A New Achievable Region for Gaussian Multiple Descriptions Based on Subset Typicality

Kumar Viswanatha, Emrah Akyol and Kenneth Rose
ECE Department, University of California - Santa Barbara
{kumar,eakyol,rose}@ece.ucsb.edu

***Abstract*—This paper addresses the $L-$channel multiple descriptions problem for a Gaussian source under mean squared error (MSE) distortion metric. We focus on particular cross-sections of the general rate-distortion region where a subset the $2^L - 1$ distortion constraints are not imposed. Specifically, we assume that certain descriptions are never received simultaneously at the decoder and thereby the transmitted codewords require joint typicality only within prescribed subsets. We derive a new encoding scheme and an associated rate-distortion region wherein joint typicality of codewords only within the prescribed subsets is maintained. We show that enforcing joint typicality of all the codewords entails strict suboptimality in the achievable rate-distortion region. Specifically, we consider a 3 descriptions scenario wherein descriptions 1 and 3 are never received simultaneously at the decoder and show that a strictly larger achievable region is obtained when the encoder maintains joint typicality of codewords only within the required subsets. To prove these results, we derive a lemma called the 'subset typicality lemma' which plays a critical role in establishing the new achievable region.**

***Index Terms*—Subset typicality lemma, Gaussian multiple descriptions under MSE, $L-$channel multiple descriptions**

I. INTRODUCTION

The Multiple Descriptions (MD) problem has been studied extensively since the late seventies, see for eg. [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11]. It was originally designed as a source-channel coding strategy to cope with channel failures, where multiple source descriptions are generated and sent over different paths. In the $L-$descriptions setting, the encoder generates L packets which are transmitted over L available channels. It is assumed that the decoder receives a subset of the descriptions perfectly and the remaining are lost. The objective is to characterize the set of achievable rate-distortion tuples for the L description rates and the 2^L-1 distortions, one for each possible combination of the received descriptions.

One of the first achievable regions for the 2-channel MD problem was derived by El-Gamal and Cover in 1982 [1]. It is given by the convex closure over all tuples $(R_1, R_2, D_1, D_2, D_{12})$ for which there exist auxiliary random variables $\hat{X}_1$, $\hat{X}_2$ jointly distributed with X taking values over arbitrary alphabets and functions $\psi_{\mathcal{K}},\ \mathcal{K} \in \{1, 2, 12\}$ such that[1],

$$\begin{aligned} R_i &\geq I(X; \hat{X}_i) \\ R_1 + R_2 &\geq I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2) \\ D_{\mathcal{K}} &\geq E\left[d_{\mathcal{K}}(X, \psi_{\mathcal{K}}(\{\hat{X}\}_{\mathcal{K}})\right] \ \mathcal{K} \in \{1, 2, 12\} \end{aligned} \tag{1}$$

where $\{\hat{X}\}_{\mathcal{K}}$ denotes the set $\{\hat{X}_i : i \in \mathcal{K}\}$. Ozarow [4] showed that for the 2 descriptions setting, the El-Gamal Cover region is complete for a Gaussian source under mean squared error (MSE) distortion metric. Specifically, he showed that the complete region can be achieved using a 'correlated quantization' scheme which imposes the following joint distribution for $\hat{X}_1, \hat{X}_2$:

$$\hat{X}_1 = X + W_1, \qquad \hat{X}_2 = X + W_2 \tag{2}$$

where W_1 and W_2 are zero mean Gaussian random variables independent of X with covariance matrix $K_{W_1W_2}$ and the functions $\psi_{\mathcal{K}}(\{\hat{X}\}_{\mathcal{K}})$ are given by the respective minimum mean squared error (MMSE) optimal estimators, for eg., $\psi_{12}(\hat{X}_1, \hat{X}_2) = E\left[X|\hat{X}_1, \hat{X}_2\right]$. The covariance matrix $K_{W_1W_2}$ is set to satisfy all the distortion constraints. We denote the complete Gaussian-MSE 2-descriptions region by $\mathcal{RD}_2$. It has been shown in the literature that the El-Gamal Cover region is complete for general sources and distortion measures in the no-excess rate regime [3]. However, Zhang and Berger showed in [2] that this region is not complete in general. Specifically, for a binary source and Hamming distortion measure, they showed that sending a common codeword among the two descriptions can lead to a strictly larger achievable region.

There have been several publications on extending the Gaussian 2-channel results to the $L-$channel framework [5], [10], [11], [12]. A natural extension of the 'correlated quantization' coding scheme for the $L-$channel setting (see for example [5]) is described as follows. Let $W_1, W_2, \ldots, W_L$ be zero mean Gaussian random variables independent of X with covariance matrix $K_{\mathbf{W}}$ and let $\hat{X}_i$ be such that:

$$\hat{X}_i = X + W_i,\ i \in \mathcal{L} \tag{3}$$

[1]Note that the original characterization of El-Gamal and Cover also includes a refinement random variable $\hat{X}_{12}$. However, it was shown recently in [7] that $\hat{X}_{12}$ can be set to a constant without loss of optimality.

