Approximating the Rate-Distortion Region of the Distributed Source Coding for Three Jointly Gaussian Tree-Structured Sources

Mohammad Ali Maddah-Ali
Dept. of Electrical Engineering and Computer Sciences,
University of California at Berkeley,
Email: maddah-a@eecs.berkeley.edu

David N. C. Tse
Dept. of Electrical Engineering and Computer Sciences,
University of California at Berkeley,
Email: dtse@eecs.berkeley.edu

***Abstract*—The rate-distortion region for the distributed source coding of the three jointly-Gaussian tree-structured sources with the quadratic distortion measure, is characterized within a constant gap. As a simplified counterpart of the Gaussian problem, we first investigate the rate region of a three binary-expanded sources where each pair of the sources have a certain number of the most-significant bits in common, and the central decoder needs to reconstruct each source with a target resolution. Motivated by the result of binary-expansion model, we prove that the achievable region of the quantize-and-binning scheme and the outer-bound of the cooperative scheme has a bounded gap of 2.4771 bits.**[1]

I. INTRODUCTION

In this paper, we investigate a system with three jointly Gaussian sources and three non-cooperative encoders, where each encoder observes one of the sources and sends an index message to a central decoder. The decoder requires reconstructing each source with a specific quadratic distortion. The objective is to approximate the ultimate boundaries for the rates of the messages, such that the decoder is able to reproduce the sources with at most the target distortions.

The lossless version of this problem with discrete sources was solved by Slepian and Wolf [2] using a random-binning approach. In [2], it is shown that for lossless scenarios, cooperation among encoders does not improve the sum-rate boundary. The slephian-Wolf result suggests a natural encoding scheme for the lossy distributed source coding. In this scheme, first the sources are quantized and transformed to discrete sources and then Slepian-Wolf scheme efficiently exploits the correlation of the quantized sources to minimize the rates. In [3], it is shown that for two Gaussian sources, this scheme of quantization-binning achieves some parts of the rate-distortion boundaries. In addition, in [4], [5], it is proven that this scheme is optimal for special forms of distributed source coding problems known as CEO problems. This results has been extended to the case, where the observation noise of the CEO problem are correlated with a specific tree structure [6]. Recently, in [7], it is proven that the quantization-binning scheme indeed achieves the entire rate-distortion region for two jointly Gaussian sources. The converse is based on coupling the problem with a Gaussian CEO problem. However, extending the converse proof to more that two sources does not look straight-forward. Recently, in [8], a linear finite-field deterministic channel model has been proposed to simplify understanding the key elements of the multi-user information theory. It is also suggested that the binary expansion model is used for the source coding problems [8]. In [9], the connection of the deterministic (lossless) and the non-deterministic models (lossy) has been exploited to approximate the rate region of the Gaussian symmetric multiple description coding.

In this paper, we first investigate a simplified counterpart of the Gaussian problem by considering three binary-expanded sources, where a certain number of the most-significant bits (MSBs) of each pair of the sources are common, and each source is required at the central decoder with a target resolution. The common bits and the target resolutions respectively represent the correlation and the target distortion of the Gaussian sources. We then argue that in the deterministic binary-expansion model, the cooperative (cut-set) outer-bound is achievable. This result motivates us to develop a simple version of the quantization-binning method as the achievable scheme and also use the cooperative (cut-set) bound as the outer-bound. We show that for the tree-structured sources, the inner-bound and the outer-bound has a gap less than or equal to 2.4771 bits.

The rest of the paper is organized as follows. In Section II, the problem is formulated. In Section III, the rate-region of the binary-expansion model is discussed. In Sections IV and V, the the achievable region and outer-bound are formulated. The bounded Gap for the sum-rate is established in Section VI. The general result is proven in Section VII.

II. PROBLEM FORMULATION

Let $\{y_1(t), y_2(t), y_3(t)\}_{t=1}^{n}$ be a sequence of independent and identically distributed (i.i.d.) Gaussian random variables. It is assumed that $(y_1(t), y_2(t), y_3(t))$ has zero mean and $\mathbf{K}_y$

[1]The detail proofs are available in [1].

covariance matrix where,

$$\mathbf{K}_y = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}. \qquad (1)$$

In the problem of the distributed source coding, there are three non-cooperative encoders, where the j^{th} encoder observes $\{y_j(t)\}_{t=1}^n$, and sends a message, chosen from $\{1, 2, \ldots, M_j\}$, to the centralized decoder. The decoder receives the messages from the three encoders and estimates $\{\hat{y}_1(t), \hat{y}_2(t), \hat{y}_3(t)\}_{t=1}^n$, for all $j \in \{1, 2, 3\}$.

