An improved lattice-based scheme for lossy distributed compression of linear functions

Yang Yang and Zixiang Xiong
Dept of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843
Email: yangyang@tamu.edu and zx@ece.tamu.edu

***Abstract*—It has recently been showed that lattice/structured codes can outperform random/unstructured codes in a number of scenarios of distributed source coding. One such instance involves lossy distributed compression of linear functions of Gaussian sources, which is the focus of this paper. Existing structured schemes employ "good" lattices for quantization and binning, however, the remaining correlation between the lattice coset indices are not exploited in an efficient way, leading to suboptimal performance when the target distortion is small. This paper proposes a new lattice-based scheme that is capable of eliminating the redundancy among coset indices and achieving a smaller sum-rate than existing schemes. The main novelty lies in the use of a hidden relationship between the *coset planes* of the quantization indices, and the enlarged set of choices for the quantizers as well as the linear estimation coefficients.**

I. Introduction

Distributed source coding (DSC) studies problems where multiple sources are to be separately compressed such that the decoder can reconstruct them subject to certain distortion constraints. This paper focuses on a special case of DSC where the decoder tries to reconstruct a linear function $Z = Y_1 + cY_2$ of two jointly Gaussian sources Y_1 and Y_2 subject to a mean squared error distortion constraint. This case is of particular interest not only because it is a building block of the general DSC problem, but also due to fact that in its binary counterpart considered by Körner and Marton [1], the conventional quantization plus binning scheme, referred to as the *Berger-Tung (BT)* scheme in this paper, is suboptimal. Hence improvement in this special case can lead to better understanding of the general DSC problem, including the multiterminal source coding introduced by Berger [2] and Tung [3].

Wagner [4] proved that when the correlation coefficient ρ between Y_1 and Y_2 has the same sign as c (this case is often referred to as the quadratic Gaussian μ-sum problem), the conventional scheme that quantizes both sources and employs random binning on the quantization indices is indeed optimal. On the other hand, when $c\rho < 0$, Krithivasan and Pradhan [5] proposed a lattice-based scheme that directly targets at reconstructing the linear function, and showed that as the lattice dimension goes to infinity, the asymptotic performance of the proposed scheme is strictly better than that of the BT scheme in certain (ρ, c) region. They also generalized the lattice construction algorithm to construct any finite number of lattices that are sequentially nested with each being "good" (as defined in [6]) for both source coding and channel coding. Later, Wagner [7], [8] proposed an improved lattice-based scheme that can strictly decrease the sum-rate of Krithivasan and Pradhan's scheme by using the same basic idea of Körner and Marton's scheme [1]. Utilizing a gene-aided outer bound on the rate-distortion region, it is also proved by Wagner that Krithivasan and Pradhan's scheme performs within one bit per sample of the rate-distortion region under certain conditions. Although conjectured to be smaller than the time-sharing sum-rate between the BT scheme and the one in [5], the total sum-rate of Wagner's scheme cannot be given in closed-form. In a more recent paper [9], the rate region achievable by lattice-based schemes is characterized and shown to be within a bounded (but possibly large) gap from the rate-distortion region.

In this paper, we propose a new lattice-based scheme with $calculable$ sum-rate that outperforms Krithivasan and Pradhan [5] while being not worse than the BT scheme.

First, we quantize a length-n block of Gaussian sources Y_1^n and Y_2^n into discrete sources Q_1^n and Q_2^n using (dithered) lattice quantizers with second moments q_1 and q_2, respectively. This way, the original DSC problem over jointly Gaussian sources is transformed into one over discrete sources, among which the correlation is still asymptotically "jointly Gaussian" in certain sense if the lattice quantizers are "good". We then decompose lattice quantization indices into *coset-planes* using a sequence of nested lattices whose nesting ratios are carefully designed, with each coset-plane corresponding to the binning index between two adjacent nested lattices. Assuming the quantized sources Q_1^n and Q_2^n are available at the decoder, one can use a linear combination of them to reconstruct Z^n as $\hat{Z}^n = c_1Q_1^n + c_2Q_2^n$ such that the distortion constraint is satisfied between Z^n and $\hat{Z}^n$. Note that there is one degree of freedom left in the coefficients (c_1, c_2) for fixed (q_1, q_2). Now the quantizers and reconstructor are parameterized by (q_1, q_2, c_1, c_2), and the remaining problem is to find out what the minimum transmission rate is in order for $\hat{Z}^n$ to be reliably reconstructed.