The work was supported by the NSF under grants CCF - 1016861 and CCF-1118075

where $\mathcal{L} = \{1, 2, \ldots, L\}$. Let $\forall \mathcal{K} \subseteq \mathcal{L}$, $\psi_{\mathcal{K}}(\{\hat{X}\}_{\mathcal{K}}) = E\left[X|\{\hat{X}\}_{\mathcal{K}}\right]$ and the covariance matrix $K_{\mathbf{W}}$ be such that:

$$\text{Var}\left(X|\{\hat{X}\}_{\mathcal{K}}\right) \leq D_{\mathcal{K}} \quad (4)$$

i.e., all the $2^L - 1$ distortion constraints are satisfied. Then any rate tuple $(R_1, R_2, \ldots, R_L)$ satisfying the following conditions is achievable:

$$\sum_{i \in \mathcal{K}} R_i \geq \sum_{i \in \mathcal{K}} h(\hat{X}_i) - h(\{\hat{X}\}_{\mathcal{K}}|X) \quad \forall \mathcal{K} \subseteq \mathcal{L} \quad (5)$$

The convex closure of such achievable rate tuples over all $K_{\mathbf{W}}$ satisfying the distortion constraints is an achievable region for the $L-$channel Gaussian MD problem and is denoted hereafter by $\mathcal{RD}_L$. The coding scheme involves generating L independent codebooks, one each for $\hat{X}_i$ $i \in \mathcal{L}$, at respective rates of 2^{nR_i}. Upon observing a typical sequence X^n, the encoder looks for one codeword from each codebook such that they are all jointly typical with the observed sequence X^n. If the encoder fails to find one such codeword tuple, it randomly picks a tuple. The index of the codeword from codebook i is sent in description i $\forall i \in \mathcal{L}$. The decoders estimate the source using MMSE estimators based on the subset of the codewords they receive. If the encoding is error-free, all the prescribed distortion constraints are met as the covariance matrix $K_{\mathbf{W}}$ satisfies (4). It can be shown using standard typicality arguments (see for example, the proof of Theorem 1 in [5]) that the probability of error at the encoder asymptotically approaches zero if the rates satisfy (5).

It was shown recently that $\mathcal{RD}_L$ is complete when one imposes only individual and central distortion constraints [10], i.e., when only $D_1, D_2, \ldots, D_L$ and $D_{\{12\ldots L\}}$ are imposed. It was further shown in [12] that the above strategy, in conjunction with a random binning mechanism used to further exploit the correlations among the descriptions, is sum rate optimal under symmetric distortion constraints when distortions at only two levels of receivers are imposed, i.e, when $D_{\mathcal{K}:|\mathcal{K}|=k}$ and $D_{\{12\ldots L\}}$ are imposed, where $|\mathcal{K}|$ denotes the cardinality of set $\mathcal{K}$.

Our objective in this paper is to derive a new coding scheme and an associated achievable region for the $L-$channel Gaussian MD problem which *strictly* improves upon $\mathcal{RD}_L$ for particular cross-sections of the rate-distortion region where we impose only a particular subset of the $2^L - 1$ distortion constraints. Specifically, let $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_M$ be M proper (strict) subsets of $\mathcal{L}$. We focus on cross-sections wherein only distortions $\{D\}_{2^{\mathcal{K}_1}}, \{D\}_{2^{\mathcal{K}_1}}, \ldots, \{D\}_{2^{\mathcal{K}_1}}$ are imposed, where $2^{\mathcal{K}}$ denotes the set of all subsets (power set) of $\mathcal{K}$ and $\{D\}_{2^{\mathcal{K}_i}} = \{D_{\mathcal{S}} : \mathcal{S} \in 2^{\mathcal{K}_i}\}$. It is critical to observe that, under these distortion constraints, joint typicality of codewords within subsets $\{\hat{X}\}_{\mathcal{K}_1}, \{\hat{X}\}_{\mathcal{K}_2}, \ldots, \{\hat{X}\}_{\mathcal{K}_M}$ is infact sufficient to achieve the prescribed distortions and imposing joint typicality of all the transmitted codewords is an unnecessary restriction. In this paper, we derive a new coding scheme which maintains joint typicality only within the required subsets and show that enforcing joint typicality of all the codewords entails strict suboptimality in the achievable rate-distortion region. In preparation, we derive a critical lemma called the subset typicality lemma, which plays a pivotal role in achieving the new region. We note that these results have potential broader implications in related problems in multi-terminal information theory which are beyond the scope of this paper. We also note that the principles underlying the proposed scheme can be easily extended to incorporate refinement layer random variables similar to [5] in conjunction with binning modules as indicated in [6], [12] and to include common layer random variables as shown in [2], [8], [9]. However, in general, the optimal joint distributions of the auxiliary random variables are hard to establish for these schemes and explicit characterizations of the achievable regions are not easily tractable.

The rest of the paper is organized as follows. In section II, we state the subset typicality lemma which plays an important role in the proofs that follow. In section III, we formally state the problem and derive a new achievable rate-distortion region. In section IV, we consider a particular cross-section of the 3-descriptions MD setup and show that the new rate-distortion region strictly improves upon $\mathcal{RD}_3$.

