Improved Achievable Regions in Networked Scalable Coding Problems

Emrah Akyol
Electrical & Computer Eng.
Binghamton-SUNY, US,
eakyol@binghamton.edu

Urbashi Mitra
Electrical Eng.
USC, CA, US
ubli@usc.edu

Ertem Tuncel
Electrical Eng.
UC Riverside, CA, US
ertem@ee.ucr.edu

Kenneth Rose
Electrical & Computer Eng.
UC Santa Barbara, CA, US
rose@ece.ucsb.edu

***Abstract*—In this paper, we present new results on the achievable rate-distortion regions in networked scalable compression problems, based on a flexible codebook generation and binning method. First, we consider the problem of scalable coding in the presence of decoder side information, for which the prior work analyzed the two important cases the degraded side information where source X and the side information variables (Y_1, Y_2) form a Markov chain in the order of either $X - Y_1 - Y_2$ or $X - Y_2 - Y_1$. First, we present an example non-Markov side information scenario where the proposed coding strategy achieves a strictly larger rate-distortion region compared to prior work. We then consider the problem of multi-user successive refinement where different users that are connected to a central server via links with different noiseless capacities strive to reconstruct the source in a progressive fashion. It is shown that a prior rate-distortion region is suboptimal in general, albeit its optimality for a Gaussian source with MSE distortion, and the proposed coding scheme achieves points beyond the achievable region of prior work.**

I. Introduction

In this paper, we present improved achievable regions associated with two networked scalable coding (SC) problems. Our first result pertains to a decoder side information setting, depicted in Figure 1, where the base decoder reconstructs the source, X, at distortion D_1, with the help of side information Y_1. Similarly, the refinement decoder, with the help of Y_2, reconstructs X at distortion D_2.

Steinberg and Merhav in [1] solved this problem when $X - Y_2 - Y_1$ forms a Markov chain in this order. The main mathematical coding tools in [1] are conditional codebook encoding for SC part of the problem, as done in similar problems in the absence of a side information [2], [3], used in conjunction with binning to utilize the decoder side information [4]. Tian and Diggavi studied the dual of this problem in [5], where the base decoder has the better side information *i.e.,* $X - Y_1 - Y_2$. Authors introduced a new coding method, *i.e.,* nested binning a common codebook whose codeword is decoded at both decoders. They showed that this coding scheme achieves the complete rate-distortion region in the important case of jointly Gaussian X, Y_1, Y_2 and mean squared error distortion.

It is well understood that SC can be viewed as a special case of the more general multiple description coding (MDC) problem. Hence, the coding tools developed for MDC can be employed for SC problems. Motivated by this fact, in [6], we proposed a variation of a coding method originally developed for the $L > 2$ channel MDC problem in [7] (binned combinatorial message sharing (CMS)) for the aforementioned SC problems with decoder side informations. For $L > 2$ channel MDC, both CMS (without binning) coding [8] and CMS with binning [7] improve the achievable region in [9] which was then the benchmark for this problem. More recently, in [10], Shirani and Pradhan further enhanced the achievable region of $L > 2$ channel MDC problem utilizing coset codes, among other mathematical tools, within the binned CMS framework of [7].

The second problem we consider here is multi-user successive refinement formulated in [11] as a "unique interplay between the concepts of multi-resolution source coding and source coding with side information". In this problem, there

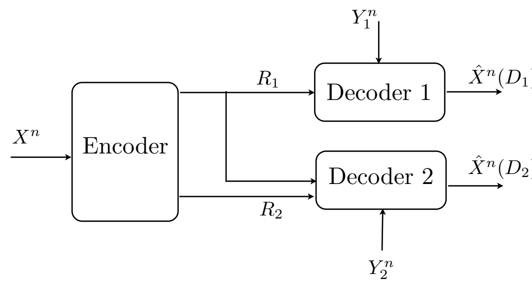

Fig. 1. Problem-1 setup: scalable coding with decoder side information. The case where $X - Y_2 - Y_1$ is denoted as SRWZ setting while $X - Y_1 - Y_2$ is called the SISC setting.

