

Data Processing Lower Bounds for Scalar Lossy Source Codes with Side Information at the Decoder

Avraham Reani and Neri Merhav
Department of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 32000, Israel
Emails: [avire@tx,merhav@ee].technion.ac.il

Abstract—In this paper, we derive lower bounds on the distortion of scalar fixed-rate codes for lossy compression with side information available at the receiver. These bounds are derived by presenting the relevant random variables as a Markov chain and applying generalized data processing inequalities a la Ziv and Zakai.

I. INTRODUCTION

The Wyner–Ziv (WZ) problem has received very much attention during the last three decades. There were several attempts to develop practical schemes for lossy coding in the WZ setting, by using codes with certain structures that facilitate the encoding and the decoding. Most notably, these studies include nested structures of linear coset codes (in the role of bins) for discrete sources, and nested lattice structures for continuous valued sources, see e.g., [2], [3]. Other directions of introducing structure into WZ coding are associated with trellis/turbo/LDPC designs ([4] and references therein) and with progressive coding, i.e., successive refinement with layered code design [5], [6]. The case of scalar source codes for the WZ problem was also handled in several papers, e.g. [7] and [8]. Zero-delay coding strategies for the WZ problem were also introduced in [9]. In [10] and [11] it was shown that under high-resolution assumptions, the optimal quantization level density is conjectured to be periodic. In addition, zero-delay schemes for specific source-side information correlation were presented in [10], [11] and [12].

In this paper, we develop lower bounds for the distortion in the scalar WZ setting. We generalize the results of [13] and [14], concerning functionals satisfying a data-processing theorem, to this setting. In [13] it was shown that the rate-distortion (RD) bound ($R(d) \leq C$) remains true when $-\log x$, in the definition of mutual information, is replaced by an arbitrary convex, non-increasing function satisfying some technical conditions. For certain choices of the convex function, the bounds obtained were better than the classical RD bounds. These results were substantially generalized in [14] to apply to even more general information measures. The methods of [13] were used in [15], [16] and [17]. In these papers, lower bounds for the distortion of delay constrained joint source-channel coding were given. These bounds were obtained by combining the Rényi information measure [18] with the generalized data processing theorem of [13], and under high-resolution and high SNR approximations. Another

related work is [19], where certain degrees of freedom of the Ziv-Zakai generalized mutual information were exploited in order to get better bounds.

We start by presenting the relevant random variables of the WZ problem as a Markov chain. Then, using a data processing theorem, we get lower bounds on the distortion in this case. We show that replacing the logarithmic function with other functions, may give better bounds for the distortion of delay-limited coding (in particular, for scalar coding) in the WZ setting. Examples of non-trivial lower bounds for scalar coding in this setting, were obtained using the convex function $Q(t) = t^{-s}$, $s > 0$, which is equivalent to using the Rényi information measure. These results will be given in detail in Section II. The importance of such bounds stems from the fact that finding the optimal scalar code in the WZ setting is in general an hard problem. In fact, it is a problem of finding an optimal partition of the source alphabet and this partition does not necessarily correspond to intervals. A main objective will be to use these bounds for studying the performance of concrete coding schemes. In the next step, a possible direction will be to extend our setting to more general scenarios, for example, to scenarios where the output of the encoder is transmitted over a noisy channel (instead of the noiseless one in the WZ setting), and to scenarios of coding with memory.

II. PROBLEM FORMULATION AND RESULTS

In this section, we present the relevant random variables of the WZ problem as a Markov chain and establish a generalized Data Processing Theorem (DPT) for this setting, using the method of [13]. We begin with notation conventions. Capital letters represent scalar random variables (RV), specific realization of them are denoted by the corresponding lower case letters and their alphabets - by calligraphic letters. For $k \geq 1$ (k is a positive integer), x^k will denote the vector (x_1, \dots, x_k) . We consider a memoryless source producing a random sequence X_1, X_2, \dots, X_n , $X_i \in \mathcal{X}$, $i = 1, 2, \dots, n$, where \mathcal{X} is a finite alphabet with cardinality $|\mathcal{X}|$. Without loss of generality, we define this alphabet to be the set $\{1, 2, \dots, |\mathcal{X}|\}$. The probability mass function of X , $P_X(x)$, is known. The encoder maps X^n into a channel symbol Z which takes on values in the set $\{1, 2, \dots, M\}$, $M \leq |\mathcal{X}|^n$. The decoder, in addition to Z , has access to a sequence Y^n , dependent on X^n via a known DMC, defined by the

