Capacity Optimality of Lattice Codes in Common Message Gaussian Broadcast Channels with Coded Side Information

Lakshmi Natarajan, Yi Hong, and Emanuele Viterbo

***Abstract*—Lattices possess elegant mathematical properties which have been previously used in the literature to show that structured codes can be efficient in a variety of communication scenarios. We consider the family of single-transmitter multiple-receiver Gaussian channels where the source transmits a set of common messages to all the receivers (multicast scenario), and each receiver has *coded side information*, i.e., prior information in the form of linear combinations of the messages. This channel model is motivated by applications to multi-terminal networks where the nodes may have access to coded versions of the messages from previous signal hops or through orthogonal channels. The capacity of this channel is known and follows from the work of Tuncel (2006), which is based on random coding arguments. In this paper, following the approach introduced by Erez and Zamir, we show that lattice codes are capacity-optimal for this family of channels. The structured coding scheme proposed in this paper is derived from Construction A lattices designed over prime fields, and utilizes *algebraic binning* at the decoders to expurgate the channel code and obtain good lattice subcodes, for every possible set of linear combinations available as side information.**

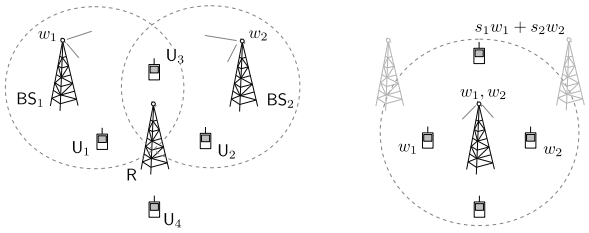
I. Introduction

Information-theoretic results often rely on random coding arguments to prove the existence of good codes. Usually, the codebook is constructed by randomly choosing the components of each codeword independently and identically from a judiciously chosen probability distribution. While this technique is powerful, the resulting codebooks do not exhibit any structure that may be of practical interest. One such desirable structure is linearity, which allows complexity reductions at the encoder and decoder by utilizing efficient algebraic processing techniques. Further, in certain communication scenarios, coding schemes based on linear codes yield a larger achievable rate region than random code ensembles, as was shown by Körner and Marton [1] for a distributed source coding problem. Structured coding schemes, such as those based on linear codes and lattices, have been widely studied in the literature, see [2], [3] and references therein. Lattice codes, in particular, are known to be efficient for communication in the additive white Gaussian noise (AWGN) channel [4]–[8], dirty-paper coding [2], [9], Wyner-Ziv coding [2] and relay networks [10]–[13], to name only a few.

Dr. Natarajan is with the Department of Electrical Engineering, Indian Institute of Technology Hyderabad, email: lakshminatarajan@iith.ac.in.

Dr. Hong and Prof. Viterbo are with the Department of Electrical and Computer Systems Engineering, Monash University, Clayton, VIC 3600, Australia, email:{yi.hong,emanuele.viterbo}@monash.edu.

This work was supported by the Australian Research Council under Grant Discovery Project No. DP160101077, and the Department of Science and Technology, Government of India, through INSPIRE research grant DST/INSPIRE/04/2015/002094.



(a) Multiple-access phase (b) Multicast phase

Fig. 1. A relay network instance where one encounters a common message broadcast channel with coded side information at the receivers.

In this paper we present capacity-achieving lattice strategies for communication in common message Gaussian broadcast channels, which we refer to as the *multicast channel*, where receivers have prior side information about the messages being transmitted. In particular, we assume that the transmitter is multicasting K message symbols $w_1, \ldots, w_K$ from a finite field $\mathbb{F}_p$, of prime size p, to all the receivers, and each receiver may have *coded side information* about the messages: the prior knowledge of the values of (possibly multiple) $\mathbb{F}_p$-linear combinations of $w_1, \ldots, w_K$. The number of linear combinations available as side information and the coefficients of these linear combinations can differ from one receiver to the next. The capacity of this channel is known and follows from the results of Tuncel [14], where the achievability part utilizes an ensemble of codebooks generated using Gaussian distribution. The channel model considered in this paper is motivated by applications in multi-terminal communication networks.

Example 1. Consider a wireless network with two base stations BS_1 and BS_2, that hold message symbols w_1 and w_2, respectively. The base stations are required to multicast w_1 and w_2 to four user nodes $\mathsf{U}_1, \ldots, \mathsf{U}_4$ through the relay node R, see Fig. 1. In the first phase of the protocol, BS_1 and BS_2 encode the data symbols w_1 and w_2, and transmit the resulting codewords simultaneously. In order to exploit the resulting signal interference at U_3, the base stations employ the encoding technique used in the compute-and-forward protocol [12]. Now, U_3 reliably decodes some linear combination $s_1w_1 + s_2w_2$, $s_1, s_2 \in \mathbb{F}_p$, from the received noisy superposition of the two transmit signals. On the other hand, R has a higher signal-to-noise ratio and successfully decodes both w_1 and w_2 by behaving as a multiple-access receiver. Further, there is no signal interference at U_1 and U_2, and these two nodes reliably decode w_1 and w_2, respectively. We observe that the second phase of the protocol is a common message broadcast channel

978-1-5090-4096-4/17/$31.00 ©2017 IEEE

with coded side information at the receivers: the relay needs to multicast w_1, w_2 to four user nodes, the first three users $\mathsf{U}_1, \mathsf{U}_2, \mathsf{U}_3$ have prior knowledge of the linear combinations $w_1 + 0w_2$, $0w_1 + w_2$ and $s_1 w_1 + s_2 w_2$, respectively, while the fourth user has no such side information. ■