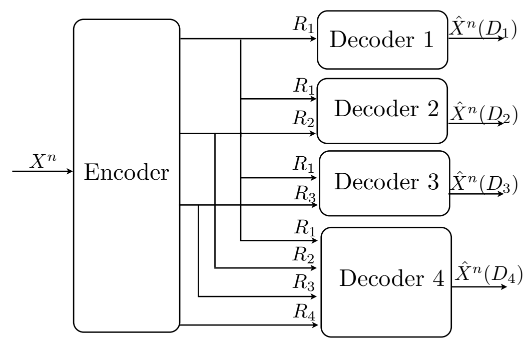

Fig. 2. Problem-2 setup: Multiuser successive refinement (MSR).

are two users (strong and weak) who strive to reconstruct the source in a progressive manner. The problem is depicted in Figure 2, where the first and second receivers decode the base layers of the weak and strong users, respectively. The third decoder receives the base and refinement layers for the weak user, and finally the fourth decoder receives all available bitstreams as well as a refinement layer for the strong user. This problem is studied in detail in [12] where the rate-distortion region for a Gaussian source under MSE distortion is derived.

In this paper, building on our earlier work in [6], we show that the binned CMS method achieves points outside the known regions for two example problems. Here, we reiterate that the methods we use as benchmarks are designed for the Gaussian-MSE scenarios for which they achieve the complete region. Nevertheless, for these examples, they are the only available benchmark coding strategies. Contributions in this paper are summarized as follows:

- We consider the SC problem with side informations $Y_i = [Z_i, W_i]$ for $i = 1, 2$ where $X - Z_1 - Z_2$ and $X - W_2 - W_1$ form Markov chains. The key idea here is that while a coding scheme based on a single nested binning would treat side informations as Y_i and cannot utilize the Markov dependence within Z_i and W_i variables in the encoding stage, binned CMS would generate a codebook for every combination of side informations W_i and Z_i and bin them accordingly, and hence fully utilize this overall (*i.e.*, in Y) non-Markovian dependence, in contrast with nested binning a common codebook which is proven to be sufficient for Markov side information settings.
- We show that in the MSR problem, binned CMS achieves points outside the achievable region reported in [12]. This result is due to the subset common codewords in the CMS codebook structure, paralleling the improvements in the $L > 2$ channel MDC problem, as discussed in [7], [8], [10].

II. Preliminaries

A. Notation

Let $\{X_t\}_{t=1}^{\infty}$, $X_t \in \mathcal{X}$, be a discrete memoryless source (DMS) with generic distribution $P(X)$. The vector $[X(1), X(2), ..., X(n)]$ is compactly denoted by x^n. Let $\mathcal{Z}$ denote the reproduction alphabet. We employ $H(X)$ to denote the entropy of a discrete random variable X, or differential entropy if X is continuous. For an arbitrary set $\mathcal{A}$, we use $2^{\mathcal{A}}$ to denote the set of all subsets of $\mathcal{A}$, *i.e.,*

$$2^{\mathcal{A}} \triangleq \{\mathcal{S} : \mathcal{S} \subseteq \mathcal{A}\}.$$

Assume a single-letter, bounded, and additive distortion measure $d : \mathcal{X} \times \mathcal{Z} \longrightarrow [0, \infty)$, *i.e.,*

$$d(x^n, z^n) = \frac{1}{n} \sum_{t=1}^{n} d(x_t, z_t) \ . \tag{1}$$

A scalable block code pair (f_1, f_2, g_1, g_2) consists of an encoding function

$$\begin{aligned} f_1 &: \mathcal{X}^n \longrightarrow \mathcal{M}_1 \\ f_2 &: \mathcal{X}^n \longrightarrow \mathcal{M}_2 \end{aligned}$$

and decoders

$$\begin{aligned} g_1 &: \mathcal{M}_1 \times Y_1^n \longrightarrow \mathcal{Z}^n \\ g_2 &: \mathcal{M}_1 \times \mathcal{M}_2 \times Y_2^n \longrightarrow \mathcal{Z}^n \ . \end{aligned}$$

A quadruple (R_1, R_2, D_1, D_2) is called *achievable* if for every $\delta > 0$ and sufficiently large n, there exists a block code (f_1, f_2, g_1, g_2) such that

$$\begin{aligned} \frac{1}{n} \log |\mathcal{M}_1| &\leq R_1 + \delta \\ \frac{1}{n} \log |\mathcal{M}_1||\mathcal{M}_2| &\leq R_1 + R_2 + \delta \\ \mathbb{E}\{d(X^n, g_1(f_1(X^n), Y_1^n))\} &\leq D_1 + \delta \\ \mathbb{E}\{d(X^n, g_2(f_1(X^n), f_2(X^n), Y_2^n))\} &\leq D_2 + \delta \ . \end{aligned}$$

We omit similar formal definitions of the coding system for the MSR problem to avoid repetition, see e.g., [12].

