Distributed Source Coding in Absence of Common Components

Farhad Shirani Chaharsooghi
EECS Department
University of Michigan
Ann Arbor,USA
Email: fshirani@umich.edu

Aria Ghasemian Sahebi
EECS Department
University of Michigan
Ann Arbor,USA
Email: ariaghs@umich.edu

S. Sandeep Pradhan
EECS Department
University of Michigan
Ann Arbor,USA
Email: pradhanv@umich.edu

Abstract—We introduce a scheme for the binary one-help-one distributed source coding problem using two layers of codes. The primary code is of constant finite block-length and the secondary code has a block-length approaching infinity. The achievable rate-distortion region for this scheme is derived for the binary one-help-one problem. It is shown that the scheme achieves the common component rate-distortion region in the case when the sources have a common component, while if a common component is not present (i.e. replaced with highly correlated functions of the two inputs) it improves upon existing achievable bounds. We show that as the block-length of the primary code is increased, the transmission rate required in the scheme decreases, reaches its minimum at some finite value and then increases. This phenomenon is not typically seen in traditional schemes used in multi-terminal source coding.

I. INTRODUCTION

In the classic lossy distributed source coding problem, two distributed encoders observe the outputs of two correlated sources and communicate a compressed version of their corresponding source sequences to a joint decoder. The decoder then wishes to reconstruct a lossy version of the two source outputs to within some distortion level. The distortion is measured as the average of a single letter distortion criterion for each of the sources. This scenario has been depicted in Figure 1. The parameters of interest are the rate-pairs for which the reconstruction is possible given a distortion level. The problem of characterizing the rate-distortion region in the general case turns out to be difficult and has remained open for several decades. The main challenge is in devising a scheme to optimally employ the correlation between the two sources while not requiring the encoders to communicate with each other.

This problem has been studied in the literature extensively in several variations [1][2][4]. It was solved by Slepian and Wolf in [1] for the case when the sources are to be reconstructed losslessly at the decoder. It is shown in [1] that the resulting achievable rate region in lossless distributed source coding is the same as the lossless centralized source coding problem. However, this is not generally the case. It is shown in [3] that for the lossy version of this problem, there is always a rate-loss due to the distributed nature of the encoders.
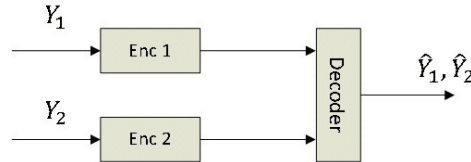
Fig. 1. General Lossy Distributed Source Coding

An achievable rate-distortion region is found in [2]. The Berger-Tung approach in [2] is to quantize the two sources independently and to bin the quantized strings in order to reduce redundancy. One of the characteristics of the Berger-Tung coding scheme is the presence of the so-called long Markov chain. Roughly speaking, it says that the single-letter distributions of the quantization noises of the two sources are independent. Recently it was pointed out in [4] that the achievable rate-distortion region in [2] can be improved for the one-help-one problem in the presence of common components i.e. when a function of one of the sources is equal to another function of the other source. The new approach uses the common component of the two sources to "break" the long Markov chain and thus yields a better performance. In the one-help-one problem, one of the two sources is to be reconstructed to within some distortion level while the other source acts only as a helper. In the coding scheme proposed in [4], in the first stage of communication, both encoders use the same quantizer to quantize the common part and transmit this quantized version to the decoder. In the second stage of the communication, since the quantized common component is known at all terminals, it is treated as side information and the approach of [2] is used to communicate the remaining parts of the two sources. It is shown in [4] that if instead of exact common components, the two sources have highly correlated components, this scheme does not lead to any gain in the achievable rate. Using this the authors in [4] provide a novel discontinuity argument and prove that the Berger-Tung region is not tight even in absence of common components.