We prove that in certain cases, only some of the lower coset planes of the quantization indices, denoted as $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ respectively, are required to reconstruct $\hat{Z}^n$. Thus theoretically, the total sum-rate can be as low as the joint entropy $\frac{1}{n}H(\boldsymbol{B}_1, \boldsymbol{B}_2)$ (given the dithers sequences), which, similar to Wagner's improved scheme [8], is not easy to evaluate in general. However, under certain conditions, it is proved that some coset planes in $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ equal to deterministic

functions of other ones with high probability, thus the joint entropy $\frac{1}{n}H(\boldsymbol{B}_1, \boldsymbol{B}_2)$ can be calculated in exact form. Finally, an optimization problem is formulated to minimize the sum-rate over all possible (q_1, q_2, c_1, c_2) quadruples, whose solution can be found using numerical methods. Note that in cases when $\frac{1}{n}H(\boldsymbol{B}_1, \boldsymbol{B}_2)$ is not easily computable, to simplify the optimization problem, we take one step back and use the simple scheme that blindly transmits $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ to achieve a sum-rate of $\frac{1}{n}(H(\boldsymbol{B}_1) + H(\boldsymbol{B}_2))$.

Both (the lattice implementation of) the BT and Krithivasan and Pradhan's schemes can be viewed as special cases of ours. In particular, in the BT lattice scheme, $\boldsymbol{B}_1$ is set to be the whole quantization index at the first encoder, while $\boldsymbol{B}_2$ is the binning index in Wyner-Ziv coding; in Krithivasan and Pradhan's scheme, q_1, q_2, c_1, and c_2 are all deterministic functions of ρ, c, and D, which might not be the solution to the optimization problem we formulate in this paper. We evaluate the (ρ, c) region where our new lattice-based scheme has a smaller sum-rate than the BT scheme and the scheme in [5]. However, we are unable to directly compare the sum-rate of the scheme in [8] because it is not amenable to analysis.

II. Problem setup

Let Y_1 and Y_2 be jointly Gaussian with zero mean and covariance matrix

$$\Sigma_Y \;=\; \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \tag{1}$$

and $Y_1^n = (Y_{1,1}, Y_{1,2}, ..., Y_{1,n})$, $Y_2^n = (Y_{2,1}, Y_{2,2}, ..., Y_{2,n})$ be two length-n blocks of independent drawings of Y_1 and Y_2, respectively.

Each of the two encoders sends a function of their own source, namely, $\phi_i^{(n)} : \mathbb{R}^n \mapsto \{1, 2, ..., M_i^{(n)}\}$ for $i = 1, 2$, and denote

$$\begin{aligned} W_1 &= \phi_1^{(n)}(Y_1^n), \\ W_2 &= \phi_2^{(n)}(Y_2^n). \end{aligned} \tag{2}$$

The decoder is only interested in a linear function

$$Z \;=\; Y_1 + cY_2 \tag{3}$$

of Y_1 and Y_2, and reconstruct $Z^n = (Z_{1,1}, Z_{1,2}, ..., Z_{1,n})$ using function $\psi^{(n)} : \{1, 2, ..., M_1^{(n)}\} \times \{1, 2, ..., M_2^{(n)}\} \mapsto \mathbb{R}^n$. Denote the reconstructed version of Z^n as

$$\hat{Z}^n \;=\; \psi^{(n)}(W_1, W_2). \tag{4}$$

We say a rate pair (R_1, R_2) is achievable if there exists a sequence of schemes $(\phi_1^{(n)}, \phi_2^{(n)}, \psi^{(n)})$ such that

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log_2 M_i^{(n)} &\le R_i, \quad i = 1, 2, \\ \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E[(Z_i - \hat{Z}_i)^2] &\le D. \end{aligned} \tag{5}$$

The set of all achievable rate pairs is called the achievable rate region and is denoted as $\mathcal{R}(D)$, and the *minimum sum-rate* is defined as

$$R(D) \;=\; \min_{(R_1,R_2)\in\mathcal{R}(D)} R_1 + R_2. \tag{6}$$

A. Existing results

For general values of ρ and c, one can easily apply the conventional BT scheme that employs independent Gaussian quantization followed by binning (Slepian-Wolf coding) to obtain the following inner rate region

$$\begin{aligned} \mathcal{R}_{BT}(D) \;=\; \mathrm{conv}\Big(\bigcup_{(U_1,U_2)\in\mathcal{U}(Y_1,Y_2)} \{(R_1, R_2) : \\ R_1 \ge I(Y_1; U_1|U_2),\ R_2 \ge I(Y_2; U_2|U_1), \\ R_1 + R_2 \ge I(Y_1, Y_2; U_1, U_2)\}\Big), \end{aligned} \tag{7}$$

where $\mathrm{conv}(\cdot)$ means convex closure, and $\mathcal{U}(Y_1, Y_2)$ contains all (U_1, U_2) pairs such that $U_1 = Y_1 + Q_1$, $U_2 = Y_2 + Q_2$ with Q_1 and Q_2 being independent Gaussian random variables.