Example 2. Assume a network of noiseless wired links in the form of a directed acyclic graph, where the source node v_{s} desires to multicast K independent messages $w_1, \ldots, w_K \in \mathbb{F}_p$ to a set of destination nodes. The wireline network employs a traditional *(scalar) linear network code* [15]–[17]. It is known that the maximum number of linearly-independent combinations of the K messages that can be made available at a destination node v_{d} is $\min\{\text{max-flow}(v_{\mathsf{d}}), K\}$, where $\text{max-flow}(v_{\mathsf{d}})$ is the maximum number of edge-disjoint paths from v_{s} to v_{d}, see [17]. Hence multicasting is possible if and only if $\text{max-flow}(v_{\mathsf{d}}) \geq K$ for every destination v_{d}. Now suppose there exist destination nodes with max-flow less than K, i.e., the communication demands are beyond the wireline network's capacity. A solution to meet the demands is to broadcast a wireless signal from the source to fill the capacity deficiency of the wired network. At each destination, the $\mathbb{F}_p$-linear combinations obtained from the wireline network serve as side information to decode the wireless broadcast signal. ■

A special case of coded side information is the Gaussian multicast channel where each receiver has prior knowledge of the values of some subset of the K messages. The known capacity-achieving coding schemes for this special case are based on random coding using i.i.d. (independent and identically distributed) codewords [14], [18]–[22]. Existence of lattice based capacity-achieving coding schemes were proved in [19], [23] for the special case where the number of messages and receivers are two and each receiver has the knowledge of one of the messages.

The objective of this paper is to prove that lattice codes achieve the capacity of the common message Gaussian broadcast channels with coded side information. We use the information-theoretic framework set by Erez and Zamir [5] to this end. The proposed coding scheme uses an ensemble of Construction A lattices and the decoding scheme involves *algebraic binning* [2] where the receiver side information is used to expurgate the channel code and obtain a lower rate subcode. The algebraic structure of the coding scheme facilitates the performance analysis by decomposing the original channel into multiple independent point-to-point AWGN channels – one corresponding to each receiver – where each of the point-to-point AWGN channels uses a (possibly) different lattice code for communication. Unlike [5], where achievability in a point-to-point AWGN channel was proved using error exponent analysis, our proof technique is based only on simple counting arguments. Full proofs of all the results appearing in this paper are available in [24].

II. Channel Model and Lattice Preliminaries

Channel Model: We consider a (non-fading) common message Gaussian broadcast channel with a single transmitter and finitely many receivers, where all terminals are equipped with single antennas. The K independent messages $w_1, \ldots, w_K$ assume values with a uniform probability distribution from a prime finite field $\mathbb{F}_p$. Each receiver desires to decode all the K messages while having prior knowledge of the values of some $\mathbb{F}_p$-linear combinations of the messages $w_1, \ldots, w_K$. Consider a generic receiver that has access to the values u_m, $m = 1, \ldots, M$, of the following set of M linear equations $\sum_{k=1}^{K} s_{m,k} w_k = u_m$, $m = 1, \ldots, M$. We will denote this side information configuration using the matrix $\boldsymbol{S} = [s_{m,k}] \in \mathbb{F}_p^{M \times K}$, where each row of $\boldsymbol{S}$ represents one linear equation. Any row of $\boldsymbol{S}$ that is linearly dependent on the other rows represents redundant information and can be discarded with no loss to the receiver side information, and hence, with no loss to system performance. Hence, without loss in generality, we will assume that the rows of $\boldsymbol{S}$ are linearly independent over $\mathbb{F}_p$, i.e., $\text{rank}(\boldsymbol{S}) = M$, and $M < K$. If the values u_m of M linearly independent combinations of the variables $w_1, \ldots, w_K$ are given, then the set of all possible solutions of $(w_1, \ldots, w_K)$ is a coset of a $(K - M)$ dimensional linear subspace of $\mathbb{F}_p^K$, and hence, is of cardinality $p^{(K-M)}$.

Note that the values of $\boldsymbol{S}$ and M can be different across the receivers. A receiver is completely characterized by its *(coded) side information matrix* $\boldsymbol{S}$ and the variance σ^2 of the additive noise. If we assume that the average transmit power at the source is 1, then the signal-to-noise ratio at this receiver is $\mathsf{SNR} = 1/\sigma^2$. We will denote a receiver by the tuple $(\boldsymbol{S}, \sigma^2)$. Suppose $(\boldsymbol{S}_1, \sigma_1^2), \ldots, (\boldsymbol{S}_N, \sigma_N^2)$ are the N receivers in the multicast channel. Following [14, Theorem 6], it is straightforward to show that the maximum rate at which each of the K messages can be reliably transmitted, i.e. the (symmetric) capacity of the channel, is

$$C = \min_{i \in \{1,\ldots,N\}} \frac{1}{(K - \text{rank}(\boldsymbol{S}_i))} \frac{1}{2} \log\left(1 + \frac{1}{\sigma_i^2}\right). \quad (1)$$

Lattice Codes from Linear Codes over $\mathbb{F}_p$: The lattice related terminology and notation used in this section is standard material, and is mainly based on [5], [25]–[27]. Given an n-dimensional lattice $\Lambda \subset \mathbb{R}^n$, we denote its fundamental Voronoi region by $\mathcal{V}(\Lambda)$, volume of the Voronoi region by $\text{Vol}(\Lambda)$, the closest lattice point quantizer by Q_Λ, and the modulo-Λ operation by $\mod \Lambda$. Covering and effective radii of Λ are denoted by $r_{\text{cov}}(\Lambda)$ and $r_{\text{eff}}(\Lambda)$, respectively. Note that there exists a sequence of lattices of increasing dimension n, said to be *Rogers good*, such that $r_{\text{cov}}/r_{\text{eff}} \to 1$. The closed n-dimensional ball of radius r centered at $\boldsymbol{s} \in \mathbb{R}^n$ is denoted by $\mathcal{B}(\boldsymbol{s}, r)$, and the volume of the unit ball $\mathcal{B}(\boldsymbol{0}, 1)$ by V_n.

