{v \in \bar{S}^c} C_v,$$

where C_v is the capacity of the channel $(\mathbb{F}_q, p(y_v|x_v), \mathcal{Y}_v)$.

The proof of Theorem 4 is very similar to the proof of Theorem 1. The difference is that we use linear codes instead of the nested lattice codes. We show the outline of the proof in the next subsections.

Remark 4: The capacity proof for linear finite-field additive noise networks in [12] can also be extended to the linear finite-field symmetric networks in Theorem 4. However, the proof in [12] relies on algebraic network coding, and, thus, it has a restriction on the field size, i.e., $q > |D|$. In our proof, we do not use the algebraic network coding, and the field size is not restricted.

A. Upper bound

As in the Gaussian case in Section III-A, the upper bound follows from the relaxed cut-set bound (6). In particular, for the linear finite-field symmetric network with interference, we have the Markov chain relation $(X_{\bar{S},\bar{S}^c}, X_{S^c,V}) \to X_{\bar{S}^c} \to Y_{\bar{S}^c}$, where $X_{\bar{S}^c} = \{X_v : v \in \bar{S}^c\}$. Using the data processing inequality, we have

$$\begin{aligned} I(X_{\bar{S},\bar{S}^c}; Y_{\bar{S}^c} | X_{S^c,V}) &\leq I(X_{\bar{S}^c}; Y_{\bar{S}^c} | X_{S^c,V}) \\ &\leq I(X_{\bar{S}^c}; Y_{\bar{S}^c}). \end{aligned}$$

Thus the upper bound is given by

$$\begin{aligned} R &\leq \min_{S\in\Gamma} \max_{p(x_{V,V})} I(X_{\bar{S},\bar{S}^c}; Y_{\bar{S}^c} | X_{S^c,V}) \\ &\leq \min_{S\in\Gamma} \max_{p(x_{V,V})} I(X_{\bar{S}^c}; Y_{\bar{S}^c}) \\ &= \min_{S\in\Gamma} \sum_{v\in\bar{S}^c} C_v. \end{aligned}$$

B. Achievability

Let us denote the vectors of channel input and output of the symmetric DMC $(\mathbb{F}_q, p(y_v|x_v), \mathcal{Y}_v)$ as $\mathbf{X}_v = \left[X_v^{(1)}, \ldots, X_v^{(n)}\right]^T$ and $\mathbf{Y}_v = \left[Y_v^{(1)}, \ldots, Y_v^{(n)}\right]^T$, respectively. Without loss of generality, we assume that the encoder input is given by a uniform random vector $\mathbf{W}_v \in \mathbb{F}_q^{\lfloor nR'_v \rfloor}$ for some $R'_v \leq 1$. Then we have the following lemma related to linear coding for the DMC.

Lemma 7 (Lemma 3 of [11]): For the symmetric DMC $(\mathbb{F}_q, p(y_v|x_v), \mathcal{Y}_v)$, a sequence of matrices $\mathbf{F}_v \in \mathbb{F}_q^{n\times\lfloor nR'_v \rfloor}$ and associated decoding function $g_v(\cdot)$ exist such that when $\mathbf{X}_v = \mathbf{F}_v \mathbf{W}_v$, $\Pr\{g(\mathbf{Y}_v) \neq \mathbf{W}_v\} \leq \epsilon$ for any $\epsilon > 0$ and n large enough if $R_v \triangleq R'_v \log q < C_v$.

We now consider linear encoding for nodes in the network. We let

$$\mathbf{X}_{u,v} = \beta_{u,v}^{-1} \mathbf{F}_v \mathbf{W}_{u,v},$$

and thus,

$$\mathbf{X}_v = \sum_{u\in\Delta(v)} \beta_{u,v} \mathbf{X}_{u,v} = \mathbf{F}_v \mathbf{T}_v,$$

where

$$\mathbf{T}_v \triangleq \sum_{u\in\Delta(v)} \mathbf{W}_{u,v}. \tag{23}$$

By Lemma 7, a linear code with sufficiently large dimension exists such that node v can recover $\mathbf{T}_v$ with an arbitrarily small error probability if $R_v < C_v$. Now, we can do the same as in Section III-D with (23) replacing (15), and the achievability part follows.