II. Subset Typicality Lemma

We introduce the notation before establishing the lemma. n independent and identically distributed (iid) random variables and their realizations are denoted by X_0^n and x_0^n respectively. We denote by $\mathcal{T}_\epsilon^n(P)$, the length n, ϵ-typical set corresponding to a random variable distributed according to P. Through out the paper, for any set $\mathcal{S}$, we use the shorthand $\{U\}_{\mathcal{S}}$ to denote the set $\{U_i : i \in \mathcal{S}\}$. Also, in the following Lemma, we use the notation $P(A) \doteq 2^{-nR}$ to denote $2^{-n(R+\delta(\epsilon))} \leq P(A) \leq 2^{-n(R-\delta(\epsilon))}$ for some $\delta(\epsilon) \to 0$ as $\epsilon \to 0$.

Lemma 1. ***Subset Typicality Lemma*** *:Let $(X_1, X_2, \ldots X_N)$ be N random variables taking values on arbitrary finite alphabets $(\mathcal{X}_1, \mathcal{X}_2, \ldots \mathcal{X}_N)$ with marginal distributions $P_1(X_1), P_2(X_2) \ldots, P_N(X_N)$, respectively. Let $\mathcal{S}_1, \mathcal{S}_2 \ldots \mathcal{S}_M$ be M subsets of $\{1, 2, \ldots, N\}$ and for all $j \in \{1, 2, \ldots, M\}$, let $P_{\mathcal{S}_j}(\{X\}_{\mathcal{S}_j})$ be any given joint distribution for $\{X\}_{\mathcal{S}_i}$ consistent with each other and with the given marginal distributions[2]. Generate sequences $x_1^n, x_2^n \ldots x_N^n$, each independent of the other, where x_i^n is drawn iid according to the marginal distribution $P_i(X_i)$, i.e., $x_i^n \sim \prod_{l=1}^n P_i(x_{il})$. Then,*

$$P\left(\{x\}_{\mathcal{S}_j}^n \in \mathcal{T}_\epsilon^n\left(P_{\mathcal{S}_j}\right), \ \forall j \in \{1 \ldots M\}\right) \doteq 2^{-n(\sum_{i=1}^N H(X_i) - H(P^*))} \quad (6)$$

where P^* is a distribution over $(\mathcal{X}_1, \mathcal{X}_2 \ldots, \mathcal{X}_N)$ which satisfies:

$$P^* = \arg\max_{\tilde{P}} H\left(\tilde{P}\right) \quad (7)$$

subject to $\tilde{P}(\{X\}_{\mathcal{S}_j}) = P_{\mathcal{S}_j}(\{X\}_{\mathcal{S}_j})$ $\forall j \in \{1 \ldots M\}$.

[2]To avoid resolvable but unnecessary complications, we assume that there exists at least one joint distribution consistent with the prescribed per subset distributions for subsets $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_M$.

Proof: The proof follows directly from Sanov's Theorm. We omit the details here and refer to [13] for a detailed analysis. ■

In [13] we also establish a conditional version of Lemma 1. This conditional version plays an important role in several related network information theoretic setups. However in this paper, the above lemma is sufficient to establish the new region. We also note that a particular instance of Lemma 1 was derived in [14]. However, as it turns out, for the setup they consider, this Lemma does not help in deriving an improved achievable region.

III. Problem Setup and A New Achievable Region

In this section, we first formally state the problem being addressed in this paper and then derive a new achievable rate-distortion region for the $L-$channel MD setup. We assume the source to be a zero mean, unit variance Gaussian random variable, i.e., $X \sim \mathcal{N}(0,1)$. In the $L-$channel MD setup, there are L encoding functions, $f_l(\cdot)$ $l \in \mathcal{L}$, which map X^n to the descriptions $J_l = f_l(X^n)$, where J_l takes on values in the set $\{1, \ldots B_l\}$. The rate of description l is defined as $R_l = \frac{1}{n}\log_2(B_l)$. Description l is sent over channel l and is either received at the decoder error free or is completely lost. Let $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_M$ be M proper (strict) subsets of $\mathcal{L}$ and let $\mathcal{S}$ denote the union over all $i \in \{1, \ldots, M\}$ of the set of all subsets of $\mathcal{K}_i$, i.e., $\mathcal{S} = \{\mathcal{K} : \mathcal{K} \subseteq \mathcal{K}_i \text{ for some } i \in \{1, 2, \ldots, M\},\ \mathcal{K} \neq \phi\}$, where ϕ denotes the null set. In this paper, we assume that the decoder always receives a subset $\mathcal{K}$ of the transmitted descriptions such that $\mathcal{K} \in \mathcal{S}$.