In the special case when $c\rho \ge 0$, it is proved that the above BT scheme is indeed optimal [4], i.e.,

$$\mathcal{R}(D) \;=\; \mathcal{R}_{BT}(D), \quad \text{if } c\rho \ge 0. \tag{8}$$

However, when $c\rho < 0$ (assuming $\rho > 0$ and $c < 0$ without loss of generality), the achievable rate region $\mathcal{R}(D)$ is not characterized yet. Krithivasan and Pradhan [5] provided an inner rate region defined as

$$\mathcal{R}_{KP}(D) = \left\{(R_1, R_2) : 2^{-2R_1} + 2^{-2R_2} \le \frac{D}{1 + 2\rho c + c^2}\right\}, \tag{9}$$

which is obtained by a lattice-based scheme that directly reconstructs the linear function $Z = Y_1 + cY_2$. On the other hand, Wagner [8] provided an outer rate region for $\mathcal{R}(D)$ when $\rho > 0$ and $c < 0$,

$$\mathcal{R}_o(D) \;=\; \bigcup_{\rho\le\alpha\le\frac{1}{\rho}} \mathcal{R}_o^{\alpha}(D), \tag{10}$$

where

$$\begin{aligned} \mathcal{R}_o^{\alpha}(D) \;=\; \Big\{(R_1, R_2) : \theta(\alpha)2^{-2R_1} + (1 - \theta(\alpha))2^{-2R_2} \\ \le \frac{D}{1 - \alpha\rho + c^2(1 - \rho/\alpha)}\Big\}, \end{aligned} \tag{11}$$

with $\theta(\alpha) = \frac{1-\alpha\rho}{1-\alpha\rho+c^2(1-\rho/\alpha)}$. Then for $\rho > 0$ and $c < 0$, we have

$$\mathcal{R}_{BT}(D) \cup \mathcal{R}_{KP}(D) \;\subseteq\; \mathcal{R}(D) \;\subseteq\; \mathcal{R}_o(D). \tag{12}$$

The minimum sum-rates of the BT and Krithivasan and Pradhan's achievable rate regions are [5]

$$\begin{aligned} R_{BT}(D) &= \min_{(R_1,R_2)\in\mathcal{R}_{BT}(D)} R_1 + R_2 \\ &= \begin{cases} \frac{1}{2}\log_2 \frac{-4c((1-\rho^2)c+\rho D)}{D^2} & D \le \frac{2(1-\rho^2)c^2}{1-\rho c} \\ \frac{1}{2}\log_2 \frac{(1+\rho c)^2}{D-(1-\rho^2)c^2} & D > \frac{2(1-\rho^2)c^2}{1-\rho c} \end{cases} \end{aligned} \tag{13}$$

$$\begin{aligned} R_{KP}(D) &= \min_{(R_1,R_2)\in\mathcal{R}_{KP}(D)} R_1 + R_2 \\ &= \log_2 \frac{2(1 + c^2 + 2\rho c)}{D}. \end{aligned} \tag{14}$$

In addition, Wagner [8] provided a lower sum-rate bound

$$R(D) \;\ge\; \max_{\boldsymbol{\mathcal{K}}\in\boldsymbol{\Omega}} \frac{1}{2} \log_2^+ \frac{c^2|\boldsymbol{\mathcal{K}}|}{D^2}, \tag{15}$$

where $\log_2^+(x) = \max(0, \log_2(x))$, and $\boldsymbol{\Omega}$ contains all 2×2 matrices $\boldsymbol{\mathcal{K}}$ such that

$$\boldsymbol{0} \preceq \boldsymbol{\mathcal{K}} \preceq \Sigma_Y, \quad [1\ c] \cdot \boldsymbol{\mathcal{K}} \cdot [1\ c]^T \ge \min(\boldsymbol{\mathcal{K}}_{1,1}, c^2\boldsymbol{\mathcal{K}}_{2,2}). \tag{16}$$

III. Proposed lattice-based scheme

In this section, we describe our proposed lattice-based scheme and give a new upper bound on the minimum sum-rate $R(D)$. All entropies and mutual informations are conditioned on the dither sequences known to both the encoders and the decoder.

First fix a pair $(q_1, q_2) \in \mathbb{R}_+^2$. Let Λ_{Q_1} and Λ_{Q_2} be two lattices with second moments $\sigma(\Lambda_{Q_1}) = q_1$ and $\sigma(\Lambda_{Q_2}) = q_2$, respectively, and assume that both are "good" for source coding. Using Λ_{Q_1} and Λ_{Q_2} to quantize Y_1^n and $-cY_2^n$ with dithers U_1^n and U_2^n uniformly distributed over the Voronoi region of Λ_{Q1} and Λ_{Q2}, we obtain the dequantized versions

$$\begin{aligned} Q_1^n &= Q_{\Lambda_{Q_1}}(Y_1^n + U_1^n) - U_1^n, \\ Q_2^n &= Q_{\Lambda_{Q_2}}(-cY_2^n + U_2^n) - U_2^n, \end{aligned}$$

then due to the "goodness" of the lattices, we may assume (with little loss in the Kullback-Leibler divergence sense) that Q_1^n and Q_2^n are n i.i.d. drawings from

$$Q_1 = Y_1 + N_1 \text{ and } Q_2 = -cY_2 + N_2, \tag{17}$$

respectively, with $N_1 \sim \mathcal{N}(0, q_1)$ and $N_2 \sim \mathcal{N}(0, q_2)$ independent of each other and (Y_1^n, Y_2^n).