single-letter transition probability matrix $\{P_{Y|X}(y|x)\}$, whose entries are the conditional probabilities of the different channel output symbols given the channel input symbols. Based on Z and Y^n , the decoder produces the reconstructed sequence \hat{X}^n . This setting is depicted in Figure 1. For simplicity, we assume that X_i, Y_i and \hat{X}_i , all take on values in the same finite alphabet \mathcal{X} . Let $Q(t)$, $0 \leq t < \infty$, be a real-valued convex function, where $\lim t \cdot Q(1/t) = 0$ as $t \rightarrow 0$. This requirement implies that $Q(t)$ is non-increasing, as was shown in [13]. We define $0 \cdot Q(r/0) = 0$, for all $0 \leq r < \infty$. The generalized mutual information relative to the function $Q(t)$ is defined by:

$$I^Q(X; Y) = \sum_{x,y} p(x,y) Q\left(\frac{p(x)}{p(x|y)}\right) \quad (1)$$

We apply the generalized DPT [13, Theorem 3] in the

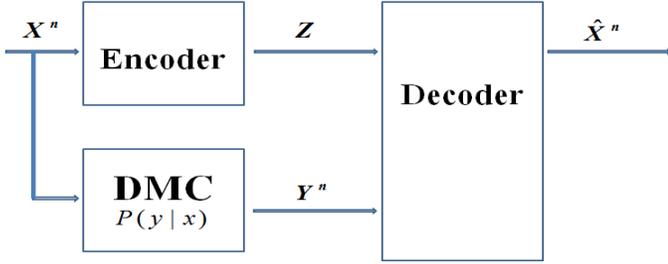


Fig. 1. The WZ Setting

following way:

$$I^Q(X^n, \hat{X}^n) \leq I^Q(X^n; Y^n, Z) \quad (2)$$

where we used the fact that $X^n \leftrightarrow (Y^n, Z) \leftrightarrow \hat{X}^n$ is a Markov chain. Since $Z \leftrightarrow X^n \leftrightarrow Y^n$ is also a Markov chain, we have:

$$p(x^n, y^n, z) = p(x)p(y^n|x^n)p(z|x^n) \quad (3)$$

and $I^Q(X^n; Y^n, Z)$ is given by:

$$\begin{aligned} I^Q(X^n; Y^n, Z) &= \sum_{x^n, y^n, z} p(x^n, y^n) p(z|x^n) \\ &\quad Q\left(\frac{p(y^n)p(z|y^n)}{p(y|x^n)p(z|x^n)}\right) \\ &= \sum_{x^n, y^n, z} p(x^n, y^n) p(z|x^n) \\ &\quad Q\left(\frac{p(y^n) \sum_{\tilde{x}^n \in \mathcal{X}^n} p(\tilde{x}^n|y^n) p(z|\tilde{x}^n)}{p(y|x^n)p(z|x^n)}\right) \end{aligned} \quad (4)$$

Assuming the encoder is given by a deterministic function $f : \mathcal{X}^n \rightarrow \{1, \dots, M\}$, we get:

$$I^Q(X^n; Y^n, Z) = \sum_{x^n, y^n} p(x^n, y^n) Q\left(\frac{p(y^n) \sum_{\tilde{x}^n \in A_z} p(\tilde{x}^n|y^n)}{p(y^n|x^n)}\right) \quad (5)$$

where $z = f(x^n)$ and $A_z \equiv \{\tilde{x}^n : f(\tilde{x}^n) = z\}$. Using $Q(t) = -\log(t)$ in (5), thus turning back to the classical DPT, we get the following result:

$$n(R(d) - I(X; Y)) \leq \sup \{H(Z|Y^n)\} \quad (6)$$

where the supremum is taken over all partitions of X^n into M disjoint subsets. The proof is given in [21]. This inequality stems from the specific Markovity of the WZ problem. We see that given some fixed-rate $R = \log M$, we should find the encoder that maximizes $H(Z|Y^n)$. In Subsection II.C, we show some examples of scalar coding, i.e., $n = 1$, where this result gives us lower bounds for the distortion, which are better than the bounds obtained from the trivial inequality $R_{WZ}(d) \leq \log M$, where $R_{WZ}(d)$ is the WZ RD function.