B. Prior Work on Scalable Coding with Decoder Side Information

Steinberg and Merhav studied the case of degraded side information in the order of $X - Y_2 - Y_1$. The encoding scheme is intuitive: generate a codebook $\mathcal{C}_1$ with marginal distribution of U_1 and then conditionally generate a codebook $\mathcal{C}_2$ for each codeword u_1^n with the conditional density $P(U_2|U_1^n = u_1^n)$. Next, bin $\mathcal{C}_1$ so that the codeword u_1^n can be decoded with the help of side information Y_1. Next, bin all the conditional codebooks C_2 so that the codewords u_1^n and u_2^n can both be decoded at the decoder with the help of the better side information Y_2. We reproduce the following rate-distortion region achievable by this scheme.

Theorem 1 ([1]). *An achievable region for this setting, $\mathcal{RD}_{SRWZ}$ is the convex hull of quadruples (R_1, R_2, D_1, D_2) for*

$$\begin{aligned} R_1 &\geq I(X; U_1|Y_1) \\ R_1 + R_2 &\geq I(X; U_2|U_1, Y_2) + I(X; U_1|Y_1) \end{aligned}$$

for a conditional distribution $P(U_1, U_2|X)$ and deterministic decoding functions g_1, g_2 which satisfy

$$D_i \geq \mathbb{E}\{d_i(X, g_i(U_i, Y_i))\} \quad i = 1, 2$$

and the Markov chain $(U_1, U_2) - X - Y_2 - Y_1$.

Tian and Diggavi studied the dual of this problem in the sense that $X - Y_1 - Y_2$ forms a Markov chain. The key coding method here is the nested binning of a common codebook, say $\mathcal{C}_0$ generated with the marginal distribution of U_0 whose codewords, u_0^n, are placed into two kinds of bins: coarse and fine. There are many coarse bins in a codebook, many (and nearly the same number of) fine bins in each coarse bin, and many (and approximately the same number) codewords in

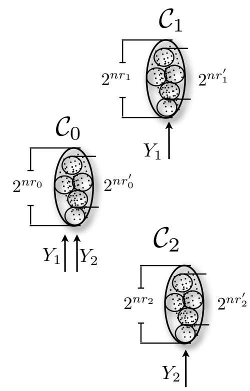

Fig. 3. The overview of the encoding scheme in [6], also used in problem 1.

each fine bin. Instead of directly sending the codeword index, the encoder sends the index of coarse bin which contains the codeword to the base decoder. Upon receiving the index of the coarse bin, the base decoder picks the codeword from coarse bin that is jointly typical with side information sequence y_1^n. The refinement decoder, which has a weaker side information sequence due to Markov chain $X - Y_1 - Y_2$, cannot find a unique u_0^n in the coarse bin. Hence, the encoder sends the index of the fine bin within the coarse bin which contains u_0^n as the refinement layer. The codeword u_0^n is decoded by finding the unique codeword in the fine bin which is jointly typical with y_2^n. Depending on the distortion requirements, the encoder also sends an additional codeword from $\mathcal{C}_1$ or $\mathcal{C}_2$ conditionally generated for each u_0^n with the conditional density $P(U_i|U_0^n = u_0^n)$ for $i = 1, 2$ depending on whether $D_1 \lesseqgtr D_2$, to be binned with respect to side information of and decoded only at one of the decoders.

We note that nested binning a common codebook, as described above, achieves the entire rate-distortion region for Gaussian-MSE setting not only for the two-decoder setting, but for any number of decoders as shown in [5]. The R-D region for two decoder case is reproduced in the following theorem.