In this paper, we address the one-help-one problem in which the two sources have highly correlated components but no common components. We note that in the absence of exact common components, quantization in [4] is done

independently by the two encoders. If we employ the same technique for the new problem, the high correlation between the components will be lost due to the independence of the two quantizers. Similar to the technique proposed in [4], one might be tempted to use the same quantizer for both sources hoping to get gains due to the correlation between the two components. In section IV we argue no matter how highly correlated the two components are, the quantization noises of the two sources approach two independent random vectors as the quantization block-length goes to infinity. However, if the block-length is kept finite, it turns out that the quantization noises remain correlated and we can benefit from their dependence in the next stage of encoding. Therefore, in our scheme we use finite length quantization in the first stage (quantizing the highly correlated variables) and then use large block-length quantizers for the second stage. An interesting implication is that in order to get gains in terms of the achievable rate, the length of the first quantization stage cannot be too small or too high; meaning that for some finite length, the scheme achieves its best performance.

Since we use finite-length quantizers, a characterization for the finite-length performance of codes is needed. The exact characterization of the rate-distortion region as a function of the quantization block-length is unknown even in the binary case; however, several upper bounds are provided in the literature for finite-length quantization rate with a constant distortion [5]. Using these results, we show that the method presented in this paper achieves a better rate-distortion region than other known results.

The rest of the paper is as follows: In section II, we present our notation along with a formal definition of the problem and some bounds we use throughout the paper. In section III, we present our scheme and derive an achievable rate-distortion region for the binary one-help-one problem. In section IV, it is shown that using the same linear quantizer with large block-length to quantize two highly correlated binary random variables yields close to independent quantization noises. In section V, simulation results for using a family of Hamming codes as our finite quantizer are given. In section VI we conclude the paper.

II. PRELIMINARIES

In this section we present a formal statement of the lossy distributed source coding problem, then we restrict this to the binary one-help-one problem which is the main example discussed in this paper. For the rest of the paper, a sequence of length n is denoted by $x(1:n)$, its ith element is denoted by $x(i)$, and the subsequence consisting of its ith element to its jth element is shown by $x(i:j)$. A two dimensional matrix of size $m \times n$ is denoted by $x(1:m,1:n)$. Random variables are shown by capital letters and their realizations are denoted by small letters. Let $\{Y_1(i)\}$ and $\{Y_2(i)\}$ be two source sequences from the alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ for the sources shown in Figure 1. Let the sources be $i.i.d$ samples of a joint PMF on $\mathcal{Y}_1 \times \mathcal{Y}_2$ given by $P_{Y_1,Y_2}(y_1,y_2)$. Let the functions $d_i : \mathcal{Y}_i \times \mathcal{Y}_i \to R_{\geq 0}, i = 1, 2$ be the distortion criteria

for the sources. Without loss of generality, the reconstruction alphabets are assumed to be the same as source alphabets. A $(2^{nR_1}, 2^{nR_2}, n)$ code consists of: 1) two encoding functions $m_i : \mathcal{Y}_i^n \to [1 : 2^{nR_i}], i = 1, 2$, and 2) A decoding function $r : [1 : 2^{nR_1}] \times [1 : 2^{nR_2}] \to \mathcal{Y}_1^n \times \mathcal{Y}_2^n$. Let $\hat{Y}_i(1 : n) = r(m_1(Y_1(1 : n)), m_2(Y_2(1 : n)))$ be the reconstruction of the two sources. A quadruple (R_1, R_2, D_1, D_2) is said to be achievable if there exists a sequence of codes $(2^{nR_1}, 2^{nR_2}, n)$ such that $\limsup_{n\to\infty} E(d(Y_i, \hat{Y}_i)) \leq D_i, i = 1, 2$.

For the binary one-help-one problem, let X, Z and E be Bernoulli random variables with $P(X = 1) = 0.5, P(Z = 1) = p, P(E = 1) = \epsilon$ where $p, \epsilon \in [0, 0.5]$. These random variables are assumed to be mutually independent. Let $Y_1 = X + E$ and $Y_2 = (X, Z)$, also $d_1(y_1, \hat{y}_1) = 0, y_1, \hat{y}_1 \in \{0, 1\}$ and $d_2(y_2, \hat{y}_2) = d_H(\hat{x} + \hat{z}, x + z)$ where $\hat{y}_2 = (\hat{x}, \hat{z}), y_2 = (x, z)$ and d_H represents the Hamming distance. Note that since the distortion function for Y_1 is always 0 irrespective of the reconstruction, the decoder is only interested in reconstructing Y_1 in hopes of achieving a lower distortion in reconstructing Y_2.

