On Scalar Quantizer Design with Decoder Side Information

Zhenyu Tu, Tiffany Jing Li and Rick S. Blum

***Abstract*— In this work, we investigate the design of scalar quantizers for the Wyner-Ziv problem, where the decoder has access to a noisy copy of the source. Necessary conditions are given for the design of optimum quantizers with a fixed number of partitions for a criterion consisting of a linear combination of distortion and rate. A simple iterative algorithm is developed to tackle the new quantizer design problem. The new algorithm generalizes the well-known Lloyd type I algorithm. The new approach will, in general, yield different scalar quantizers than the traditional approaches and these new quantizers provide some advantages. We also develop approximations to simplify quantizer design for high rate cases. With such approximations, the traditional optimal fixed-rate and entropy-constrained scalar quantizers are shown to remain optimum under certain conditions.**

I. Introduction

Scalar quantizers have been of interest to both academic and industrial researchers for decades due to their simple implementation. Typically, the rate of a fixed-rate scalar quantizer is determined by the number of partitions only. The design of such quantizers leads to finding the partitions and the associated reproductions which result in minimum distortion between the source and its quantized version. The well known class of Lloyd-Max iterative algorithms [1], [2] provide efficient methods to obtain these optimum partitions and their reproductions. The rate can be reduced if entropy coding is applied to the partition indexes, while the distortion level is kept constant. Thus a better rate-distortion trade-off is achieved at the price of increased complexity due to the entropy coding. Fixed-rate scalar quantizer design is generalized to allow entropy coding in [3], [4], [5] yielding what is often referred to as entropy-constrained quantizer design.

It has been shown in [6] that the rate of the quantizer can be reduced, when the decoder has access to a noisy version of the source, by exploiting the statistical correlation between the source and the decoder side information. Such rate reduction provides a new rate and distortion trade-off for quantizers. Additionally, the decoder side information can help in reconstructing the source so that smaller distortion can be achieved.

Quantizer design for this new scenario, which is referred to as Wyner-Ziv (WZ) problem, has been a focused research area for some time. Early attempts to design quantizers for the Wyner-Ziv problem were based on high dimensional nested lattices, followed by either fixed-rate coding or entropy coding of the quantization indices (see for example [7][8]). Recently, the application of scalar quantizers followed by Slepian-Wolf binning was proposed for the Wyner-Ziv problem in [9]. Although a heuristic approach was suggested, the focus was not on optimum quantizer design. In [10][11], an objective function formed by a linear combination of distortion and rate is proposed for quantizer design. An iterative Lloyd algorithm based on local optimization is proposed for quantizer design in a network source coding problem. The encoding/decoding mapping function is iteratively chosen to minimize an objective function assuming all other encoding/decoding mapping functions are fixed. It is generalized to allow more general rate measures in [12].

This existing work has shed useful insight into the design of WZ quantizers. Researchers have come to a consensus that a good quantizer for the Wyner-Ziv problem needs to be designed with explicit consideration of its correlation with the decoder side information as well as the nature of the succeeding index encoder. However, little attempt has been made to analytically formulate the quantizer design problem as an optimization over its parameters, e.g. its partition boundaries, directly. Note this is not tractable in a general framework. However, we may be able to do this for a specific type of quantizer design.

In this work, we focus on the scenario described in the original work by Wyner and Ziv [6], with emphasis on theoretically motivated algorithms for optimal (at least locally optimal) scaler quantizer design. The simplicity of scaler quantizers enables us to explicitly formulate a rate-distortion objective function which unifies diverse cases and which can be numerically solved using practical iterative algorithms. Hence, our approach is different from those in [12], [11], [10], since we avoid the discretization approximations and training data methods in the design process.

The rest of the paper is organized as follows. Section II formulates the problem we consider. Section III discusses the conditions for optimality, along with an iterative algorithm for numerical computations. High rate approximations are also studied. Section III-E provides some design examples. Finally, Section IV summarizes the paper.

II. Problem Formulation

In this work, we use upper-case letters to represent random variables, e.g. X. We use the corresponding lower-case letters to represent realizations, e.g. x. We use a superscript to denote a sequence, e.g. $x^n = x_1, x_2, \ldots, x_n$.