A. Generalized DPT for scalar coding

A fixed-rate scalar source code with rate $R = \log M$, partitions \mathcal{X} into M disjoint subsets (A_1, A_2, \dots, A_M) . The encoder f is given by a function $f : \mathcal{X} \rightarrow \{1, 2, \dots, M\}$, that is, $z_i = f(x_i)$. The decoder g receives z_i , together with the side information y_i , and generates \hat{x}_i , using a decoding function $g : \{1, 2, \dots, M\} \times \mathcal{X} \rightarrow \mathcal{X}$, i.e., $\hat{x}_i = g(z_i, y_i)$. This is the setting described in Figure 1, with block length $n = 1$. We define the following vectors $\{\vec{p}_z\}_{z=1}^M$:

$$\vec{p}_z = [1_{1 \in A_z}, 1_{2 \in A_z}, \dots, 1_{|\mathcal{X}| \in A_z}] \quad (7)$$

where 1_B is the indicator function for the event B . The j th coordinate of \vec{p}_z is 1 if $j \in A_z$ and 0 else. By definition of $\{\vec{p}_z\}_{z=1}^M$, we have the following property:

$$\sum_{z=1}^M \vec{p}_z = [1, 1, \dots, 1] \quad (8)$$

Using these vectors, we can rewrite (5) in the following way:

$$\begin{aligned} I^Q(X; Y, Z) &= \sum_{x,y} p(x,y) Q\left(\frac{p(y) \cdot [\vec{p}_{z(x)} \cdot \vec{p}_{\hat{x}|y}]}{p(y|x)}\right) \\ &= \sum_y p(y) \sum_z \sum_{x \in A_z} p(x|y) \\ &\quad Q\left(\frac{p(y) \cdot [\vec{p}_z \cdot \vec{p}_{\hat{x}|y}]}{p(y|x)}\right) \\ &= \sum_y p(y) \sum_z \vec{p}_z \cdot \vec{q}_{z,y} \\ &= \sum_y p(y) \sum_z G_y(\vec{p}_z) \end{aligned} \quad (9)$$

where we have defined the following $|\mathcal{X}|$ -vectors:

$$\begin{aligned}\vec{p}_{\hat{x}|y} &= [p(x_1|y), p(x_2|y), \dots] \\ \vec{q}_{z,y} &= \left[p(x_1|y)Q\left(\frac{p(y) \cdot [\vec{p}_z \cdot \vec{p}_{\hat{x}|y}]}{p(y|x_1)}\right), \right. \\ &\quad \left. p(x_2|y)Q\left(\frac{p(y) \cdot [\vec{p}_z \cdot \vec{p}_{\hat{x}|y}]}{p(y|x_2)}\right), \dots \right]\end{aligned}\quad (10)$$

and the set of functions $\{G_y\}_{y=1}^{|\mathcal{X}|}$, $G_y : \mathbb{R}^{|\mathcal{X}|} \rightarrow \mathbb{R}$:

$$G_y(\vec{p}_z) = \vec{p}_z \cdot \vec{q}_{z,y} \quad (11)$$

Notice that the vector $\vec{p}_{\hat{x}|y}$ depends only on y and that the dot product $[\vec{p}_z \cdot \vec{p}_{\hat{x}|y}]$ is a function of z and y . Applying the RD bound [13, Theorem 4], we get:

$$R^Q(d) \leq I^Q(X, \hat{X}) \leq I^Q(X; Y, Z) \leq C^Q \quad (12)$$

where:

$$\begin{aligned}R^Q(d) &= \inf I^Q(X, \hat{X}) \\ C^Q &= \sup I^Q(X; Y, Z) \\ &= \sup_y \sum_y p(y) \sum_z G_y(\vec{p}_z)\end{aligned}\quad (13)$$