It was shown in [4] that when $\epsilon = 0$ the following rate-distortion quadruples are achievable by the scheme given in that paper but not by the Berger-Tung scheme in [2]:

$$(r_1, r_2, d_1, d_2) = (1 - h_b(\delta), h_b(p * \delta) - h_b(\delta_1), 0, \delta_1) \quad (1)$$

Also it was shown that (1) is not achievable when $\epsilon \neq 0$.

We address the one-help-one problem for the case where $\epsilon \neq 0$ and show that our scheme presented in Section III achieves a larger rate distortion region than other known schemes. According to [5], for quantizing an equiprobable binary source to Hamming distortion δ with a quantizer with block-length n, the following rate is achievable:

$$R(n, \delta) = R(\delta) + \frac{1}{2}\frac{\log n}{n} + O(\frac{1}{n})$$

where $R(\delta) = 1 - h_b(\delta)$ is the binary rate distortion function. Define $\theta_n = \frac{1}{2}\frac{\log n}{n} + O(\frac{1}{n})$. Note that $\{\theta_n\}$ is a sequence of positive numbers converging to 0.

III. THE BINARY ONE-HELP-ONE PROBLEM

A. Finite Length Quantizer Scheme

In this section, we introduce a coding scheme for the binary one-help-one problem when $\epsilon \neq 0$. The rest of this section is dedicated to proving the following theorem:

Theorem 1: For the binary one-help-one problem, the following rate-distortion region is achievable for any positive integer n.

$$R_1 \geq 1 - h_b(\delta) + \theta_n \quad (2)$$

$$R_2 \geq h_b(p * \delta) - h_b(\delta_1) \quad (3)$$

$$D_2 \leq \delta_1 * ((1 - (1 - \epsilon)^n)(\delta + \frac{\epsilon}{(1 - (1 - \epsilon)^n)} * \delta)) \quad (4)$$

where $p * \delta \geq \delta_1$, $a * b = a(b - 1) + b(a - 1)$, and $\{\theta_n\}$ is defined in section II.

Remark 1: Before we proceed to the proof of the theorem, note that this bound is continuous in ϵ and simplifies to the
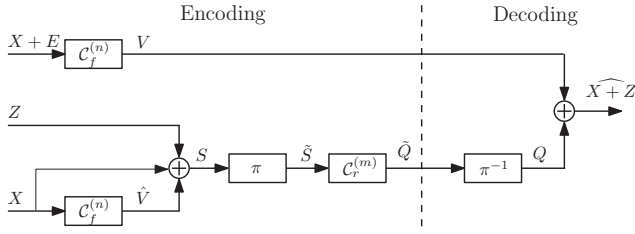
Fig. 2. A block-diagram of the scheme

one given in [4] when $\epsilon = 0$, whereas the bound in [4] is not continuous in ϵ.

Proof: To achieve (4) for some fixed n, we use two quantization codes. The first code $\mathcal{C}_f^{(n)}$ is a finite length quantizer for a binary symmetric source (BSS) with codewords of length n and average distortion δ and rate $R_f^{(n)} = 1 - h_b(\delta) + \theta_n$. The existence of such codes and bounds on θ_n are discussed in [5]. The second code $\mathcal{C}_r^{(m)}$ has codewords of length m, and it is chosen from a family of codes suitable for quantization of a Bernoulli source with parameter $p * \delta$ with average distortion δ_1. The rate of the code converges to $h_b(p * \delta) - h_b(\delta_1)$ as m tends to infinity. The existence of such codes is given by Shannon's rate distortion theorem in [6]. Let Π be the set of permutations on the set $[1 : n]$. We choose permutations $\pi_i, i \in [1 : m]$ randomly and uniformly from Π. These permutations are also made available to the second encoder and the decoder beforehand.