Consider the source coding problem with side information at the decoder. The source X^n is assumed to be a sequence of i.i.d. zero-mean Gaussian random variables, each with variance σ_X^2. The decoder is assumed to have access to a noisy version of X^n which is characterized by

$$Y^n = X^n + Z^n \tag{1}$$

where Z^n is a sequence of i.i.d. zero-mean Gaussian variables, each with variance σ_Z^2, and n is the block length. We assume that X^n is independent of Z^n. We define the correlation SNR as $10 \log_{10} \frac{\sigma_X^2}{\sigma_Z^2}$ dB.

We adopt the system diagram proposed in [9] which is depicted in Figure 1. The encoder consists of a scalar quantizer followed by an index encoder. The scalar quantizer determines the partition of the source vector.

The index encoder maps a vector of partition indexes to a data stream which is sent to the decoder. The index decoder recovers the partition indexes with the help of the side information y^n. The optimum estimator outputs the reproduction vector for the source based on the partition indexes and y^n.

The goodness of the reproduction is measured by some distortion function quantifying the difference between x^n and $\hat{x}^n$. We adopt the mean square error distortion measure

$$d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{x}_i)^2 \tag{2}$$

throughout the paper. The expected distortion is given by

$$D = E[d(X^n, \hat{X}^n)] \tag{3}$$

where the expectation is with respect to the distribution of X^n and Z^n.

Let R denote the rate of the source code. We form the objective function as a weighted sum of distortion and rate as considered in [12], [11]:

$$J = (1 - \lambda) D + \lambda R \tag{4}$$

where $\lambda \in (0, 1)$.

III. Optimal Quantizer Design

In our optimum quantizer design, we seek to obtain quantizers with minimum J for a fixed number of partitions (and fixed λ). Toward this goal, we compute D and R in (4).

Let $u_i, \ i = 1, \ldots, N + 1$, denote the end points of the partitions with $u_1 = -\infty$ and $u_{N+1} = \infty$; Let $v_i, \ i = 1, \ldots, N$, denote the reproduction for partition (u_i, u_{i+1}). Let X denote a random variable with the common probability density function (pdf) of each of $X_1, \ldots, X_n$. Similarly, let Y denote a random variable with the common pdf of each of $Y_1, \ldots, Y_n$. The mean-square-error optimal estimator (reproduction) for a partition given by the interval (s, t) is shown in (5) (see [13]) where $c = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2}$ and $\sigma = \sqrt{c}\sigma_Z$.

To simplify notation later, we define the expected distortion conditioned on $Y = y$ and $X \in (s, t)$ as

$$f(y, s, t) = E[X^2 | Y = y, X \in (s, t)] - v^2(s, t). \tag{6}$$

The first term of the right hand side of (6) can be computed as in (7).

Similarly, we define the expected distortion conditioned on $X \in (s, t)$ as

$$F(s, t) = \int_{-\infty}^{\infty} f(y, s, t) p_Y(y | X \in (s, t)) dy, \tag{8}$$

where $p_Y(y|X \in (s, t))$ denotes the conditional pdf of Y given $X \in (s, t)$ (we employ similar notation for different variables later).

A. Evaluating D and R

Using (8), we can compute (3) as

$$\begin{aligned} D &= E[(X - \hat{X})^2] \\ &= \sum_{i=1}^{N} \int_{u_i}^{u_{i+1}} p_X(x) dx F(u_i, u_{i+1}). \end{aligned} \tag{9}$$

Since X is independent of Z, the pdf in (8) can be computed as the convolution of two known pdfs: $p_Y(x|X \in (s, t)) = p_X(x|X \in (s, t)) * p_Z(x))$. Without loss of generality, we assume $\sigma_X = 1$ and compute the convolution in (10).