V. Conclusion

In this paper, we considered the multicast problem for relay networks with interference and examined roles of some structured codes for the networks. Initially, we showed that nested lattice codes can achieve the multicast capacity of Gaussian relay networks with interference within a constant gap determined by the network topology. We also showed that linear codes achieve the multicast capacity of linear finite-field symmetric networks with interference. Finally, we should note that this work is an intermediate step toward more general networks. As an extension to multiple source networks, we showed that the same lattice coding scheme considered in this work can achieve the capacity of the Gaussian two-way relay channel within $\frac{1}{2}$ bit [15], [17]. As another direction of extension, we can consider applying structured codes to networks with non-orthogonal broadcast channels. There is a recent work on the interference channel [26] which is related to this issue.

APPENDIX

A. Proof of Theorem 2

Consider a lattice (more precisely, a sequence of lattices) Λ_1^n with $\sigma^2(\Lambda_1^n) = P_1$, which is simultaneously Rogers-good and Poltyrev-good (simultaneously good shortly). In [20], it was shown that such a lattice always exists. Then, by the argument in [24], we can find a fine lattice Λ_2^n such that $\Lambda_1^n \subseteq \Lambda_2^n$ and Λ_2^n is also simultaneously good. We let the partitioning ratio be

$$\left(\frac{\mathrm{Vol}(\Lambda_1^n)}{\mathrm{Vol}(\Lambda_2^n)}\right)^{\frac{1}{n}} = \left(\frac{P_1}{P_2 - \delta'}\right)^{\frac{1}{2}} \left(\frac{1}{2\pi e G(\Lambda_1^n)}\right)^{\frac{1}{2}} \tag{24}$$

for some $\delta' > 0$. Since the partitioning ratio can approach an arbitrary value as n tends to infinity, for any $\delta > 0$, n' exists such that we can choose $\delta' \leq \delta$ when $n \geq n'$. We now have

$$\begin{aligned}\sigma^2(\Lambda_2^n) &= G(\Lambda_2^n) \cdot \mathrm{Vol}(\Lambda_2^n)^{\frac{2}{n}} \\ &= G(\Lambda_2^n) \cdot 2\pi e(P_2 - \delta'),\end{aligned}$$

where the second equality follows from (24). Since Λ_2^n is Rogers-good, n'' exists such that $1 \leq 2\pi e G(\Lambda_2^n) \leq \frac{P_2}{P_2 - \delta'}$, for $n \geq n''$. Thus, for $n \geq \max\{n', n''\}$, we have

$$P_2 - \delta \leq \sigma^2(\Lambda_2^n) \leq P_2.$$

By repeating the same procedure, we obtain a lattice chain $\Lambda_1^n \subseteq \Lambda_2^n \subseteq \cdots \subseteq \Lambda_K^n$, where Λ_i^n, $1 \leq i \leq K$, are simultaneously good and $P_i - \delta \leq \sigma^2(\Lambda_i^n) \leq P_i$ for sufficiently large n.

Moreover, by Theorem 5 of [19], if Λ_K^n is simultaneously good, a Poltyrev-good lattice Λ_C^n exists such that $\Lambda_K^n \subseteq \Lambda_C^n$ and the coding rate R_K can be arbitrary as $n \to \infty$, i.e.,

$$R_K = \frac{1}{n} \log \left(\frac{\mathrm{Vol}(\Lambda_K^n)}{\mathrm{Vol}(\Lambda_C^n)}\right) = \gamma + o_n(1).$$

Given R_K, the coding rates R_i, $1 \leq i \leq K-1$, are given by

$$\begin{aligned}R_i &= \frac{1}{n} \log \left(\frac{\mathrm{Vol}(\Lambda_i^n)}{\mathrm{Vol}(\Lambda_C^n)}\right) \\ &= \frac{1}{n} \log \left(\frac{\mathrm{Vol}(\Lambda_i^n)}{\mathrm{Vol}(\Lambda_K^n)}\right) + R_K \\ &= \frac{1}{2} \log \left(\frac{\sigma^2(\Lambda_i^n)}{\sigma^2(\Lambda_K^n)}\right) + R_K + o_n(1) \\ &= \frac{1}{2} \log \left(\frac{P_i}{P_K}\right) + R_K + o_n(1),\end{aligned}$$

where the third equality follows by the fact that Λ_i^n and Λ_K^n are both Rogers-good, and the fourth follows by the fact that $\sigma^2(\Lambda_i^n) = P_i - o_n(1)$. $\square$