Theorem 2 ([5]). *An achievable region for this setting, $\mathcal{RD}_{SISC}$ is the convex hull of quadruples (R_1, R_2, D_1, D_2) for*

$$R_1 \geq I(X; U_0, U_1|Y_1)$$
$$R_1 + R_2 \geq I(X; U_0, U_2|Y_2) + I(X; U_1|Y_1, U_0)$$

for a conditional distribution $P(U_0, U_1, U_2|X)$ and deterministic decoding functions g_1, g_2 which satisfy

$$D_i \geq \mathbb{E}\{d_i(X, g_i(U_i, Y_i))\} \quad i = 1, 2$$

and the Markov chain $(U_0, U_1, U_2) - X - Y_1 - Y_2$.

In [6], we introduced a coding strategy for both of the aforementioned problems, solely based on binning as follows. The encoder generates a codebook for each combination of side informations, as shown in Figure 3, which corresponds to three codebooks $\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2$ with marginal distributions of auxiliary random variables U_0, U_1 and U_2 independently. The encoding is done as follows: Given a source-word x^n, the encoder find codewords u_0^n, u_1^n, u_2^n that are jointly typical with x^n. To guarantee the existence of such jointly typical codewords, there is a set of conditions on coding rates from mutual covering lemma [13]. The codeword from $\mathcal{C}_0$, i.e., u_0^n is decoded with both y_1^n and y_2^n, while u_1^n and u_2^n is be decoded with y_1^n and y_2^n. All codebooks are binned so that at each decoder, given the bin indices, the decoder can find a unique codeword tuple jointly typical with their respective side informations in each bins. This dictates another set of conditions that involve binning and codebook rates. Application of Fourier-Moltzkin elimination over these set of equations, in conjunction with rate transfer arguments [14] yields rate distortion regions identical those of the prior work, as shown in [6].

Theorem 3 ([6]). *Both $\mathcal{RD}_{SISC}$ and $\mathcal{RD}_{SRWZ}$ can be obtained via binned CMS encoding.*

Remark 1. *In [6], the region obtained by binned CMS is reported to be possibly larger due to the weaker Markov chain conditions. At the time of publication, we had left open the question of strict improvement, due to this weaker Markov chain conditions. In fact, we can now show that this potential degree of freedom in Markov chains does not provide any improvement in the R-D region. We defer this discussion to the extended version of this paper [15] due to space constraints.*

C. Prior work on two descriptions problem

The first two descriptions coding scheme was proposed in the seminal paper by El Gamal and Cover [16], where the encoder generates two random codebooks according to $P(U_1)$ and $P(U_2)$. On observing a typical sourceword x^n, the encoder finds one codeword from each codebook that are jointly typical with x^n. The encoder has to generate sufficient number of codewords to guarantee that it can find a jointly typical of the codewords u_1^n, u_2^n with the source-word x^n, as well as it has to guarantee that u_1^n and u_2^n are jointly typical.

The **EGC Region** region, denoted by $\mathcal{R}_{EGC}$, is the convex closure of all rate-distortion tuples satisfying:

$$D_{\mathcal{A}} \geq \mathbb{E}\{d(X, g_{\mathcal{A}}(U_{2_1^{\mathcal{A}}}))\}, \quad \mathcal{A} = \{1, 2, 12\}$$
$$R_i \geq I(X; U_i), \quad i = 1, 2$$
$$R_1 + R_2 \geq I(X; U_1, U_2) + I(U_1, U_2)$$

over $P(X, U_1, U_2)$. We note that the original region in [16] includes a refinement layer codeword, and does not involve the decoding functions $g_{\mathcal{A}}(\cdot)$ as above (although a precursor does, see *e.g.*, [17]), however as demonstrated in [18] the refinement layer can be removed for this setting without any performance degradation, and we use the decoding functions $g_{\mathcal{A}}(\cdot)$ (which do not improve the rate region for this problem) to be consistent with other regions throughout the paper.

