On the Role of Common Codewords in Quadratic Gaussian Multiple Descriptions Coding

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

***Abstract*—This paper focuses on the problem of $L-$channel quadratic Gaussian multiple description (MD) coding. We recently introduced a new encoding scheme in [1] for general $L-$channel MD problem, based on a technique called 'Combinatorial Message Sharing' (CMS), where every subset of the descriptions shares a distinct common message. The new achievable region subsumes the most well known region for the general problem, due to Venkataramani, Kramer and Goyal (VKG) [2]. Moreover, we showed in [3] that the new scheme provides a strict improvement of the achievable region for any source and distortion measures for which some 2-description subset is such that the Zhang and Berger (ZB) scheme achieves points outside the El-Gamal and Cover (EC) region. In this paper, we show a more surprising result: CMS outperforms VKG for a general class of sources and distortion measures, which includes scenarios where for all 2-description subsets, the ZB and EC regions coincide. In particular, we show that CMS strictly extends VKG region, for the L-channel quadratic Gaussian MD problem for all $L \geq 3$, despite the fact that the EC region is complete for the corresponding 2-descriptions problem. Using the encoding principles derived, we show that the CMS scheme achieves the complete rate-distortion region for several asymmetric cross-sections of the $L-$channel quadratic Gaussian MD problem, which have not been considered earlier.**

***Index Terms*—Multiple description coding, Combinatorial message sharing, Quadratic Gaussian multiple descriptions**

I. Introduction

The multiple descriptions (MD) problem has been studied extensively, yielding a series of advances , ranging from achievability [4], [5], [2], [6], [1], [3], [7] to converse results [8], [9], [10]. In the general MD setup, the encoder generates L-descriptions of the source for transmission over L available channels and it is assumed that the decoder receives a subset of the descriptions perfectly and the remaining are lost. The objective is to quantify the set of all achievable rate-distortion (RD) tuples for the $L-$rates $(R_1, \ldots, R_L)$ and distortion levels corresponding to the $2^L - 1$ possible description loss patterns $(D_{\mathcal{K}}, \mathcal{K} \subseteq \{1, \ldots, L\})$. One of the first achievable regions for the 2-channel MD problem was derived by El-Gamal and Cover (EC) in 1982 [4]. It was shown by Ozarow in [8] that the EC region is complete when the source is Gaussian and the distortion measure is mean squared error (MSE). Zhang and Berger (ZB), however, later showed in [5] that the EC coding scheme is strictly sub-optimal in general. In particular, for a binary source under Hamming distortion, sending a common codeword within the two descriptions can achieve points that are strictly outside the the EC region. The converse to the ZB scheme in still not known for general sources and distortion measures.

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

Since then several researchers have worked on extending the EC and ZB approaches to the $L-$channel MD problem [2], [6], [9], [10]. An achievable scheme, due to Venkataramani, Kramer and Goyal (VKG) [2], directly builds on EC and ZB, and introduces a combinatorial number of refinement codebooks, one for each subset of the descriptions. Motivated by ZB, a *single* common codeword is also shared by all the descriptions, which assists in better coordination of the messages, improving the RD trade-off. We recently introduced a new coding scheme called 'Combinatorial Message Sharing' (CMS) in [1], wherein a distinct common codeword is shared by members of each subset of the transmitted descriptions. The new achievable RD region subsumes the VKG region for general sources and distortion measures. Moreover, we demonstrated in [3] that CMS achieves a strictly larger region than VKG for all $L > 2$, if there exists a 2-description subset for which ZB achieves points strictly outside the EC region. In particular, CMS achieves strict improvement for a binary source under Hamming distortion.

Ozarow's converse result [8] motivated researchers to seek extended results for the $L-$channel quadratic Gaussian MD problem [9], [10]. It was shown in [9] that a special case of the VKG coding scheme, called the 'correlated quantization' scheme (a generalization of Ozarow's encoding mechanism to $L-$channels), where *no common codewords are sent,* achieves the complete rate region, when only the individual and the central distortion constraints are imposed. A different and important line of attack focused on a practically interesting cross-section of the general MD problem, called the 'symmetric MD problem' (see [6]), based on encoding principles derived from Slepian and Wolf's random binning techniques. In fact, CMS principles can be extended to incorporate such random binning techniques, to utilize the underlying symmetry in the problem setup as illustrated recently in [7]. However, in this paper, we restrict ourselves to the general asymmetric setup to demonstrate the potential gains of using the common codewords of CMS for the quadratic Gaussian MD problem.

Optimality of EC for the 2-descriptions setup has led to a natural conjecture that common codewords do not play a necessary role in quadratic Gaussian MD coding, and all the achievable regions characterized so far neglect the common layer codewords (see eg., [2], [9], [10]). In this paper, we show that, surprisingly CMS strictly outperforms VKG for a Gaussian source under MSE distortion. More generally, we show that strict improvement holds for a general class of sources and distortion measures, which includes several scenarios in which, for every 2-description subset, ZB and EC lead to the same achievable region. We also show that the common codewords of CMS play a critical role in achieving the complete RD region for several asymmetric cross-sections of the $L-$channel quadratic Gaussian MD problem.

We note that, due to severe space constraints, in this paper, we avoid restating all the prior results and refer the readers to [1] and [3] for a brief description of EC, ZB and VKG schemes. In the following section, we begin with a brief description of the CMS coding scheme.

II. Formal definition and CMS coding scheme

A source produces a sequence of n iid random variables, denoted by $X^n = \left(X^{(1)}, X^{(2)} \ldots, X^{(n)}\right)$. We denote $\mathcal{L} = \{1, \ldots, L\}$. 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 $J_l \in \{1, \ldots B_l\}$ for some $B_l > 0$. The rate of description l is defined as $R_l = \log_2(B_l)$. Each of the descriptions are sent over a separate channel and are either received at the decoder error free or are completely lost. There are $2^L - 1$ decoding functions for each possible received combination of the descriptions $\hat{X}^n_{\mathcal{K}} = \left(\hat{X}^{(1)}_{\mathcal{K}}, \ldots, \hat{X}^{(n)}_{\mathcal{K}}\right) = g_{\mathcal{K}}(J_l : l \in \mathcal{K})$, $\forall \mathcal{K} \subseteq \mathcal{L}, \mathcal{K} \neq \phi$, where $\hat{X}_{\mathcal{K}}$ takes on values on a finite set $\hat{\mathcal{X}}_{\mathcal{K}}$, and ϕ denotes the null set. When a subset $\mathcal{K}$ of the descriptions are received at the decoder, the distortion is measured as $D_{\mathcal{K}} = E\left[\frac{1}{N}\sum_{t=1}^{n} d_{\mathcal{K}}(X^{(t)}, \hat{X}^{(t)}_{\mathcal{K}})\right]$ for some bounded distortion measures $d_{\mathcal{K}}(\cdot)$ defined as $d_{\mathcal{K}} : \mathcal{X} \times \hat{\mathcal{X}}_{\mathcal{K}} \rightarrow \mathcal{R}$. A RD tuple $(R_i, D_{\mathcal{K}} : i \in \mathcal{L}, \mathcal{K} \subseteq \mathcal{L}, \mathcal{K} \neq \phi)$ is achievable if there exit L encoding functions with rates $(R_1 \ldots, R_L)$ and $2^L - 1$ decoding functions yielding distortions $D_{\mathcal{K}}$. The closure of the set of all achievable RD tuples is defined as the '*L-channel multiple descriptions RD region*'.

In what follows, $2^{\mathcal{S}}$ denotes the set of all subsets (power set) of any set $\mathcal{S}$ and $|\mathcal{S}|$ denotes the set cardinality. Note that $|2^{\mathcal{S}}| = 2^{|\mathcal{S}|}$. $\mathcal{S}^c$ denotes the set complement. For two sets $\mathcal{S}_1$ and $\mathcal{S}_2$, we denote the set difference by $\mathcal{S}_1 - \mathcal{S}_2 = \{\mathcal{K} : \mathcal{K} \in \mathcal{S}_1, \mathcal{K} \notin \mathcal{S}_2\}$. We use the shorthand $\{U\}_{\mathcal{S}}$ for $\{U_{\mathcal{K}} : \mathcal{K} \in \mathcal{S}\}$[1]. Before describing CMS, we define the following subsets of $2^{\mathcal{L}}$:

$$\begin{aligned} \mathcal{I}_W &= \{\mathcal{S} : \mathcal{S} \in 2^{\mathcal{L}},\ |\mathcal{S}| = W\} \\ \mathcal{I}_{W+} &= \{\mathcal{S} : \mathcal{S} \in 2^{\mathcal{L}},\ |\mathcal{S}| > W\} \end{aligned} \tag{1}$$

Let $\mathcal{B}$ be any non-empty subset of $\mathcal{L}$ with $|\mathcal{B}| \leq W$. We define the following subsets of $\mathcal{I}_W$ and $\mathcal{I}_{W+}$:

$$\begin{aligned} \mathcal{I}_W(\mathcal{B}) &= \{\mathcal{S} : \mathcal{S} \in \mathcal{I}_W,\ \mathcal{B} \subseteq \mathcal{S}\} \\ \mathcal{I}_{W+}(\mathcal{B}) &= \{\mathcal{S} : \mathcal{S} \in \mathcal{I}_{W+},\ \mathcal{B} \subseteq \mathcal{S}\} \end{aligned} \tag{2}$$

[1]Note the difference between $\{U\}_{\mathcal{S}}$ and $U_{\mathcal{S}}$. $\{U\}_{\mathcal{S}}$ is a set of variables, whereas $U_{\mathcal{S}}$ is a single variable.