B. Proof of Theorem 3

Let r_i^{cov} and r_i^{eff} denote the covering and effective radii of Λ_i, respectively. Then the second moment per dimension of $r_i^{\mathrm{cov}}\mathcal{B}$ is given by

$$\sigma_i^2 \triangleq \sigma^2(r_i^{\mathrm{cov}}\mathcal{B}) = \frac{(r_i^{\mathrm{cov}})^2}{n+2}.$$

Next, we define independent Gaussian random variables

$$\mathbf{Z}_i \sim \mathcal{N}(\mathbf{0}, \sigma_i^2\mathbf{I}),\ i = 1, \ldots, K,$$

and

$$\mathbf{Z}^* = (1-\alpha)\sum_{j=1}^{K}\mathbf{Z}_j + \alpha\mathbf{Z}.$$

Then, we have the following lemmas.

Lemma 8: The variance of Z^*, each element of $\mathbf{Z}^*$, is denoted by $\mathrm{Var}(Z^*)$ and satisfies

$$\begin{aligned}\mathrm{Var}(Z^*) &= (1-\alpha)^2\sum_{j=1}^{K}\sigma_j^2 + \alpha^2 \\ &\leq \max_j \left(\frac{r_j^{\mathrm{cov}}}{r_j^{\mathrm{eff}}}\right)^2 \cdot \frac{\sum_{j=1}^{K} P_j}{\sum_{j=1}^{K} P_j + 1}.\end{aligned}$$

Lemma 9: The pdf of $\tilde{\mathbf{Z}}$, denoted by $p_{\tilde{\mathbf{Z}}}(\mathbf{x})$ satisfies

$$p_{\tilde{\mathbf{Z}}}(\mathbf{x}) \leq e^{n\sum_{j=1}^{K}\epsilon_j} \cdot p_{\mathbf{Z}^*}(\mathbf{x}),$$

where

$$\epsilon_j = \log\left(\frac{r_j^{\mathrm{cov}}}{r_j^{\mathrm{eff}}}\right) + \frac{1}{2}\log 2\pi e G(\mathcal{B}) + \frac{1}{n}.$$

The above two lemmas are slight modifications of Lemmas 6 and 11 in [19]. The proofs also follow immediately from [19].

Now, we bound the error probability by

$$\begin{aligned}p_e &= \Pr\left\{\tilde{\mathbf{Z}} \bmod \Lambda_1 \notin \mathcal{R}_C\right\} \\ &\leq \Pr\left\{\tilde{\mathbf{Z}} \notin \mathcal{R}_C\right\} \\ &\leq e^{n\sum_{j=1}^{K}\epsilon_j} \cdot \Pr\left\{\mathbf{Z}^* \notin \mathcal{R}_C\right\}, \qquad (25)\end{aligned}$$

where (25) follows from Lemma 9. Note that $\mathbf{Z}^*$ is a vector of i.i.d. zero-mean Gaussian random variables, and the VNR of Λ_C relative to $\mathbf{Z}^*$ is given by

$$\begin{aligned}\mu &= \frac{(\mathrm{Vol}(\Lambda_C))^{2/n}}{2\pi e\mathrm{Var}(Z^*)} \\ &\geq \frac{(\mathrm{Vol}(\Lambda_1))^{2/n}/2^{2R_1}}{2\pi e\cdot\frac{\sum_{j=1}^{K}P_j}{\sum_{j=1}^{K}P_j+1}} - o_n(1) \qquad (26) \\ &= \frac{1}{2^{2R_1}}\cdot\frac{1}{2\pi e G(\Lambda_1)}\cdot\left(\frac{P_1}{\sum_{j=1}^{K}P_j} + P_1\right) - o_n(1) \qquad (27) \\ &= \frac{1}{2^{2\bar{R}_1}}\cdot\left(\frac{P_1}{\sum_{j=1}^{K}P_j} + P_1\right) - o_n(1), \qquad (28)\end{aligned}$$

where (26) follows from Lemma 8 and the fact that Λ_i, $1 \leq i \leq K$, are Rogers-good, (27) from the definition of $G(\Lambda_1)$, and (28) from the fact that Λ_1 is Rogers-good and $R_1 = \bar{R}_1 + o_n(1)$. When we consider the Poltyrev exponent, we are only interested in the case that $\mu > 1$. Thus, from the definition of R_1^* and (28), we can write

$$\mu = 2^{2(R_1^* - \bar{R}_1)} - o_n(1),$$

for $\bar{R}_1 < R_1^*$. Finally, from (25) and by the fact that Λ_C is Poltyrev-good, we have

$$\begin{aligned} p_e &\leq e^{n\sum_{j=1}^{K}\epsilon_j} \cdot e^{-nE_P(\mu)} \\ &= e^{-n\left(E_P\left(2^{2(R_1^*-\bar{R}_1)}\right)-o_n(1)\right)}. \end{aligned}$$