The infimum is taken over all conditional distributions $P(\hat{x}|x)$ that satisfy the distortion constraint $\mathbb{E}d(X, \hat{X}) \leq d$, where $d(x, \hat{x})$ is a distortion measure. The supremum should be taken over all scalar encoders with a fixed-rate $R = \log M$. Alternatively, we can carry out a continuous optimization by taking the supremum over all sets of positive vectors $\{\vec{p}_z\}_{z=1}^M$ that satisfy (8), i.e., over all conditional distributions $P(z|x)$. The result will be of course greater than or equal to C^Q . We end this subsection with an upper bound on C^Q for a specific convex function $Q(t)$.

Lemma 1: For the convex function $Q(t) = t^{1-\alpha}$, $1 < \alpha < 2$, we have the following upper bound:

$$C^Q \leq M^{\alpha-1} \cdot \sum_y \left(\sum_{x'} p(x') \cdot p(y|x')^{\frac{1}{2-\alpha}} \right)^{2-\alpha} \quad (14)$$

The proof is given in [21]. The usefulness of this result stems from its generality. It holds for any source distribution and any transition probability matrix $\{P_{Y|X}(y|x)\}$. This result is used in Subsection II.C, along with tighter bounds on the capacity that can be achieved in several special cases.

B. The generalized rate-distortion function for uniform source distribution

In this subsection, we introduce the generalized rate-distortion function of uniformly distributed sources w.r.t general weakly symmetric distortion measures. We refer to a distortion measure $d(x, \hat{x})$ as weakly symmetric if the rows of the distortion matrix, $\{d(x, \hat{x})\}$, are permutations of each other. A uniformly distributed source is a source for which $P_X(x) = \frac{1}{|\mathcal{X}|}$, $\forall x \in \mathcal{X}$.

Lemma 2: Consider a discrete source X , uniformly distributed over a finite alphabet \mathcal{X} , and let $Q(t)$, $0 \leq t < \infty$

be any real-valued convex function. Then, $R^Q(d)$ w.r.t to any weakly symmetric distortion measure is given by:

$$R^Q(d) = \sum_{k=1}^{|\mathcal{X}|} p_k \cdot Q\left(\frac{1}{|\mathcal{X}|p_k}\right) \quad (15)$$

where $\{p_k\}_{k=1}^{|\mathcal{X}|}$ is a probability measure which is given by the following equations ($k = 1, \dots, |\mathcal{X}|$):

$$Q\left(\frac{1}{p_k}\right) - \frac{1}{p_k}Q'\left(\frac{1}{p_k}\right) + \lambda_1 + \lambda_2 d_k - \mu_k = 0 \quad (16)$$

where $\{d_k\}_{k=1}^{|\mathcal{X}|}$ are the elements of each row of the matrix $\{d(x, \hat{x})\}$ (the same elements appear in each row) and λ_1 , λ_2 , $\{\mu_k\}_{k=1}^{|\mathcal{X}|}$ are constants that are chosen such that:

$$\begin{aligned}\sum_{k=1}^{|\mathcal{X}|} p_k &= 1 \\ \sum_{k=1}^{|\mathcal{X}|} p_k d_k &= d \\ \mu_k &= \begin{cases} 0, & \text{if } p_k > 0 \\ 1, & \text{if } p_k = 0 \end{cases}\end{aligned}\quad (17)$$

For the Hamming distortion measure, defined by:

$$d(x, \hat{x}) = \begin{cases} 0 & x = \hat{x} \\ 1 & x \neq \hat{x} \end{cases} \quad (18)$$

we get the following result:

Corollary 1: Consider a discrete source X , uniformly distributed over a finite alphabet \mathcal{X} , and let $Q(t)$, $0 \leq t < \infty$ be any real-valued convex function. Then, $R^Q(d)$ w.r.t to the Hamming distortion measure is given by:

$$R^Q(d) = (1-d) \cdot Q\left(\frac{1}{|\mathcal{X}|} \left(\frac{1}{1-d}\right)\right) + d \cdot Q\left(\frac{(|\mathcal{X}|-1)}{|\mathcal{X}|} \frac{1}{d}\right) \quad (19)$$

The proof is given in [21]. The general form of $R^Q(d)$ enables the use of any convex function $Q(t)$. These results make the Ziv-Zakai mechanism much more useful, at least for the case of uniform sources.