Once the quantizers are fixed, we first assume that Q_1^n and Q_2^n are available at the decoder, which reconstructs Z^n as $\hat{Z}^n = c_1 Q_1^n + c_2 Q_2^n$ using coefficients $(c_1, c_2) \in \mathbb{R}^2$. In order to meet the desired distortion constraint D, we need

$$\frac{1}{n}\sum_{i=1}^{n} E((Z_i - \hat{Z}_i)^2) \quad \le \quad D, \tag{18}$$

which can be simplified using (17) as

$$c_1^2 q_1 + c_2^2 q_2 + (1-c_1)^2 + c^2(1+c_2)^2 + 2c\rho(1-c_1)(1+c_2) \le D. \tag{19}$$

We denote the set of all quadruples $\mathbb{T} = (q_1, q_2, c_1, c_2)$ satisfying (19) as $\mathcal{T}(D)$.

Obviously, for any $\mathbb{T} \in \mathcal{T}(D)$, the encoders can simply transmit at a sum-rate of (as $n \to \infty$)

$$\begin{aligned} r_{BT}^{\mathbb{T}}(D) &= \frac{1}{n} E(H(Q_{\Lambda_{Q_1}}(Y_1^n + U_1^n), Q_{\Lambda_{Q_2}}(-cY_2^n + U_2^n))) \\ &= \frac{1}{2}\log_2 \frac{(1+q_1)(c^2+q_2) - c^2\rho^2}{q_1 q_2} \end{aligned} \tag{20}$$

using Slepian-Wolf coding. To achieve this sum-rate, lattice binning can be used. Specifically, one encoder (say the first one) transmits the whole quantization index using rate $\frac{1}{n}E[H(Q_{\Lambda_{Q_1}}(Y_1^n + U_1^n))]$, while the other encoder only transmits the bin index $Q_{\Lambda_{Q_2}}(-cY_2^n + U_2^n) \bmod \Lambda_B$, where $\sigma(\Lambda_B) = \mathrm{var}(Q_2|Q_1)$. If Λ_B is "good" for channel coding, then the decoder should be able to reconstruct Q_2^n with vanishing probability of error.

However, the above conventional BT lattice scheme is not always optimal. The main reason is because to reconstruct $\hat{Z}^n = c_1 Q_1^n + c_2 Q_2^n$, the decoder does not need the whole indices $Q_{\Lambda_{Q_1}}(Y_1^n + U_1^n)$ and $Q_{\Lambda_{Q_2}}(-cY_2^n + U_2^n)$. In fact, using similar reasonings as those in [5], it can be proved that in order to reliably reconstruct $\hat{Z}^n$, the decoder only needs to know

$$\begin{aligned} \boldsymbol{B}_1 &= c_1 Q_{\Lambda_{Q_1}}(Y_1^n + U_1^n) \bmod \Lambda_C, \\ \boldsymbol{B}_2 &= -c_2 Q_{\Lambda_{Q_2}}(-cY_2^n + U_2^n) \bmod \Lambda_C \end{aligned} \tag{21}$$

where Λ_C is a "good" channel lattice code with second moment of

$$\sigma(\Lambda_C) = \mathrm{var}(\hat{Z}) = (1+q_1)c_1^2 + (c^2+q_2)c_2^2 - 2cc_1c_2\rho, \tag{22}$$

and we assume that $c_1\Lambda_{Q_1}$ and $-c_2\Lambda_{Q_2}$ are both nested in Λ_C. If we follow Krithivasan and Pradhan's scheme [5] to directly transmit $(\boldsymbol{B}_1, \boldsymbol{B}_2)$ without exploring the correlation in between, we obtain a sum-rate of

$$r_{KP}^{\mathbb{T}}(D) = \frac{1}{2}\log_2 \frac{\left((1+q_1)c_1^2 + (c^2+q_2)c_2^2 - 2cc_1c_2\rho\right)^2}{q_1 q_2}. \tag{23}$$