We define Δ_j as $\Delta_j = \frac{1}{n}\sum_{t=1}^n E[(y_j^n(t) - \hat{y}_j^n(t))^2]$, $\forall j \in \{1, 2, 3\}$. The rate-distortion tuple $(R_1, R_2, R_3, d_1^*, d_2^*, d_3^*)$ is admissible, if for every $\epsilon > 0$ and for a sufficiently large n, there exists a seven-tuple $(n, M_1, M_2, M_3, \Delta_1, \Delta_2, \Delta_3)$, such that $\frac{1}{n}\log(M_j) \leq R_j + \epsilon$ and $\Delta_j \leq d_j^* + \epsilon$ for $j = 1, 2, 3$.

In this paper, the objective is to approximate the rate-distortion region Ω within a constant gap for the tree-structured sources. For this class of sources, there is an i.i.d. random sequence $\{x(t)\}_{t=1}^n$, where $x(t) \sim \mathcal{N}(0, 1)$, and

$$y_j(t) = c_j x(t) + n_j(t), \; j = 1, 2, 3. \qquad (2)$$

$\{n_j(t)\}_{t=1}^n$ are i.i.d. mutually independent Gaussian random variables with zero mean and $1-c_j^2$ variances, and independent of $\{x(t)\}_{i=1}^n$. c_j is a constant, $0 \leq c_j \leq 1$, for $j = 1, 2, 3$. It is easy to see that $\rho_{jk} = c_j c_k$. For simplicity and without loss of generality, we assume $c_1 \leq c_2 \leq c_3$, and therefore, $\rho_{12} \leq \rho_{13} \leq \rho_{23}$. It is easy to see that $\mathbf{K}_y = \mathrm{diag}([1 - c_1^2, 1 - c_2^2, 1 - c_3^2]) + [c_1, c_2, c_3][c_1, c_2, c_3]'$. Using the identity that $|\mathbf{X} + \mathbf{ab}'| = |\mathbf{X}|(1 + \mathbf{b}'\mathbf{X}^{-1}\mathbf{a})$, we have $|\mathbf{K}_y| = (1 + \sum_{i=1}^3 \frac{c_i^2}{1-c_i^2})\prod_{i=1}^3(1 - c_i^2)$. Then, it is easy to show that

$$(1 - \rho_{12})(1 - \rho_{23}) \leq |\mathbf{K}_y|. \qquad (3)$$

Using (3), we have

$$\frac{(1 - \rho_{23}^2)(1 - \rho_{13}^2)}{|\mathbf{K}_y|} \leq \frac{(1 - \rho_{23}^2)(1 - \rho_{12}^2)}{|\mathbf{K}_y|} \leq 4. \qquad (4)$$

On the other hand, for general $\mathbf{K}_y$ with the form (1), we can show that

$$|\mathbf{K}_y| \leq (1 - \rho_{23}^2)(1 - \rho_{13}^2). \qquad (5)$$

Remark: We note that if the sources have other form of tree structure like what is depicted in Fig. 1, we can simply transform it to the fork-shape tree. For example, for the tree shown in Fig. 1, we can take node A as the root and transform the noise statistics accordingly to have the fork-shape tree.

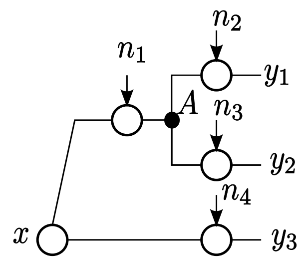

Fig. 1. An Example of the Tree Structure

III. Distributed Coding for The Binary-Expansion Model

We notice that $I(x; y_j) = -.5\log_2(1 - c_j^2)$. Roughly speaking, it means that the $-.5\log_2(1 - c_j^2)$ most significant bits (MSBs) of y_j and x are the same. Because of the independent additive noise n_j, the remaining least significant bit's (LSB's) of y_j and x are independent. Therefore for every two sources y_j and y_k, the most significate bits are the same and the rest are independent. Based on this observation, we suggest a simple binary-expansion model of the Gaussian problem. In this model, we assume y_j has a uniform distribution over $[0, 1]$ with the following binary expansion:

$$y_j = 0.y_j^1 y_j^2 y_j^3 \ldots, \qquad (6)$$

where y_j^l is the $l^t h$ bit in the binary expansion of y_j, and is zero or one, equally likely. We assume that y_j and y_k have the m_{jk} MSBs in common, where $\rho_{jk} = \delta(1 - 2^{-2m_{jk}})$ and $\delta = \frac{1}{12}$. For simplicity and without loss of the generality, we assume that $m_{12} \leq m_{13} \leq m_{23}$, then it is easy to see that $m_{12} = m_{13}$. In addition assume that the decoder requires the first b_j bits of source y_j, for $j = 1, 2, 3$. Therefore, LSBs after $y_j^{b_j}$ are irrelevant. Since the source has been transformed to a discrete sources, then we can conclude from Slepian-Wolf theorem that cooperative outer-bound is tight. However, for these specific sources, we can presents an insightful explicit achievablity and converse proof.