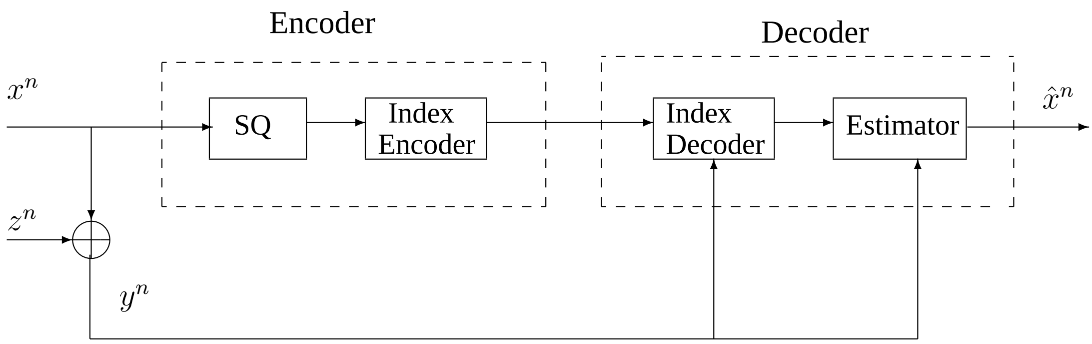

Fig. 1. Encoder-Decoder Structure

$$
\begin{aligned}
v(s,t) &= E[X|Y=y, X\in(s,t)] \\
&= cy + \frac{2}{\text{erf}(\frac{t-cy}{\sqrt{2}\sigma}) - \text{erf}(\frac{s-cy}{\sqrt{2}\sigma})} \frac{\sigma}{\sqrt{2\pi}} (e^{-\frac{(s-cy)^2}{2\sigma^2}} - e^{-\frac{(t-cy)^2}{2\sigma^2}})
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
E[X^2|Y=y, X\in(s,t)] &= \sigma^2 + c^2y^2 + \frac{2}{\text{erf}(\frac{t-cy}{\sqrt{2}\sigma}) - \text{erf}(\frac{s-cy}{\sqrt{2}\sigma})} \cdot \\
&\quad \frac{(\frac{s-cy}{\sqrt{2}\sigma}\sigma^2 + \sqrt{2}cy\sigma)e^{-(\frac{s-cy}{\sqrt{2}\sigma})^2} - (\frac{t-cy}{\sqrt{2}\sigma}\sigma^2 + \sqrt{2}cy\sigma)e^{-(\frac{t-cy}{\sqrt{2}\sigma})^2}}{\sqrt{\pi}}
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
p_X(x|X\in(s,t)) * p_Z(x) &= \int_s^t \frac{1}{Pr(X\in(s,t))} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} \frac{1}{\sqrt{2\pi}\sigma_Z} e^{-(x-w)^2/(2\sigma_Z^2)} dw \\
&= \frac{2}{\text{erf}(t/\sqrt{2}) - \text{erf}(s/\sqrt{2})} \frac{\sqrt{2}e^{-x^2/(2(\sigma_Z^2+1))}}{4\sqrt{\pi(\sigma_Z^2+1)}} \\
&\quad \left(\text{erf}\left(\frac{-(1+\sigma_Z^2)s + x}{\sqrt{2\sigma_Z^2(1+\sigma_Z^2)}}\right) - \text{erf}\left(\frac{-(1+\sigma_Z^2)t + x}{\sqrt{2\sigma_Z^2(1+\sigma_Z^2)}}\right)\right).
\end{aligned} \tag{10}
$$

It has been shown in [14] that when ideal SWC is assumed, the achievable rate R is given by $H(B|Y) = H(B) - I(B;Y)$, where B denotes the partition index. Note that ideal SWC performs entropy coding and exploits the statistical correlation between B and Y at the same time. On the other hand, practical SWC schemes, e.g. [15], [16], [17], are syndrome-based with linear channel codes which exploit the statistical correlation but give up entropy coding. The achievable rate of those schemes is given by $\log_2(N) - C(B;Y)$ assuming that linear codes are capable of approaching the capacity of the virtual channel $B \mapsto p(y|B)$ which is denoted by $C(B;Y)$. The computation of $C(B;Y)$ could be quite complicated in general. However, we can lower bound it with $I(B;Y)$ which denotes the mutual information for the fixed partition and distribution of X. Thus assigning $H(B|Y)$ and $\log_2(N) - I(B;Y)$ to R provides lower and upper bounds on the operational achievable rate.