We also define $\mathcal{J}(\mathcal{K}) = \bigcup_{l \in \mathcal{K}} \mathcal{I}_{1+}(l)$. Note that $\mathcal{J}(\mathcal{L}) = 2^{\mathcal{L}} - \phi$.

Next, we briefly describe the CMS encoding scheme in [1]. Recall that, unlike VKG, CMS allows for 'combinatorial message sharing', i.e a common codeword is sent in each (non-empty) subset of the descriptions. The shared random variables are denoted by 'V'. The base and the refinement layer random variables are denoted by 'U'. First, the codebook for $V_{\mathcal{L}}$ is generated. Then, the codebooks for $V_{\mathcal{S}}$, $|S| = W$ are generated in the order $W = L-1, L-2 \ldots 2$. $2^{nR''_{\mathcal{Q}}}$ codewords of $V_{\mathcal{Q}}$ are independently generated conditioned on each codeword tuple of $\{V\}_{\mathcal{I}_{W+}(\mathcal{Q})}$. This is followed by the generation of the base layer codebooks, i.e. U_l, $l \in \mathcal{L}$. Conditioned on each codeword tuple of $\{V\}_{I_{1+}(l)}$, $2^{nR'_l}$ codewords of U_l are generated independently. Then the codebooks for the refinement layers are formed by generating a single codeword for $U_{\mathcal{S}}$, $|\mathcal{S}| > 1$ conditioned on every codeword tuple of $(\{V\}_{\mathcal{J}(\mathcal{S})}, \{U\}_{2^{\mathcal{S}} - \mathcal{S}})$. Observe that the base and the refinement layers in the CMS scheme are similar to that in the VKG scheme, except that they are now generated conditioned on a subset of the shared codewords.

The encoder employs joint typicality encoding, i.e., on observing a typical sequence x^n, it tries to find a jointly typical codeword tuple, one from each codebook. As with VKG, the codeword index of U_l (at rate R'_l) is sent in description l. However, now the codeword index of $V_{\mathcal{S}}$ (at rate $R''_{\mathcal{S}}$) is sent in *all* the descriptions $l \in \mathcal{S}$. Therefore the rate of description l is:

$$R_l = R'_l + \sum_{\mathcal{K} \in \mathcal{J}(l)} R''_{\mathcal{K}} \tag{3}$$

We next formally state the achievable RD region. Let $\mathcal{Q}$ be any subset of $2^{\mathcal{L}}$. Then, we say that $\mathcal{Q} \in \mathcal{Q}^*$ if it satisfies the following property:

$$\mathcal{K} \in \mathcal{Q} \ \Rightarrow\ \mathcal{I}_{|\mathcal{K}|+}(\mathcal{K}) \subset \mathcal{Q} \tag{4}$$

$\forall \mathcal{K} \in \mathcal{Q}$. Further, we denote by $[\mathcal{Q}]_1$ the set of all elements of $\mathcal{Q}$ of cardinality 1, i.e.,:

$$[\mathcal{Q}]_1 = \{\mathcal{K} : \mathcal{K} \in Q,\ |\mathcal{K}| = 1\} \tag{5}$$

Let $(\{V\}_{\mathcal{J}(\mathcal{L})}, \{U\}_{2^{\mathcal{L}} - \phi})$ be any set of $2^{L+1} - L - 2$ random variables jointly distributed with X. For any set $\mathcal{Q} \in \mathcal{Q}^*$ we define:

$$\begin{aligned} \alpha(\mathcal{Q}) \;=\; & \sum_{\mathcal{K} \in \mathcal{Q} - [\mathcal{Q}]_1} H\left(V_{\mathcal{K}} | \{V\}_{\mathcal{I}_{|\mathcal{K}|+}(\mathcal{K})}\right) \\ & + \sum_{\mathcal{K} \in 2^{[\mathcal{Q}]_1} - \phi} H\left(U_{\mathcal{K}} | \{V\}_{\mathcal{I}_{1+}(\mathcal{K})}, \{U\}_{2^{\mathcal{K}} - \phi - \mathcal{K}}\right) \\ & - H\left(\{V\}_{\mathcal{Q} - [\mathcal{Q}]_1}, \{U\}_{2^{[\mathcal{Q}]_1} - \phi} | X\right) \end{aligned} \tag{6}$$

We follow the convention $\alpha(\phi) = 0$. Next we state the rate-distortion region achievable by the CMS scheme for the $L-$descriptions framework.

Theorem 1. *Let $(\{V\}_{\mathcal{J}(\mathcal{L})}, \{U\}_{2^{\mathcal{L}} - \phi})$ be any set of $2^{L+1} - L - 2$ random variables jointly distributed with X, where $U_{\mathcal{S}}$ and $V_{\mathcal{S}}$ take values in some finite alphabets $\mathcal{U}_{\mathcal{S}}$ and $\mathcal{V}_{\mathcal{S}}$,*

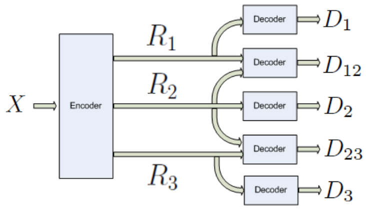

Figure 1. The cross-section that we consider in order to prove that CMS achieves points outside the VKG region for a general class of source and distortion measures. CMS achieves the the complete RD region for this setup for several distortion regimes for the quadratic Gaussian MD problem.

respectively $\forall \mathcal{S}$. Let $\mathcal{Q}^*$ be the set of all subsets of $2^{\mathcal{L}} - \phi$ satisfying (4) and let $R^{''}_{\mathcal{S}}$, $\mathcal{S} \in \mathcal{I}_{1+}$ and $R^{'}_l$, $l \in \mathcal{L}$ be $2^L - 1$ auxiliary rates satisfying:

$$\sum_{\mathcal{S}\in\mathcal{Q}-[\mathcal{Q}]_1} R^{''}_{\mathcal{S}} + \sum_{l\in[\mathcal{Q}]_1} R^{'}_l > \alpha(\mathcal{Q}) \quad \forall \mathcal{Q} \in \mathcal{Q}^* \tag{7}$$

Then, the RD region for the $L-$channel MD problem contains the rates and distortions for which there exist functions $\psi_{\mathcal{S}}(\cdot)$, such that,

$$R_l = R^{'}_l + \sum_{\mathcal{S}\in\mathcal{J}(l)} R^{''}_{\mathcal{S}} \tag{8}$$

$$D_{\mathcal{S}} \geq E\left[d_{\mathcal{S}}\left(X, \psi_{\mathcal{S}}\left(U_{\mathcal{S}}\right)\right)\right] \tag{9}$$

The closure of the achievable tuples over all such $2^{L+1} - L - 2$ random variables is denoted by $\mathcal{RD}_{CMS}$.

Remark 1. $\mathcal{RD}_{CMS}$ can be extended to continuous random variables and well-defined distortion measures using techniques similar to [11]. We omit the details here and assume that the above region continues to hold even for well behaved continuous random variables (for example, a Gaussian source under MSE).

Remark 2. $\mathcal{RD}_{CMS}$ is convex, as a time sharing random variable can be embedded in $V_{\mathcal{L}}$.