Zhang and Berger [17] considered the addition of a common codeword which is sent in both of the descriptions within the

EGC coding scheme. The **ZB Region**, denoted by $\mathcal{R}_{ZB}$, is the convex closure of all rate-distortion tuples satisfying

$$D_{\mathcal{A}} \geq \mathbb{E}\{d(X, g_{\mathcal{A}}(U_{2_1^{\mathcal{A}}}))\}, \quad \mathcal{A} = \{1, 2, 12\}$$
$$R_i \geq I(X; U_{12}, U_i) \quad i = 1, 2$$
$$R_1 + R_2 \geq 2I(X; U_{12}) + I(X; U_1, U_2|U_{12}) + I(U_1; U_2|U_{12})$$

over $P(X, U_1, U_2, U_{12})$. Clearly, $\mathcal{R}_{ZB}$ subsumes $\mathcal{R}_{EGC}$ by construction, hence the critical question is whether $\mathcal{R}_{ZB}$ includes points outside $\mathcal{R}_{EGC}$. Zhang and Berger showed that for binary symmetric source (BSS) under the Hamming distortion measure, the answer is affirmative, as reproduced in the following corollary.

Corollary 1 ([17]). *$\mathcal{R}_{EGC} \subset \mathcal{R}_{ZB}$. Particularly, for a BSS under Hamming distortion, the rate-distortion vector $RD^* \triangleq [R_1, R_2, D_1, D_2, D_{12}] = [0.629, 0.629, 0.11, 0.11, 0] \in \mathcal{R}_{ZB}$, but $RD^* \notin \mathcal{R}_{EGC}$.*

III. Result-I: Scalable Coding with Non-Markov Side Information

In this section, we consider the problem of scalable coding with decoder side information as shown in Figure 1. Here, we note that when X, Y_1, Y_2 are jointly Gaussian (not necessarily Markov), there exists a jointly Gaussian X, Y_1', Y_2' triple with $P(XY_1') = P(XY_1)$ and $P(XY_2') = P(XY_2)$ that forms a Markov chain in the order of $X - Y_1' - Y_2'$ or $X - Y_2' - Y_1'$. Since all the relevant rate and distortion expressions only depend on the pairwise marginals $P(XY_1)$ and $P(XY_2)$, but not $P(X, Y_1, Y_2)$, one can equivalently consider X, Y_1', Y_2' instead of X, Y_1, Y_2 in the problem formulation. However, this special consideration is unique to jointly Gaussian variables e.g., for binary variables the same stochastic degradedness argument does not hold. Motivated by the fact that most practical sources are indeed non-Markov processes, but involve Markov components, e.g., a hidden Markov processes is not Markov in itself but involves an underlying Markov component, we consider stochastic processes that are composed of Markov components but overall themselves are not Markov. Particularly, we take $Y_i = [Z_i, W_i]$ for $i = 1, 2$ where $X - Z_1 - Z_2$ and $X - W_2 - W_1$ to show that nested binning a single common codebook is not sufficient for optimality as follows.

The encoder will generate a codebook $\mathcal{C}_0$, whose codeword u_0^n is decoded at both decoders, with the help of Z_i, W_i at decoder i for $i = 1, 2$ via nested binning. Since it cannot utilize W_i and Z_i individually, let us assume that it generates the $\mathcal{C}_1$ and $\mathcal{C}_2$ codebooks conditioned on u_0^n so that u_1^n can be decoded with w_1^n, z_1^n at the base decoder and u_2^n can be decoded with w_2^n, z_2^n at the refinement decoder. This requires that any u_1^n codeword to be also jointly typical with the associated w_1^n and similarly u_2^n to be also jointly typical with z_2^n to be decodable at their respective decoders. Due to Markov chains, $X - Z_1 - Z_2$ and $X - W_2 - W_1$ these codewords can also be decoded at both decoders, and hence they are actually part of u_0^n. Hence, the benchmark encoder will not bin the index of the $\mathcal{C}_1$ and $\mathcal{C}_2$ with respect to any side information. With

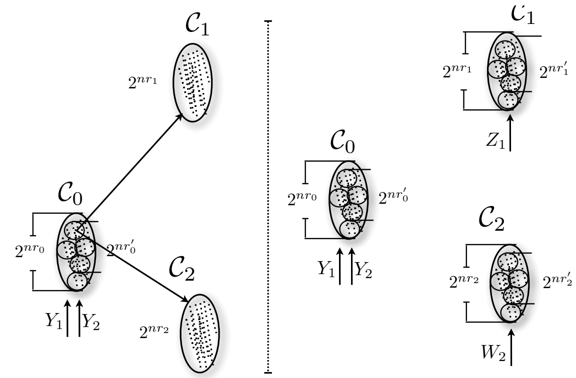

Fig. 4. Benchmark coding strategy (left) and binned CMS (right) for problem 1.

these observations at hand, we present the following region for the benchmark encoder.