C. Applications

In this subsection, we use the results of the previous subsections to get lower bounds on the distortion in several cases. Non-trivial bounds were obtained using the convex function $Q(t) = t^{1-\alpha}$, $\alpha > 1$, which is equivalent to using the Rényi information measure. Taking $t \rightarrow 1$, we turn back to the logarithmic function, thus to the classical DPT. We assume that the source is uniformly distributed, and that the DMC is symmetric. A channel is said to be symmetric if the rows of the channel transition matrix, $\{P_{Y|X}(y|x)\}$, are permutations of each other, and the columns are permutations of each other. Since the source is uniformly distributed and the channel is symmetric, Y is also uniformly distributed. Under these conditions, Eq. (11) becomes:

$$G_y(\vec{p}_z) = |\mathcal{X}|^{\alpha-1} \frac{\vec{p}_z \cdot \vec{p}_{y|\hat{x}}^\alpha}{(\vec{p}_z \cdot \vec{p}_{y|\hat{x}})^{\alpha-1}} \quad (20)$$

where we have defined the following $|\mathcal{X}|$ -vectors:

$$\begin{aligned}\vec{p}_{y|\bar{x}} &= [p(y|x_1), p(y|x_2), \dots] \\ \vec{p}_{y|\bar{x}}^\alpha &= [p(y|x_1)^\alpha, p(y|x_2)^\alpha, \dots]\end{aligned}\quad (21)$$

Applying (19) and (20) in (12), we get:

$$\begin{aligned}|\mathcal{X}|^{\alpha-1} \left((1-d)^\alpha + \frac{d^\alpha}{(|\mathcal{X}|-1)^{\alpha-1}} \right) &\leq \\ |\mathcal{X}|^{\alpha-2} \sup \left\{ \sum_{y,z} \frac{\vec{p}_z \cdot \vec{p}_{y|\bar{x}}^\alpha}{(\vec{p}_z \cdot \vec{p}_{y|\bar{x}})^\alpha} \right\} &= C^Q\end{aligned}\quad (22)$$

where the supremum is taken over all sets of positive vectors $\{\vec{p}_z\}_{z=1}^M$ that satisfy (8). Notice that optimizing over α will produce the best lower bound for the distortion. Generally, the functions $\{G_y(\vec{p}_z)\}$ in this case are neither convex nor concave. However, at least for some cases, C^Q can be calculated directly, as shown in the following examples. We can also upper bound C^Q as was done in (14). This upper bound may give us non-trivial bounds, as shown in example 2. Additional examples, with more general distortion measures, will be given in [21].

1) *Example 1:* The symmetric DMC is defined by:

$$p(y|x) = \begin{cases} \mu & y = x \\ \epsilon & \text{else} \end{cases}\quad (23)$$

where $\mu, \epsilon \in [0, 1]$, $\mu > \epsilon$, and $\mu + (|\mathcal{X}|-1)\epsilon = 1$. The distortion measure we use, is the Hamming distortion, defined in (18).

Lemma 3: Consider the WZ setting with uniformly distributed source and the DMC defined in (27). Then, the minimal achievable distortion w.r.t to the Hamming distortion measure, of a scalar source code with a fixed-rate $R = \log M$, is:

$$d(M) = \epsilon(|\mathcal{X}| - M)\quad (24)$$

The proof is given in [21]. Having the exact solution in this case, we can compare it to our bound to examine its quality. The generalized capacity (13) for this channel is given by:

$$\begin{aligned}C^Q &= |\mathcal{X}|^{\alpha-2} \cdot \sup \left\{ \sum_{y,z} \frac{\vec{p}_z \cdot \vec{p}_{y|\bar{x}}^\alpha}{(\vec{p}_z \cdot \vec{p}_{y|\bar{x}})^\alpha} \right\} \\ &= |\mathcal{X}|^{\alpha-2} \epsilon \cdot \sup \left\{ \sum_z M_z \cdot \frac{(M_z + \mu^\alpha/\epsilon^\alpha - 1)}{(M_z + \mu/\epsilon - 1)^{\alpha-1}} \right. \\ &\quad \left. + (|\mathcal{X}| - M_z)M_z^{2-\alpha} \right\} \\ &= |\mathcal{X}|^{\alpha-2} \epsilon \cdot \sup \left\{ \sum_z q_\alpha(M_z) \right\}\end{aligned}\quad (25)$$

where $M_z, M_z \in \{1, \dots, |\mathcal{X}| - M + 1\}$, is the cardinality of A_z , i.e., the number of source symbols that are encoded to z . Obviously, $\sum_z M_z = |\mathcal{X}|$. Notice that the supremum is taken over all feasible encoders, where each encoder is represented

by a specific set $\{\vec{p}_z\}_{z=1}^M$ as defined in (7). The second equality is explained in [21]. The function $q_\alpha(M_z)$ is concave for $1 < \alpha \leq 2$, as shown in [21], and may be concave also out of this range, with dependence on the channel parameters. When $q_\alpha(M_z)$ is concave, we can bound the supremum by taking equal M_z 's, i.e., $M_z = |\mathcal{X}|/M, \forall z$, and we get:

$$\begin{aligned}C^Q &\leq |\mathcal{X}|^{\alpha-1} \epsilon \left(\frac{(|\mathcal{X}|/M + \mu^\alpha/\epsilon^\alpha - 1)}{(|\mathcal{X}|/M + \mu/\epsilon - 1)^{\alpha-1}} + \right. \\ &\quad \left. (M-1) \left(\frac{|\mathcal{X}|}{M} \right)^{2-\alpha} \right)\end{aligned}\quad (26)$$

If M divides $|\mathcal{X}|$, this bound is achieved by any feasible encoder that partitions the source alphabet into equally-sized subsets, thus the optimization is exact. An example for specific values of μ and ϵ is presented in Figure 2. The bound is compared with the bound obtained from the classical DPT (6), the bound obtained from the trivial inequality $R_{WZ}(d) \leq \log M$, the bound obtained by using (14) and the exact solution of Theorem 3. $R_{WZ}(d)$ was calculated using the Blahut-Arimoto-type algorithm presented in [20]. Eq. (22) was optimized over α , for each $M \leq |\mathcal{X}|$, as to get the best lower bound on the distortion. We see that even the classical

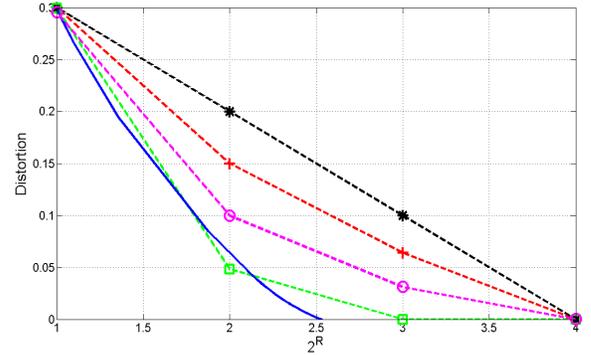


Fig. 2. $|\mathcal{X}| = 4, \mu = 0.7$. Plus - the lower bound obtained from (26), Circle - the lower bound obtained from the classical DPT (6), Star - the exact solution, Solid line - the lower bound obtained from $R_{WZ}(d)$, Square - the lower bound obtained from (14).

DPT gives us non-trivial lower bounds and that the lower bound obtained from (26) is much better than the trivial bound obtained from $R_{WZ}(d)$. The lower bound obtained from (14) is not useful in this case. There is a gap between the exact solution and the best bound, even for $M = 2$, where the optimization (13) is exact.