We will rely on the following methodology used in [5], [12] to construct a pair of nested lattices $\Lambda_{\text{c}} \subset \Lambda$. Let $g(\cdot)$ denote the natural map that embeds $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ into $\mathbb{Z}$. Let $\mathscr{C} \subset \mathbb{F}_p^n$ be a linear code of rank L, $1 \leq L \leq n$, $\mathscr{C} = \{\boldsymbol{G}\boldsymbol{w} \,|\, \boldsymbol{w} \in \mathbb{F}_p^L\}$, where $\boldsymbol{G}$ is the $n \times L$ generator matrix with full column rank. The set $g(\mathscr{C}) + p\mathbb{Z}^n$ is the *Construction A* lattice of the linear code $\mathscr{C}$ [25]. We obtain the fine lattice Λ by scaling the Construction A lattice by p^{-1} and transforming

it by the generator matrix $\boldsymbol{B}_{\mathrm{c}}$ of the coarse lattice Λ_{c}

$$\Lambda = \boldsymbol{B}_{\mathrm{c}} p^{-1} \left(g(\mathscr{C}) + p\mathbb{Z}^n \right) = \boldsymbol{B}_{\mathrm{c}} p^{-1} g(\mathscr{C}) + \Lambda_{\mathrm{c}}.$$

The following key property of the (nested) lattice code $\Lambda/\Lambda_{\mathrm{c}} = \Lambda \cap \mathcal{V}(\Lambda_{\mathrm{c}})$ will be useful in proving our main theorem.

Lemma 1. *The map* $\boldsymbol{w} \to \left[\boldsymbol{B}_{\mathrm{c}}\, p^{-1} g\left(\boldsymbol{G}\boldsymbol{w}\right)\right] \mod \Lambda_{\mathrm{c}}$ *is a bijection between* $\mathbb{F}_p^L$ *and* $\Lambda/\Lambda_{\mathrm{c}}$.

III. Main Theorem and the Lattice Coding Scheme

In order to rigorously state the main result, we consider a non-zero tolerance $\epsilon > 0$ that determines the gap to capacity.

Theorem 1 (Main theorem)**.** *Let the number of messages K, design rate R and tolerance $\epsilon > 0$ be given. For every sufficiently large prime integer p, there exists a sequence of lattice codes of increasing dimension n that encode K message vectors over $\mathbb{F}_p$ such that the rate of transmission of each message is at least $(R-\epsilon)$ b/dim and the probability of error at a receiver $(\boldsymbol{S}, \sigma^2)$ decays exponentially in n if*

$$\frac{1}{(K - \mathrm{rank}(\boldsymbol{S}))} \cdot \frac{1}{2} \log\left(1 + \frac{1}{\sigma^2}\right) \geq (R+\epsilon). \tag{2}$$

Note that our coding scheme considers the scenario where the K messages are to be transmitted at the same rate R, resulting in sum rate KR. Combining Theorem 1 with a simple union bound argument, we immediately deduce that lattice codes achieve the symmetric capacity (1) of the common message Gaussian broadcast channel with coded side information if the number of receivers is finite and the prime p is large enough. Also, when $K = 1$ and $\mathrm{rank}(\boldsymbol{S}) = 0$, we observe that the proposed coding scheme is capacity-optimal for the single-user AWGN channel with no receiver side information.

We now describe the construction of lattice code ensemble, encoder and decoder used to prove Theorem 1. A sketch of the proof of Theorem 1 is given in Section IV.

A. Random lattice code ensemble

Given the design rate R, number of messages K and tolerance $\epsilon > 0$, let p be any prime integer satisfying $p \geq \max\left\{2^{2KR}, (2^{\epsilon/4} - 1)^{-1} 2^{-R}\right\}$. Note that this implies $p \geq (2^{\epsilon/4} - 1)^{-1} 2^{-(R+\epsilon)(K-M)}$ for any $0 \leq M \leq K-1$. Rearranging the terms in this inequality we obtain

$$\left(p\, 2^{(R+\epsilon)(K-M)}\right)^{-1} + 1 \leq 2^{\epsilon/4}. \tag{3}$$

Once p is fixed, we choose ℓ as the largest integer that satisfies

$$\ell/n \log p \leq R. \tag{4}$$

The left-hand side in the above inequality is the actual rate at which the lattice code encodes each message, while R is the design rate. The difference between the two is at the most $\log p/n$ which converges to 0 as $n \to \infty$. It follows that the code rate $\ell/n \log_2 p \geq R - \epsilon$ for all sufficiently large n.