□

C. Proof of Lemma 3

For notational simplicity, we prove this lemma in the standard MAC in Section III-C. We assume that the uniform random mapping is done at each input node of the standard MAC, as was done in the network. Let A and A^c be nonempty partitions of $\{1,\ldots,K\}$, i.e., $A\cup A^c=\{1,\ldots,K\}$, and $A\cap A^c=\emptyset$. We assume that A implies the set of nodes that can distinguish W and W', and A^c implies the set of nodes that cannot. For node $i\in A$, $\mathbf{W}_i(W)$ and $\mathbf{W}_i(W')$ are uniform over $\mathcal{C}_i$ and independent of each other due to the uniform random mapping. However, for node $i\in A^c$, we always have $\mathbf{W}_i(W)=\mathbf{W}_i(W')$. Thus, if $A=\emptyset$, $\mathbf{T}(W)=\mathbf{T}(W')$ always holds, i.e.,

$$\Pr\left\{\mathbf{T}(W)=\mathbf{T}(W')|\mathcal{D}=A,\bar{\mathcal{D}}=A^c\right\}=1.$$

If $A\neq\emptyset$, given $\mathcal{D}=A$ and $\bar{\mathcal{D}}=A^c$, the event $\mathbf{T}(W)=\mathbf{T}(W')$ is equivalent to $\tilde{\mathbf{T}}(W)=\tilde{\mathbf{T}}(W')$, where

$$\tilde{\mathbf{T}}(W)=\left[\sum_{j\in A}\left(\mathbf{W}_j(W)-Q_j(\mathbf{W}_j(W)+\mathbf{U}_j)\right)\right]\bmod\Lambda_1,$$

and $\tilde{\mathbf{T}}(W')$ is given accordingly. Now, let $l\triangleq\min A$, then

$$\begin{aligned}\mathbf{T}'(W)&\triangleq\tilde{\mathbf{T}}(W)\bmod\Lambda_l\\ &=\left[\mathbf{W}_l(W)+\sum_{\substack{j\in A\\ \setminus\{l\}}}\left(\mathbf{W}_j(W)-Q_j(\mathbf{W}_j(W)+\mathbf{U}_j)\right)\right]\bmod\Lambda_l,\end{aligned}$$

which follows from the fact that $\Lambda_1\subseteq\Lambda_l$, and ,thus, $(\mathbf{x}\bmod\Lambda_1)\bmod\Lambda_l=\mathbf{x}\bmod\Lambda_l$. Note that, due to the crypto-lemma and the uniform random mapping, $\mathbf{T}'(W)$ and $\mathbf{T}'(W')$ are uniform over $\mathcal{C}_l$ and independent of each other. Therefore,

$$\begin{aligned}\Pr\left\{\mathbf{T}(W)=\mathbf{T}(W')|\mathcal{D}=A,\bar{\mathcal{D}}=A^c\right\}&=\Pr\left\{\tilde{\mathbf{T}}(W)=\tilde{\mathbf{T}}(W')|\mathcal{D}=A\right\}\\ &\leq\Pr\left\{\mathbf{T}'(W)=\mathbf{T}'(W')|\mathcal{D}=A\right\}\\ &=\frac{1}{|\mathcal{C}_l|}=2^{-nR_l}.\end{aligned}$$

Thus, by changing notations properly to those of the network, we complete the proof. □

D. Proof of Lemma 4

In the time-expanded network, there are two types of cuts, steady cuts and wiggling cuts [5]. The steady cut separates the nodes in different layers identically. That is, for a steady cut S_{TE}, $v[k]\in S_{\mathrm{TE}}$ for some k if and only if $v[1],\ldots,v[L+1]\in S_{\mathrm{TE}}$. Let us denote the set of all steady cuts as $\tilde{\Gamma}_{\mathrm{TE}}$. Then, since $\tilde{\Gamma}_{\mathrm{TE}}\subseteq\Gamma_{\mathrm{TE}}$,