First we give a summary of the scheme and then we present the formal scheme and a proof of achievability. As shown in Figure 2, the first encoder uses $\mathcal{C}_f^{(n)}$ to quantize a block of length n of its input source $X + E$. It then transmits the quantized version to the decoder. The second encoder guesses the quantized codeword sent by the first encoder by quantizing the block of length n of source X into $\hat{V}$. This guess would be correct with high probability if $\epsilon \ll \frac{1}{n}$ since the expected number of bits where E is 1 in one block of length n is $n\epsilon$ which is small under this condition. Then using $\hat{V}$ it calculates the quantization noise $X + \hat{V}$ of the source X. This quantization noise is correlated with the quantization noise $X + E + V$ of the source $X + E$ observed at the first encoder. Note that in the case when $\epsilon = 0$, we have $\hat{V} = V$ and hence $X + \hat{V}$ captures the uncertainty at the decoder about X. The second encoder sends a quantized version of $X + \hat{V} + Z$ to the decoder to refine the description it received from the first encoder. When $\epsilon \neq 0$ we make the second encoder use $X + \hat{V}$ as an approximation to $X + V$ and add it to Z to get S. This process is repeated for m blocks each of length n. The encoder first uses the permutations $\pi_i, i \in [1 : m]$ described above to turn S into an i.i.d Bernoulli source $\tilde{S}$ with parameter $p * \delta$. This will be explained in more depth later. Then this source is quantized to distortion δ_1 and sent to the decoder. Let the quantized version of $\tilde{S}$ be denoted by $\tilde{Q}$. The decoder having received $\tilde{Q}$, calculates Q which is a quantized version of S. Finally the decoder declares $Q + V$ as the reconstruction of

$X + Z$.

Now we proceed to formally present the scheme and prove the theorem. The first encoder receives a string of nm bits of $Y_1 = X + E$ and breaks them into m blocks of length n. We denote each bit in this string as $X(i,j) + E(i,j), i \in [1 : m], j \in [1 : n]$ where i indicates the block containing that bit and j indicates the index of the bit in that block. The encoder uses $C_f^{(n)}$ to quantize each block of length n. In other words for $i = 1, ..., m$, it finds $V(i, 1 : n)$ such that:

$$V(i, 1:n) = argmin_{v(1:n) \in \mathcal{C}_f^{(n)}} \{d_H(Y_1(i, 1:n), v(1:n))\}.$$

Then it transmits the index corresponding to $V(i, 1 : n)$ in $C_f^{(n)}$ as the quantized version of the ith block. The rate of transmission for this encoder is $R_1 = R_f^{(n)} = 1 - h_b(\delta) + \theta_n$. Now we proceed to the second encoder. The second encoder receives nm bits from sources X and Z. It divides them into m blocks of length n as explained in the case of the first encoder. Then it quantizes each block of $X(i, 1:n)$ in the same manner as in the first encoder. Let $\hat{V}(i, 1:n)$ be the quantized codeword corresponding to $X(i, 1:n)$. The encoder computes $S(1:m, 1:n) = X(1:m, 1:n) + \hat{V}(1:m, 1:n) + Z(1:m, 1:n)$. Let $\tilde{S}(i, j) = S(i, \pi_i(j)), i \in [1 : m], j \in [1 : n]$, a permuted version of S.

Lemma 1: $\tilde{S}(1:m, j)$ is a string of i.i.d Bernoulli random variables with parameter $p * \delta$.

The proof of the lemma is given in the appendix. As an implication of this lemma, $\tilde{S}$ will be treated as a Bernoulli source with parameter $p * \delta$. The encoder quantizes each $\tilde{S}(1 : m, j)$ using the code $C_r^{(m)}$ with the usual typicality quantization criteria. Let $\tilde{Q}(1:m, j)$, be the quantized version of $\tilde{S}(1:m, j)$. The encoder transmits the index of $\tilde{Q}(1:m, j)$ in $C_r^{(m)}$ to the decoder. Let $\tilde{T}(1 : m, 1 : n) = \tilde{Q}(1 : m, 1 : n) + \tilde{S}(1:m, 1:n)$ be the quantization noise. We know that this noise becomes $i.i.d.$ Bernoulli with parameter δ_1 and $\tilde{T}(i, j)$ becomes independent of $\tilde{S}(i, j)$ as m tends to infinity. Also, define $T(i, j) = \tilde{T}(i, \pi_i^{-1}(j)), i \in [1 : m], j \in [1 : n]$. The rate of transmission in the second encoder is $R_2 = R_r^{(m)}$ which approaches $h_b(p * \delta) - h_b(\delta_1)$ as m goes to infinity.