We will choose a Rogers-good lattice of sufficiently large dimension n as Λ_{c} with $r_{\mathrm{cov}}(\Lambda_{\mathrm{c}})/r_{\mathrm{eff}}(\Lambda_{\mathrm{c}}) \leq 2^{\epsilon/4}$, and to satisfy the transmit power constraint scale it so that $r_{\mathrm{cov}}(\Lambda_{\mathrm{c}}) = \sqrt{n}$. It follows that $r_{\mathrm{eff}}(\Lambda_{\mathrm{c}}) \geq 2^{-\epsilon/4}\, r_{\mathrm{cov}}(\Lambda_{\mathrm{c}}) = 2^{-\epsilon/4}\sqrt{n}$. Using the definition of the effective radius,

$$\mathrm{Vol}(\Lambda_{\mathrm{c}}) = V_n\, r_{\mathrm{eff}}^n(\Lambda_{\mathrm{c}}) \geq V_n\, n^{n/2}\, 2^{-n\epsilon/4}. \tag{5}$$

To construct the fine lattice Λ, we consider a linear code $\mathscr{C}$ of length n and rank $L = K\ell$, which is the number of message symbols to be encoded by the lattice code. Note that this requires that $K\ell < n$ be true. Using (4) and the property $p \geq 2^{2KR}$, we have $K\ell \leq nKR/\log p \leq nKR/2KR = n/2$, which ensures that $K\ell < n$. We then set $\Lambda = \boldsymbol{B}_{\mathrm{c}} p^{-1} g(\mathscr{C}) + \Lambda_{\mathrm{c}}$. We will choose $\boldsymbol{G}$ uniformly random over the set of all $n \times K\ell$ matrices of $\mathbb{F}_p$, resulting in a random ensemble of fine lattices Λ. Probability that $\boldsymbol{G}$ is not full-rank is at the most $p^{-(n-K\ell)}$ [12]. This in turn is upper bounded by $2^{-n/2}$ since $K\ell \leq n/2$ and $p \geq 2$.

Finally, we will assume that the dither $\boldsymbol{d}$ is distributed uniformly in $\mathcal{V}(\Lambda_{\mathrm{c}})$ and is chosen independently of $\boldsymbol{G}$.

B. Encoding

The encoder first concatenates K length-ℓ message vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K \in \mathbb{F}_p^\ell$ into the vector $\boldsymbol{w} = \left(\boldsymbol{w}_1^{\mathsf{T}}, \cdots, \boldsymbol{w}_K^{\mathsf{T}}\right)^{\mathsf{T}}$, maps it to a point $\boldsymbol{t} \in \mathbb{R}^n$ using Construction A as follows

$$\boldsymbol{t} = \left[\boldsymbol{B}_{\mathrm{c}}\, p^{-1} g\left(\boldsymbol{G}\boldsymbol{w}\right)\right] \mod \Lambda_{\mathrm{c}}. \tag{6}$$

From the discussion in Section III-A, we know that $\boldsymbol{B}_{\mathrm{c}} p^{-1} g(\boldsymbol{G}\boldsymbol{w}) \in \Lambda$, and hence, $\boldsymbol{t} \in \Lambda/\Lambda_{\mathrm{c}}$, i.e., the lattice code $\Lambda \cap \mathcal{V}(\Lambda_{\mathrm{c}})$. The transmit codeword $\boldsymbol{x}$ is generated by dithering,

$$\boldsymbol{x} = [\boldsymbol{t} - \boldsymbol{d}] \mod \Lambda_{\mathrm{c}} = \left[\boldsymbol{B}_{\mathrm{c}}\, p^{-1} g\left(\boldsymbol{G}\boldsymbol{w}\right) - \boldsymbol{d}\right] \mod \Lambda_{\mathrm{c}}. \tag{7}$$

Since $r_{\mathrm{cov}}(\Lambda_{\mathrm{c}}) = \sqrt{n}$, the power constraint $\|\boldsymbol{x}\|^2 \leq r_{\mathrm{cov}}^2(\Lambda_{\mathrm{c}}) = n$ is satisfied. From Lemma 1, no two distinct message tuples are mapped to the same codeword if $\boldsymbol{G}$ is full-rank. We already observed in Section III-A that the probability that $\boldsymbol{G}$ is not full-rank is exponentially small in n.

C. Decoding

The receiver employs a two stage decoder: in the first stage the receiver identifies the subcode of the lattice code corresponding to the available side information. And in the second stage it decodes the channel output to a point in this subcode using the standard MMSE scaling and infinite lattice decoding [5].

Expurgation using Side Information: The side information at $(\boldsymbol{S}, \sigma^2)$ over a block of ℓ realizations of the K messages is of the form $\sum_{k=1}^{K} s_{m,k} \boldsymbol{w}_k = \boldsymbol{u}_m$, $m = 1, \ldots, M$. Using the notation $\boldsymbol{u} = \left(\boldsymbol{u}_1^{\mathsf{T}}, \cdots, \boldsymbol{u}_M^{\mathsf{T}}\right)^{\mathsf{T}} \in \mathbb{F}_p^{M\ell}$, the side information can be rewritten compactly as

$$\left(\boldsymbol{S} \otimes \boldsymbol{I}_\ell\right) \boldsymbol{w} = \boldsymbol{u}, \tag{8}$$

where $\otimes$ denotes the Kronecker product of matrices and $\boldsymbol{I}_\ell$ is the $\ell \times \ell$ identity matrix over $\mathbb{F}_p$. Let $\boldsymbol{A_S}$ be any $K\ell \times (K-M)\ell$ matrix whose columns form a basis of the null space of $\boldsymbol{S} \otimes \boldsymbol{I}_\ell$. Observe that (8) is an under-determined system of linear equations, and the set of solutions is a coset of the null space of $\boldsymbol{S} \otimes \boldsymbol{I}_\ell$. The set of all solutions to (8)

is $\boldsymbol{v}+\{\boldsymbol{A_S}\tilde{\boldsymbol{w}}\,|\,\tilde{\boldsymbol{w}}\in\mathbb{F}_p^{(K-M)\ell}\}$, where $\boldsymbol{v}$ is the coset leader. From (6), the undithered codeword must be of the form

$$\boldsymbol{t}=\left[\boldsymbol{B}_{\rm c}p^{-1}g\left(\boldsymbol{Gv}+\boldsymbol{GA_S}\tilde{\boldsymbol{w}}\right)\right]\bmod\Lambda_{\rm c},\ \tilde{\boldsymbol{w}}\in\mathbb{F}_p^{(K-M)\ell}. \quad (9)$$