We argue that in certain cases, one can further explore the correlation between $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ to achieve a sum-rate that is smaller than both $r_{BT}^{\mathbb{T}}(D)$ and $r_{KP}^{\mathbb{T}}(D)$. For the sake of simplicity, for any quadruple $\mathbb{T} = (q_1, q_2, c_1, c_2)$, define

$$\begin{aligned} v_1 &= \mathrm{var}(Q_1) = 1 + q_1, \\ v_2 &= \mathrm{var}(Q_2) = c^2 + q_2, \\ v_{12} &= \mathrm{var}(Q_1|Q_2) = v_1 - \frac{\rho^2}{v_2}, \\ v_{21} &= \mathrm{var}(Q_2|Q_1) = v_2 - \frac{\rho^2}{v_1}, \\ t_1 &= \frac{\mathrm{var}(\hat{Z})}{c_1^2} = v_1 + \frac{c_2^2}{c_1^2}v_2 - 2\rho c\frac{c_2}{c_1}, \\ t_2 &= \frac{\mathrm{var}(\hat{Z})}{c_2^2} = v_2 + \frac{c_1^2}{c_2^2}v_1 - 2\rho c\frac{c_1}{c_2}, \\ t_{12} &= t_2\rho(Q_1|Q_2)^2 = \frac{t_2\rho^2}{v_2^2}, \\ t_{21} &= t_1\rho(Q_2|Q_1)^2 = \frac{t_1\rho^2}{v_1^2}, \end{aligned}$$

where $\mathrm{var}(Q_i|Q_j)$ is the conditional variance of Q_i given Q_j, and $\rho(Q_i|Q_j)$ denotes the optimal linear coefficient in estimating Q_i given Q_j.

Define a *coset plane* of a length-n vector X^n with respect to three nested lattices $\Lambda_A \subset \Lambda_B \subset \Lambda_C$ as

$$\begin{aligned} &\mathbb{b}_{\Lambda_A \subset \Lambda_B \subset \Lambda_C}(X^n) \\ &= \mathbb{i}\big[(Q_{\Lambda_C}(X^n) \bmod \Lambda_A) - (Q_{\Lambda_C}(X^n) \bmod \Lambda_B)\big], \end{aligned} \tag{24}$$

where $\mathbb{i}(\cdot)$ is any index function that maps all lattice points of Λ_B inside the Voronoi region of Λ_A to $\{0, 1, ..., \frac{\sigma(\Lambda_A)^n}{\sigma(\Lambda_B)^n} - 1\}$.

Let Λ_{V_i}, Λ_{T_i}, $\Lambda_{V_{ij}}$, $\Lambda_{T_{ij}}$ for $i, j \in \{1, 2\}, i \neq j$ be "good" nested lattices with second moments v_i, t_i, v_{ij}, and t_{ij}, respectively. Existence of such nested lattices is proved in [5]. Note that the nesting order depends on the values of v_i, t_i, v_{ij}, and t_{ij}. We now calculate the minimum sum-rate for a given quadruples $\mathbb{T} = (q_1, q_2, c_1, c_2)$.

1) When $t_1 > v_1$, we have

$$\boldsymbol{B}_1 \overset{w.h.p}{=} c_1 Q_{\Lambda_{Q_1}}(Y_1^n + U_1^n),$$

hence the whole quantization index $Q_{\Lambda_{Q_1}}(Y_1^n + U_1^n)$ is required to reconstruct $\hat{Z}^n$. Then

$$\begin{aligned} H(\boldsymbol{B}_1, \boldsymbol{B}_2) &= H(\boldsymbol{B}_1) + H(\boldsymbol{B}_2|\boldsymbol{B}_1) \\ &= \frac{n}{2}\log_2 \frac{v_1}{q_1} + \frac{n}{2}\log_2 \frac{\min(t_2, v_{21})}{q_2}. \end{aligned} \tag{25}$$

2) Similarly, when $t_2 > v_2$, we have

$$H(\boldsymbol{B}_1, \boldsymbol{B}_2) = H(\boldsymbol{B}_2) + H(\boldsymbol{B}_1|\boldsymbol{B}_2) = \frac{n}{2}\log_2\frac{v_2}{q_2} + \frac{n}{2}\log_2\frac{\min(t_1, v_{12})}{q_1}. \quad (26)$$

3) When $t_1 \le v_1$ and $t_2 \le v_2$, assume that the second encoder directly transmit $\boldsymbol{B}_2$ using rate $H(\boldsymbol{B}_2)$, while the first encoder tries to compress $\boldsymbol{B}_1$ given decoder side information $\boldsymbol{B}_2$ to achieve a rate of $H(\boldsymbol{B}_1|\boldsymbol{B}_2)$. Using the relationship $Q_1 = \rho(Q_1|Q_2)Q_2 + N_{12}$ with $N_{12} \sim \mathcal{N}(0, v_{12})$, the decoder can tell that $Q_1^n = \rho(Q_1|Q_2)Q_2^n + \gamma^n + N_{12}^n$ for some lattice point $\gamma^n \in \Lambda_{T_{12}}$. Now consider four subcases.