Now we discuss the scheme for decoding the quantized version of $X + Z$ and calculate the resulting distortion. The decoder computes $Q(i, j) = \tilde{Q}(i, \pi_i^{-1}(j))$, that is the decoder undoes the permutation. Note that $E(d_H(Q(i, j), S(i, j))) = E(d_H(\tilde{Q}(i, j), \tilde{S}(i, j))) = E(w_H(T(i, j))) = \delta_1$. The decoder declares $Q(1:m, 1:n) + V(1:m, 1:n)$ as the reconstruction of the source $X + Z$. The resulting average distortion is:

$$D = \frac{1}{mn} E\{d_H((X + Z)(1:m, 1:n), (Q + V)(1:m, 1:n))\}$$

We have:

$$E\{d_H((X + Z)(1:m, 1:n), (Q + V)(1:m, 1:n))\}$$
$$= E\{w_H((X + Z + S + T + V)(1:m, 1:n))\}$$
$$= E\{w_H((\hat{V} + V + T)(1:m, 1:n))\}$$
$$\stackrel{a}{=} mn(\delta_1 * \frac{1}{mn} E\{w_H((\hat{V} + V)(1:m, 1:n))\})$$

Now we calculate $E\{w_H((\hat{V}+V)(1:m,1:n))\}$:

$$\sum_{i=1}^{m} E\{w_H((\hat{V}+V)(i,1:n))\}$$

$$\overset{b}{=} m(E\{w_H((\hat{V}+V)(1:n)|E(1:n)=0)P(E(1:n)=0)\}$$
$$+ E\{w_H((\hat{V}+V)(1:n)|E(1:n)\neq 0)P(E(1:n)\neq 0))\})$$
$$\overset{c}{=} m(E\{w_H((\hat{V}+V)(1:n)|E(1:n)\neq 0)P(E(1:n)\neq 0))$$
$$= m(1-(1-\epsilon)^n)E\{w_H((\hat{V}+V)(1:n))|E(1:n)\neq 0\}$$
$$= m(1-(1-\epsilon)^n)E\{w_H((X+\hat{V}+X+V)(1:n))|E(1:n)\neq 0\}$$
$$\leq m(1-(1-\epsilon)^n)(E\{w_H((X+\hat{V})(1:n))|E(1:n)\neq 0\}$$
$$+ E\{w_H((X+E+V+E)(1:n))|E(1:n)\neq 0\})$$
$$\overset{d}{=} m(1-(1-\epsilon)^n)(E\{w_H((X+\hat{V})(1:n))\}$$
$$+ \frac{\epsilon}{(1-(1-\epsilon)^n)} * E\{w_H((X+E+V)(1:n))\})$$
$$\overset{e}{=} mn(1-(1-\epsilon)^n)(\delta + \frac{\epsilon}{(1-(1-\epsilon)^n)} * \delta)$$
$$\to D \leq \delta_1 * ((1-(1-\epsilon)^n)(\delta + \frac{\epsilon}{(1-(1-\epsilon)^n)} * \delta))$$

Note that (a) is true since T becomes independent of all the other variables as m tends to infinity; (b) is true since each block is quantized identically and hence the expected value is equal for all blocks; (c) is correct since if $E(1:n)=0$ then $V(1:N) = \hat{V}(1:N)$ since they are both quantized versions of $X(1:n)$; (d) holds since $(X+E+V)(1:n)$ is a function of $(X+E)(1:n)$ also $(X+E)(1:n)$ is independent of $E(1:n)$ since X is Bernoulli with parameter 0.5, and finally (e) holds since the average distortion of the finite length quantizer was assumed to be δ. This completes the proof of theorem 1. ∎

B. Comparison with the Common Component Scheme

Proposition 1: The scheme presented here achieves a larger rate distortion region than the one presented in [4].