Note that for any $\boldsymbol{a},\boldsymbol{b}\in\mathbb{F}_p^n$, $g(\boldsymbol{a}+\boldsymbol{b})=g(\boldsymbol{a})+g(\boldsymbol{b})\bmod p$. Therefore, $g(\boldsymbol{Gv}+\boldsymbol{GA_S}\tilde{\boldsymbol{w}})=g(\boldsymbol{Gv})+g(\boldsymbol{GA_S}\tilde{\boldsymbol{w}})+p\boldsymbol{c}$ for some $\boldsymbol{c}\in\mathbb{Z}^n$. Using this and the fact $\boldsymbol{B}_{\rm c}\boldsymbol{c}\in\Lambda_{\rm c}$ in (9),

$$\begin{aligned}\boldsymbol{t}&=\left[\boldsymbol{B}_{\rm c}p^{-1}g\left(\boldsymbol{Gv}\right)+\boldsymbol{B}_{\rm c}p^{-1}g\left(\boldsymbol{GA_S}\tilde{\boldsymbol{w}}\right)+\boldsymbol{B}_{\rm c}\boldsymbol{c}\right]\bmod\Lambda_{\rm c}\\&=\left[\boldsymbol{B}_{\rm c}p^{-1}g\left(\boldsymbol{Gv}\right)+\left[\boldsymbol{B}_{\rm c}p^{-1}g\left(\boldsymbol{GA_S}\tilde{\boldsymbol{w}}\right)\right]\bmod\Lambda_{\rm c}\right]\bmod\Lambda_{\rm c}. \quad (10)\end{aligned}$$

Since the receiver knows $\boldsymbol{v}$, the component of $\boldsymbol{t}$ unavailable from the side information is

$$\tilde{\boldsymbol{t}}=\left[\boldsymbol{B}_{\rm c}p^{-1}g\left(\boldsymbol{GA_S}\tilde{\boldsymbol{w}}\right)\right]\bmod\Lambda_{\rm c}. \quad (11)$$

Let $\mathscr{C}_{\boldsymbol{S}}\subset\mathbb{F}_p^n$ be the subcode of $\mathscr{C}$ with generator matrix $\boldsymbol{GA_S}$, and $\Lambda_{\boldsymbol{S}}=\boldsymbol{B}_{\rm c}p^{-1}g(\mathscr{C}_{\boldsymbol{S}})+\Lambda_{\rm c}$. Using $\boldsymbol{GA_S}$ instead of $\boldsymbol{G}$ in Lemma 1, we see that $\tilde{\boldsymbol{t}}\in\Lambda_{\boldsymbol{S}}/\Lambda_{\rm c}$ and that (11) is a one-to-one correspondence between $\tilde{\boldsymbol{w}}\in\mathbb{F}_p^{(K-M)\ell}$ and $\tilde{\boldsymbol{t}}\in\Lambda_{\boldsymbol{S}}/\Lambda_{\rm c}$ if $\boldsymbol{GA_S}$ is full rank. The decoding problem at the second stage is to estimate $\tilde{\boldsymbol{t}}$, or equivalently $\tilde{\boldsymbol{w}}$, from the channel output.

MMSE Scaling and Lattice Decoding: Let the channel output at the receiver $(\boldsymbol{S},\sigma^2)$ be $\boldsymbol{y}=\boldsymbol{x}+\boldsymbol{n}$, where $\boldsymbol{n}$ is a Gaussian vector with zero mean and variance σ^2 per dimension. Scaling the received vector by the MMSE coefficient $\alpha=1/(1+\sigma^2)$,

$$\alpha\boldsymbol{y}=\alpha\boldsymbol{x}+\alpha\boldsymbol{n}=\boldsymbol{x}+\alpha\boldsymbol{n}-(1-\alpha)\boldsymbol{x}=\boldsymbol{x}+\boldsymbol{z},$$

where $\boldsymbol{z}=\alpha\boldsymbol{n}-(1-\alpha)\boldsymbol{x}$ is the effective noise term. Using the facts that $\boldsymbol{n}$ is independent of $\boldsymbol{x}$ and $\|\boldsymbol{x}\|\le n$, it is easy to show that the power $\mathbb{E}\,\|\boldsymbol{z}\|^2/n$ of the effective noise $\boldsymbol{z}$ is at the most $\sigma^2/(1+\sigma^2)$. Denoting $\sigma^2/(1+\sigma^2)$ by $\sigma_{\boldsymbol{z}}^2$, the lower bound (2) on signal-to-noise ratio can be rewritten as

$$\sigma_{\boldsymbol{z}}^2\le 2^{-2(R+\epsilon)(K-M)}. \quad (12)$$

From (7), (10) and (11), $\boldsymbol{x}=\tilde{\boldsymbol{t}}+\boldsymbol{B}_{\rm c}p^{-1}g(\boldsymbol{Gv})-\boldsymbol{d}+\boldsymbol{\lambda}_{\rm c}$ for some $\boldsymbol{\lambda}_{\rm c}\in\Lambda_{\rm c}$. After MMSE scaling, the decoder removes the contributions of $\boldsymbol{d}$ and $\boldsymbol{B}_{\rm c}p^{-1}g(\boldsymbol{Gv})$ from $\alpha\boldsymbol{y}$ to obtain

$$\boldsymbol{y'}=\alpha\boldsymbol{y}-\boldsymbol{B}_{\rm c}p^{-1}g(\boldsymbol{Gv})+\boldsymbol{d}=\tilde{\boldsymbol{t}}+\boldsymbol{\lambda}_{\rm c}+\boldsymbol{z}.$$