Theorem 4. *An achievable region via this coding scheme for this problem, $\mathcal{RD}_{NM}$ is the convex hull of rate-distortion vectors $(R_1, R_2, D_1, D_2, D_3)$ for*

$$R_1 \geq I(X; U_0|W_1, Z_1) + I(X; U_1|U_0)$$
$$R_1 + R_2 \geq I(X; U_0|W_2, Z_2) + I(X; U_1|U_0) + I(X; U_2|U_0)$$

for a conditional distribution $P(U_0, U_1, U_2|X)$ and deterministic decoding functions g_1, g_2 which satisfy

$$D_i \geq \mathbb{E}\{d_i(X, g_i(U_i, W_i, Z_i))\} \quad i = 1, 2$$

Proof. We provide a sketch of the proof here, and defer the details to the full paper [15]. The encoder generates a common codebook $\mathcal{C}_0$ with the marginal distribution of U_0 (with the assumption of $I(U_0; W_1, Z_1) \geq I(U_0; W_2, Z_2)$). The common codeword, u_0^n is decoded at both decoders via nested binning. The u_1^n and u_2^n are decoded only at their respective decoders as refinement, without using any side information which results in the region above. □

Here, the proposed binned CMS coding scheme generates and bins a codebook for each combination of side informations. For the example scenario, this implies that there exists a common codeword u_0^n to be decoded at both decoders, another codeword u_1^n to to be decoded with the help of z_1^n, and another one u_2^n only to be decoded at the second decoder with w_2^n. This way, the encoder can fully utilize the Markov dependencies as done in degraded side information scenarios described earlier. The coding methods are depicted on Figure 4. The achievable region by this encoding scheme is the following.

Theorem 5. *An achievable region $\mathcal{RD}_{NM}^*$ is the convex hull of rate-distortion vectors $(R_1, R_2, D_1, D_2, D_3)$ for*

$$R_1 \geq I(X; U_0|W_1, Z_1) + I(X; U_1|U_0, Z_1)$$
$$R_1 + R_2 \geq I(X; U_0|W_2, Z_2) + I(X; U_1|U_0) + I(X; U_2|U_0, W_2)$$

for a conditional distribution $P(U_0, U_1, U_2|X)$ and deterministic decoding functions g_1, g_2 which satisfy

$$D_i \geq \mathbb{E}\{d_i(X, g_i(U_i, W_i, Z_i))\} \quad i = 1, 2$$

Proof. We again provide a sketch of the proof here, and defer the details to [15]. The encoder generates a common codebook $\mathcal{C}_0$ in an identical manner to the encoding scheme of $\mathcal{RD}_{NM}$. However, here u_1^n and u_2^n are decoded with the help of z_1^n and w_2^n at the base and refinement decoders respectively since the encoder can generates a codebook for each combination of side informations. The coding and binning rates are determined via mutual civering and decoding conditions, as done in [6], and via standard manipulations and rate transfer arguments yield $\mathcal{RD}^*_{NM}$. □

Corollary 2. $\mathcal{RD}_{NM} \subset \mathcal{RD}^*_{NM}$.

Proof. We note that distortion and rate expressions are identical for both regions except four terms which yield: $I(X;U_1|U_0) - I(X;U_1|U_0,Z_1) = I(U_1;Z_1|U_0) \geq 0$ and $I(X;U_2|U_0) - I(X;U_2|U_0,W_2) = I(W_2;U_2|U_0) \geq 0$, hence for the same D_1 and D_2 values, $\mathcal{RD}^*_{NM}$ achieves lower rates than those of $\mathcal{RD}_{NM}$. □

IV. Result-II: An Improved R-D Region for Multiuser Successive Refinement

In [12], the following R-D region, denoted here as $\mathcal{RD}_{MSR}$ is obtained for the MSR problem, using a coding scheme inspired by the EGC two description coding.