Defining the conditional differential entropy as

$$
g(s,t) = h(Y|X\in(s,t))), \tag{11}
$$

we find

$$
I(B;Y) = h(Y) - \sum_{i=1}^{N} g(u_i, u_{i+1}) \int_{u_i}^{u_{i+1}} p_X(x)dx. \tag{12}
$$

Since Y is Gaussian with zero mean and variance $\sigma_X^2 + \sigma_Z^2$, we have $h(Y) = \frac{1}{2}\log_2(2\pi e(\sigma_X^2 + \sigma_Z^2))$. Further

$$H(B) = -\sum_{i=1}^{N}\left(\int_{u_i}^{u_{i+1}} p_X(x)dx\right)\log_2(\int_{u_i}^{u_{i+1}} p_X(x)dx). \quad (13)$$

For given λ and N, the necessary conditions for minimum J are

$$\frac{\partial J}{\partial u_j} = 0, \quad j = 2, \ldots, N. \quad (14)$$

From (13), we find

$$\frac{\partial H(B)}{\partial u_j} = p_X(u_j)\log_2\frac{\text{erf}(u_{j+1}/\sqrt{2}) - \text{erf}(u_j/\sqrt{2})}{\text{erf}(u_j/\sqrt{2}) - \text{erf}(u_{j-1}/\sqrt{2})}$$

for $j = 2, \ldots, N$. From (12), we obtain[1] (15) for $j = 2, \ldots, N$. From (9), we obtain (16) for $j = 2, \ldots, N$.

Using these results in (4), we can rewrite (14) for the two cases we consider here as

1) Ideal SWC: $R = H(B|Y)$.

$$\begin{aligned}\frac{\partial J}{\partial u_j} &= (1-\lambda)\frac{\partial D}{\partial u_j} + \lambda\frac{\partial H(B)}{\partial u_j} - \lambda\frac{\partial I(B;Y)}{\partial u_j} \\ &= 0\end{aligned} \quad (17)$$

2) Practical SWC: $R = \log_2(N) - I(B;Y)$.

$$\begin{aligned}\frac{\partial J}{\partial u_j} &= (1-\lambda)\frac{\partial D}{\partial u_j} - \lambda\frac{\partial I(B;Y)}{\partial u_j} \\ &= 0\end{aligned} \quad (18)$$

Note that the simultaneous equations obtained from (14) are complicated nonlinear equations. There seems to be no easy way to solve them analytically in closed-form. Therefore we propose a modified Lloyd iterative algorithm [2] to solve those equations numerically.

B. Sufficient Conditions for Optimality

In the previous section, we have given, in general, only necessary conditions in (14) for minimum J. They may not be sufficient conditions. For example, when entropy coding is allowed, the conditions in (14) may lead to maximum J instead of minimum J. In such cases, we found that minimizing J sometimes leads to convergence of neighboring end points, resulting in a degenerate quantizer design with $N-1$ partitions instead of N partitions. One way to address the problem, besides varying N, is to vary λ. For example, decreasing λ in (4) can be helpful. In general, as discussed in [1], (14) yields sufficient conditions for a local minimum if the matrix whose (i,j)th entry is $\frac{\partial^2 J}{\partial u_i \partial u_j}$ is positive definite at the solution point. Such a check can be easily implemented numerically.

C. Modified Lloyd Type I Algorithm

In this section, we propose a modified Lloyd type I algorithm to solve the equations in (14) numerically. The idea is to adjust partition end points sequentially. We start with an initial partition $(u_1, \cdots, u_{N+1})$. Then we assume that u_{j-1} and u_{j+1} are fixed, and update u_j to satisfy (17) and (18) respectively for $j = 2$ to N. These updates are repeated until convergence of J is observed. Further, due to the fact that $p_X(x)$ is symmetric, we focus on symmetric partitions to reduce the number of variables from $N-1$ to $N/2-1$ for even N, and $(N-1)/2$ for odd N. Note that our algorithm does not require symmetry of the partitions.

Now we analyze this algorithm. First, we consider the update of u_j via (18). Recall that $g(s,t)$ is the conditional differential entropy as defined in (11). Increasing the width of the interval (s,t) yields higher uncertainty of Y, thus yielding higher entropy. Hence we have $\frac{\partial g}{\partial s} < 0$ and $\frac{\partial g}{\partial t} > 0$.