The decoder proceeds by quantizing $\boldsymbol{y'}$ to the lattice $\Lambda_{\boldsymbol{S}}$ and reducing the result modulo $\Lambda_{\rm c}$. If the noise $\boldsymbol{z}$ is sufficiently 'small', then this sequence of operations will yield $\left[Q_{\Lambda_{\boldsymbol{S}}}(\tilde{\boldsymbol{t}}+\boldsymbol{\lambda}_{\rm c}+\boldsymbol{z})\right]\bmod\Lambda_{\rm c}=\left[\tilde{\boldsymbol{t}}+\boldsymbol{\lambda}_{\rm c}\right]\bmod\Lambda_{\rm c}=\tilde{\boldsymbol{t}}$. Given $\tilde{\boldsymbol{t}}$, the receiver uses (10) to obtain the undithered codeword $\boldsymbol{t}$, and hence the message vector $(\boldsymbol{w}_1^{\mathsf{T}},\ldots,\boldsymbol{w}_K^{\mathsf{T}})^{\mathsf{T}}$ as $\boldsymbol{t}=\left[\boldsymbol{B}_{\rm c}p^{-1}g(\boldsymbol{Gv})+\tilde{\boldsymbol{t}}\right]\bmod\Lambda_{\rm c}$. To conclude, the decoder obtains the estimate of the undithered codeword $\boldsymbol{t}$ from the received vector $\boldsymbol{y}$ as

$$\hat{\boldsymbol{t}}=\left[\left[Q_{\Lambda_{\boldsymbol{S}}}(\boldsymbol{y'})\right]\bmod\Lambda_{\rm c}+\boldsymbol{B}_{\rm c}p^{-1}g(\boldsymbol{Gv})\right]\bmod\Lambda_{\rm c}.$$

Note that a decoding error occurs if and only if $\boldsymbol{z}$ is closer to a point in $\Lambda_{\boldsymbol{S}}\backslash\Lambda_{\rm c}$ than any other vector in the coarse lattice $\Lambda_{\rm c}$, i.e., if and only if the following event occurs

$$\mathcal{E}: Q_{\Lambda_{\boldsymbol{S}}}(\boldsymbol{z})\in\Lambda_{\boldsymbol{S}}\backslash\Lambda_{\rm c}. \quad (13)$$

IV. Proof of Main Theorem

We first present two lemmas which will be used to show that the error probability at a given fixed receiver $(\boldsymbol{S},\sigma^2)$ is small. We then complete the proof by showing that the error probability at every receiver of the multicast channel is simultaneously small. The first lemma, which is a direct generalization of [6, Lemma 1] and [7, Lemma 2.3], bounds the number of lattice points lying inside a ball.

Lemma 2. *For any $\boldsymbol{s}\in\mathbb{R}^n$, $r>0$ and any n-dimensional lattice $\Lambda_{\rm c}$, $|\Lambda_{\rm c}\cap\mathcal{B}(\boldsymbol{s},r)|\le\left(r_{\rm cov}(\Lambda_{\rm c})+r\right)^n V_n/\mathrm{Vol}(\Lambda_{\rm c})$, where V_n is the volume of a unit ball in $\mathbb{R}^n$.*

As in [6], [7], [28], we will also rely on the following result which shows that with very high probability, the norm of the noise $\boldsymbol{z}$ is not much larger than $\sqrt{n\sigma_{\boldsymbol{z}}^2}$.

Lemma 3. *Let $\boldsymbol{x}$ be uniformly distributed in $\mathcal{V}(\Lambda_{\rm c})$ and $\delta>0$. Then the probability $\mathrm{P}\left(\|\boldsymbol{z}\|^2>n\sigma_{\boldsymbol{z}}^2(1+\delta)\right)$ is at the most $\exp\left(-\frac{n(\delta-\ln(1+\delta))}{2}\right)+\exp\left(-\frac{n\sigma^2\delta^2}{4}\right)$.*

Error probability at a single receiver: We now derive an upper bound on the error probability $\mathrm{P}_{\boldsymbol{S}}$ at $(\boldsymbol{S},\sigma^2)$ when averaged over the ensemble of lattice codes. From the *Crypto lemma* [5, Lemma 1], we know that $\boldsymbol{t}$ is statistically independent of $\boldsymbol{x}=[\boldsymbol{t}-\boldsymbol{d}]\bmod\Lambda_{\rm c}$, and hence, the codeword $\boldsymbol{t}$ and effective noise $\boldsymbol{z}=\alpha\boldsymbol{n}-(1-\alpha)\boldsymbol{x}$ are independent. Hence the error event $\mathcal{E}$ in (13) is independent of the transmit message.