Theorem 6 ([12]). *An achievable region for the MSR, $\mathcal{RD}_{MSR}$ is the convex hull of rate-distortion vectors $(R_1,R_2,R_3,R_4,D_1,D_2,D_3,D_4)$ for*

$$R_1 \geq I(X;U_1),\ \sum_{i=1}^{2} R_i \geq I(X;U_{\mathcal{A}_2}),\ \sum_{i=1,3} R_i \geq I(X;U_{\mathcal{A}_3}),$$
$$\sum_{i=1}^{3} R_i \geq I(X;U_{\mathcal{A}_3},U_2) + I(U_2;U_3|U_1)$$
$$\sum_{i=1}^{4} R_i \geq I(X;U_{\mathcal{A}_4}) + I(U_2;U_3|U_1)$$

for a conditional distribution $P(U_1,U_2,U_3,U_4|X)$ and deterministic decoding functions g_i which satisfy

$$D_i \geq \mathbb{E}\{d_i(X,g_i(U_{\mathcal{A}_i}))\} \quad i=1,2,3,4$$

where $\mathcal{A}_1 = \{1\}$, $\mathcal{A}_2 = \{1,2\}$, $\mathcal{A}_3 = \{1,3\}$, and $\mathcal{A}_4 = \{1,2,3,4\}$.

In the following theorem, we present the region achievable via the binned CMS coding, denoted as $\mathcal{RD}^*_{MSR}$.

Theorem 7. *An achievable region for the MSR problem, $\mathcal{RD}^*_{MSR}$ is the convex hull of rate-distortion vectors $(R_1,R_2,R_3,R_4,D_1,D_2,D_3,D_4)$ for*

$$R_1 \geq I(X;U_1)\ \sum_{i=1}^{2} R_i \geq I(X;U_{\mathcal{A}_2}),\ \sum_{i=1,3} R_i \geq I(X;U_{\mathcal{A}_3})$$
$$\sum_{i=1}^{3} R_i \geq I(X;U_{\mathcal{A}_3},U_2) + I(U_2;U_3|U_1,U_{12})$$
$$\sum_{i=1}^{4} R_i \geq I(X;U_{\mathcal{A}_4}) + I(U_2;U_3|U_1,U_{12})$$

for a conditional distribution $P(U_{\mathcal{A}_4}|X)$ and deterministic decoding functions g_i which satisfy

$$D_i \geq \mathbb{E}\{d_i(X,g_i(U_{\mathcal{A}_i}))\} \quad i=1,2,3,4$$

where $\mathcal{A}_1 = \{1\}$, $\mathcal{A}_2 = \{1,2,\{12\}\}$, $\mathcal{A}_3 = \{1,\{12\},3\}$, and $\mathcal{A}_4 = \{1,2,\{12\},3,4\}$.

Proof. The coding scheme works very similar to the coding scheme in [6]. The encoder independently generates codebooks $\mathcal{C}_0,\mathcal{C}_1,\mathcal{C}_{12},\mathcal{C}_2,\mathcal{C}_3,\mathcal{C}_4$ with the marginal distributions of auxiliary random variables $U_0,U_1,U_2,U_{12},U_3,U_4$ respectively. The codebooks sizes are selected so that the encoder can find a jointly typical codeword tuple $u_0^n,\ldots,u_4^n$ with the source-word x^n (see, e.g., mutual covering lemma [13]). Then, these codebooks are binned where the bin sizes are selected to guarantee that there is a unique codeword tuple jointly typical with the respective side informations at every decoder. The remaining part of the derivation, as in [6], follows from standard manipulations and rate transfer arguments (see e.g. [14]) whose details can be found in [15]. □

We next present our main result for this setup.

Corollary 3. $\mathcal{RD}_{MSR} \subset \mathcal{RD}^*_{MSR}$.