2) *Example 2:* The symmetric DMC is defined by:

$$p(y|x) = \begin{cases} 1/l & y \in \{x \bmod |\mathcal{X}|, \dots, (x+l-1) \bmod |\mathcal{X}|\} \\ 0 & \text{else} \end{cases}\quad (27)$$

where l is an integer, $0 < l < |\mathcal{X}|$, and $x \bmod |\mathcal{X}|$ is defined to be x . Given an input x , the channel produces one of l values with equal probability. The generalized capacity (13)

for this channel is given by:

$$\begin{aligned}
C^Q &= |\mathcal{X}|^{\alpha-2} \cdot \sup \left\{ \sum_{y,z} \frac{\vec{p}_z \cdot \vec{p}_{y|\bar{x}}^\alpha}{(\vec{p}_z \cdot \vec{p}_{y|\bar{x}})^{\alpha-1}} \right\} \\
&= |\mathcal{X}|^{\alpha-2} \cdot \sup \left\{ \sum_{y,z} \frac{M_{y,z}(1/l)^\alpha}{(M_{y,z}(1/l))^{\alpha-1}} \right\} \quad (28) \\
&= |\mathcal{X}|^{\alpha-2} \cdot l^{-1} \cdot \sup \left\{ \sum_{y,z} M_{y,z}^{2-\alpha} \right\}
\end{aligned}$$

where $M_{y,z} = l \cdot [\vec{p}_z \cdot \vec{p}_{y|\bar{x}}]$. It is easy to see that $\sum_z M_{y,z} = l$. For $1 < \alpha < 2$, the function $M_{y,z}^{2-\alpha}$ is concave in $M_{y,z}$. Thus, the supremum is achieved by setting $M_{y,z} = l/M, \forall \{y, z\}$:

$$C^Q = |\mathcal{X}|^{\alpha-2} \cdot l^{-1} \cdot |\mathcal{X}| \cdot M \cdot (l/M)^{2-\alpha} = |\mathcal{X}|^{\alpha-1} \cdot (M/l)^{\alpha-1} \quad (29)$$

If M divides l , equal $M_{y,z}$'s can be obtained by the following feasible encoder:

$$z = f(x) = 1 + x \bmod M \quad (30)$$

Therefore, in this case, the optimization is exact. For $\alpha > 2$, C^Q is infinite, because we can always set some of the $M_{y,z}$'s to 0 by appropriate choice of encoder. Thus, this range of α does not lead to a useful bound. An example for specific values of $|\mathcal{X}|$ and l is presented in Figure 3. The lower bound for the distortion, which coincides with the bound obtained from (14), is compared with the bound obtained from the classical DPT (6) and the bound obtained by the trivial inequality $R_{WZ}(d) \leq \log M$. Eq. (22) was optimized over α , for each $M \leq |\mathcal{X}|$, as to get the best lower bound on the distortion. We see that in this case, the generalized DPT leads to bounds that are better than the trivial bound, whereas the classical DPT does not lead to a useful bound. We also present the exact distortion of the encoder defined in (30), which is of course an upper bound on the distortion. Thus, the distortion of the optimal encoder must be in the range between this upper bound and our highest lower bound. For $M = l$, zero distortion can indeed be achieved using the encoder defined in (30), thus our lower bound in this point is tight.

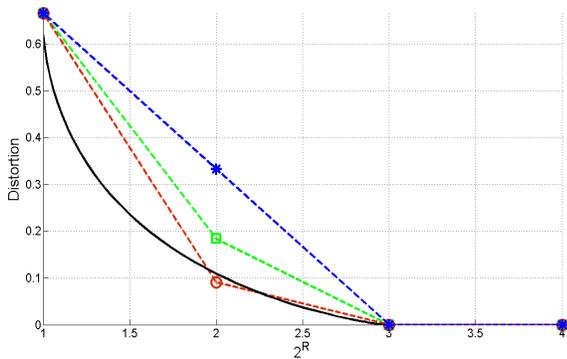


Fig. 3. $|\mathcal{X}| = 4, l = 3$. Square - the lower bound obtained from (26) and (14), Circle - the lower bound obtained from the classical DPT (6), Solid line - the lower bound obtained from $R_{WZ}(d)$, Star - the exact distortion of the encoder (30).