A. Achievable Scheme

The achievable scheme consists of two stages:

Stage 1- Quantization Stage: In this stage, the first b_j bits of y_j are selected to form u_j as follows:

$$u_j = 0.y_j^1 y_j^2 y_j^3 \ldots y_j^{b_j} 1, \qquad (7)$$

'1' is added as the last digit of u_j, so that u_j is an unbiased estimation of y_j. It is easy to see $E[(y_j - u_j)^2] \leq d_j^*$, where $d_j^* = \delta 2^{-2b_j}$. Therefore, if all bits of u_j, $j = 1, 2, 3$, are available at the decoder, it can reproduce y_j with the distortion of at most $d_j^* = 2^{-2b_j}$.

Stage 2- Redundancy Elimination: As mentioned, u_j's have some common bits, induced by the correlation of y_j's. Therefore, the non-cooperative encoders agree on a transmission protocol in which if a bit is in common between any two or three of them, only one encoder is in charge to send it to the decoder. One simple transmission protocol is presented in the converse proof.

B. Cooperative Outer Bound

In what follows, we use the cooperative (cut-set) outer-bound on $\sum_{j\in S} R_j$, for any $S \subset \{1, 2, 3\}$. To derive this outer-bound, we assume that a central encoder observes the sources in S, and sends an index to the decoder such that all y_j, $j \in S$, are able to be reconstructed with the target resolutions, while all y_i, $i \in S^c$, are perfectly available at the decoder as the side information. For any set S, $S \subset$

$E = \{1,2,3\}$, we define $F(S)$ as the set of all bits in $\{y_j^l, l = 1, \ldots, b_j,\ j = 1,2,3\}$ which are *exclusively* available at the observation nodes S. In addition, we define $g(S)$ as the cardinality of $F(S)$, i.e. $g(S) = |F(S)|$. Then, the cooperative (cut-set) outer-bound is formulated as follows.

$$C^{out} = \{(R_1, R_2, R_3) \mid \sum_{j \in S} R_j \geq g(S),\ \forall S \subset E\}. \quad (8)$$

Theorem 1 *In the binary expansion model, the cooperative outer-bound C^{out} is achievable.*

Sketch of the proof: First we prove that C^{out} is a contra-polymatroid. Consequently, C^{out} has six corner points corresponding to the different permutations of E. Let π be a permutation of the set E. Then, the following distributed scheme achieve the corner point of C^{out} corresponding to the permutation π. In this scheme, node $\pi(1)$ sends all bits in $F(\{\pi(1)\})$, node $\pi(2)$ sends all bits in $F(\{\pi(1), \pi(2)\}) - F(\{\pi(2)\})$, and node $\pi(3)$ sends all bits in $F(\{\pi(1), \pi(2), \pi(3)\}) - F(\{\pi(1), \pi(2)\})$. Since all corner points of the outer-bound is achievable, and since the outer-bound is a contra-polymatroid, then all points in C^{out} is achievable. ■

The above explicit proof, rather than the probabilistic proof of Slepian-Wolf, shows that network coding is not needed to achieve the rate region. Note that in some source coding problems, such as the non-symmetric multiple description coding, network coding is a necessary part of the achievable scheme [10]. Moreover, this proof of the converse leads us to develop a similar approach for the Gaussian sources to establish the bounded gap.

IV. Achievable Scheme for the Gaussian Sources

Similar to what has been presented for the binary expansion model, the encoding scheme for Gaussian sources includes two stages: (I) Quantization and (II) Redundancy Elimination (Random Binning). In the first stage, the sources are quantized using the test channels

$$u_j = \eta_j y_j + z_j,\ j = 1,2,3, \quad (9)$$

where $z_j \sim \mathcal{N}(0, 1 - \eta_j^2)$, and z_j is independent of y_j. The encoder j chooses one codeword $\{u_j(t)\}_{t=1}^n$ which is jointly typical with $\{y_j(t)\}_{t=1}^n$. In the second stage, we use the random binning scheme to eliminate the redundancy.