Proof: Refer to [1]. ∎

III. Strict Improvement for a General Class of Sources and Distortion Measures

We begin by defining $\mathcal{Z}_{ZB}$, the set of all sources (for given distortion measures at the decoders), for which there exists an operating point $(R_1, R_2, D_1, D_2, D_{12})$ that *cannot* be achieved by an 'independent quantization' mechanism using the ZB coding scheme. More specifically, $X \in \mathcal{Z}_{ZB}$, if there exists a strict suboptimality in the ZB region when the closure is defined only over joint densities for the auxiliary random variables satisfying the following conditions:

$$\begin{aligned} P(U_1, U_2|X, V_{12}) &= P(U_1|X, V_{12})P(U_2|X, V_{12}) \\ E\left[d_{\mathcal{K}}(X, \psi_{\mathcal{K}}(U_{\mathcal{K}}))\right] &\leq D_{\mathcal{K}}, \quad \mathcal{K} \in \{1, 2, 12\} \\ U_{12} &= f(U_1, U_2, V_{12}) \end{aligned} \tag{10}$$

where f is any deterministic function. We will show in Theorem 2 that $\forall X \in \mathcal{Z}_{ZB}$, $\mathcal{RD}_{VKG} \subset \mathcal{RD}_{CMS}$.

Before stating the result we describe the particular cross-section of the RD region that we will use to prove strict improvement in Theorem 2. Consider a 3-descriptions MD setup for a source X wherein we impose constraints only on distortions $(D_1, D_2, D_3, D_{12}, D_{23})$ and set the rest of the distortions, (D_{13}, D_{123}) to ∞. This cross-section is schematically shown in Fig. 1. To illustrate the gains underlying CMS, here we restrict ourselves to the setting wherein we further impose $D_1 = D_3$ and $D_{12} = D_{23}$. The points in this cross-section, achievable by VKG and CMS, are denoted by $\overline{\mathcal{RD}}_{VKG}(X)$ and $\overline{\mathcal{RD}}_{CMS}(X)$, respectively. We note that the symmetric setting is considered *only* for simplicity. The arguments can be easily extended to the asymmetric framework.

This particular symmetric cross-section of the 3-descriptions MD problem is equivalent to the corresponding 2-descriptions problem, in the sense that, one could use any coding scheme to generate bit-streams for descriptions 1 and 2, respectively. Description 3 would then carry a replica (exact copy) of the bits sent in description 1. Due to the underlying symmetry in the problem setup, the distortion constraints at all the decoders are satisfied. Hence an achievable region based on the ZB coding scheme can be derived as follows. Let $(G_{12}, F_1, F_2, F_{12})$ be any random variables jointly distributed with X and taking values over arbitrary finite alphabets. Then the following RD-region is achievable for which there exist functions $(\psi_1, \psi_2, \psi_{12})$ such that $R_1 = R_3$, $D_1 = D_3$, $D_{12} = D_{23}$ and:

$$\begin{aligned} R_1 \geq I(X; F_1, G_{12}),\ & R_2 \geq I(X; F_2, G_{12}) \\ R_1 + R_2 \geq 2I(X; G_{12}) &+ H(F_1|G_{12}) + H(F_2|G_{12}) \\ -H(F_1, F_2, F_{12}|X, G_{12}) &+ H(F_{12}|F_1, F_2, G_{12}) \\ D_{\mathcal{K}} \geq E\left[d_{\mathcal{K}}(X, \psi_{\mathcal{K}}(F_{\mathcal{K}}))\right],\ & \mathcal{K} \in \{1, 2, 12\} \end{aligned} \tag{11}$$

The closure of achievable RD-tuples over all random variables $(G_{12}, M_1, M_2, M_{12})$ is denoted by $\overline{\mathcal{RD}}(X)$. In the following theorem, we will show that $\overline{\mathcal{RD}}(X) \subseteq \overline{\mathcal{RD}}_{CMS}(X)$. We also show that the VKG coding scheme *cannot* achieve the above RD region, i.e., $\overline{\mathcal{RD}}_{VKG}(X) \subset \overline{\mathcal{RD}}(X)$, if $X \in \mathcal{Z}_{ZB}$. We note that in Theorem 2, we focus only on the 3-descriptions setting. However, the results can be easily extended to the general $L-$descriptions scenario. Also note that $\overline{\mathcal{RD}}_{CMS}(X)$ could be strictly larger than $\overline{\mathcal{RD}}(X)$, in general.

Theorem 2. *(i) For the setup shown in Fig. 1 the CMS scheme achieves $\overline{\mathcal{RD}}(X)$, i.e., $\overline{\mathcal{RD}}(X) \subseteq \overline{\mathcal{RD}}_{CMS}(X)$.*

*(ii) If $X \in \mathcal{Z}_{ZB}$, then there exists points in $\overline{\mathcal{RD}}(X)$ that **cannot** be achieved by the VKG encoding scheme, i.e., $\overline{\mathcal{RD}}_{VKG} \subset \overline{\mathcal{RD}}(X)$,*

Remark 3. It directly follows from (i) and (ii) that $\mathcal{RD}_{VKG} \subset \mathcal{RD}_{CMS}$ for the $L-$channel MD problem $\forall L \geq 3$, if $X \in \mathcal{Z}_{ZB}$.

Proof: We first provide an intuitive argument to justify the claim and then follow it up with a formal argument. Due to the underlying symmetry in the setup CMS introduces common layer random variables $V_{123} = G_{12}$ and $V_{13} = F_1$. It then sends the codeword of V_{13} is both descriptions 1 and 3 (i.e.,

$U_1 = U_3 = V_{13}$). Hence it is sufficient for the encoder to generate enough codewords of $U_2 = F_2$ (conditioned on V_{123}) to maintain joint typicality with the codewords of $V_{13} = F_1$. However VKG is forced to set the common layer random variable V_{13} to a constant. Thus, in this case, the encoder needs to generate enough number of codewords of U_2 so as to maintain joint typicality individually with the codewords of U_1 and U_3, which are now generated independently conditioned on V_{123}, entailing some excess rate for U_2[2].

Part (i) of the theorem is straightforward to prove. We set $V_{123} = G_{12}$, $V_{13} = F_1$, $U_2 = F_2$, $U_{12} = U_{23} = F_{12}$ and $U_1 = U_3 = V_{13}$ and the rest of the random variables to constants in the CMS achievable region in [1]. This leads to the RD region in (11).

We next prove (ii). We consider one particular boundary point of (11) and show that this cannot be achieved by VKG. Let D_1, D_2 and D_{12} be fixed. Consider the following quantity:

$$R^*_{VKG}(D_1, D_2, D_{12}) = \inf \Big\{ R_2 : R_1 = R_3 = R_X(D_1) \quad (12)$$
$$(R_1, R_2, R_3, D_1, D_2, D_1, D_{12}, D_{12}) \in \overline{\mathcal{RD}}_{VKG}(X) \Big\}$$

Note that the corresponding quantity achievable using $\overline{\mathcal{RD}}_{CMS}(X)$ is given by the solution to the following optimization problem:

$$R^*_{CMS}(D_1, D_2, D_{12}) = \inf \Big\{ I(U_2; X, U_1|V_{123})$$
$$I(X; V_{123}) + I(U_{12}; X|V_{123}, U_1, U_2) \Big\} \qquad (13)$$

where the infimum is over all joint densities $P(V_{123}, U_1, U_2, U_{12}|X)$, where $P(V_{123}, U_1|X)$ is any joint density for which there exists a function $\psi_1(\cdot)$ such that:

$$I(X; V_{123}, U_1) = R(D_1), \qquad E\left[d_1(X, \psi_1(U_1))\right] = D_1 \quad (14)$$

i.e., (V_{123}, U_1) leads to an RD-optimal reconstruction of X at D_1 and $P(U_{12}, U_2|X, U_1, V_{123})$ is any distribution for which there exists function $\psi_2(\cdot)$ and $\psi_{12}(\cdot)$ satisfying the distortion constraints for D_2 and D_{12}, respectively. We will show that $R^*_{VKG} > R^*_{CMS}$. We next specialize and restate $\overline{\mathcal{RD}}_{VKG}(X)$ for the considered cross-section. Let $(V_{123}, U_1, U_2, U_3, U_{12}, U_{23})$ be any random variables jointly distributed with X taking values on arbitrary alphabets. Then, $\overline{\mathcal{RD}}_{VKG}$ contains all rates and distortions for which there exist functions $\psi_1(\cdot), \psi_2(\cdot), \psi_3(\cdot), \psi_{12}(\cdot), \psi_{23}(\cdot)$, such that:

$$\begin{aligned}
R_i &\geq I(X; U_i, V_{123}), \; i \in \{1,2,3\} \\
R_i + R_2 &\geq 2I(X; V_{123}) + I(U_i; U_2|V_{123}) \\
&\quad + I(X; U_i, U_2, U_{i2}|V_{123}), \; i \in \{1,3\} \\
R_1 + R_3 &\geq 2I(X; V_{123}) + H(U_1|V_{123}) \\
&\quad + H(U_3|V_{123}) - H(U_1, U_3|X, V_{123}) \\
R_1 + R_2 + R_3 &\geq 3I(X; V_{123}) + \sum_{i=1}^{3} H(U_i|V_{123}) \\
&\quad + \sum_{\mathcal{K} \in \{12,23\}} H(U_{\mathcal{K}}|\{U\}_{\mathcal{K}}, V_{123}) \\
&\quad - H(U_1, U_2, U_3, U_{12}, U_{23}|X, V_{123}) \qquad (15)
\end{aligned}$$