Proof. Here, we first note that $\mathcal{RD}_{MSR} \subseteq \mathcal{RD}^*_{MSR}$ since by simply setting $U_{12} = \Phi$ in $\mathcal{RD}^*_{MSR}$, where Φ is deterministic, we obtain $\mathcal{RD}_{MSR}$. Strict improvement, i.e., the fact that inclusion of U_{12} is not redundant can be shown via setting $R_1 = R_4 = 0$ in the problem formulation, which transforms the problem into a two description coding setting. Then, $\mathcal{RD}^*_{MSR}$ simplifies to $\mathcal{RD}_{ZB}$ due to the common codeword U_{12}, while $\mathcal{RD}_{MSR}$ simplifies to $\mathcal{RD}_{EGC}$. Due to Theorem 1, we obtain the strict improvement, i.e., for a BSS with Hamming distortion, $RD^{**} \triangleq [R_1,R_2,R_3,R_4,D_1,D_2,D_3,D_4] = [0,0.629,0.629,0,0.5,0.11,0.11,0] \in \mathcal{RD}^*_{MSR}$ but $RD^{**} \notin \mathcal{RD}_{MSR}$. □

V. Discussions

In this paper, we have presented new results on networked scalable coding problems using the coding method in [6]. We have first shown that nested binning may not be sufficient for optimality in the case of scalable coding with non-Markov side information. We have then obtained a new achievable region for the MSR problem which strictly subsumes that of [12]. Details of the proofs can be found in the expanded version of our paper [15].

REFERENCES

[1] Y. Steinberg and N. Merhav, "On successive refinement for the Wyner-Ziv problem," *IEEE Trans. on Information Theory*, vol. 50, no. 8, pp. 1636–1654, 2004.

[2] V. Koshelev, "Hierarchical coding of discrete sources," *Problemy peredachi informatsii*, vol. 16, no. 3, pp. 31–49, 1980.

[3] B. Rimoldi, "Successive refinement of information: Characterization of the achievable rates," *IEEE Trans. on Information Theory*, vol. 40, no. 1, pp. 253–259, 1994.

[4] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Information Theory*, vol. 22, no. 1, pp. 1–10, 1976.

[5] C. Tian and S. Diggavi, "Side-information scalable source coding," *IEEE Trans. on Information Theory*, vol. 54, no. 12, pp. 5591–5608, 2008.

[6] E. Akyol, U. Mitra, E. Tuncel, and K. Rose, "On scalable coding in the presence of decoder side information," in *2014 IEEE International Symposium on Information Theory*. IEEE, 2014, pp. 2052–2056.

[7] E. Akyol, K. Viswanatha, and K. Rose, "Combinatorial message sharing and random binning for multiple descriptions," in *IEEE International Symp. on Information Theory*, 2012.

[8] K. B. Viswanatha, E. Akyol, and K. Rose, "Combinatorial message sharing and a new achievable region for multiple descriptions," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 769–792, Feb 2016.

[9] R. Venkataramani, G. Kramer, and V. Goyal, "Multiple description coding with many channels," *IEEE Trans. on Information Theory*, vol. 49, no. 9, pp. 2106–2114, 2003.

[10] F. Shirani and S. S. Pradhan, "An achievable rate-distortion region for multiple descriptions source coding based on coset codes," *IEEE Transactions on Information Theory*, vol. 64, no. 5, pp. 3781–3809, 2018.

[11] S. Pradhan and K. Ramchandran, "Multiuser successive refinement," EECS Department, University of California, Berkeley, Tech. Rep. UCB/ERL M00/20, 2000. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2000/3834.html

[12] C. Tian, J. Chen, and S. N. Diggavi, "Multiuser successive refinement and multiple description coding," *IEEE Transactions on Information Theory*, vol. 54, no. 2, pp. 921–931, 2008.

[13] A. El Gamal and Y. Kim, *Network information theory*. Cambridge University Press, 2011.

[14] E. Tuncel, "The rate transfer argument in two-stage scenarios: When does it matter?" in *IEEE International Symposium on Information Theory, 2009.* IEEE, 2009, pp. 41–45.

[15] E. Akyol, U. Mitra, E. Tuncel, and K. Rose, "On Networked Scalable Source Coding," https://www.dropbox.com/s/l8zx851d56phgi1/workingdraft.pdf?dl=0, 2020, [Working draft].

[16] A. El Gamal and T. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. on Information Theory*, vol. 28, no. 6, pp. 851–857, 1982.

[17] Z. Zhang and T. Berger, "New results in binary multiple descriptions," *IEEE Trans. on Information Theory*, vol. 33, no. 4, pp. 502–521, 1987.

[18] 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 Trans. on Information Theory*, vol. 57, no. 3, pp. 1443–1456, 2011.