Proof: We shall prove there exists p and ϵ such that the rate-distortion region in theorem 1 strictly contains the rate-distortion region in [4]. It was shown in [4] that when $\epsilon = 0$ the Berger-Tung bound does not include the set of quadruples $(r_1, r_2, d_1, d_2) = (1-h_b(\delta), h_b(p*\delta)-h_b(\delta_1), 0, \delta_1)$ when $\delta \in (0, 0.5)$ and $\delta_1 < p*\delta$. Also it is stated that the rate region in [4] reduces to the standard Berger-Tung bound for $\epsilon \neq 0$ since there is no common component between Y_1 and Y_2 in that case. Since the Berger-Tung scheme must perform worse when $\epsilon \neq 0$ as compared to the case when $\epsilon = 0$, we infer that it cannot achieve $(1-h_b(\delta), h_b(p*\delta)-h_b(\delta_1), 0, \delta_1)$ when $\epsilon \neq 0$. This means that for a given δ and δ_1 there exists a radius $\gamma > 0$ for which no quadruple in the set $B((r_1, r_2, d_1, d_2), \gamma) = \{(R_1, R_2, 0, D_2) : d_E((R_1, R_2, D_2), (r_1, r_2, d_2)) \leq \gamma\}$ is achievable by the scheme in [4]. Note that d_E is just the Euclidean distance in the three dimensional space. Note that for a given ϵ and n we showed that $(r'_1, r'_2, 0, d'_2) = (1-h_b(\delta)+\theta_n, h_b(p*\delta)-h_b(\delta_1), 0, \delta_1*((1-(1-\epsilon)^n)(\delta + \frac{\epsilon}{(1-(1-\epsilon)^n)}*\delta))$

is achievable by our scheme. We have:

$$d_E((r'_1, r'_2, d'_2), (r_1, r_2, d_2)) = \qquad (5)$$
$$\sqrt{\theta_n^2 + (\delta_1 * ((1-(1-\epsilon)^n)(\delta + \frac{\epsilon}{(1-(1-\epsilon)^n)}\delta)) - \delta_1)^2}$$

Since θ_n is converging to 0, one can take n to be large enough so that θ_n is less than $\frac{\gamma}{2}$. Since (5) is a continuous function of ϵ which is less than $\frac{\gamma}{2}$ as ϵ goes to 0, there exists non-negative ϵ for which $(r'_1, r'_2, 0, d'_2) \in B((r_1, r_2, 0, d_2)$ for n described as above. Hence the point $(r'_1, r'_2, 0, d'_2)$ is achievable by the scheme purposed here while it is not achievable by [4]. This shows that the rate-distortion region in theorem 1 strictly contains the one in [4] for non-zero ϵ. ∎

IV. LARGE BLOCK-LENGTH QUANTIZATION OF BINARY VARIABLES

In this section, we show that minimum distance quantization of two highly correlated BSS's using linear codes results in quantization noises that behave similar to independent random variables. We use linear codes here since the congruence of the Voronoi regions of a linear code facilitates our analysis. Consider two strings of binary random variables $\{x_i\}$ and $\{x_i + e_i\}$ generated by sources X and $X+E$ described in section II. We will choose a family of linear codes $\{\mathcal{C}_{\mathcal{G}}^{(n)}\}$ randomly and uniformly from all possible linear codes such that the generator matrix $\mathcal{G}$ is non-singular and their block-length is n and their rate is $\frac{1}{n}[n(1-h_b(\delta))]$ where $[a]$ is the largest integer smaller than a. It is well known that as n approaches infinity the average distortion of a randomly chosen linear code with rate close to $1-h_b(\delta)$ approaches δ for the binary symmetric source with Hamming distortion (as an example one can refer to [7]). We show that as n goes to infinity for any fixed ϵ, the average Hamming distance between the quantization noises approaches the ϵ-vicinity of $\delta*\delta*\epsilon$. Let $Q(x(1:n)) = argmin_{c \in \mathcal{C}_{\mathcal{G}}^n}(d_H(x(1:n), c))$ be the quantized version of $x(1:n)$ and let $s(1:n) = Q(x(1:n)) + x(1:n)$ and $t(1:n) = Q((x+e)(1:n)) + (x+e)(1:n)$ be the quantization noises. Let the Voronoi region for the 0 codeword be P_0. It is relatively straightforward to show that:

$$p(S(1:n) = s(1:n), T(1:n) = t(1:n)) \to \qquad (6)$$
$$\frac{1}{2^{nh_b(\delta)}}P((s+t+E)(1:n) \in \mathcal{C}_{\mathcal{G}}^{(n)}), (s,t)(1:n) \in P_0 \times P_0$$

as n goes to infinity. The proof is omitted because of space limitations. Then we have:

$$E(w_H((S+T)(1:n))) \to \qquad (7)$$
$$= \frac{1}{2^{nh_b(\delta)}} \sum_{(s,t) \in P_0 \times P_0, c \in \mathcal{C}_{\mathcal{G}}^{(n)}} w_H((s+t)(1:n))P(E(1:n) = (s+t+c)(1:n))$$

Define $u = t+c$, then $Q(u(1:n)) = c(1:n)$ and $(u+Q(u))(1:n) = t(1:n)$. (7) becomes:

$$\frac{1}{2^{nh_b(\delta)}} \sum_{t \in P_0, u(1:n) \in \{0,1\}^n}$$
$$w_H((s+u)(1:n) + Q(u(1:n)))P(E(1:n) = (s+u)(1:n))$$

Simplifying the above it can be shown that Equations (6) and (7) yield (as n tends to infinity):

$$\Big|\frac{1}{2^{nh_b(\delta)}}\sum_{t\in P_0, u(1:n)\in\{0,1\}^n} w_H(Q((s+u)(1:n)))P(E(1:n)=u(1:n))$$
$$- E(w_H((S+T)(1:n)))\Big| \leq n\epsilon \quad (8)$$

Using (6) and (8) we get:

$$\Big|E(w_H(Q((S+E)(1:n)))) - E(w_H((S+T)(1:n)))\Big| \leq n\epsilon$$

First, note that $(S+E)(1:n)$ has average weight $n(\delta*\epsilon)$ as n goes to infinity. One can derive the probability distribution of $Q((S+E)(1:n))$ using standard random coding arguments to get $P(Q(s+e)(1:n) = q(1:n)) \to \frac{1}{2^{nh_b(\delta)}}, q(1:n) \in B(s+e,n\delta)$. Note that here the randomness is coming from the choice of $\mathcal{C}_{\mathcal{G}}^{(n)}$. Taking the average weight of $Q((S+E)(1:n))$ with this distribution gives $\delta*\delta*\epsilon$. Therefor if we were to use the scheme in section III with this large block-length quantizer, we would get quantization noise in ϵ vicinity of $\delta_1*\delta*\delta*\epsilon$ which is worse than what finite length quantizers would achieve. Also, note that the average Hamming distance between the quantization noises is not continuous with ϵ since if ϵ was 0, the distance would be 0 because both quantizers are quantizing the exact same sequence. However, if $\epsilon \neq 0$ then the distance is bigger than $\delta*\delta$.

V. SIMULATIONS FOR HAMMING CODES

In this section, we present our results for the case where the first encoder in section III uses a Hamming code as its finite block-length quantizer. Hamming codes are perfect codes of block-length $2^r - 1$ and rate $1 - \frac{r}{2^r-1}$. They have minimum distance of 3. Using (4), one can compare the performance of the scheme presented here for $\epsilon \neq 0$ with (1). As stated before, (1) contains the rate-distortion region of the binary one-help-one problem for any ϵ. Figure 3 shows the two bounds along with the time-sharing bound which is described next. One strategy in this setting is for the first encoder to transmit $X+E$ losslessly and the other encoder to send a quantized version of Z and for the decoder to add them together. Another strategy is for the second encoder to quantize $X + Z$ and transmit it while the first encoder does not send anything. The third bound in Figure 3 illustrates the bound resulting from time-sharing between these two strategies. This time-sharing strategy seems to be a good strategy for the Berger-Tung approach since if we use independent quantization for two encoders simultaneously the quantizations noises will add to each other and we will get a worse distortion than the time-sharing strategy. It can be seen from the plot that when we use finite block-length Hamming codes for quantization, we achieve better results than the time-sharing bound.