$$E\left(d_{\mathcal{K}}(X, \psi_{\mathcal{K}}(U_{\mathcal{K}}))\right) \leq D_{\mathcal{K}}, \; \mathcal{K} \in \{1,2,3,12,13\} \quad (16)$$

where $R_1 = R_3$, $D_1 = D_3$ and $D_{12} = D_{23}$. Observe that the random variables U_{13} and U_{123} have been set to constants as we do not impose distortion constraints D_{13} and D_{123}, respectively. We can further restrict the joint density $P(V_{123}, U_1, U_2, U_3, U_{12}, U_{23}|X)$ to satisfy:

$$P(U_{12}, U_{23}|X, V_{123}, U_1, U_2, U_3) =$$
$$P(U_{12}|X, V_{123}, U_1, U_2) P(U_{23}|X, V_{123}, U_2, U_3) \qquad (17)$$

without any loss of optimality.

Next imposing $R_1 = R_3 = R_X(D_1)$ in (15), enforces the joint density $P(V_{123}, U_1, U_3|X)$ to satisfy the following constraints:

$$\begin{aligned}
I(X; V_{123}, U_i) &= R(D_1), \; i \in \{1,3\} \\
E\left[d_i(X, \psi_i(V_{123}, U_i))\right] &= D_1, \; i \in \{1,3\} \qquad (18) \\
P(U_1, U_3|X, V_{123}) &= P(U_1|X, V_{123}) \times P(U_3|X, V_{123})
\end{aligned}$$

where the last condition is required to satisfy the constraint on $R_1 + R_3$ in (15). Therefore, using (15) and (17) we have:

$$R^*_{VKG} = \inf \Big\{ I(X; V_{123}) + I(U_2; U_1, U_3, X|V_{123})$$
$$+ I(X; U_{12}|U_1, U_2, V_{123}) + I(X; U_{23}|U_2, U_3, V_{123}) \Big\} \qquad (19)$$

where the infimum is over all joint densities $P(V_{123}, U_1, U_2, U_3, U_{12}, U_{23}|X)$ satisfying (18) for which there exist functions $\psi_2(\cdot), \psi_{12}(\cdot), \psi_{23}(\cdot)$ satisfying the distortion constraints in (16).

From (19) and (13) it follows that R^*_{VKG} is equal to R^*_{CMS} if and only if the two quantities on the RHS of (19) and (13), respectively, are equal. However for any joint density, we have $I(U_2; U_1, U_3, X|V_{123}) \geq I(U_2; U_1, X|V_{123})$ and $I(X; U_{23}|V_{123}, U_2, U_3) \geq 0$. Also note that the constraints in (14) are a subset of the constraints in (18). Hence for R^*_{VKG} to be equal to R^*_{CMS}, any joint density which achieves R^*_{VKG} must satisfy the following conditions:

(i) The joint density of $(X, V_{123}, U_1, U_2, U_{12})$ must be the same as the corresponding joint density which achieves R^*_{CMS} (in (13)).

(ii)$I(U_2; U_3|V_{123}, U_1, X) = 0$, $I(X; U_{23}|V_{123}, U_2, U_3) = 0$. The constraint $I(X; U_{23}|V_{123}, U_2, U_3) = 0$ implies that

[2] It might be tempting to conclude that the suboptimality in VKG is due to conditions for joint typicality of all the codewords, while for this cross-section, joint typicality of codewords of U_1 and U_3 is unnecessary. However, it is possible to show that common codewords provide strict improvement even when joint typicality only within prescribed subsets is imposed. The details are omitted here.

X and U_{23} are independent given V_{123}, U_2 and U_3. Equivalently this constraint implies that the reconstruction $\hat{X}_{23}$ can be written as a deterministic function of V_{123}, U_2 and U_3, i.e., for R^*_{VKG} to be equal to R^*_{CMS}, there must exist a function $\tilde{\psi}_{23}(V_{123}, U_2, U_3)$ such that $E\left(d_{23}(X, \tilde{\psi}_{23}(V_{123}, U_2, U_3))\right) \leq D_{23} = D_{12}$. On the other hand, the constraint $I(U_2; U_3|V_{123}, U_1, X) = 0$ implies that $H(U_3|V_{123}, U_1, X) = H(U_3|V_{123}, U_1, U_2, X)$. However, the joint density of (X, V_{123}, U_1, U_3) must satisfy (18) for $R_1 = R_3 = R_X(D_1)$ to hold, i.e., $H(U_3|V_{123}, U_1, X) = H(U_3|V_{123}, X)$. Hence for R^*_{VKG} to be equal to R^*_{CMS}, we require:

$$H(U_3|V_{123}, X) = H(U_3|V_{123}, U_1, U_2, X) \tag{20}$$

which implies that $U_2 \leftrightarrow (X, V_{123}) \leftrightarrow U_3$ must hold. Recall that the joint density $P(U_3, V_{123}|X)$ is RD-optimal at D_1 and the joint density $P(U_2, V_{123}|X, U_1)$ must be identical to the joint density which achieves R^*_{CMS} (from condition (i)). Hence, it follows that, if $X \in \mathcal{Z}_{ZB}$, there exists at least one RD tuple in $\overline{\mathcal{RD}}(X)$ that cannot be achieved if we constrain the joint density to simultaneously satisfy both the conditions (i) and (ii), proving the theorem. ∎

Discussion: A direct consequence of the above theorem is that, if $X \in \mathcal{Z}_{ZB}$, then the common layer codewords of CMS provide strict improvement in the achievable region as compared not using them, i.e., if $X \in \mathcal{Z}_{ZB}$, $\mathcal{RD}_{VKG}\Big|_{V_{\mathcal{L}}=\Phi} \subseteq \mathcal{RD}_{VKG} \subset \mathcal{RD}_{CMS}$, where $\mathcal{RD}\Big|_{V_{\mathcal{L}}=\Phi}$ denotes the VKG region when the common layer random variable (denoted by $V_{\mathcal{L}}$) is set to a constant Φ[3]. In fact, it is possible to show that, whenever $X \in \mathcal{Z}_{EC}$, $\mathcal{RD}\Big|_{V_{\mathcal{L}}=\Phi} \subset \mathcal{RD}_{CMS}$, where $\mathcal{Z}_{EC}$ is defined as the set of all sources for which there exists an operating point (with respect to the given distortion measures) that *cannot* be achieved by an 'independent quantization' mechanism using the EC coding scheme, i.e., if there exists an operating point that *cannot* be achieved by EC using a joint density for the auxiliary random variables satisfying:

$$\begin{aligned} P(U_1, U_2|X) &= P(U_1|X)P(U_2|X) \\ E\left[d_{\mathcal{K}}(X, \psi_{\mathcal{K}}(U_{\mathcal{K}}))\right] &\leq D_{\mathcal{K}}, \;\; \mathcal{K} \in \{1, 2, 12\} \\ U_{12} &= f(U_1, U_2) \end{aligned} \tag{21}$$

where f is any deterministic function. Note that the set $\mathcal{Z}_{ZB}$ is a subset of $\mathcal{Z}_{EC}$. Also observe that if $X \notin \mathcal{Z}_{EC}$, the concatenation of two independent optimal quantizers is optimal in achieving a joint reconstruction. While this condition could be satisfied for specific values of D_1, D_2 and D_{12}, it is seldom achieved *for all* values of (D_1, D_2, D_{12}). Though such sources are of some theoretical interest, the multiple descriptions encoding for such sources is degenerate. Hence with some trivial exceptions, it can be asserted that the common layer codewords in CMS can be used to achieve a strictly larger region (compared to not using any common codewords) for all sources and distortion measures, $\forall L \geq 3$.

[3]Note that setting $V_{\mathcal{L}}$ to a constant in VKG is equivalent to setting all the common layer random variables to constants in CMS.

IV. Gaussian MSE Setting

In the following theorem we show that, under MSE, a Gaussian source belongs to $\mathcal{Z}_{ZB}$.

Theorem 3. *(i) CMS achieves the **complete** RD region for the symmetric 3-descriptions quadratic Gaussian setup shown in Fig. 1.*

(ii) The VKG encoding scheme cannot achieve all the points in the region, i.e., $\overline{\mathcal{RD}}_{VKG} \subset \overline{\mathcal{RD}}_{CMS}$.