There is a decoding function for each possible received combination of the descriptions $\hat{X}^n_{\mathcal{K}} = \left(\hat{X}^{(1)}_{\mathcal{K}}, \hat{X}^{(2)}_{\mathcal{K}} \ldots, \hat{X}^{(n)}_{\mathcal{K}}\right) = g_{\mathcal{K}}(J_l : l \in \mathcal{K})$, $\forall \mathcal{K} \in \mathcal{S}$, where $\hat{X}_{\mathcal{K}}$ takes on real values. The distortion (MSE) at the decoder when a subset $\mathcal{K}$ of the descriptions is received is measured as:

$$D_{\mathcal{K}} = E\left[\frac{1}{n}\sum_{t=1}^{n}(X^{(t)} - \hat{X}^{(t)}_{\mathcal{K}})^2\right] \tag{8}$$

We say that a rate-distortion tuple $(R_i, D_{\mathcal{K}} : i \in \mathcal{L}, \mathcal{K} \in \mathcal{S})$ is achievable if there exit L encoding functions with rates $(R_1 \ldots, R_L)$ and $|\mathcal{S}|$ decoding functions yielding respective distortions $D_{\mathcal{K}}$. The closure of the set of all achievable rate-distortion tuples is defined as the rate-distortion region for the set $\mathcal{S}$. Note that, this region has $L + |\mathcal{S}|$ dimensions and is a cross-section of the general rate-distortion region for the $L-$channel Gaussian MD problem. In the following theorem, we give an achievable region for any given subsets $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_M$.

Theorem 1. *Let $W_1, W_2, \ldots, W_L$ be zero mean Gaussian random variables independent of X with covariance matrix $K_{\mathbf{W}}$ and let $\hat{X}_i$ be defined as:*

$$\hat{X}_i = X + W_i \;\; i \in \{1, 2, \ldots, L\} \tag{9}$$

Let $\psi_{\mathcal{K}}(\{\hat{X}\}_{\mathcal{K}}) = E\left[X|\{\hat{X}\}_K\right]$ and the covariance matrix $K_{\mathbf{W}}$ be such that:

$$Var\left(X|\{\hat{X}\}_{\mathcal{K}}\right) \leq D_{\mathcal{K}} \;\; \forall \mathcal{K} \in \mathcal{S} \tag{10}$$

Then any rate tuple $(R_1, R_2, \ldots, R_L)$ satisfying the following conditions is achievable:

$$\sum_{i\in\mathcal{K}} R_i \geq \sum_{i\in\mathcal{K}} h(\hat{X}_i) - h^*(\{\hat{X}\}_{\mathcal{K}}|X) \;\; \forall \mathcal{K} \subseteq \mathcal{L} \tag{11}$$

where,

$$h^*\left(\{\hat{X}\}_{\mathcal{K}}\middle|X\right) = \max_{P(\tilde{X}_{\mathcal{K}}|X)} h\left(\{\tilde{X}\}_{\mathcal{K}}|X\right) \tag{12}$$

where $P(\tilde{X}_{\mathcal{K}}|X)$ is subject to:

$$P\left(\{\tilde{X}\}_{\mathcal{K}\cap\mathcal{K}_j}\middle|X\right) = P\left(\{\hat{X}\}_{\mathcal{K}\cap\mathcal{K}_j}\middle|X\right) \;\; \forall j \in \{1 \ldots M\} \tag{13}$$

The convex closure over all such $K_{\mathbf{W}}$ satisfying (10) is an achievable rate-distortion region and is denoted by $\mathcal{RD}^*_L$.

Proof: We again omit the details of the proof here as it follows almost directly from Theorem 1 in [13]. ■

Remark 1. Observe that for any $\mathcal{K} \in \mathcal{S}$, the conditions in (11) are the same as the conditions in (5). However, when $\mathcal{K} \notin \mathcal{S}$, the above theorem asserts that all the constraints in (5) can be set to their respective optima simultaneously. This is precisely what provides the strict improvement over $\mathcal{RD}_L$ as each of the constraints are typically optimized by a different $K_{\mathbf{W}}$ and hence it is impossible to achieve simultaneous optimality of all the constraints in $\mathcal{RD}_L$.

Remark 2. We note that the rate-distortion region can be enlarged, in general, by defining a convex closure of the achievable rate-distortion tuples over all joint densities $P\left(\{\hat{X}\}_{\mathcal{L}}\middle|X\right)$ and functions $\psi_{\mathcal{K}}(\{\hat{X}\}_{\mathcal{K}})$ that satisfy the distortion constraints. However, it can be shown using the contra-polymatroidal nature of the achievable region that for the Gaussian source under MSE distortion metric, the form considered in the theorem for $P\left(\{\hat{X}\}_{\mathcal{L}}\middle|X\right)$ and $\psi_{\mathcal{K}}(\{\hat{X}\}_{\mathcal{K}})$ is sufficient to achieve all the corner points.

IV. Strict Improvement

In the following theorem, we show that $\mathcal{RD}^*_L$ is strictly larger than $\mathcal{RD}_L$ for any $L > 2$.

Theorem 2. *(i) $\mathcal{RD}^*_L \supseteq \mathcal{RD}_L$, i.e., in general, $\mathcal{RD}^*_L$ subsumes $\mathcal{RD}_L$.*

(ii) $\forall L > 2$, There exists scenarios for which,

$$\mathcal{RD}^*_L \supset \mathcal{RD}_L \tag{14}$$

i.e $\mathcal{RD}^*_L$ contains points that are strictly outside $\mathcal{RD}_L$ $\forall L >$ 2.