a) When $t_{12} > v_{12} > t_1$, the first encoder transmits

$$K_{1a} = \boldsymbol{B}_1 \bmod \Lambda_{T_1}. \quad (27)$$

b) When $t_{12} > t_1 > v_{12}$, the first encoder needs to send

$$K_{1b} = \boldsymbol{B}_1 \bmod \Lambda_{V_{12}}. \quad (28)$$

c) When $t_1 > t_{12} > v_{12}$, the first encoder needs to transmit both K_{1b} and the coset plane

$$K_{1c} = \mathbb{b}_{\Lambda_{T_1} \subset \Lambda_{T_{12}} \subset \Lambda_{Q_1}}(Q_1^n + U_1^n). \quad (29)$$

d) If none of above three holds, the first encoder may choose from the following two options. It can simply transmit $\boldsymbol{B}_1$ itself, to achieve a sum-rate of $H(\boldsymbol{B}_1, \boldsymbol{B}_2)$ (which equals to $r_{KP}^{\mathbb{T}}(D)$), or ask the second encoder to transmit an extra coset plane

$$K_{2a} = \mathbb{b}_{\Lambda_{V_2} \subset \Lambda_{T_2} \subset \Lambda_{Q_2}}(Q_2^n + U_2^n), \quad (30)$$

before it sends K_{1b} to achieve a sum-rate of $r_{BT}^{\mathbb{T}}(D)$.

We thus have the following theorem, which is the main result of this paper. The detailed proof is omitted.

Theorem 1: For any pair (ρ, c), it holds that

$$R(D) \le R_{new}(D) \triangleq \inf_{\mathbb{T} \in \mathcal{T}(D)} r_{sum}^{\mathbb{T}}, \quad (31)$$

where

$$r_{sum}^{\mathbb{T}} = \begin{cases} \frac{1}{2}\log_2\frac{v_1 \min(t_2, v_{21})}{q_1 q_2} & t_1 > v_1 \\ \frac{1}{2}\log_2\frac{v_2 \min(t_1, v_{12})}{q_1 q_2} & t_2 > v_2 \\ \frac{1}{2}\log_2\frac{t_2 \min(t_1, v_{12})}{q_1 q_2} & t_1 \le v_1 \ \& \ t_2 \le v_2, \ t_{12} > v_{12} \ \& \ t_{12} > t_1 \\ \frac{1}{2}\log_2\frac{t_1 \min(t_2, v_{21})}{q_1 q_2} & t_1 \le v_1 \ \& \ t_2 \le v_2, \ t_{21} > v_{21} \ \& \ t_{21} > t_2 \\ \frac{1}{2}\log_2\frac{t_1 t_2 v_{12}}{t_{12} q_1 q_2} & t_1 \le v_1, t_2 \le v_2, \ t_1 > t_{12} > v_{12} \\ \frac{1}{2}\log_2\frac{t_1 t_2 v_{21}}{t_{21} q_1 q_2} & t_1 \le v_1, t_2 \le v_2, \ t_2 > t_{21} > v_{21} \\ \frac{1}{2}\log_2\frac{\min(t_1 t_2, v_1 v_{21})}{q_1 q_2} & \text{otherwise.} \end{cases} \quad (32)$$

The following corollary of Theorem 1 gives a suboptimal but easy-to-compute upper bound on the sum-rate $R(D)$.

Corollary 1: For any pair (ρ, c), it holds that

$$R(D) \le \min\{r_{sum}^{\mathbb{T}_1}, r_{sum}^{\mathbb{T}_2}\}, \quad (33)$$

where

$$\mathbb{T}_1 = \left(\frac{Dc(1-\rho^2)}{2c(1-\rho^2) - D(c-\rho)}, \frac{D(1-\rho^2)}{2c^2(1-\rho^2) - D(1-c\rho)}, \frac{2c(1-\rho^2) - D(c-\rho)}{2c(1-\rho^2)}, \frac{D(1-c\rho) - 2c^2(1-\rho^2)}{2c(1-\rho^2)}\right),$$
$$\mathbb{T}_2 = \left(\frac{D\sigma_Z^2}{2(\sigma_Z^2 - D)}, \frac{D\sigma_Z^2}{2(\sigma_Z^2 - D)}, \frac{\sigma_Z^2 - D}{\sigma_Z^2}, \frac{\sigma_Z^2 - D}{\sigma_Z^2}\right). \quad (34)$$

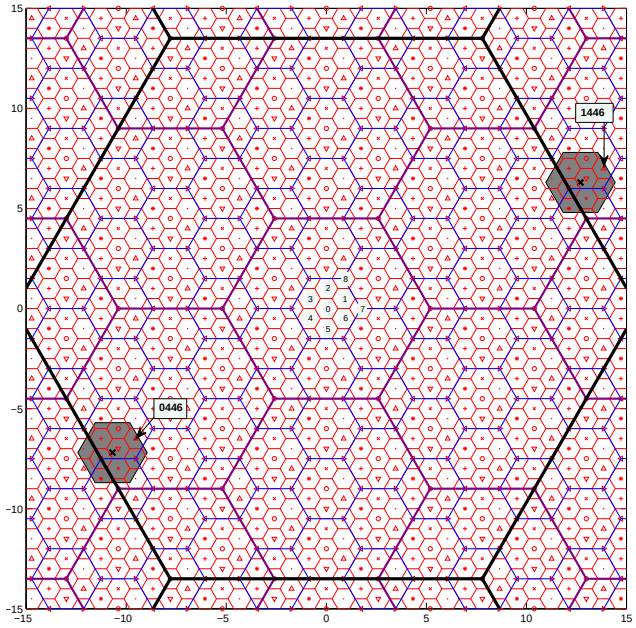

Fig. 1. A graphical illustration of our lattice-based scheme.