Remark 4. It follows from (i) and (ii) that $\mathcal{RD}_{VKG} \subset \mathcal{RD}_{CMS}$ for the $L-$channel quadratic Gaussian MD problem $\forall L > 2$.

Proof: Proof of (i) is straightforward and follows directly from the proof of Theorem 1. Hence, we only prove (ii). Specifically, we show that, a Gaussian random variable, under MSE, belongs to $\mathcal{Z}_{ZB}$. (ii) then follows directly from Theorem 1.

Consider the 2-description quadratic Gaussian problem. It follows from Ozarow's results (see also [4]) that, if $D_{12} \leq D_1 + D_2 - 1$, then the following rate region is achievable (and complete):

$$R_{\mathcal{K}} \geq \frac{1}{2}\log\frac{1}{D_{\mathcal{K}}}, \;\; \mathcal{K} \in \{1, 2, 12\} \tag{22}$$

i.e., there is no excess rate incurred due to encoding the source using two descriptions. Observe that the excess sum rate term in the ZB must be set to zero to achieve the above rate-region. We will show that, if we restrict the optimization to conditionally independent joint densities, then it is impossible to simultaneously satisfy all the distortions and achieve $I(U_1; U_2|V_{12}) = 0$. Recall that the ZB region achievable using any joint density $P(X, V_{12}, U_1, U_2, U_{12})$ is given by:

$$\begin{aligned} R_i &\geq I(X; V_{12}, U_i) \;\; i \in \{1, 2\} \\ R_1 + R_2 &\geq I(X; V_{12}) + I(U_1; U_2|V_{12}) \\ &\quad + I(X; V_{12}, U_1, U_2, U_{12}) \\ D_{\mathcal{K}} &\geq E\left[(X - \psi_{\mathcal{K}}(U_{\mathcal{K}}))\right], \;\; \mathcal{K} \in \{1, 2, 12\} \end{aligned} \tag{23}$$

Let us consider the corner point $P_0 \triangleq (R_1, R_2) = (\frac{1}{2}\log\frac{1}{D_1}, \frac{1}{2}\log\frac{D_1}{D_{12}})$ for some (D_1, D_2, D_{12}) satisfying $D_{12} \leq D_1 + D_2 - 1$ and show that this point is not achievable by the ZB scheme when we restrict the joint densities to satisfy (10). First, as $I(X; V_{12}, U_1, U_2, U_{12}) \geq \frac{1}{2}\log\frac{1}{D_{12}}$, P_0 can be achieved only by joint densities that satisfy $I(X; V_{12}) = 0$. Hence, to prove the theorem, it is sufficient to show that P_0 is not achievable when we restrict the optimization to joint densities satisfying (10) and $I(X; V_{12}) = 0$.

Let $P(V_{12}, U_1, U_2, U_{12}, X)$ be any such joint density and let $\mathcal{V}_{12}$ be the alphabet for V_{12}. Then the associated ZB achievable region can be rewritten as:

$$\begin{aligned}
R_i &\geq \sum_{v_{12}\in\mathcal{V}_{12}} P(v_{12}) I(X;U_i|V_{12}=v_{12}),\ i\in\{1,2\}\\
R_1+R_2 &\geq \sum_{v_{12}\in\mathcal{V}_{12}} P(v_{12})\Big[I(U_1;U_2|V_{12}=v_{12})\\
&\qquad +I(X;U_1,U_2,U_{12}|V_{12}=v_{12})\Big]\\
D_{\mathcal{K}} &\geq E\left[(X-\psi_{\mathcal{K}}(U_{\mathcal{K}}))^2\right],\ \mathcal{K}\in\{1,2,12\}\\
&= E\left[E\left[(X-\psi_{\mathcal{K}}(U_{\mathcal{K}}))^2\Big|V_{12}\right]\right] \qquad (24)
\end{aligned}$$

We will next show that the optimization can be further restricted to joint densities $P(X,V_{12})Q(\tilde{U}_1,\tilde{U}_2,\tilde{U}_{12}|X,V_{12})$ such that $(X,\tilde{U}_1,\tilde{U}_2,\tilde{U}_{12})$ are jointly Gaussian given $V_{12}=v_{12}$, $\forall v_{12}\in\mathcal{V}_{12}$ and satisfying $Q(\tilde{U}_1,\tilde{U}_2|X,V_{12}) = Q(\tilde{U}_1|X,V_{12})Q(\tilde{U}_2|X,V_{12})$. First, note that, as $I(X;V_{12})=0$, $P(X|V_{12}=v_{12})$ is Gaussian $\forall v_{12}\in\mathcal{V}_{12}$. Next, recall that P_0 is obtained by first minimizing R_1 followed by minimizing R_2 given R_1 subject to all the distortion constraints. From (24), we have $R_1 = \min\sum_{v_{12}\in\mathcal{V}_{12}} P(v_{12})I(X;U_1|V_{12}=v_{12})$, where the minimization is over all joint densities $P(X,V_{12},U_1)$ satisfying the distortion constraint for D_1.

Let $P(X,V_{12},U_1)$ be any joint density satisfying the distortion constraint for D_1. Consider the joint density generated as $Q(X,V_{12},\tilde{U}_1) = P(X,V_{12})Q(\tilde{U}_1|X,V_{12})$ where $(X,\tilde{U}_1)$ are jointly Gaussian given $V_{12}=v_{12}$ and $K_{X,\tilde{U}_1|V_{12}=v_{12}} = K_{X,U_1|V_{12}=v_{12}}$, $\forall v_{12}\in\mathcal{V}_{12}$. Observe that $Q(\cdot)$ also satisfies the distortion constraint for D_1. As a Gaussian density over the relevant random variables maximizes the conditional entropy for a fixed covariance matrix [12], it follows that $I(X;U_1|V_{12}=v_{12}) \geq I(X;\tilde{U}_1|V_{12}=v_{12})$, $\forall v_{12}\in\mathcal{V}_{12}$. Hence, to achieve minimum R_1, it is sufficient to restrict the optimization to densities wherein (X,U_1) are jointly Gaussian given V_{12}.

Next consider minimizing R_2 given R_1. From (24), we have:

$$\begin{aligned}
R_2 &= \min\Big\{\sum_{v_{12}\in\mathcal{V}_{12}} P(v_{12})\Big[I(\tilde{U}_1;U_2|V_{12}=v_{12})\\
&\qquad +I(X;\tilde{U}_1,U_2,U_{12}|V_{12}=v_{12})\Big]-R_1\Big\} \qquad (25)
\end{aligned}$$

where the minimization is over all joint densities $P(X,V_{12},U_1)P(U_2,U_{12}|X,V_{12},U_1)$ satisfying (10) and $I(X;V_{12})=0$ and where (X,U_1) are jointly Gaussian given $V_{12}=v_{12}$, $\forall v_{12}\in\mathcal{V}_{12}$ (required to minimize R_1). It is easy to show using similar arguments that the above minimization is again achieved by a joint density where (X,U_1,U_2,U_{12}) are jointly Gaussian given $V_{12}=v_{12}$ and such that $Q(U_1,U_2|X,V_{12}) = Q(U_1|X,V_{12})Q(U_2|X,V_{12})$ $\forall v_{12}\in\mathcal{V}_{12}$. Hence, to achieve P_0 using a joint density that satisfies (10), it is sufficient to optimize the rates over joint densities satisfying the following properties:

1) (X,U_1,U_2,U_{12}) are jointly Gaussian given $V_{12}=v_{12}$ $\forall v_{12}\in\mathcal{V}_{12}$
2) $I(X;V_{12})=0$
3) $I(U_1;U_2|X,V_{12})=0$
4) $P(X,V_{12},U_1,U_2,U_{12})$ satisfies all the distortion constraints

Observe that, for any joint density that satisfies all the above properties, it is impossible to achieve $I(U_1;U_2|V_{12})=0$. Therefore, the excess sum rate term in the ZB scheme is non-zero, concluding that P_0 is not achievable by the ZB scheme using any independent quantization mechanism. Therefore, a Gaussian random variable under MSE belongs to $\mathcal{Z}_{ZB}$, proving the theorem. ■

Note that, as $\mathcal{Z}_{ZB}\subseteq\mathcal{Z}_{EC}$, a Gaussian source under MSE belongs to $\mathcal{Z}_{EC}$. Hence, the 'correlated quantization' scheme (an extreme special case of VKG) which has been proven to be complete for several cross-sections of the $L-$descriptions quadratic Gaussian MD problem [9], is strictly suboptimal in general.