Proof: We note that part (i) of the theorem follows directly as the RHS in (11) is smaller than or equal to the RHS in (5) for each hyperplane. The challenging task to the prove part (ii) of the theorem. Also note that to prove (ii) it is sufficient for us to show the result for $L = 3$. We consider the following cross-section of the 3-descriptions rate-distortion region wherein we impose distortions D_1, D_2, D_3, D_{12} and D_{23}. Observe that for this cross-section, we have $\mathcal{K}_1 = \{1,2\}$ and $\mathcal{K}_2 = \{2,3\}$ and hence it is sufficient to maintain joint typicality of sequences $(\hat{x}_1^n, \hat{x}_2^n, x^n)$ and $(\hat{x}_2^n, \hat{x}_3^n, x^n)$ respectively. We will show that imposing joint typicality of all the sequences leads to a strict sub-optimality.

We first specialize and restate the regions $\mathcal{RD}_3$ and $\mathcal{RD}_3^*$ for $\mathcal{K}_1 = \{1,2\}$, $\mathcal{K}_2 = \{2,3\}$. Let W_1, W_2, W_3 be zero mean Gaussian random variables independent of X with covariance matrix $K_{\mathbf{W}}$ and let $\hat{X}_i = X + W_i$. Let $K_{\mathbf{W}}$ be such that all the prescribed distortion constraints are satisfied. Then all rate tuples satisfying the following conditions are part of $\mathcal{RD}_3$:

$$\begin{aligned} R_i &\geq I(X;\hat{X}_i) \; i \in \{1,2,3\} \\ R_1 + R_2 &\geq I(X;\hat{X}_1,\hat{X}_2) + I(\hat{X}_1;\hat{X}_2) \\ R_2 + R_3 &\geq I(X;\hat{X}_2,\hat{X}_3) + I(\hat{X}_2;\hat{X}_3) \\ R_1 + R_3 &\geq h(\hat{X}_1) + h(\hat{X}_3) - h(\hat{X}_1,\hat{X}_3|X) \\ R_1 + R_2 + R_3 &\geq h(\hat{X}_1) + h(\hat{X}_2) + h(\hat{X}_3) \\ &\quad - h(\hat{X}_1,\hat{X}_2,\hat{X}_3|X) \end{aligned} \tag{15}$$

The corresponding constraints for $\mathcal{RD}_3^*$ are given by:

$$\begin{aligned} R_i &\geq I(X;\hat{X}_i) \; i \in \{1,2,3\} \\ R_1 + R_2 &\geq I(X;\hat{X}_1,\hat{X}_2) + I(\hat{X}_1;\hat{X}_2) \\ R_2 + R_3 &\geq I(X;\hat{X}_2,\hat{X}_3) + I(\hat{X}_2;\hat{X}_3) \\ R_1 + R_2 + R_3 &\geq h(\hat{X}_1) + h(\hat{X}_2) + h(\hat{X}_3) \\ &\quad - h^*(\hat{X}_1,\hat{X}_2,\hat{X}_3|X) \\ &=^{(a)} I(\hat{X}_2;X,\hat{X}_1) + I(\hat{X}_2;X,\hat{X}_3) \\ &\quad + I(X;X_2) \end{aligned} \tag{16}$$

The convex closure of such rates over all $K_{\mathbf{W}}$ satisfying the distortion constraints defines the respective regions $\mathcal{RD}_3$ and $\mathcal{RD}_3^*$. Observe that the condition on $R_1 + R_3$ is not imposed in (16). This is because the term $h(\hat{X}_1, \hat{X}_3|X)$ in (15) is set to its maximum entropy subject to the pairwise distributions of $(X, \hat{X}_1)$ and $(X, \hat{X}_3)$ (refer to (13) in Theorem 1), i.e., $h(\hat{X}_1, \hat{X}_3|X) = h(\hat{X}_1|X) + h(\hat{X}_3|X)$ and hence the constraint on $R_1 + R_3$ is redundant. Equality (a) follows from the fact that the maximum entropy term, $h^*(\hat{X}_1, \hat{X}_2, \hat{X}_3|X)$, subject to the given distributions within subsets $(X, \hat{X}_1, \hat{X}_2)$ and $(X, \hat{X}_3, \hat{X}_2)$ is achieved by a joint density of the form $\hat{X}_1 \leftrightarrow (X, \hat{X}_2) \leftrightarrow \hat{X}_3$ and hence $h^*(\hat{X}_1, \hat{X}_2, \hat{X}_3|X)$ can be written as $h(\hat{X}_2|X) + h(\hat{X}_1|X, \hat{X}_2) + h(\hat{X}_3|X, \hat{X}_2)$. We will show that in (16), the $K_{\mathbf{W}}$ which achieves the optimum sum rate (which of course satisfies $\hat{X}_1 \leftrightarrow (X, \hat{X}_2) \leftrightarrow \hat{X}_3$) does not satisfy the condition $\hat{X}_1 \leftrightarrow X \leftrightarrow \hat{X}_3$ and hence leads to a suboptimal bound on $R_1 + R_3$. This argument is in fact sufficient to show that $\mathcal{RD}_3^*$ strictly subsumes $\mathcal{RD}_3$.