A. A graphical illustration

Fig. 1 shows an illustration of the nested lattices used in the third case of our scheme when $v_1 = 81$, $t_{12} = 27$, $t_1 = 9$, $v_{12} = 3$, and $q_1 = 1$. Note that five two-dimensional ($n = 2$) self-similar hexagonal lattices with second moments of $81, 27, 9, 3$, and 1 are used for the sake of clarity, while $n \to \infty$ are required to achieve the best performance. The nesting ratio between each pair of adjacent nested lattices is nine (with bin indices $\{0, 1, ...8\}$ assigned according to the center of Fig. 1) and is the same for each nested pair. Then any finest lattice point can be indexed by $a_1a_2a_3a_4$ with a_i being the index between the i-th and $(i+1)$-th coarsest lattices.

Once $\boldsymbol{B}_2$ is available, the decoder can tell that $Q_1^n = \rho(Q_1|Q_2)Q_2^n + \gamma^n + N_{12}^n$ for some lattice point $\gamma^n \in \Lambda_{T_{12}}$ where $N_{12} \sim \mathcal{N}(0, 3)$ and $\Lambda_{T_{12}}$ has a second moment of $t_{12} = 27$. For illustration, we assume that with the knowledge of $\boldsymbol{B}_2$, Q_1^n must be within the hexagonal areas shaded in Fig. 1 (note that these hexagons are disjoint since $v_{12} = 3 < 27 = t_{12}$). Now in order for the decoder to know $\boldsymbol{B}_1 = a_3a_4$, the first encoder only needs to specify the lowest coset plane a_4. For example, in Fig. 1, if the first encoder sends an index "$a_4 = 6$", the decoder realizes (based on the knowledge of $\boldsymbol{B}_2$) that Q_1^n must be indexed by a_1446 for some $a_1 \in \{0, 1, .., 8\}$, and the desired $\boldsymbol{B}_1$ is correctly reconstructed as 46. Note that as $n \to \infty$, good nested lattices exist to ensure a vanishing decoding error.

IV. Comparison with other schemes

Comparing $R_{new}(D)$ in (31) with $R_{BT}(D)$ and $R_{KP}(D)$ defined in (13) and (14) respectively, we have the following lemma, which readily follows from discussions in the previous section.

Lemma 1: For any pair (ρ, c), it holds that

$$R_{new}(D) \leq \min(R_{BT}(D), R_{KP}(D)). \quad (35)$$

We now provide the (ρ, c) region where our lattice-based scheme improves the BT scheme for certain target distortion D, i.e., $R_{new}(D) < R_{BT}(D)$.

Lemma 2: For any $\rho \in [\frac{3}{5}, 1]$, let $\theta(\rho)$ be the unique solution of x to

$$x^4 - 4\rho x^3 + (1 + 5\rho^2)x^2 - 4\rho x + 1 = 0, \quad (36)$$

in the range $(0, 1]$, where the uniqueness can be easily verified by checking the boundary conditions at $\rho = \frac{3}{5}$ and $\rho = \frac{\sqrt{2}}{2}$. Then there exists $D \in \mathbb{R}_+$ such that $R_{new}(D) < R_{BT}(D)$ if

$$\rho \in [\frac{3}{5}, \frac{\sqrt{2}}{2}] \quad \text{and} \quad \theta(\rho) < -c < \frac{1}{\theta(\rho)},$$

$$\text{or } \rho \in [\frac{\sqrt{2}}{2}, 1) \quad \text{and} \quad \frac{1}{2\rho} < -c < 2\rho. \quad (37)$$

Fig. 2 plots the region defined in Lemma 2 (we are only interested in the region where $-c \leq 1$, since a problem defined by (ρ, c, D) is equivalent to that by $(\rho, \frac{1}{c}, \frac{D}{c^2})$) inside which $R_{new}(D) < R_{BT}(D)$ for small enough D. For comparison, we highlight the (ρ, c) region (in shade) where there exists $D \in \mathbb{R}_+$ such that our lattice scheme outperforms both the BT scheme and that in [5], i.e., $R_{new}(D) < \min(R_{KP}(D), R_{BT}(D))$.