Combining these results with (15) and using the fact that the function erf is monotonically increasing, we find that $\lim_{u_j \to u_{j-1}} \frac{\partial I(B;Y)}{\partial u_j} > 0$ and $\lim_{u_j \to u_{j+1}} \frac{\partial I(B;Y)}{\partial u_j} < 0$. Also note that $F(s,t)$ is the conditional expected distortion as defined in (8). Decreasing the width of the interval (s,t) yields smaller distortion. Hence we have $\frac{\partial F}{\partial s} < 0$ and $\frac{\partial F}{\partial t} > 0$. Following the same argument above, we have $\lim_{u_j \to u_{j-1}} \frac{\partial D}{\partial u_j} < 0$ and $\lim_{u_j \to u_{j+1}} \frac{\partial D}{\partial u_j} > 0$. Combining these results with (18), we find that $\lim_{u_j \to u_{j-1}} \frac{\partial J}{\partial u_j} < 0$ and $\lim_{u_j \to u_{j+1}} \frac{\partial J}{\partial u_j} > 0$. Hence we can guarantee that a solution for u_j in the interval (u_{j-1}, u_{j+1}) can be found. We call this bracketing u_j. If there is only one solution for u_j in (u_{j-1}, u_{j+1}), it must be a local minimum for J due to the fact that $\frac{\partial J}{\partial u_j}$ is negative at u_{j-1} and positive at u_{j+1}. In many of the cases we considered, our numerical study seemed to suggest that (18) had only one solution in (u_{j-1}, u_{j+1}). Since we have bracketed u_j in (u_{j-1}, u_{j+1}), we can apply the bisection search to solve (18) numerically.

Next, we consider the update of u_j via (17). Note that $\lim_{u_j \to u_{j-1}} \frac{\partial H(B)}{\partial u_j} \to \infty$ and $\lim_{u_j \to u_{j+1}} \frac{\partial H(B)}{\partial u_j} \to -\infty$. With this dominant term, we have $\lim_{u_j \to u_{j-1}} \frac{\partial J}{\partial u_j} > 0$ and $\lim_{u_j \to u_{j+1}} \frac{\partial J}{\partial u_j} < 0$. Due to this, the solution to (17) could lead to a local maximum of J rather than a local minimum, and in this case the local minimum of J is

[1] Here $\frac{\partial F}{\partial t}|_{(u_{j-1},u_j)}$ means to set the variables $(s,t) = (u_{j-1}, u_j)$ after the derivative where s, t are as used in (11)

$$\frac{\partial I(B;Y)}{\partial u_j} = p_X(u_j)(g(u_j,u_{j+1}) - g(u_{j-1},u_j)) - \frac{\partial g}{\partial t}\bigg|_{(u_{j-1},u_j)} \frac{\text{erf}(u_j/\sqrt{2}) - \text{erf}(u_{j-1}/\sqrt{2})}{2}$$
$$- \frac{\partial g}{\partial s}\bigg|_{(u_j,u_{j+1})} \frac{\text{erf}(u_{j+1}/\sqrt{2}) - \text{erf}(u_j/\sqrt{2})}{2} \tag{15}$$

$$\frac{\partial D}{\partial u_j} = p_X(u_j)(F(u_{j-1},u_j) - F(u_j,u_{j+1})) + \frac{\text{erf}(u_j/\sqrt{2}) - \text{erf}(u_{j-1}/\sqrt{2})}{2} \frac{\partial F}{\partial t}\bigg|_{(u_{j-1},u_j)}$$
$$+\frac{\text{erf}(u_{j+1}/\sqrt{2}) - \text{erf}(u_j/\sqrt{2})}{2} \frac{\partial F}{\partial s}\bigg|_{(u_j,u_{j+1})} \tag{16}$$

achieved at one end point. This is not surprising considering the fact that reducing the number of partitions leads to a smaller entropy.