More precisely, we consider one point on the boundary of $\mathcal{RD}_3^*$ and prove that this point does not lie in $\mathcal{RD}_3$. Define $P_{min} = \{\frac{1}{2}\log(\frac{1}{D_1}), R_{min}, \frac{1}{2}\log(\frac{1}{D_3})\}$ where R_{min} is given by:

$$R_{min} = \inf R_2 : \Big\{ R_1 = \frac{1}{2}\log(\frac{1}{D_1}), \;\; R_3 = \frac{1}{2}\log(\frac{1}{D_3}) \\ (R_1, R_2, R_3) \in \mathcal{RD}_3(D_1, D_2, D_3, D_{12}, D_{23}) \Big\} \tag{17}$$

Similarly we define $P_{min}^* = \{\frac{1}{2}\log(\frac{1}{D_1}), R_{min}^*, \frac{1}{2}\log(\frac{1}{D_3})\}$ where R_{min}^* is obtained by replacing $\mathcal{RD}_3$ by $\mathcal{RD}_3^*$ in (17). Observe that strict improvement is proved if we show that $R_{min}^* < R_{min}$.

Recall that $\frac{1}{2}\log(\frac{1}{D})$ is the rate distortion function at distortion D for a Gaussian source under MSE distortion measure and the unique rate-distortion optimal reconstruction random variable is jointly Gaussian with the source. Hence imposing $R_1 = \frac{1}{2}\log(\frac{1}{D_1})$ and $R_3 = \frac{1}{2}\log(\frac{1}{D_3})$ in (15), enforces the random variables $\hat{X}_1$ and $\hat{X}_3$ to be independent given X such that $P(\hat{X}_1|X)$ and $P(\hat{X}_2|X)$ are the respective RD-optimal channels. This implies that W_1 and W_3 are independent zero mean Gaussian random variables with respective variances $D_1/(1-D_1)$ and $D_3/(1-D_3)$. Hence the minimization for R_{min} in (17) is over all positive definite $K_{\mathbf{W}}$ of the form:

$$K_{\mathbf{W}} = \begin{bmatrix} \sigma_{W_1}^2 & \rho_{12}\sigma_{W_1}\sigma_{W_2} & 0 \\ \rho_{12}\sigma_{W_1}\sigma_{W_2} & \sigma_{W_2}^2 & \rho_{23}\sigma_{W_2}\sigma_{W_3} \\ 0 & \rho_{23}\sigma_{W_2}\sigma_{W_3} & \sigma_{W_3}^2 \end{bmatrix} \tag{18}$$

where $\sigma_{W_1}^2 = \frac{D_1}{1-D_1}$ and $\sigma_{W_3}^2 = \frac{D_3}{1-D_3}$ and where $\sigma_{W_2}^2, \rho_{12}, \rho_{23}$ achieve the other three distortion constraints. However, observe that R_{min}^* is obtained by optimizing over all positive definite $K_{\mathbf{W}}$ of the form:

$$K_{\mathbf{W}} = \begin{bmatrix} \sigma_{W_1}^2 & \rho_{12}\sigma_{W_1}\sigma_{W_2} & \rho_{13}\sigma_{W_1}\sigma_{W_3} \\ \rho_{12}\sigma_{W_1}\sigma_{W_2} & \sigma_{W_2}^2 & \rho_{23}\sigma_{W_2}\sigma_{W_3} \\ \rho_{13}\sigma_{W_1}\sigma_{W_3} & \rho_{23}\sigma_{W_2}\sigma_{W_3} & \sigma_{W_3}^2 \end{bmatrix} \tag{19}$$

where $\sigma_{W_1}^2 = \frac{D_1}{1-D_1}$ and $\sigma_{W_3}^2 = \frac{D_3}{1-D_3}$ and where $\sigma_{W_2}^2, \rho_{12}, \rho_{23}, \rho_{13}$ achieve the remaining distortion constraints. It is this extra degree of freedom that provides leeway to achieve a strictly lower R_{min}^*. Hence, all that remains for us to show is that R_{min}^* is not achieved by a $K_{\mathbf{W}}$ which has $\rho_{13} = 0$.