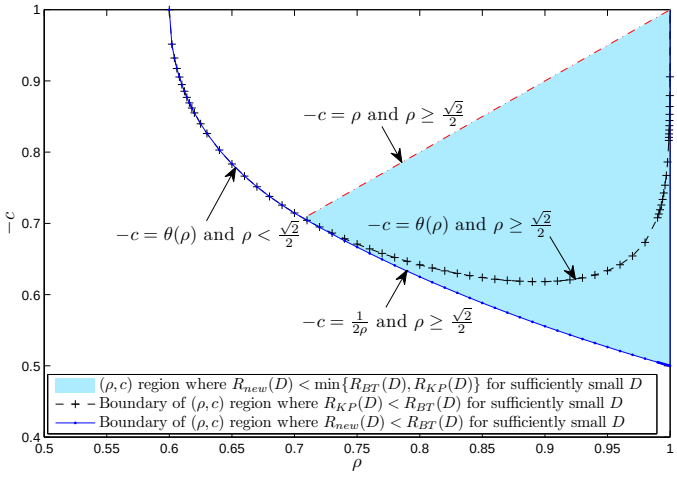

Fig. 2. (ρ, c) region where the lattice scheme outperforms the BT scheme for small enough D.

Fig. 3 is a comparison of sum-rate-distortion functions among the BT scheme, Krithivasan and Pradhan's scheme, and our scheme for $\rho = 0.9$, $-c = 0.7$. We observe that for relatively small D values, $R_{new}(D)$ strictly improve both $R_{BT}(D)$ and $R_{KP}(D)$. While in the high distortion regime, our scheme degenerates to the BT scheme with $R_{new}(D) = R_{BT}(D)$. It is noted that $R_{new}(D)$ itself is not a convex function in D, which means that it can be further reduced for medium D values by exploiting time-sharing.

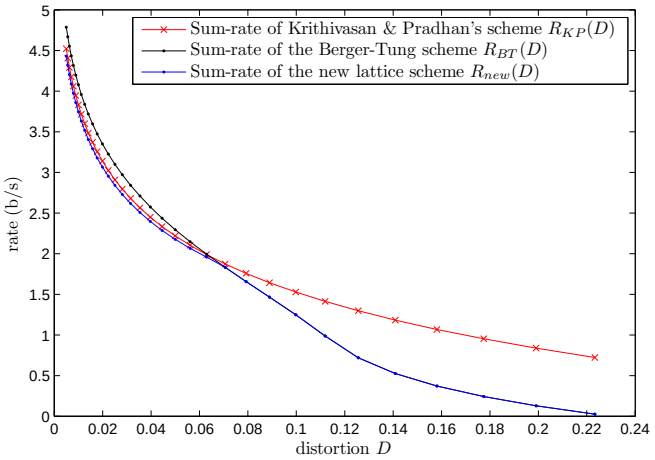

Fig. 3. Comparison among the sum-rate-distortion functions of three schemes for $\rho = 0.9$ and $-c = 0.7$.

V. Conclusions

We have described a lattice-based scheme for lossy distributed compression of a linear function of two jointly Gaussian sources. The minimum sum-rate of our new scheme strictly improves those of the BT scheme and Krithivasan and Pradhan's scheme under certain conditions. The reason for improvement is due to more flexible choices of quantizers and linear combination coefficients, while exploring the remaining correlation among the lattice binning indices. There is still a gap between our sum-rate upper bound and the lower bound in [7], [8] when $c\rho < 0$. We are currently looking into new sum-rate lower bounds and leave these results for a future publication [10].

References

[1] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 2, pp, 219-221, March 1979.

[2] T. Berger, "Multiterminal source coding", *The Inform. Theory Approach to Communications,* G. Longo, Ed., New York: Springer-Verlag, 1977.

[3] S. Tung, *Multiterminal Rate-distortion Theory,* Ph. D. Dissertation, School of Electrical Engineering, Cornell University, Ithaca, NY, 1978.

[4] A. Wagner, S. Tavildar, and P. Viswanath, "The rate region of the quadratic Gaussian two-terminal source-coding problem," *IEEE Trans. Inform. Theory*, vol. 54, pp. 1938-1961, May 2008.

[5] D. Krithivasan, S. S. Pradhan, "Lattices for distributed source coding: jointly Gaussian sources and reconstruction of a linear function," *IEEE Trans. Inform. Theory*, vol. 55, No. 12, pp. 5628-5651, Dec. 2009.

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

[7] A. B. Wagner, "An outer bound for distributed compression of linear functions," Proc. CISS'08, pp. 435-440, Princeton, NJ, Mar. 2008.

[8] A. B. Wagner, "On distributed compression of linear functions," Proc. Allerton'08, pp. 1546-1553, Monticello, IL, Oct. 2008.

[9] M. A. Maddah-Ali and D. Tse, "Interference neutralization in distributed lossy source coding," Proc. ISIT'10, Jan. 2010.

[10] Y. Yang and Z. Xiong, "Distributed compression of linear functions: Partial sum-rate tightness and gap to optimal sum-rate," submitted to ISIT'11, St Petersburg, Russia, July 2011.