Motivated by the results of the binary-expansion model, we choose η_j, $j = 1,2,3$, such that

$$E[(y_j - E(y_j|u_j))^2] \leq d_j^*. \quad (10)$$

This means that we simply ignore the correlation of y_1, y_2, and y_3 in constructing the quantizers. In the random-binning stage, we use Slepian-Wolf scheme to eliminate the redundancy in the output of the quantizers. We note that in a more efficient scheme, we should build the quatizers such that $E[(y_j - E(y_j|u_1, u_2, u_3))^2] \leq d_j^*$. Our results confirm that the simplified version of quantizers (10) costs the system at most around 2.4771 bits, while it makes the analysis tractable. Using the results on the rate-distortion theory of the single Gaussian source, we conclude that in the simplified scheme, η_j is selected such that $I(y_j; u_j) = \frac{1}{2}\log\frac{1}{d_j^*}$, i.e. $\eta_j = \sqrt{1 - d_j^*},\ j = 1,2,3$. It is proven that if

$$\sum_{j \in S} R_j \geq I(y(S); u(S)|u(S^c)),\ \forall S \subset \{1,2,3\}, \quad (11)$$

then we can develop encoding scheme such that the decoder can recover the sources with the distortion matrix, arbitrary close to $\mathbf{D}_I$, where

$$\mathbf{D}_I^{-1}(d_1^*, d_2^*, d_3^*) = \mathbf{K}_y^{-1} + \text{diag}[z_1^*, z_2^*, z_3^*], \quad (12)$$

and $z_j^* = \frac{1-d_j^*}{d_j^*}$. It is easy to see that the diagonal elements of $\mathbf{D}_I$ are upper-bounded with d_j^*. Let us define $R^{in}_{\sum_{j \in S}}(d_1^*, d_2^*, d_3^*)$ as $I(y(S); u(S)|u(S^c))$, for $S \subset E$, and contra-polymatroid $\mathcal{B}(d_1^*, d_2^*, d_3^*)$ as the set of all triples (R_1, R_2, R_3) satisfying $\sum_{j \in S} R_j \geq R^{in}_{\sum_{j \in S} j}(d_1^*, d_2^*, d_3^*)$. Regarding the Markovian structure, governing the relation of y_j and u_j, $j = 1,2,3$, we can show that $I(y(S); u(S)|u(S^c)) = I(y(E); u(E)) - I(y(S^c); u(S^c))$. Then we can show that

$$R^{in}_{1+2+3}(d_1^*, d_2^*, d_3^*) = \frac{1}{2}\log\frac{|\mathbf{K}_y|}{|\mathbf{D}_I(d_1^*, d_2^*, d_3^*)|}, \quad (13)$$

$$R^{in}_{i+j}(d_1^*, d_2^*, d_3^*) = R^{in}_{1+2+3}(d_1^*, d_2^*, d_3^*) - \frac{1}{2}\log(1 + z_k^*), \quad (14)$$

$$R^{in}_{i}(d_1^*, d_2^*, d_3^*) = R^{in}_{1+2+3}(d_1^*, d_2^*, d_3^*) - \frac{1}{2}\log\frac{|\mathbf{K}_{(y_j, y_k)}|}{|\mathbf{D}_I(d_j^*, d_k^*)|},$$

and

$$\mathbf{D}_I^{-1}(d_j^*, d_k^*) = \mathbf{K}^{-1}_{(y_j, y_k)} + \text{diag}[z_j^*, z_k^*]. \quad (15)$$

The base of the logarithms is two. If the triple (R_1, R_2, R_3) belongs to $\mathcal{B}(d_1, d_2, d_3)$, defined above, then we can develop encoding scheme such that the decoder can recover the sources with the distortion matrix, arbitrary close to $\mathbf{D}_I(d_1^*, d_2^*, d_3^*)$, where $\mathbf{D}_I^{-1}(d_1, d_2, d_3)$ is defined in (12). We define, $\mathcal{C}(d_1^*, d_2^*, d_3^*)$ as

$$\mathcal{C}(d_1^*, d_2^*, d_3^*) = \text{Convex Hull} \bigcup_{d_1 \leq d_1^*, d_2 \leq d_2^*, d_3 \leq d_3^*} \mathcal{B}(d_1, d_2, d_3).$$

If $(R_1, R_2, R_3) \in \mathcal{C}(d_1^*, d_2^*, d_3^*)$, the sources can be reproduced by the distortions of at most d_1^*, d_2^*, d_3^*.