We assume that the decoder declares an error if $\boldsymbol{G}$ is not full-rank since this implies that the encoding is not one-to-one. We have already observed in Section III-A that the probability of this event $P\left(\mathrm{rank}(\boldsymbol{G}<K\ell)\right)\le 2^{-n/2}$. Now given the value of ϵ, we set $\delta=2^{\epsilon/2}-1$, $r_{\boldsymbol{z}}=\sqrt{n(1+\delta)\sigma_{\boldsymbol{z}}^2}$ which is the radius of the typical noise $\boldsymbol{z}$, and $\mathcal{B}_{r_{\boldsymbol{z}}}=\mathcal{B}(\boldsymbol{0},r_{\boldsymbol{z}})$. Then,

$$\begin{aligned}\mathrm{P}(\mathcal{E})&=\mathrm{P}(\boldsymbol{z}\in\mathcal{B}_{r_{\boldsymbol{z}}})\,\mathrm{P}(\mathcal{E}|\boldsymbol{z}\in\mathcal{B}_{r_{\boldsymbol{z}}})+\mathrm{P}(\boldsymbol{z}\notin\mathcal{B}_{r_{\boldsymbol{z}}})\,\mathrm{P}(\mathcal{E}|\boldsymbol{z}\notin\mathcal{B}_{r_{\boldsymbol{z}}})\\&\le\mathrm{P}(\mathcal{E}|\boldsymbol{z}\in\mathcal{B}_{r_{\boldsymbol{z}}})+\mathrm{P}(\boldsymbol{z}\notin\mathcal{B}_{r_{\boldsymbol{z}}}). \quad (14)\end{aligned}$$

Lemma 3 provides an exponential upper bound on $\mathrm{P}(\boldsymbol{z}\notin\mathcal{B}_{r_{\boldsymbol{z}}})$. In the following theorem we show that $\mathrm{P}(\mathcal{E}|\boldsymbol{z}\in\mathcal{B}_{r_{\boldsymbol{z}}})$ is also exponentially small in n. The proof of this result uses the technique of [6], [7] to bound the number of lattice points lying in an n-dimensional ball.

Theorem 2. *For any receiver $(\boldsymbol{S},\sigma^2)$ that satisfies $1/2\log\left(1+1/\sigma^2\right)>(R+\epsilon)\left(K-\mathrm{rank}(\boldsymbol{S})\right)$, and for all large enough n, $\mathrm{P}(\mathcal{E}|\boldsymbol{z}\in\mathcal{B}_{r_{\boldsymbol{z}}})\le 2^{-n\epsilon/4}$ when averaged over the ensemble of random lattice codes.*

Proof Sketch: From (13), we note that the decoder is in error when $\boldsymbol{z}$ is closer to some coset $\boldsymbol{t'}+\Lambda_{\rm c}$, with $\boldsymbol{t'}\in\Lambda_{\boldsymbol{S}}/\Lambda_{\rm c}$ and $\boldsymbol{t'}\neq\boldsymbol{0}$, than any point in $\Lambda_{\rm c}$. The proof of Proposition 1 of [6] provides an upper bound on $\mathrm{P}(\mathcal{E}|\boldsymbol{z}\in\mathcal{B}_{r_{\boldsymbol{z}}})$ averaged over the ensemble of the lattice codes, when the coarse lattice is chosen as $\mathbb{Z}^n$ (the inequality preceeding (23) in [6]). An adaptation of the steps in this proof to our choice of coarse lattice $\Lambda_{\rm c}$ yields

$$\mathrm{P}(\mathcal{E}|\boldsymbol{z}\in\mathcal{B}_{r_{\boldsymbol{z}}})\le p^{-n}p^{(K-M)\ell}\,\mathbb{E}\left(|\Lambda_{\rm c}\cap\mathcal{B}(p\boldsymbol{z},pr_{\boldsymbol{z}})|\,\big|\,\boldsymbol{z}\in\mathcal{B}_{r_{\boldsymbol{z}}}\right).$$

Using Lemma 2, we can bound $|\Lambda_{\rm c} \cap \mathcal{B}(p\boldsymbol{z}, pr_{\boldsymbol{z}})|$, and obtain

$$\mathrm{P}(\mathcal{E}|\boldsymbol{z} \in \mathcal{B}_{r_{\boldsymbol{z}}}) \leq p^{-n} p^{(K-M)\ell} \tfrac{V_n}{\mathrm{Vol}(\Lambda_{\rm c})} \left(r_{\rm cov}(\Lambda_{\rm c}) + pr_{\boldsymbol{z}}\right)^n .$$

Manipulating the above inequality using the identities $p^{(K-M)\ell} \leq 2^{nR(K-M)}$, from (4); $\mathrm{Vol}(\Lambda_{\rm c}) \geq V_n\, n^{n/2}\, 2^{-n\epsilon/4}$, from (5); $\sigma_{\boldsymbol{z}} \leq 2^{-(R+\epsilon)(K-M)}$, from (12); $r_{\rm cov}(\Lambda_{\rm c}) = \sqrt{n}$, $r_{\boldsymbol{z}} = \sqrt{n(1+\delta)\sigma_{\boldsymbol{z}}^2}$, and $1 + \delta = 2^{\epsilon/2}$, we obtain

$$\mathrm{P}(\mathcal{E}|\boldsymbol{z} \in \mathcal{B}_{r_{\boldsymbol{z}}}) \leq 2^{-n\epsilon/2} \left(\left(p\, 2^{(R+\epsilon)(K-M)}\right)^{-1} + 1 \right)^n .$$

Using (3) in the above inequality, $\mathrm{P}(\mathcal{E}|\boldsymbol{z} \in \mathcal{B}_{r_{\boldsymbol{z}}}) \leq 2^{-n\epsilon/4}$. ■

Using Lemma 3, Theorem 2, the bound $P(\mathrm{rank}(\boldsymbol{G}) < K\ell) \leq 2^{-n/2}$ along with (14), we conclude that the error probability at a given receiver $(\boldsymbol{S}, \sigma^2)$ averaged over the code ensemble is exponentially small in n, i.e., $P_{\boldsymbol{S}} \leq 2^{-n\varepsilon}$ for some $\varepsilon > 0$. The constant ε is independent of $\boldsymbol{S}$ as long as the receiver $(\boldsymbol{S}, \sigma^2)$ satisfies the hypothesis (2) of Theorem 1. Hence, there exists a choice of lattice code (which is chosen for the given matrix $\boldsymbol{S}$) with a small error probability at this receiver. To prove Theorem 1, we must establish a slightly stronger result, viz., there exists a lattice code such that the decoding error probability for every possible side information matrix $\boldsymbol{S}$ is small as long as the receiver SNR is large enough.