Towards this end, we first find R_2 from (16) for a fixed $K_{\mathbf{W}}$ which has the form (19):

$$R_2 = \frac{1}{2}\log_2\left(\frac{(1+\sigma_{W_2}^2)}{\sigma_{W_2}^2(1-\rho_{12}^2-\rho_{23}^2-\rho_{13}^2+2\rho_{12}\rho_{23}\rho_{13})}\right) \tag{20}$$

To find R_{min}^*, we need to minimize the RHS of (20) over all $\sigma_{W_2}^2, \rho_{12}, \rho_{23}, \rho_{13}$ subject to the distortion constraints, D_2, D_{12} and D_{23}. Observe that ρ_{13} does not contribute to any of the distortion constraints and hence can be optimized by setting the differential of the RHS in (20) to 0 for fixed $\sigma_{W_2}^2, \rho_{12}$ and ρ_{23}. We obtain the optimum ρ_{13}^* as:

$$\rho_{13}^* = \rho_{12}\rho_{23} \tag{21}$$

Note that if $\rho_{13}^* = \rho_{12}\rho_{23}$, then $K_{\mathbf{W}}$ is always positive definite for any $\rho_{12}, \rho_{23} : |\rho_{12}| < 1, |\rho_{23}| < 1$. Also observe that $\rho_{13}^* = \rho_{12}\rho_{23}$ induces the condition $\hat{X}_1 \leftrightarrow (X, \hat{X}_2) \leftrightarrow \hat{X}_3$

which is necessary for sum rate optimality in (16). Now, the RHS in (20) can be rewritten as:

$$R_2 = \frac{1}{2}\log_2\left(\frac{(1+\sigma_{W_2}^2)}{\sigma_{W_2}^2(1-\rho_{12}^2)(1-\rho_{23}^2)}\right) \qquad (22)$$

and R^*_{min} is obtained by minimizing (22) over all $\sigma^2_{W_2}, \rho_{12}, \rho_{23}$ subject to the distortion constraints. We first denote $\frac{\sigma^2_{W_2}}{(1+\sigma^2_{W_2})}$ by $\tilde{D}_2$. Clearly, we require $\tilde{D}_2 \leq D_2$ to satisfy the distortion constraint for $\hat{X}_2$. Next, observe that for a fixed $\tilde{D}_2$, R_2 decreases as ρ^2_{12} ,ρ^2_{23} decrease. Hence, using arguments similar to [4], [14], it follows that ρ_{12} and ρ_{23} should be set equal to the respective largest values in the range $(-1, 0)$, such that distortions D_{12} and D_{23} are satisfied with equality. It can be easily shown that (see [15]):

$$\rho^*_{ij} = \begin{cases} -\frac{\sqrt{\pi_{ij}D^2_{ij}+\gamma_{ij}}-\sqrt{\pi_{ij}D^2_{ij}}}{(1-D_{ij})\sqrt{D_i\tilde{D}_j}} & \bar{D}_{ij} \leq \tilde{D}_j \leq 1+D_{ij}-D_i \\ -\sqrt{\frac{\pi_{ij}}{D_i\tilde{D}_j}} & \tilde{D}_j \geq 1+D_{ij}-D_i \\ 0 & \tilde{D}_{ij} \leq \bar{D}_j \end{cases} \qquad (23)$$

for $i \in \{1,3\}$ and $j = 2$, where $\pi_{ij} = (1-D_i)(1-\tilde{D}_j)$ and $\gamma_{ij} = (1-D_{ij})\left[(D_i - D_{ij})(\tilde{D}_j - D_{ij}) + D_{ij}D_i\tilde{D}_j - D^2_{ij}\right]$ and where $\bar{D}_{ij} = \frac{D_{ij}D_i}{D_i - D_{ij} + D_i D_{ij}}$. We choose D_2 to be such that $D_1 + D_2 < 1 + D_{12}$ and $D_3 + D_2 < 1 + D_{23}$ hold. Then it is sufficient for us to show that the optimum $\tilde{D}_2$ is strictly greater than $\bar{D}_{12}$ and $\bar{D}_{32}$, i.e., $\tilde{D}_j > \bar{D}_{ij}$ for both $i = 1, 3$. Then we will have ρ^*_{12} and ρ^*_{23} strictly less than 0, implying $\rho^*_{13} \neq 0$. On substituting the optimum values for ρ^*_{ij} (assuming $\bar{D}_{ij} < \tilde{D}_j \leq 1 + D_{ij} - D_i$ holds for both $i = 1, 3$), we have the following expression for R_2:

$$R_2 = \frac{1}{2}\log_2\left(\tilde{D}_2\Delta_{12}\Delta_{32}\right) \qquad (24)$$

$$\Delta_{ij} = \frac{D_i(1-D_{ij})^2}{D_{ij}\left[(1-D_{ij})^2 - \left(\sqrt{(1-D_i)(1-\tilde{D}_j)} - \sqrt{(D_i - D_{ij})(\tilde{D}_j - D_{ij})}\right)^2\right]}$$