V. Cooperative Outer Bound

For the binary-expansion model, it is shown that the gap between the achievable scheme and the cooperative outer-bound is zero. This result motivates us to use the same outer-bound for the Gaussian sources. Since the source is jointly Gaussian, the cooperative outer-bound on sum-rate, denoted by $R^{out}_{1+2+3}(d_1^*, d_2^*, d_3^*)$, is equal to

$$R^{out}_{1+2+3} = \frac{1}{2}\log\frac{|\mathbf{K}_y|}{\Gamma_E}. \quad (16)$$

where

$$\Gamma_E = \max_{\mathbf{D}_{co}} |\mathbf{D}_{co}| \quad (17)$$

$$s.t. \quad \mathbf{0} \leq \mathbf{D}_{co} \preceq \mathbf{K}_y \quad (18)$$

$$\mathbf{D}_{co}(j,j) \leq d_j,\ j=1,2,3 \quad (19)$$

The optimization (17) is a convex optimization problem and can be solved numerically. However, to compute the gap between the sum-rate of the centralized and the distributed source coding, we need an explicit expression for R^{out}_{1+2+3}, which is, to the best of our knowledge, not available. Alternatively, we propose the following upper-bound on Γ_E.

Theorem 2 *Let $\mathbf{D}^*_{co}$ be the optimizing matrix in* (17)*, with diagonal entries $\hat{d}_j$, $\hat{d}_j \leq d_j$, $j=1,2,3$. Then,*

$$\Gamma_E = |\mathbf{D}^*_{co}| \leq \hat{\Gamma}(\hat{d}_1, \hat{d}_2, \hat{d}_3), \quad (20)$$

where

$$\hat{\Gamma}(\hat{d}_1, \hat{d}_2, \hat{d}_3) = \hat{d}_1\hat{d}_2\hat{d}_3 \left(1-\theta^2_{13}(\hat{d}_1,\hat{d}_3)\right)\left(1-\theta^2_{23}(\hat{d}_2,\hat{d}_3)\right),$$

and

$$\theta_{jk}(\hat{d}_j,\hat{d}_k) = \max\left\{0, \frac{\rho_{12} - \sqrt{(1-\hat{d}_j)(1-\hat{d}_k)}}{\sqrt{\hat{d}_j\hat{d}_k}}\right\}.$$

In addition, $\hat{d}_j$ satisfies the inequalities $\hat{d}_j \leq d^*_j$, for $j = 1,2,3$, and $\theta_{jk}(\hat{d}_j,\hat{d}_k) \leq 1$.

$\theta_{jk}(\hat{d}_1,\hat{d}_2)$ has the following important property.

Lemma 3 *For any pair $(\hat{d}_j,\hat{d}_k)$, $1 \neq 2$, $\hat{d}_j \leq 1$, $\hat{d}_j \leq 1$, if $\theta_{jk}(\hat{d}_j,\hat{d}_k) \leq 1$, then we have,*

$$d_j\left(1-\theta^2_{jk}(\hat{d}_j,\hat{d}_k)\right) \leq 1-\rho^2_{jk}. \quad (21)$$

Combining Lemma 3 and Theorem 2, we can develop several upper-bounds. For example, we can show that $\Gamma_E \leq \min\{d^*_1,d^*_2,d^*_3\}(1-\rho^2_{13})(1-\rho^2_{23})$, or $\Gamma_E \leq d^*_1 d^*_2(1-\rho^2_{23})$. Note that unlike the outer-bound introduced in [11], this outer-bound captures the variation of both the target distortions and the cross-correlations.

Consider a permutation π of the set $E=\{1,2,3\}$. Using the cooperative (cut-set) outer-bound, we can derive outer-bound on $R_{\pi(1)}+R_{\pi(2)}$ and $R_\pi(1)$, denoted respectively by $R^{out}_{\pi(1)+\pi(2)}(d^*_1,d^*_2,d^*_3)$ and $R^{out}_{\pi(1)}(d^*_1,d^*_2,d^*_3)$. For simplicity, we define $\Gamma_{(\pi(1),\pi(2))|\pi(3)}$ and $\Gamma_{(\pi(1)|(\pi(2),\pi(3))}$, such that

$$R^{out}_{\pi(1)+\pi(2)}(d^*_1,d^*_2,d^*_3) = \frac{1}{2}\log\frac{|\mathbf{K}_y|}{\Gamma_{(\pi(1),\pi(2))|\pi(3)}}, \quad (22)$$

$$R^{out}_{\pi(1)}(d^*_1,d^*_2,d^*_3) = \frac{1}{2}\log\frac{|\mathbf{K}_y|}{\Gamma_{\pi(1)|(\pi(2),\pi(3))}}. \quad (23)$$

$\Gamma_{(\pi(1),\pi(2))|\pi(3)}$ and $\Gamma_{(\pi(1)|(\pi(2),\pi(3))}$ can be formulated as a convex optimization based on Wyner-Ziv result on rate-distortion region with side information.

VI. Gap between Sum-Rate of Inner and Outer bounds

In this section, we show that the gap between the sum-rate of the achievable scheme and the outer-bound is bounded.