Acknowledgments

This research is supported by the Israeli Science Foundation (ISF), grant no. 208/08.

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pt. I, pp. 379–423, 1948; pt. II, pp. 623–656, 1948.
- [2] J. Kusuma, "Slepian–Wolf coding and related problems," preprint 2001. Available at: www.mit.edu/~6.454/www_fall_2001/kusuma/summary.pdf
- [3] S. D. Servetto, "Quantization with side information: lattice codes, asymptotics, and applications to sensor networks," vol. 53, no. 2, pp. 714–731, February 2007.
- [4] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, pp. 80–94, September 2004.
- [5] Y. Steinberg and N. Merhav, "On successive refinement for the Wyner–Ziv problem," *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1636–1654, August 2004.
- [6] S. Cheng and Z. Xiong, "Successive refinement for the Wyner–Ziv problem and layered code design," in *Proc. DCC 2004*, Snowbird, UT, 2004.
- [7] J. Kusuma, L. Doherty and K. Ramchandran, "Distributed compression for sensor networks," in *Proc. ICIP 2001*, vol. 1, pp. 82–85, Thessaloniki, Greece, Oct. 2001.
- [8] D. Muresan and M. Effros, "Quantization as histogram segmentation: Optimal scalar quantizer design in network systems," *IEEE Trans. Inform. Theory*, vol. 54, pp. 344–366, Jan. 2008.
- [9] D. Teneketzis, "On the structure of optimal real-time encoders and decoders in noisy communication," *IEEE Trans. Inform. Theory*, vol. 52, pp. 4017–4035, Sept. 2006.
- [10] J. Nayak, E. Tuncel, "Low-delay quantization for source coding with side information," *Proc. ISIT 2008*, pp. 2732–2736, Toronto, Canada, Aug. 2008.
- [11] X. Chen, E. Tuncel, "Low-delay prediction and transform-based Wyner–Ziv Coding," *IEEE Trans. Signal Processing*, vol. 59, no. 2, pp. 653–666, Nov. 2010.
- [12] X. Chen, E. Tuncel, "High-resolution predictive Wyner–Ziv coding of Gaussian sources," *Proc. ISIT 2009*, pp. 1204–1208, Seoul, Korea, Aug. 2009.
- [13] J. Ziv and M. Zakai, "On functionals satisfying a data-processing theorem," *IEEE Trans. Inform. Theory*, vol. 19, no. 3, pp. 275–283, May 1973.
- [14] M. Zakai and J. Ziv, "A generalization of the rate-distortion theory and applications," in: *Information Theory New Trends and Open Problems*, edited by G. Longo, Springer-Verlag, pp. 87–123, 1975.
- [15] I. Leibowitz, R. Zamir, "A Ziv–Zakai–Rényi lower bound on distortion at high resolution," *Proc. ITW 2008*, pp. 174–178, Porto, Portugal, May. 2008.
- [16] S. Tridenski, R. Zamir, "Bounds for joint source-channel coding at high SNR," *Proc. ISIT 2011*, pp. 771–775, St. Petersburg, Russia, Aug. 2011.
- [17] A. Ingber, I. Leibowitz, R. Zamir, M. Feder, "Distortion lower bounds for finite dimensional joint source-channel coding," *Proc. ISIT 2008*, pp. 1183–1187, Toronto, Canada, Aug. 2008.
- [18] A. Rényi, "On measures of entropy and information," *Proc. 4th Berk. Symp. Math., Stat. and Prob.*, pp. 547–561, Univ. of Calif. Press, 1961.
- [19] N. Merhav, "Data Processing Theorems and the Second Law of Thermodynamics," *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 4926–4939, Aug. 2011.
- [20] F. Dupuis, W. Yu, F.M.J. Willems, "Blahut–Arimoto algorithms for computing channel capacity and rate-distortion with side information," *Proc. ISIT 2004*, pp. 179, Chicago, USA, June 2004.
- [21] A. Reani, N. Merhav, "Data processing lower bounds for scalar lossy source codes with side information at the decoder," in preparation.