As we just showed, finding minimum J by solving (14) could be problematic. To avoid such difficulty, we apply tuning to adjust the iterative algorithm when necessary. We take discrete samples between u_{j-1} and u_{j+1} and try to bracket u_j in an interval where $\frac{\partial J}{\partial u_j}$ is negative at the left end point and positive at the right end point. If it succeeds, the bisection search is applied to solve for u_j. Otherwise, we keep u_j unchanged and move to update u_{j+1}.

We apply another level of control to help guide the algorithm. Before u_j is updated, we compute the current J and the J after the update. If J decreases, then we finish the update. Otherwise, we abandon the update. Thus J cannot increase after any update. If the algorithm fails to update any u_j, we decrease λ and try the updates again.

Since J is lower bounded by 0, the iterative algorithm will terminate with a solution which may not be optimum. Then we check if all the equations in (14) are balanced. If there are unbalanced equations, we decrease λ, which reduces the impact of rate. Otherwise, we check if the matrix with (i,j)th entry $\frac{\partial^2 J}{\partial u_i \partial u_j}$ is positive definite. If it is, we have found a local optimum. If it is not, we decrease λ.

D. High Rate Approximation

By high rate, we mean that we assume large N. In such cases, the interval (u_j, u_{j+1}) (here we exclude consideration of the two intervals at the boundaries) becomes so small that we can apply certain approximations to simplify the task of quantizer design. With N being large, we ignore the effect of the two boundary partitions hereafter.

We apply two approximations for large N. Firstly, note that the performance gain of knowing y^n in the optimum estimator (see Figure 1) diminishes as N increases. Hence we can use a fixed set of reproductions and simplify the computation of (3) to

$$D = \sum_{i=1}^{N} \int_{u_i}^{u_{i+1}} (x - v_i)^2 p_X(x) dx. \tag{19}$$

Secondly, we apply the following approximation for the error function when $\triangle x$ is small

$$\text{erf}(x + \triangle x) - \text{erf}(x) \approx \frac{2}{\sqrt{\pi}} e^{-(x+\triangle x/2)^2} \triangle x. \tag{20}$$

Assuming s and t are close enough, we rewrite (10) as

$$p_X(x|X \in (s,t)) * p_Z(x) \approx \frac{1}{\sqrt{2\pi}\sigma_Z} e^{-\frac{1}{2\sigma_Z^2}(x - \frac{s+t}{2})^2} \tag{21}$$

With the approximation in (21), we can compute the conditional differential entropy as

$$g(s,t) = \frac{1}{2} \log_2(2\pi e \sigma_Z^2). \tag{22}$$

Note that with (22), we have $I(B;Y) = \frac{1}{2}\log_2(1 + \frac{\sigma_X^2}{\sigma_Z^2})$ which is a constant for given correlation SNR. With $I(B;Y)$ being constant, we find that minimizing J in (4) is equivalent to minimizing $(1-\lambda)D + \lambda H(B)$ when ideal SWC is assumed, or equivalent to minimizing D when SWC based on a linear code is assumed. The former corresponds to the conventional entropy-constrained quantizer design and the latter corresponds to the fixed-rate quantizer design. In other words, for large N, the optimal entropy-constrained and fixed-rate quantizers remain optimal for the problem considered here.

E. Numerical results

Note that the conventional fixed-rate and entropy-constrained scalar quantizer designs are special cases of the problem described in the previous section for the case where the correlation SNR goes to $-\infty$. Numerical results indicate that the proposed modified Lloyd type I algorithm, which aims to obtain the optimal partition directly, yields the same fixed-rate quantizers reported in [1]. We have verified this when starting from a uniform initial partition.

Next, we consider the case where the decoder has access to a noisy version of the source. We report some design examples for cases where practical SWC is assumed in Table I-A[2]. Design examples for cases where ideal SWC is assumed are provided in Table I-B. The correlation SNR is set to be 10 db. All the quantizers reported in this work satisfy the necessary conditions in (14) and sufficient condition for locally optimality discussed in Section III-B.