Theorem 4 *The Gap between $R^{in}_{1+2+3}(d^*_1,d^*_2,d^*_3)$ and $R^{out}_{1+2+3}(d^*_1,d^*_2,d^*_3)$ is less than or equal to $\frac{1}{2}\log 25$.*

Sketch of the proof: We have,

$$R^{in}_{1+2+3} - R^{out}_{1+2+3} = \frac{1}{2}\log\frac{|\mathbf{K}_y|}{|\mathbf{D}_I|} - \frac{1}{2}\log\frac{|\mathbf{K}_y|}{\Gamma_E} \quad (24)$$
$$\leq \frac{1}{2}\log\hat{\Gamma}(\hat{d}_1,\hat{d}_2,\hat{d}_3)|\mathbf{D}_I^{-1}|.$$

Regarding (12), we have

$$|\mathbf{D}_I^{-1}(d^*_1,d^*_2,d^*_3)| = \frac{1}{|\mathbf{K}_y|}\{1+\sum_{S\subset E, S\neq\emptyset}|\mathbf{K}_{y(S)}|\prod_{j\in S} z^*_j\}$$
$$\leq \frac{1}{|\mathbf{K}_y|}\{\sum_{S\subset E, S\neq\emptyset}|\mathbf{K}_{y(S)}|\prod_{j\in S}\frac{1}{d^*_j}\}.$$

Then, we can use the above inequality, (24), Lemma 3, (4), and (5) to prove the theorem.

■

VII. Gaps from other Boundaries

In this section, we aim to extend this result, and show that $\mathcal{C}(d^*_1,d^*_2,d^*_3)$ has a bounded gap with other outer bounds $R^{out}_{j+k}(d^*_1,d^*_2,d^*_3)$ and $R^{out}_j(d^*_1,d^*_2,d^*_3)$, for all $j,k\in\{1,2,3\}$, $j \neq k$. For simplicity, we first review the outline of the proof for a simpler scenario which involves only two jointly Gaussian sources. In Fig 2, a typical rate region and the outer-bounds for two Gaussian sources have been shown. The boundaries of the achievable region $\mathcal{C}(d^*_1,d^*_2)$ consists of one straight and two curvy lines. In this figure, the outer-bounds have been shown by blue lines, and $\mathcal{B}(d^*_1,d^*_2)$ has been shown by dashed red lines. Similar to Theorem (4), we can prove that $\mathcal{B}(d^*_1,d^*_2)$ has a bounded gap from $R^{out}_{1+2}(d^*_1,d^*_2)$. However, there is no guarantee that $\mathcal{B}(d^*_1,d^*_2)$ has a bounded gap from the other two outer-bounds i.e. $R^{out}_1(d^*_1,d^*_2)$ and $R^{out}_2(d^*_1,d^*_2)$. We explain this by focusing on the point C^{in} of the achievable region. Roughly speaking, at this point, R_1 is the minimum rate that the decoder needs to reconstruct the first source with distortion d^*_1, while it has the second source *with the distortion* d^*_2 as the side information. However, R^{out}_1 is the minimum required rate where the decoder can reproduce the first source with distortion d^*_1, while it has the second source *perfectly* as the side information. One can construct examples for which the gap between these two rates is unbounded. However, it turns out that if the decoder has access to the second source with the distortion d^N_2, which is slightly less than d^*_2, it can reconstruct the first source with the distortion d^*_1 and with rate R_1, while $R_1 - R^{out}_1$ is bounded. In Fig. 2, we focus on another contra-polymatroid $\mathcal{B}(d^*_1,d^N_2) \subset \mathcal{C}(d^*_1,d^*_2)$ and the point A^{in}, where $d^N_2 \leq d^*_2$. At this point we have the the second source with the distortion of at most d^N_2 and the first source with the distortion of at most d^*_1. As we reduce

d_2^N, $R_1^{in}(d_1^*, d_2^N) - R_1^{out}(d_1^*, d_2^*)$ becomes smaller, however at the same time $R_{1+2}^{in}(d_1^*, d_2^N) - R_{1+2}^{out}(d_1^*, d_2^*)$ becomes larger. Noting that the final gap of the inner and outer bounds is the maximum of all gaps, we choose d_2^N such that $R_1^{in}(d_1^*, d_2^N) - R_1^{out}(d_1^*, d_2^*)$ and $R_{1+2}^{in}(d_1^*, d_2^N) - R_{1+2}^{out}(d_1^*, d_2^*)$ are as close as possible. We choose d_2^N in a two-stage procedure. First, we select d_2^T such that

$$R_1^{in}(d_1^*, d_2^T) - R_1^{out}(d_1^*, d_2^*) = R_{1+2}^{in}(d_1^*, d_2^T) - R_{1+2}^{out}(d_1^*, d_2^*),$$