$i \in \{1, 3\}$ and $j = 2$. R^*_{min} is then obtained by minimizing the RHS of (24) with respect to $\tilde{D}_2$ subject to the constraint that $\tilde{D}_2 \leq D_2$. In general, it is hard to derive a closed form solution for the minimizing $\tilde{D}_2$. Instead, we plot the variation of $\frac{1}{2}\log_2\left(\tilde{D}_2\Delta_{12}\Delta_{32}\right)$ as a function of $\tilde{D}_2$ for a particular D_1, D_3, D_{12} and D_{23} and show that $\exists \tilde{D}^*_2 > \bar{D}_{12} : \{R_2|_{\tilde{D}_2=\bar{D}_{12}} > R_2|_{\tilde{D}_2=\tilde{D}^*_2}\}$. Let us set $D_1 = D_3 = 0.2$ and $D_{12} = D_{23} = 0.05$ and $D_2 \leq 0.85$. Fig. 1 shows the variation of $\frac{1}{2}\log_2\left(\tilde{D}_2\Delta_{12}\Delta_{32}\right)$ in the range $\bar{D}_{12} \leq \tilde{D}_2 \leq 1 + D_{12} - D_1$. It can be verified that $R_2|_{\tilde{D}_2=\bar{D}_{12}=0.063} \approx 1.99$ and $R_2|_{\tilde{D}_2=0.36} \approx 1.73$. Hence the optimum $\tilde{D}_2$ is strictly greater than $\bar{D}_{12}$ and $\bar{D}_{23}$. Therefore, it follows that ρ^*_{12} and ρ^*_{23} are strictly less than 0 and consequently $\rho^*_{13} \neq 0$ proving the theorem. ∎

V. Conclusion

In this paper, we focused on the $L-$channel Gaussian MD problem where it is known apriori that certain descriptions are not received simultaneously at the decoder. This assumption allows us to design encoding schemes wherein joint typicality of codewords is required only within prescribed subsets. We derived a new achievable rate-distortion region for this setup by establishing a lemma called the subset typicality lemma. We showed using a 3 descriptions setting that a strictly larger rate-distortion region is obtained by maintaining joint typicality only within the required subsets as opposed a scheme which maintains joint typicality of all the transmitted codewords.

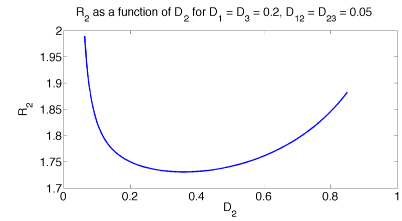

Figure 1. Variation of R_2 as a function of $\tilde{D}_2$

References

[1] A. El Gamal and T. M. Cover, “Achievable rates for multiple descriptions,” IEEE Trans. Inf. Theory, vol. IT-28, pp. 851–857, Nov. 1982.

[2] Z. Zhang and T. Berger, “New results in binary multiple descriptions,” IEEE Trans. Inf. Theory, vol. IT-33, pp. 502–521, July 1987.

[3] R. Ahlswede, “The rate-distortion region for multiple descriptions without excess rate," IEEE Trans. Inf. Theory, vol. 31, pp. 721-726, Nov. 1985.

[4] L. Ozarow, “On a source-coding problem with two channels and three receivers,” Bell Syst. Tech. J., vol. 59, no. 10, pp. 1909-1921, Dec. 1980.

[5] R. Venkataramani, G. Kramer, V.K. Goyal, “Multiple description coding with many channels”, IEEE Trans. on Information Theory, vol.49, no.9, pp. 2106- 2114, Sept 2003.

[6] R. Puri, S. S. Pradhan, and K. Ramchandran, “n-channel symmetric multiple descriptions-part II: an achievable rate-distortion region”, IEEE Trans. Information Theory, vol. 51, pp. 1377-1392, Apr. 2005.

[7] J. Wang, J. Chen, L. Zhao, P. Cuff and H. Permuter, “On the role of the refinement layer in multiple description coding and scalable coding,” IEEE Transactions on Information Theory, vol.57, no.3, pp.1443-1456, Mar. 2011.

[8] K. Viswanatha, E. Akyol and K. Rose, “A strictly larger achievable region for multiple descriptions using combinatorial message sharing”, in Proc. IEEE Information Theory Workshop (ITW), Oct. 2011.

[9] E. Akyol, K. Viswanatha and K. Rose, “On random binning versus conditional codebook methods in multiple descriptions coding”, in Proc. IEEE Information Theory Workshop (ITW), Sep. 2012.

[10] J. Chen, “Rate region of Gaussian multiple description coding with individual and central distortion constraints," IEEE Transactions on Information Theory, vol.55, no.9, pp.3991-4005, Sep. 2009.

[11] S. Mohajer, C. Tian and S. Diggavi, “Asymmetric multilevel diversity coding and asymmetric Gaussian multiple descriptions,” IEEE Transactions on Information Theory, vol.56, no.9, pp.4367-4387, Sep. 2010.

[12] H. Wang and P. Viswanath, , “Vector Gaussian multiple description with two levels of receivers,” IEEE Transactions on Information Theory, vol.55, no.1, pp.401-410, Jan. 2009.

[13] K. Viswanatha, E. Akyol and K. Rose, “Subset typicality lemmas and improved achievable regions in multiterminal source coding”, In Technical report, Available for download at: http://www.scl.ece.ucsb.edu/Subset_Typicality.pdf

[14] E. Perron, S. Diggavi, E. Telatar, “On the role of encoder side-information in source coding for multiple decoders," In Proc. IEEE International Symposium on Information Theory (ISIT), vol., no., pp.331-335, 9-14 Jul 2006.

[15] R. Zamir, “Gaussian codes and Shannon bounds for multiple descriptions,” IEEE Transactions on Information Theory, vol.45, no.7, pp.2629-2636, Nov. 1999.