V. Points on the Boundary

Before stating the results formally, we review Ozarow's result for the 2-descriptions MD setting. Ozarow showed that the complete region for the 2-descriptions Gaussian MD problem can be achieved using a 'correlated quantization' scheme which imposes the following joint distribution for (U_1,U_2,U_{12}) in the EC scheme:

$$\begin{aligned}
U_1 &= X+W_1\\
U_2 &= X+W_2 \qquad (26)
\end{aligned}$$

$U_{12}=E(X|U_1,U_2)$, 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 $\psi_{\mathcal{K}}(U_{\mathcal{K}})$ are given by the respective MSE optimal estimators, eg., $\psi_1(U_1)=E[X|U_1]$. The covariance matrix $K_{W_1W_2}$ is set to satisfy all the distortion constraints. Specifically, the optimum $K_{W_1W_2}$ is given by:

$$K_{W_1W_2} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \qquad (27)$$

where $\sigma_i^2 = \frac{D_i}{1-D_i}$ $i\in\{1,2\}$ and the optimum ρ_{12}, denoted by ρ_{12}^*, is given by (see [13]):

$$\begin{aligned}
\rho_{12}^* &= \begin{cases} -\frac{\sqrt{\pi D_{12}^2+\gamma}-\sqrt{\pi D_{12}^2}}{(1-D_{12})\sqrt{D_1D_2}} & D_{12}\leq D_{12}^{max}\\ 0 & D_{12}\geq D_{12}^{max}\end{cases}\\
\gamma &= (1-D_{12})\Big[(D_1-D_{12})(D_2-D_{12})\\
&\qquad +D_{12}D_1D_2-D_{12}^2\Big]\\
D_{12}^{max} &= D_1D_2/(D_1+D_2-D_1D_2)\\
\pi &= (1-D_1)(1-D_2) \qquad (28)
\end{aligned}$$

We denote the complete Gaussian-MSE L-descriptions region by $\mathcal{RD}_G^L$. The characterization of $\mathcal{RD}_G^2$ is given in [4] (see also [13]) and we omit restating it explicitly here for brevity.

In this section we show that CMS achieves the complete RD region for several cross-sections of the general quadratic Gaussian $L-$channel MD problem. We again begin with the 3-descriptions case and then extend the results to the L channel framework. Recall the setup shown in Fig. 1, i.e, a cross-section of the general 3-descriptions rate-distortion region wherein we impose constraints only on distortions $(D_1, D_2, D_3, D_{12}, D_{23})$ and set the rest of the distortions, (D_{13}, D_{123}) to 1. Here we consider the general asymmetric case, i.e. $D_1 \neq D_3$ and $D_{12} \neq D_{23}$ and show that the CMS scheme achieves the complete rate region in several distortion regimes.

In the following theorem, without loss of generality we assume that $D_1 \leq D_3$. If $D_3 \leq D_1$, then the theorem holds by interchanging '1' and '3' everywhere. Let D_{12} be any distortion such that $D_{12} \leq \min\{D_1, D_2\}$. We define $D_{23}^* = D_{23}^*(D_1, D_2, D_3, D_{12})$ as:

$$D_{23}^* = \frac{\sigma_2^2\sigma_3^2\left(1-\rho^2\right)}{\sigma_2^2\sigma_3^2\left(1-\rho^2\right)+\sigma_2^2+\sigma_3^2-2\sigma_2\sigma_3\rho} \tag{29}$$

where $\sigma_i^2 = \frac{D_i}{1-D_i}$ $i \in \{2,3\}$ and

$$\rho = \rho_{12}^* \frac{\sigma_1}{\sigma_3} \tag{30}$$

where ρ_{12}^* is defined in (28). In the following theorem, we will show that CMS achieves the complete rate-region if $D_{23} = D_{23}^*$.

Theorem 4. *For the setup shown in Fig. 1, let $D_1 \leq D_3$. Then,*

(i) CMS achieves the complete rate-region if:

$$D_{23} \quad = \quad D_{23}^*(D_1, D_2, D_3, D_{12}) \tag{31}$$

where D_{23}^* is defined in (29). The rate region is given by:

$$\begin{aligned} R_i &\geq \frac{1}{2}\log\frac{1}{D_i} \;\; i \in \{1,2,3\} \\ R_1 + R_2 &\geq \frac{1}{2}\log\frac{1}{D_1D_2} + \delta(D_1, D_2, D_{12}) \\ R_2 + R_3 &\geq \frac{1}{2}\log\frac{1}{D_2D_3} + \delta(D_2, D_3, D_{23}) \end{aligned} \tag{32}$$

where $\delta(\cdot)$ is defined by:

$$\delta(D_1, D_2, D_{12}) = \frac{1}{2}\log\left(\frac{1}{1-(\rho_{12}^*)^2}\right) \tag{33}$$

where ρ_{12}^* is defined in (28).

(ii) Moreover, CMS achieves the minimum sum-rate if one of the following hold:

(a) For a fixed D_{12}, $D_{23} \geq D_{23}^*(D_1, D_2, D_3, D_{12})$

(b) For a fixed D_{23}, $D_{12} \in \{D_{12} : \delta(D_2, D_3, D_{23}) \geq \delta(D_1, D_2, D_{12})\}$

Remark 5. We note that the above rate region *cannot* be achieved by VKG. We omit the details of the proof here as it can be proved in same lines as the proof of Theorem 3.

Remark 6. An achievable rate-distortion region can be derived for general distortions using the encoding principles we derive as part of this proof. However, it is hard to prove outer bounds if the conditions in (31) are not satisfied and hence we omit stating the results explicitly here.

Remark 7. Both the CMS and the VKG schemes achieve the complete rate region when $D_{12} \geq D_{12}^{max}$ and $D_{23} \geq D_{23}^{max}$, where D_{12}^{max} and D_{23}^{max} are defined in (28). It is easy to show that in this case an independent quantization scheme is optimal and the complete achievable rate-region is given by $R_i \geq \frac{1}{2}\log\frac{1}{D_i}$ $i \in \{1,2,3\}$.

Remark 8. It is easy to verify that CMS achieves the minimum sum-rate whenever $D_{12} = D_{23}$ for any D_1, D_3.

Proof: We begin with the proof of (i). The proof of (ii) then follows almost directly from (i). First we show the converse, which is quite obvious. Conditions on R_i follow from the converse to the source coding theorem. Conditions on $R_1 + R_2$ and $R_2 + R_3$ follow from Ozarow's result, to achieve (D_1, D_2, D_{12}) using descriptions $\{1,2\}$ and to achieve (D_2, D_3, D_{23}) using descriptions $\{2,3\}$ at the respective decoders.

We next prove that CMS achieves the rate region in (32) if (31) holds. We first give an intuitive argument to explain the encoding scheme. Description 3 carries an RD-optimal quantized version of X (which achieves distortion D_3). Description 1 carries all the bits embedded in description 3 along with 'refinement bits' which assist in achieving distortion $D_1 \leq D_3$. This entails no loss in optimality as a Gaussian source is successively refinable under MSE [14]. Description 2 then carries a quantized version of the source which is correlated with the information in descriptions 1 and 3. We will show that if $D_{23} = D_{23}^*(D_1, D_2, D_3, D_{12})$, then the correlations can be set such that description 2 is optimal with respect to both descriptions 1 and 3.

Formally, to achieve the rate region in (32), we set the auxiliary random variables in the CMS coding scheme as follows:

$$\begin{aligned} V_{13} &= X + W_1 + W_3 \\ U_3 &= V_{13} \\ U_1 &= X + W_1 \\ U_2 &= X + W_2 \\ U_{12} = \Phi \qquad & U_{23} = \Phi \end{aligned} \tag{34}$$

and the functions $\psi(\cdot)$ as the respective MSE optimal estimators, where W_1, W_2, W_3 are zero mean Gaussian random variables independent of X with a covariance matrix:

$$K_{W_1W_2W_3} = \begin{bmatrix} \tilde{\sigma}_1^2 & \rho_{12}\tilde{\sigma}_1\tilde{\sigma}_2 & 0 \\ \rho_{12}\tilde{\sigma}_1\tilde{\sigma}_2 & \tilde{\sigma}_2^2 & 0 \\ 0 & 0 & \tilde{\sigma}_3^2 \end{bmatrix} \tag{35}$$

where $\tilde{\sigma}_1^2 = \sigma_1^2 = \frac{D_1}{1-D_1}$, $\tilde{\sigma}_2^2 = \sigma_2^2 = \frac{D_2}{1-D_2}$, $\tilde{\sigma}_3^2 = \sigma_3^2 - \sigma_1^2 = \frac{D_3}{1-D_3} - \frac{D_1}{1-D_1}$. The correlation coefficient ρ_{12} is set to achieve distortion D_{12}, i.e. $\rho_{12} = \rho_{12}^*$ defined in (28). Let us denote by $W_{13} = W_1 + W_3$. Observe that the encoding for descriptions 2 and 3 resembles Ozarow's correlated quantization scheme with $U_2 = X + W_2$ and $U_3 = X + W_{13}$. Let us denote the correlation coefficient between W_2 and W_{13} be ρ. We have the following equation relating ρ_{12} and ρ (which is equivalent to (30)):

$$\rho_{12}^*\tilde{\sigma}_1 = \rho\sqrt{\tilde{\sigma}_1^2 + \tilde{\sigma}_3^2} \tag{36}$$