and then $d_2^N = \min\{d_2^*, d_2^T\}$. The second stage has been added to ensure that $d_2^N \leq d_2^*$ and therefore $\mathcal{B}(d_1^*, d_2^N) \subset \mathcal{C}(d_1^*, d_2^*)$. We can show that $\mathcal{B}(d_1^*, d_2^N)$ has a bounded gap from both $R_{1+2}^{out}(d_1^*, d_2^*)$ and $R_1^{out}(d_1^*, d_2^*)$. Similarly, we introduce d_1^N such that $\mathcal{B}(d_1^N, d_2^*) \subset \mathcal{C}(d_1^*, d_2^*)$ has a bounded gap with both $R_{1+2}^{out}(d_1^*, d_2^*)$ and $R_2^{out}(d_1^*, d_2^*)$. Consequently, the convex hull of these two contra-polymatroids has a bounded gap from all three outer bounds.

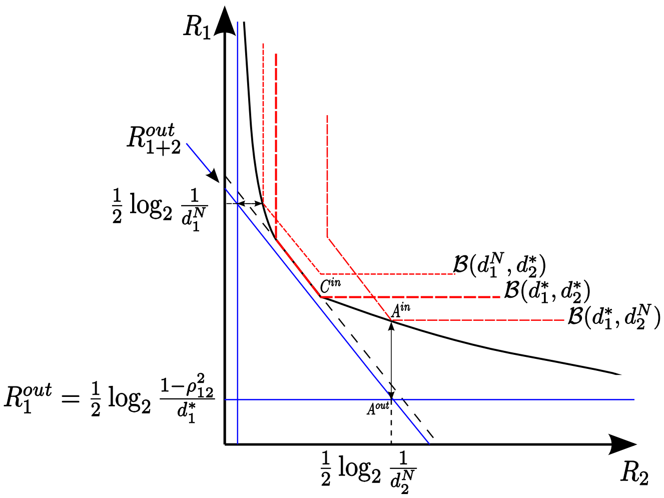

Fig. 2. $\mathcal{C}(d_1^*, d_2^*)$ for Two Jointly Gaussian Sources

We mimic the same scheme for the three-source scenario. Let π be a permutation of E. We aim to find a contra-polymatoird $\mathcal{B}(z_{\pi(1)}^*, z_{\pi(2)}^N, z_{\pi(3)}^N) \subset \mathcal{C}(z_{\pi(1)}^*, z_{\pi(2)}^*, z_{\pi(3)}^*)$ which has bounded and balanced gaps from $R_{1+2+3}^{out}(d_1^*, d_2^*, d_3^*)$, $R_{\pi(1)+\pi(2)}^{out}(d_1^*, d_2^*, d_3^*)$, and $R_{\pi(1)}^{out}(d_1^*, d_2^*, d_3^*)$. Similar to the two-source case, we choose $z_{\pi(3)}^N$ and $z_{\pi(2)}^N$ such that $R_{1+2+3}^{in}(d_1^*, d_2^N, d_3^N) - R_{1+2+3}^{out}(d_1^*, d_2^*, d_3^*)$, $R_{\pi(1)+\pi(2)}^{in}(d_1^*, d_2^N, d_3^N) - R_{\pi(1)+\pi(2)}^{out}(d_1^*, d_2^*, d_3^*)$, and $R_{\pi(1)}^{in}(d_1^*, d_2^N, d_3^N) - R_{\pi(1)}^{out}(d_1^*, d_2^*, d_3^*)$ are as close as possible. First we choose $z_{\pi(3)}^T$ such that $R_{1+2+3}^{in}(d_1^*, d_2^T, d_3^T) - R_{1+2+3}^{out}(d_1^*, d_2^*, d_3^*)$ and $R_{\pi(1)+\pi(2)}^{in}(d_1^*, d_2^T, d_3^T) - R_{\pi(1)+\pi(2)}^{out}(d_1^*, d_2^*, d_3^*)$ are equal. Using (14), we have

$$\begin{aligned}&\frac{1}{2}\log(1 + z_{\pi(3)}^T) \\ &= R_{\pi(1)+\pi(2)+\pi(3)}^{out}(d_1^*, d_2^*, d_3^*) - R_{\pi(1)+\pi(2)}^{out}(d_1^*, d_2^*, d_3^*).\end{aligned}$$