TABLE I

DESIGN EXAMPLES

(A)	WZ Scalar Quantizers Using Practical SWC			
N	λ	$\mathbf{u}$	D	R
4	0.5	$(0\ 0.837\ \infty)$	0.0457	0.776
4	0.05	$(0\ 0.758\ \infty)$	0.045	0.781
8	0.5	$(0\ 0.453\ 0.946\ 1.563\ \infty)$	0.0215	1.464
8	0.05	$(0\ 0.428\ 0.892\ 1.462\ \infty)$	0.0211	1.466
(B)	WZ Scalar Quantizers Using Ideal SWC			
N	λ	$\mathbf{u}$	D	R
4	0.05	$(0\ 1.039\ \infty)$	0.049	0.677
4	0.01	$(0\ 0.714\ \infty)$	0.045	0.786
8	0.01	$(0\ 0.457\ 0.941\ 1.508\ \infty)$	0.0213	1.379
8	0.005	$(0\ 0.431\ 0.892\ 1.441\ \infty)$	0.021	1.408

IV. SUMMARY

In this paper, we studied optimum scalar quantizer design for the Wyner-Ziv problem. Based on necessary conditions for optimality of an objective function which combines distortion and rate, we developed an iterative algorithm to find the optimum partitions numerically. It can be viewed as a generalization of the well-known Lloyd algorithm and can be used to design optimal scalar quantizers when the decoder does or does not have access to a noisy version of the source. When decoder side information is available, the complexity of the computation needed to solve the necessary conditions increases over that needed when the decoder side information is not available. Such difficulty is more evident when the number of partitions is large. For those cases, we developed high rate approximations based on the characteristics of the error function and prove that for a large number of partitions the traditional (single-source) optimal fixed-rate quantizers and entropy-constrained quantizers remain (near) optimal when decoder side information is available.

REFERENCES

[1] J. Max. Quantizing for mimimum distortion. *IRE Trans. Inform. Theory*, pages 7–12, Mar. 1960.

[2] S.P. Lloyd. Least squares quantization in PCM. *IEEE Trans. Inform. Theory*, pages 129–137, Mar. 1982.

[3] T. Berger. Optimum quantizers and permutation codes. *IEEE Trans. Inform. Theory*, pages 759–765, Nov. 1972.

[4] T. Berger. Minimum entropy quantizers and permutation codes. *IEEE Trans. Inform. Theory*, pages 149–157, Mar. 1982.

[5] R.C. Wood. On optimum quantization. *IEEE Trans. Inform. Theory*, pages 248–252, Mar. 1969.

[6] A. D. Wyner and J. Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inform. Theory*, pages 1–10, Jan. 1976.

[7] R. Zamir and S. Shamai. Nested linear/lattice codes for Wyner-Ziv encoding. *Proc. of IEEE Information Theory Workshop*, pages 92–93, June 1998.

[8] S. D. Servetto. Lattice quantization with side information. *Proc. of IEEE Data Compression Conf.*, pages 510–519, March 2000.

[9] S. S. Pradhan and K. Ramchandram. Distributed source coding using syndromes (DISCUS): Design and construction. *IEEE Tran. Inform. Theory*, pages 626–643, Mar. 2003.

[10] M. Fleming and M. Effros. Network vector quantization. *Proc. IEEE Data Compression Conference*, March 2001.

[11] M. Fleming Q. Zhao and M. Effros. Network vector quantization. *IEEE Trans. Inform. Theory*, pages 1584–1604, Aug. 2004.

[12] R. Zhang D. Rebollo-Monedero and B. Girod. Design of optimal quantizers for distributed source coding. *Proc. IEEE Data Compression Conference*, April 2003.

[13] H.V. Poor. *An introduction to signal detection and estimation*. Springer, NY, 1994.

[14] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, pages 471–480, July 1973.

[15] A. D. Liveris, Z. Xiong, and C. N. Georghiades. Distributed compression of binary sources using conventional parallel and serial concatenated convolutional codes. *Proc. of IEEE Data Compression Conference (DCC)*, Mar. 2003.

[16] A. Liveris, Z. Xiong, and C. N. Georghiades. Compression of binary sources with side information at the decoder using LDPC codes. *IEEE Communications Letters*, 6:440–442, Oct. 2002.

[17] Z. Tu, T. J. Li, and R.S. Blum. Compression of a binary source with side information using parallel concatenated convolutional codes. *Proc. of IEEE Globecom*, Dec. 2004.

[2]Since the partition is symmetric, we only list the non-negative end points here and throughout the rest of the paper.