Completing the proof of the main theorem: Consider a hypothetical multicast network that consists of one receiver for each possible choice of $\boldsymbol{S}$. The number of linearly independent equations M available at the receivers can take values in the set $\{0, 1, \ldots, K-1\}$. Hence, the set of all possible values of side information matrix $\boldsymbol{S}$ is $\mathcal{S} = \cup_{M=0}^{K-1} \mathbb{F}_p^{M \times K}$. Therefore the number of receivers in the hypothetical multicast network $|\mathcal{S}| = \sum_{M=0}^{K-1} \left|\mathbb{F}_p^{M\times K}\right| = \sum_{M=0}^{K-1} p^{MK} \leq K p^{K^2}$. We assume that each receiver $(\boldsymbol{S}, \sigma^2)$, $\boldsymbol{S} \in \mathcal{S}$, satisfies the lower bound (2) on SNR, and hence, has error probability $P_{\boldsymbol{S}} \leq 2^{-n\varepsilon}$. We say that the multicast network is in error if any of the receivers commits a decoding error. By union bound, the network error probability $\mathrm{P}(\text{network error}) \leq |\mathcal{S}|\, 2^{-n\varepsilon} \leq K p^{K^2}\, 2^{-n\varepsilon}$, is exponentially small in n since p and K are constants. Hence, there exists a sequence of lattice codes such that the decoding error probability at every receiver $(\boldsymbol{S}, \sigma^2)$, $\boldsymbol{S} \in \mathcal{S}$, is simultaneously exponentially small in the code length n.

V. Conclusion

The lattice coding schemes in [5]–[7], [12] (for single-user AWGN and relay channels), which are based on Construction A, require the size p of the prime field to grow as a function of the code length n. By using an error analysis based on counting arguments and a Rogers good coarse lattice, we showed that our lattice coding scheme provides reliable communication for all sufficiently large prime p, which is independent of the code length. However, our proof technique still requires the field size p to depend on the gap to capacity ϵ. Further work is required to devise good lattice strategies for fixed small values of p intended for practical applications.

References

[1] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources (corresp.)," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 219–221, Mar. 1979.

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

[3] B. Nazer and M. Gastpar, "The case for structured random codes in network capacity theorems," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 455–474, 2008.

[4] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 273–278, Jan. 1998.

[5] U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1+\mathrm{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct 2004.

[6] O. Ordentlich and U. Erez, "A simple proof for the existence of "good" pairs of nested lattices," in *Electrical Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of*, Nov. 2012, pp. 1–12.

[7] N. Di Pietro, "On infinite and finite lattice constellations for the additive white Gaussian Noise Channel," Theses, Université de Bordeaux, Jan. 2014. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01135575

[8] C. Ling and J.-C. Belfiore, "Achieving AWGN channel capacity with lattice Gaussian coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5918–5929, Oct. 2014.

[9] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.

[10] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within $\frac{1}{2}$ bit," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

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

[12] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[13] C. Feng, D. Silva, and F. R. Kschischang, "An algebraic approach to physical-layer network coding," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7576–7596, Nov. 2013.

[14] E. Tuncel, "Slepian-Wolf coding over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, Apr. 2006.

[15] S.-Y. Li, R. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[16] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[17] R. W. Yeung, *Information Theory and Network Coding*. New York: Springer Science+Business Media, LLC, 2008.

[18] L.-L. Xie, "Network coding and random binning for multi-user channels," in *Proc. 10th Canadian Workshop on Information Theory (CWIT)*, Jun. 2007, pp. 85–88.

[19] G. Kramer and S. Shamai, "Capacity for classes of broadcast channels with receiver side information," in *Proc. IEEE Information Theory Workshop (ITW)*, Sep. 2007, pp. 313–318.

[20] Y. Wu, "Broadcasting when receivers know some messages a priori," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2007, pp. 1141–1145.

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

[22] F. Xue and S. Sandhu, "PHY-layer network coding for broadcast channel with side information," in *Information Theory Workshop, 2007. ITW '07. IEEE*, Sep. 2007, pp. 108–113.

[23] T. Wang, S. C. Liew, and L. Shi, "A lattice approach for optimal rate-diverse wireless network coding," *arXiv preprint*, 2015. [Online]. Available: http://arxiv.org/abs/1509.07250

[24] L. Natarajan, Y. Hong, and E. Viterbo, "Lattice codes achieve the capacity of common message Gaussian broadcast channels with coded side information," *arXiv:1509.01332v2*, Jan. 2017.

[25] J. H. Conway and N. Sloane, *Sphere packings, lattices and groups*. New York: Springer-Verlag, 1999.

[26] J. Forney, G.D., "Multidimensional constellations–Part II. Voronoi constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 941–958, Aug. 1989.

[27] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*. Cambridge University Press, 2014.

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