Note that the above relation is derived using the independence of W_2 and (W_1, W_3), which follows from our choice of

$K_{W_1W_2W_3}$. Hence the minimum distortion D_{23} achievable using the above choice for the joint density of the auxiliary random variables is given by:

$$\begin{aligned} D_{23} &= \text{Var}(X|U_2,U_3,V_{13}) \\ &= \text{Var}(X|U_2,V_{13}) \\ &= D_{23}^* \end{aligned} \quad (37)$$

We next derive the rates required by this choice of $K_{W_1W_2W_3}$. Direct application of Theorem 1 using the above joint density leads to the following achievable rate region for any given distortions $D_1, D_2, D_3, D_{12}, D_{23}$:

$$\begin{aligned} R_{13}^{''} &\geq \frac{1}{2}\log\frac{1}{D_3} \\ R_2^{'} &\geq \frac{1}{2}\log\frac{1}{D_2} \\ R_1^{'}+R_{13}^{''} &\geq \frac{1}{2}\log\frac{1}{D_1} \\ R_2^{'}+R_{13}^{''} &\geq H(V_{13})+H(U_2)-H(V_{13},U_2|X) \\ &= H(U_3)+H(U_2)-H(U_3,U_2|X) \\ &= \frac{1}{2}\log\frac{1}{D_3D_2}+\frac{1}{2}\log\left(\frac{1}{1-\rho^2}\right) \\ &= \frac{1}{2}\log\frac{1}{D_3D_2}+\delta(D_2,D_3,D_{23}^*) \end{aligned}$$

$$\begin{aligned} R_1^{'}+R_2^{'}+R_{13}^{''} &\geq H(V_{13})+H(U_1|V_{13})+H(U_2) \\ &\quad -H(U_1,V_{13},U_2|X) \\ &= I(X;U_1,V_{13})+I(U_2;X,U_1,V_{13}) \\ &=^{(a)} I(X;U_1)+I(X;U_2) \\ &\quad +I(U_2;U_1,V_{13}|X) \\ &= I(X;U_1)+I(X;U_2) \\ &\quad +I(U_2;U_1,U_3|X) \\ &=^{(b)} I(X;U_1)+I(X;U_2) \\ &\quad +I(W_2;W_1,W_1+W_3) \\ &= I(X;U_1)+I(X;U_2)+I(W_2;W_1) \\ &\quad +I(W_2;W_3|W_1) \\ &=^{(c)} I(X;U_1)+I(X;U_2)+I(W_2;W_1) \\ &= \frac{1}{2}\log\frac{1}{D_1D_2}+\frac{1}{2}\log\left(\frac{1}{1-(\rho_{12}^*)^2}\right) \\ &= \frac{1}{2}\log\frac{1}{D_1D_2}+\delta(D_1,D_2,D_{12}) \end{aligned}$$

$$\begin{aligned} R_1 &= R_{13}^{''}+R_1^{'} \\ R_2 &= R_2^{'} \\ R_3 &= R_{13}^{''} \end{aligned} \quad (38)$$

where (a) follows from the Markov chain $X \leftrightarrow U_1 \leftrightarrow V_{13}$, (b) from the independence of X and (W_1,W_2,W_3) and (c) from the independence of W_3 and (W_1,W_2).

At a first glance, it might be tempting to conclude that the region for the tuple (R_1,R_2,R_3) in (38) is equivalent to the region given by (32). This is not the case in general as the equations in (38) have an implicit constraint on the auxiliary rates $R_{13}^{''},R_1^{'},R_2^{'} \geq 0$. However, we will show that if $D_3 \geq D_1$, then the two regions are indeed equivalent. We denote the rate region given in (32) by $\mathcal{R}$ and the region in (38) by $\mathcal{R}^*$. Clearly, $\mathcal{R}^* \subseteq \mathcal{R}$, as any (R_1,R_2,R_3) that satisfies (38) also satisfies (32). We need to show that $\mathcal{R}^* \supseteq \mathcal{R}$. Towards proving this claim, note that both $\mathcal{R}$ and $\mathcal{R}^*$ are convex regions bounded by hyper-planes. Hence, it is sufficient for us to show that all the corner points of $\mathcal{R}$ lie in $\mathcal{R}^*$. Clearly, $\mathcal{R}$ has 6 corner points denoted by P_{ijk} $i,j,k \in \{1,2,3\}$ defined as:

$$\begin{aligned} P_{ijk} &= \{r_i,r_j,r_k\} \\ r_i &= \min R_i \\ r_j &= \min_{R_i=r_i} R_j \\ r_k &= \min_{R_i=r_i,R_j=r_j} R_k \end{aligned} \quad (39)$$

To prove $\mathcal{R}^* \supseteq \mathcal{R}$, we need to prove that every corner point $(r_1,r_2,r_3) \in \mathcal{R}$ is achieved by some non-negative $(R_{13}^{''},R_1^{'},R_2^{'},R_1,R_2,R_3) \in \mathcal{R}^*$ such that $R_i = r_i$, $i \in \{1,2,3\}$. We set $R_{13}^{''} = R_3 = r_3$ and $R_2^{'} = R_2 = r_2$ and show that we can always find $R_1^{'} \geq 0$ satisfying (38) such that $R_1 = R_1^{'} + R_{13}^{'} = r_1$. Let us first consider the points $P_{213} = P_{231}$ given by:

$$\begin{aligned} r_1 &= \frac{1}{2}\log\frac{1}{D_1}+\delta(D_1,D_2,D_{12}) \\ r_2 &= \frac{1}{2}\log\frac{1}{D_2} \\ r_3 &= \frac{1}{2}\log\frac{1}{D_3}+\delta(D_2,D_3,D_{23}) \end{aligned} \quad (40)$$

This can be achieved by using the following auxiliary rates, $R_2^{'} = r_2$, $R_{13}^{''} = r_3$ and

$$\begin{aligned} R_1^{'} &= \frac{1}{2}\log\frac{D_3}{D_1}+\delta(D_1,D_2,D_{12}) \\ &\quad -\delta(D_2,D_3,D_{23}) \\ &= \frac{1}{2}\log\frac{(1-D_1)D_3-(\rho_{12}^*)^2D_1(1-D_3)}{(1-D_1)D_1(1-(\rho_{12}^*)^2)} \end{aligned} \quad (41)$$

It is easy to verity that $R_1^{'} \geq 0$ if $D_3 \geq D_1$. Hence $P_{213} = P_{231} \in \mathcal{R}^*$. Let us next consider the points $P_{132} = P_{312}$ given by:

$$\begin{aligned} r_1 &= \frac{1}{2}\log\frac{1}{D_1} \\ r_2 &= \frac{1}{2}\log\frac{1}{D_2} \\ &\quad +\max\{\delta(D_1,D_2,D_{12}),\delta(D_2,D_3,D_{23})\} \\ r_3 &= \frac{1}{2}\log\frac{1}{D_3} \end{aligned} \quad (42)$$

Again it is easy to show that $(R_{13}^{''},R_1^{'},R_2^{'}) = (r_3,r_1-r_3,r_2)$ belongs to $\mathcal{R}^*$. Finally, we consider the remaining two points

P_{123} and P_{321}. P_{123} is given by:

$$\begin{aligned} r_1 &= \frac{1}{2}\log\frac{1}{D_1} \\ r_2 &= \frac{1}{2}\log\frac{1}{D_2} + \delta(D_1, D_2, D_{12}) \\ r_3 &= \frac{1}{2}\log\frac{1}{D_3} \\ &\quad + (\delta(D_2, D_3, D_{23}) - \delta(D_1, D_2, D_{12}))^+ \end{aligned} \tag{43}$$

where $x^+ = \max\{x, 0\}$. Consider the following auxiliary rates: $R_{13}^{''} = r_3$, $R_2^{'} = r_2$ and $R_1^{'} = \frac{1}{2}\log\frac{D_3}{D_1}$. Clearly the first three constraints in (38) are satisfied by these auxiliary rates. The following inequalities prove that the last two constraints are also satisfied by these rates and hence $P_{123} \in \mathcal{R}^*$.