$$\begin{aligned}\min_{S_{\mathrm{TE}}\in\Gamma_{\mathrm{TE}}}\sum_{k=2}^{L+1}\sum_{\substack{v[k]\in\\ \bar{S}^c_{\mathrm{TE}}[k]}}\left(\max_{\substack{u[k-]\in\\ \Delta_{\mathrm{TE},S}(v[k])}}R_{u,v}\right)&\leq\min_{S_{\mathrm{TE}}\in\tilde{\Gamma}_{\mathrm{TE}}}\sum_{k=2}^{L+1}\sum_{\substack{v[k]\in\\ \bar{S}^c_{\mathrm{TE}}[k]}}\left(\max_{\substack{u[k-]\in\\ \Delta_{\mathrm{TE},S}(v[k])}}R_{u,v}\right)\\ &=L\cdot\min_{S\in\Gamma}\sum_{v\in\bar{S}^c}\left(\max_{u\in\Delta_S(v)}R_{u,v}\right).\end{aligned}$$

We now prove the lower bound. For any two cuts S_1 and S_2 in $\mathcal{G}$, i.e., $S_1, S_2 \in \Gamma$, define that

$$\xi(S_1, S_2) = \sum_{v \in S_2^c} \left(\max_{u \in S_1} R_{u,v} \right),$$

where $R_{u,v} = 0$ if $(u, v) \notin E$. Then, we have the following lemma

Lemma 10: Consider a sequence of non-identical cuts $S_1, \ldots, S_{L'} \in \Gamma$ and define $S_{L'+1} = S_1$. For the sequence, we have

$$\sum_{k=1}^{L'} \xi(S_k, S_{k+1}) \geq \sum_{k=1}^{L'} \xi(S'_k, S'_k),$$

where for $k = 1, \ldots, L'$,

$$S'_k = \bigcup_{\substack{\{i_1,\ldots,i_k\} \subseteq \\ \{1,\ldots,L'\}}} (S_{i_1} \cap \cdots \cap S_{i_k}).$$

The proof of Lemma 10 is tedious but straightforward. Similar lemmas were presented and proved in [5, Lemma 6.4], [8, Lemma 2], and the proof of Lemma 10 also follows similarly.

Now, since $S'_k \in \Gamma$, it follows that

$$\begin{aligned} \xi(S'_k, S'_k) &\geq \min_{S \in \Gamma} \sum_{v \in S^c} \left(\max_{u \in S} R_{u,v} \right) \\ &= \min_{S \in \Gamma} \sum_{v \in \bar{S}^c} \left(\max_{\substack{u \in \\ \Delta_S(v)}} R_{u,v} \right). \end{aligned} \tag{29}$$

Also, since $S_{\mathrm{TE}}[k]$'s correspond to cuts in V, we can rewrite

$$\min_{S_{\mathrm{TE}} \in \Gamma_{\mathrm{TE}}} \sum_{k=2}^{L+1} \sum_{\substack{v[k] \in \\ \bar{S}^c_{\mathrm{TE}}[k]}} \left(\max_{\substack{u[k^-] \in \\ \Delta_{\mathrm{TE},S}(v[k])}} R_{u,v} \right) = \min_{S_{\mathrm{TE}} \in \Gamma_{\mathrm{TE}}} \sum_{k=2}^{L+1} \xi\left(S_{\mathrm{TE}}[k^-], S_{\mathrm{TE}}[k]\right).$$

Since there are $|\Gamma| = 2^{|V|-2}$ different cuts, at least the first $L - |\Gamma| + 2$ of the sequence $S_{\mathrm{TE}}[1], \ldots, S_{\mathrm{TE}}[L+1]$ form loops, and, thus, by Lemma 10 and (29), we have

$$\min_{S_{\mathrm{TE}} \in \Gamma_{\mathrm{TE}}} \sum_{k=2}^{L+1} \xi\left(S_{\mathrm{TE}}[k^-], S_{\mathrm{TE}}[k]\right) \geq (L - |\Gamma| + 2) \cdot \min_{S \in \Gamma} \sum_{v \in \bar{S}^c} \left(\max_{\substack{u \in \\ \Delta_S(v)}} R_{u,v} \right).$$

$\square$

E. Proof of Lemma 5

We first consider the case that $1 \in A$, and the case that $1 \notin A$ afterward.