VI. CONCLUSION

We have presented a new scheme for the binary one-help-one problem which improves the existing achievable rate distortion regions. This scheme uses two layers of quantization; In the first layer, a finite block-length quantizer is used and
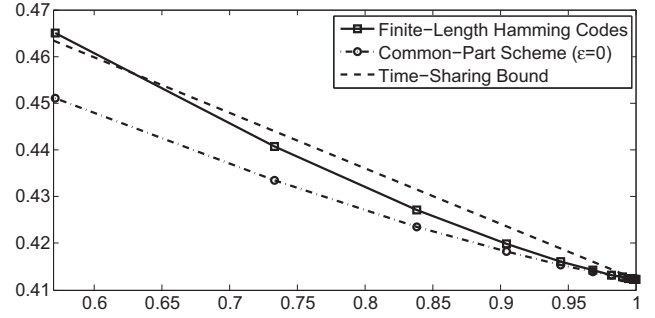

Fig. 3. Comparison between Hamming codes and the bound from [4] in the one-help-one problem when $\delta_1 = 0.1, p = 0.3, \epsilon = 10^{-10}$

the second layer is the usual typicality quantizer with block-length approaching infinity. It is shown that the optimal rate-distortion region for the new scheme is achieved for a finite block-length quantizer in the first layer and that increasing block-length of the quantizer would give strictly smaller rate distortion regions.

APPENDIX

Here we provide the proof for Lemma 1. First note that for $i \neq i'$ since $\tilde{S}(i,1:n)$ is a function of $(X(i,1:n),Z(i,1:n))$ and $\tilde{S}(i',1:n)$ is a function of $(X(i',1:n),Z(i',1:n))$, $\tilde{S}(i,1:n)$ and $\tilde{S}(i',1:n)$ are independent of each other. So we only need to prove that $S(i,j)$ are identically distributed for all i,j. We have:

$$P(\tilde{S}(i,j)=1) = P(X(i,\pi_i(j))+\hat{V}(i,\pi_i(j))+Z(i,\pi_i(j))=1)$$
$$\overset{a}{=} p * P(X(i,\pi_i(j)) + \hat{V}(i,\pi_i(j)) = 1)$$
$$\overset{b}{=} p * \frac{1}{n}\sum_{j'=1}^{n} E(w_H(X(i,j') + \hat{V}(i,j')))$$
$$\overset{c}{=} p * \delta$$

(a) is true because $Z(1:m,1:n)$ is independent of $X(1:m,1:n)$ and $X(i,\pi_i(j))+\hat{V}(i,\pi_i(j))$ is a function of $X(1:m,1:n)$, (b) is true since the choice of π_i is independent of the source sequences, and (c) is correct since the average distortion of $C_f^{(n)}$ is δ.

REFERENCES

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources", Information Theory, IEEE Transactions on , vol.19, no.4, pp. 471- 480, Jul 1973
[2] S.-Y. Tung, "Multiterminal source coding", Ph.D. Thesis, Cornell University, Ithaca, NY, 1974.
[3] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder", IEEE Trans. Inf. Theory, vol. 22, no. 1, pp. 110, 1976
[4] A.B. Wagner, B.G. Kelly and Y. Altuğ, "Distributed Rate-Distortion With Common Components," Information Theory, IEEE Transactions on , vol.57, no.7, pp.4035-4057, July 2011
[5] V. Kostina and S. Verdú, "Fixed-Length Lossy Compression in the Finite Blocklength Regime," Information Theory, IEEE Transactions on , vol.58, no.6, pp.3309-3338, June 2012
[6] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion", IRE Nat. Conv. Rec., vol. 7, no. 4, pp. 142163, Mar. 1959.
[7] T. M. Cover and J. A. Thomas, "Elements of Information Theory" Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006.