$$\begin{aligned} R_2^{'} + R_{13}^{''} &= r_2 + r_3 \\ &= \frac{1}{2}\log\frac{1}{D_2D_3} + \delta(D_1, D_2, D_{12}) \\ &\quad + (\delta(D_2, D_3, D_{23}) - \delta(D_1, D_2, D_{12}))^+ \\ &\geq \frac{1}{2}\log\frac{1}{D_2D_3} + \delta(D_2, D_3, D_{23}) \end{aligned}$$

$$\begin{aligned} R_2^{'} + R_1^{'} + R_{13}^{''} &= \frac{1}{2}\log\frac{1}{D_1D_2} + \delta(D_1, D_2, D_{12}) \\ &\quad + (\delta(D_2, D_3, D_{23}) - \delta(D_1, D_2, D_{12}))^+ \\ &\geq \frac{1}{2}\log\frac{1}{D_1D_2} + \delta(D_1, D_2, D_{12}) \end{aligned} \tag{44}$$

Next consider P_{321}:

$$\begin{aligned} r_1 &= \frac{1}{2}\log\frac{1}{D_1} \\ &\quad + (\delta(D_1, D_2, D_{12}) - \delta(D_2, D_3, D_{23}))^+ \\ r_2 &= \frac{1}{2}\log\frac{1}{D_2} + \delta(D_2, D_3, D_{23}) \\ r_3 &= \frac{1}{2}\log\frac{1}{D_3} \end{aligned} \tag{45}$$

Using same arguments as before, it is easy to show that $P_{321} \in \mathcal{R}^*$ by using the following auxiliary rates: $R_{13}^{''} = r_3$, $R_2^{'} = r_2$ and $R_1^{'} = \frac{1}{2}\log\frac{D_3}{D_1} + (\delta(D_1, D_2, D_{12}) - \delta(D_2, D_3, D_{23}))^+$. Therefore, it follows that $\mathcal{R} = \mathcal{R}^*$ and hence CMS achieves the complete rate region, proving (i).

We next prove (ii)(a). It follows from (i) that the following rate point is achievable $\forall D_{23} \geq D_{23}^*$:

$$\begin{aligned} \{R_1, R_2, R_3\} &= \Big\{\frac{1}{2}\log\frac{1}{D_1}, \frac{1}{2}\log\frac{1}{D_2} + \delta(D_1, D_2, D_{12}), \\ &\quad \frac{1}{2}\log\frac{1}{D_3}\Big\} \end{aligned} \tag{46}$$

Also observe that $\forall D_{23} \geq D_{23}^*$, $\delta(D_1, D_2, D_{12}) \geq \delta(D_2, D_3, D_{23})$ and hence a lower bound to the sum rate is $\frac{1}{D_1D_2D_3} + \delta(D_1, D_2, D_{12})$. Therefore the above point achieves the minimum sum rate $\forall D_{23} \geq D_{23}^*$.

The proof of (ii)(b) follows similarly by noting that if $D_{12} \in \{D_{12} : \delta(D_2, D_3, D_{23}) \geq \delta(D_1, D_2, D_{12})\}$, the minimum sum rate is given by $\frac{1}{D_1D_2D_3} + \delta(D_2, D_3, D_{23})$ which is achieved by the point:

$$\begin{aligned} \{R_1, R_2, R_3\} &= \Big\{\frac{1}{2}\log\frac{1}{D_1}, \frac{1}{2}\log\frac{1}{D_2} + \delta(D_2, D_3, D_{12}), \\ &\quad \frac{1}{2}\log\frac{1}{D_3}\Big\} \end{aligned} \tag{47}$$

This proves the theorem. ■

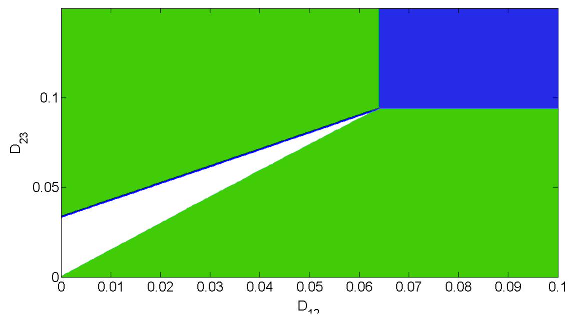

Figure 2. Example: This figure denotes the regime of distortions wherein the CMS scheme achieves the complete rate region and the minimum sum rate. Here $D_1 = 0.1, D_2 = 0.15$ and $D_3 = 0.2$. The blue points correspond to the region of distortions wherein the CMS scheme achieves the complete rate-region and the green points represent the region where the CMS scheme achieves the minimum sum rate .

It is interesting to observe that the optimum encoding scheme introduces common codewords (creates an interaction) between descriptions 1 and 3, even though these two descriptions are never received simultaneously at the decoder. While common codewords typically imply redundancy in the system, in this case, introducing them allows for better co-ordination between the descriptions leading to a smaller rate for the common branch. Similar principles can be used to show that CMS achieves the complete RD region for the $L-$channel quadratic Gaussian MD problem, for several distortion regimes.

Example 1. We consider an asymmetric setting where $D_1 = 0.1, D_2 = 0.15$ and $D_3 = 0.2$. Fig. 2 shows the regime of distortions where CMS achieves the complete rate-region and minimum sum rate. The blue region corresponds to the set of distortion pairs (D_{12}, D_{23}) wherein the CMS rate-region is complete. The green region denotes the minimum sum rate points. It is clearly evident from the figure that CMS achieves the minimum sum rate for a fairly large regime of distortions.

VI. Conclusion

In this paper, we showed that CMS achieves a strictly larger region compared to VKG for a general class of sources and distortion measures, which includes the quintessential setting of Gaussian source under mean squared error. As a consequence, it follows that the 'correlated quantization' scheme (an extreme special case of VKG), is strictly suboptimal in general. We also showed that CMS achieves the complete rate region for the 3-description symmetric cross-section and several asymmetric cross-sections of the setup shown in Fig. 1.

References

[1] K. Viswanatha, E. Akyol, and K. Rose. Combinatorial message sharing for a refined multiple descriptions achievable region. In *Proc. of IEEE ISIT*, pages 1312 –1316, Aug 2011.

[2] R. Venkataramani, G. Kramer, and V. Goyal. Multiple descriptions coding with many channels. *IEEE Trans. on Information Theory*, 49(9):2106–2114, Sep 2003.

[3] K. Viswanatha, E. Akyol, and K. Rose. A strictly improved achievable region for multiple descriptions using combinatorial message sharing. In *Proc. of IEEE ITW*, pages 100 –104, Oct 2011.

[4] A. Gamal and T.M. Cover. Achievable rates for multiple descriptions. *IEEE Trans. on Information Theory*, IT-28:851–857, Nov 1982.

[5] Z. Zhang and T. Berger. New results in binary multiple descriptions. *IEEE Trans. on Information Theory*, IT-33:502–521, Jul 1987.

[6] S.S. Pradhan, R. Puri, and K. Ramchandran. n-channel symmetric multiple descriptions - part II: An achievable rate-distortion region. *IEEE Trans. on Information Theory*, 51(4):1377 – 1392, Apr 2004.

[7] E. Akyol, K. Viswanatha, and K. Rose. On random binning versus conditional codebook methods in multiple descriptions coding. In *Proc. IEEE ITW*, Sep 2012.

[8] L. Ozarow. On a source-coding problem with two channels and three receivers. *Bell Systems Tech. Journal*, 59(10):1909–1921, Dec 1980.

[9] J. Chen. Rate region of gaussian multiple description coding with individual and central distortion constraints. *IEEE Trans. on Information Theory*, 55(9):3991 –4005, Sep 2009.

[10] Hua Wang and P. Viswanath. Vector gaussian multiple description with two levels of receivers. *IEEE Trans. on Information Theory*, 55(1):401 –410, Jan 2009.

[11] A.D. Wyner. The rate-distortion function for source coding with side information at the decoder ii : General sources. *Information and Controls*, 38:60–80, Jul 1978.

[12] J. Thomas. Feedback can at most double gaussian multiple access channel capacity (Corresp.). *IEEE Trans. on Information Theory*, 33(5):711 – 716, Sep 1987.

[13] R. Zamir. Gaussian codes and shannon bounds for multiple descriptions. *IEEE Trans. on Information Theory*, 45(7):2629 –2636, nov 1999.

[14] W.H.R. Equitz and T.M. Cover. Successive refinement of information. *IEEE Trans. on Information Theory*, 37(2):269 –275, Mar 1991.