a) $1 \in A$

In this case, $l = 1$, and the gap is

$$\begin{aligned} &\frac{1}{2} \log \left(1 + \left(\sum_{j \in A} \sqrt{P_j} \right)^2 \right) - \left[\frac{1}{2} \log \left(\left(\frac{1}{\sum_{j=1}^K P_j} + 1 \right) P_1 \right) \right]^+ \\ &\leq \frac{1}{2} \log \left(1 + \left(\sum_{j=1}^K \sqrt{P_j} \right)^2 \right) - \frac{1}{2} \log \left(\left(\frac{1}{\sum_{j=1}^K P_j} + 1 \right) P_1 \right) \\ &\leq \frac{1}{2} \log \left(1 + K^2 P_1 \right) - \frac{1}{2} \log \left(\frac{1}{K} + P_1 \right) \\ &\leq \log K. \end{aligned}$$

b) $1 \notin A$

Since $1 \notin A$, $|A| \leq K-1$. Now, the gap is given by

$$\begin{aligned}
&\frac{1}{2}\log\left(1+\left(\sum_{j\in A}\sqrt{P_j}\right)^2\right)-\left[\frac{1}{2}\log\left(\left(\frac{1}{\sum_{j=1}^{K}P_j}+1\right)P_l\right)\right]^+\\
&\leq \frac{1}{2}\log\left(1+(K-1)^2P_l\right)-\left[\frac{1}{2}\log P_l\right]^+\\
&\leq \frac{1}{2}\log(1+(K-1)^2)\\
&\leq \log K.
\end{aligned}$$

$\square$

REFERENCES

[1] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 5, pp. 572–584, Sep. 1979.

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

[3] M. R. Aref, "Information flow in relay networks," Ph.D. dissertation Stanford Univ., Stanford, CA, 1980.

[4] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2425–2432, June 2006.

[5] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow," in *Proc. 45th Annual Allerton Conference*, Sept. 2007.

[6] ——, "Approximate capacity of Gaussian relay networks," in *Proc. IEEE International Symp. Inform. Theory*, Toronto, Canada, July 2008.

[7] A. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," in *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 789–804, Mar. 2006.

[8] B. Smith and S. Vishwanath, "Unicast transmission over multiple access erasure networks: Capacity and duality," in *IEEE Information Theory Workshop*, Tahoe city, California, Sept. 2007.

[9] B. Nazer and M. Gastpar, "Computing over multiple-access channels with connections to wireless network coding," in *Proc. IEEE International Symp. Inform. Theory*, Seattle, USA, July 2006.

[10] ——, "Lattice coding increases multicast rates for Gaussian multiple-access networks," in *Proc. 45th Annual Allerton Conference*, Sept. 2007.

[11] ——, "Computation over multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.

[12] ——, "The case for structured random codes in network capacity theorems," *European Trans. Telecomm.: Special Issue on New Directions in Inform. Theory*, no. 4, vol. 19, pp. 455–474, June 2008.

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

[14] A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Inform. Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.

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

[16] K. Narayanan, M. P. Wilson, and A. Sprintson, "Joint physical layer coding and network coding for bi-directional relaying," in *Proc. 45th Annual Allerton Conference*, Sept. 2007.

[17] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian Two-way Relay Channel to within $\frac{1}{2}$ Bit," submitted to *IEEE Trans. Inform. Theory*, available at http://arxiv.org/PS_cache/arxiv/pdf/0902/0902.2438v1.pdf.

[18] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1250–1276, June 2002.

[19] 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. 50, no. 10, pp. 2293–2314, Oct. 2004.

[20] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3401–3416, Oct. 2005.

[21] H. A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1767–1773, Nov. 1997.

[22] G. Poltyrev, "On coding without restrictions for the AWGN channel," *IEEE Trans. Inform. Theory*, vol. 40, no. 2, pp. 409–417, Mar. 1994.

[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 Annual Allerton Conference*, Oct. 2003.

[24] D. Krithivasan and S. S. Pradhan, "A proof of the existence of good nested lattices," available at http://www.eecs.umich.edu/techreports/systems/cspl/cspl-384.pdf.

[25] T. Philosof, A. Khisti, U. Erez, and R. Zamir, “Lattice strategies for the dirty multiple access channel,” in *Proc. IEEE International Symp. Inform. Theory*, Nice, France, June-July 2007.

[26] S. Sridharan, A. Jafarian, S. Vishwanath, S. A. Jafar, and S. Shamai, “A layered lattice coding scheme for a class of three user Gaussian interference channels,” available at http://arxiv.org/PS_cache/arxiv/pdf/0809/0809.4316v1.pdf.

[27] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[28] R. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.