Then, $z_{\pi(3)}^N$ is chosen as $z_{\pi(3)}^N = \max\{z_{\pi(3)}^T, z_{\pi(3)}^*\}$. Due to the above choice of $z_{\pi(3)}^N$, we have,

$$\begin{aligned}R_{\pi(1)+\pi(2)}^{in}(d_1^*, d_2^T, d_3^N) - R_{\pi(1)+\pi(2)}^{out}(d_1^*, d_2^*, d_3^*) \leq \\ R_{\pi(1)+\pi(2)+\pi(3)}^{in}(d_1^*, d_2^T, d_3^N) - R_{\pi(1)+\pi(2)+\pi(3)}^{out}(d_1^*, d_2^*, d_3^*),\end{aligned}$$

irrespective of $z_{\pi(2)}^T$. This means that $R_{1+2+3}^{in}(d_1^*, d_2^T, d_3^N) - R_{1+2+3}^{out}(d_1^*, d_2^*, d_3^*)$ is the dominant gap, therefore we choose $z_{\pi(2)}^T$ such that

$$R_{\pi(1)}^{in}(d_1^*, d_2^T, d_3^N) - R_{\pi(1)}^{out}(d_1^*, d_2^*, d_3^*) = \quad (25)$$

$$R_{1+2+3}^{in}(d_1^*, d_2^T, d_3^N) - R_{1+2+3}^{out}(d_1^*, d_2^*, d_3^*). \quad (26)$$

Then, $z_{\pi(2)}^N$ is chosen as $z_{\pi(2)}^N = \max\{z_{\pi(2)}^T, z_{\pi(2)}^*\}$. We can show that $\mathcal{B}(z_{\pi(1)}^*, z_{\pi(2)}^N, z_{\pi(3)}^N) \subset \mathcal{C}(d_1^*, d_2^*, d_3^*)$ has a gap of at most $\frac{1}{2}\log 31$ from the three outer-bounds hyperplane $R_{1+2+3}^{out}(z_{\pi(1)}^*, z_{\pi(2)}^*, z_{\pi(3)}^*)$, $R_{\pi(1)+\pi(2)}^{out}(z_{\pi(1)}^*, z_{\pi(2)}^*, z_{\pi(3)}^*)$, and $R_{\pi(1)}^{out}(z_{\pi(1)}^*, z_{\pi(2)}^*, z_{\pi(3)}^*)$. Then, the convex hull of the union of the six introduced contra-polymatoid is (i) a subset of $\mathcal{C}(d_1^*, d_2^*, d_3^*)$, (ii) has a gap of at most $\frac{1}{2}\log 31 = 2.4771$ from all outer-bounds. Then, it is easy to see that for any (R_1^*, R_2^*, R_3^*) on the boundary of the union of the six contra-polymatroids, $(R_1^* - \beta, R_2^* - \beta, R_3^* - \beta)$ is not achievable, where $\beta = 2.4771 + \epsilon$, $\forall \epsilon \geq 0$.

Theorem 5 *$\mathcal{C}(d_1^*, d_2^*, d_3^*)$ has a bounded gap of at most $\frac{1}{2}\log 31$ from the cooperative outer-bounds.*

For the detailed proofs, please refer to [1].

References

[1] M.A. Maddah-Ali and D. Tse, "Approximating the rate-distortion region of the distributed source coding for three jointly Gaussian three-structured sources," Tech. Rep., 2009, available at http://eecs.berkeley.edu/∼maddah-a/Three_source.pdf.

[2] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, pp. 471–480, Jul. 1973.

[3] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Transactions on Information Theory*, vol. 43, pp. 1912–1923, Nov. 1997.

[4] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Transactions on Information Theory*, vol. 51, pp. 2577–2593, July 2005.

[5] V. Prabhakaran, D. Tse, and K. Ramachandran, "Rate region of the quadratic Gaussian CEO problem," in *Proceedings of International Symposium on Information Theory*, June 2004, p. 119.

[6] S. Tavildar, P. Viswanath, and A.B. Wagner, "The Gaussian many-help-one distributed source coding problem," *IEEE Transactions on Information Theory*, 2008, submitted.

[7] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic Gaussian two-encoder source-coding problem," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1938–1961, May 2008.

[8] Amir Salman Avestimehr, *Wireless network information flow: a deterministic approach*, Ph.D. thesis, EECS Department, University of California, Berkeley, October 2008.

[9] C. Tian, S. Mohajer, and S.N. Diggavi, "Approximating the Gaussian multiple description rate region under symmetric distortion constraints," *IEEE Transactions on Information Theory*, 2008, submitted.

[10] R.W. Yeung, "Multilevel diversity coding with distortion," *IEEE Transactions on Information Theory*, vol. 41, pp. 412–422, March 1995.

[11] R. Zamir and T. Berger, "Multiterminal source coding with high resolution," *IEEE Transactions on Information Theory*, vol. 45, pp. 106–117, Jan. 1999.