

Optimization of Multiple Description Quantizers for Stochastic and Time-Varying Loss Probabilities

Morten Holm Larsen, Kristian Dalgaard Arnbak and Søren Vang Andersen
Department of Communication Technology, Aalborg University, Denmark
emails: {mhl,kdar00,sva}@kom.aau.dk

Abstract—This paper addresses the optimization of multiple description quantizers for stochastic and time-varying loss probabilities. We propose new iterative design methods, based on rate-distortion curves, for these cases. Experimental results indicate that significant fidelity improvements are possible using these methods.

I. INTRODUCTION

In multiple description (MD) coding an encoder describes a source sequence to a user at rate R_1 over channel one and at rate R_2 over channel two. Either of the channels may fail with given probabilities. Often, the design problem is formulated as a constrained optimization. Here, d_0 , the distortion when both channels work, is minimized subject to $d_i \leq D_i$, where d_i is the distortion when only channel i is operating, for $i = 1, 2$. Rate-distortion bounds for this two-channel problem with respect to the Gaussian source and mean-square error distortion criterion have been obtained by Ozarow [1] and scalar and vector quantizer design methods have been published by numerous authors [2], [6]–[9]. Alternatively, the design problem is formulated as a minimization of mean distortion for given loss probabilities in the two channels, assuming independent losses, a possibility explored in e.g. [3]–[5].

In numerous applications of MD quantizers, the exact loss probabilities for the channels are not known at design time. Rather, these probabilities are known only with stochastic uncertainty, or they are known to be time-varying quantities during application of the coding system. One such example is the use of MD for real-time media transmission using the Internet protocol (IP) and the related real-time protocol (RTP). In this application, packet loss probabilities can be estimated at the receiver and fed back to the transmitter via the real-time control protocol (RTCP) [10]. Because of the time-varying nature of the loss probabilities in real-world networks, and because of the estimation, sparse sampling, and delay of this information in the RTCP feedback, the loss probabilities available to an encoder will in this case be both time-varying and affiliated with stochastic uncertainty.

This paper addresses MD quantizer design for stochastic and time-varying loss probabilities. In the case of stochastic loss probabilities, defined by a probability density function (pdf), we optimize the MD quantizer to minimize a performance criterion over the support region of the pdf.

In the case of time-varying loss probabilities, we propose a system in which an MD quantizer is selected from a bank of quantizers, each quantizer optimized for a range of loss probabilities. To optimize this system, the design procedure for the stochastic case is extended such that the support of the pdf for each quantizer is the adjoint range of loss probabilities. The remaining optimization thus becomes that of partitioning the range of possible loss probabilities into a predefined number of optimized intervals. In our approach to these problems we make use of theoretic rate-distortion curves to device iterative optimization procedures without the need for redesigning the actual MD quantizers at each iteration. Thereby, our approach becomes feasible, even when the training of the MD quantizer in itself is computationally complex.

The remainder of this paper is organized as follows. In Section II the rate-distortion framework is outlined. Section III describes our procedure to minimize mean distortion and Section IV describes our procedure to minimize maximum performance loss compared to the case of deterministic loss probabilities. Numerical results for the rate-distortion framework and the resulting quantizers are given in Section V and Section VI, respectively. Finally, the conclusions on this work are drawn in Section VII.

II. OPTIMIZED MD DESIGN

Ozarow has derived the set of achievable quintuples $(R_1, R_2, d_1, d_2, d_0)$, for a unit variance i.i.d. Gaussian source and mean-square error [1]. The achievable quintuples must satisfy the relations,

$$d_1 > 2^{-2R_1} \quad (1)$$

$$d_2 > 2^{-2R_2} \quad (2)$$

$$d_0 > \frac{2^{-2(R_1+R_2)}}{1 - (\sqrt{\Pi} - \sqrt{\Delta})^2} \quad (3)$$

where $\Pi = (1 - d_1)(1 - d_2)$, $\Delta = d_1d_2 - 2^{-2(R_1+R_2)}$ and $\Pi \geq \Delta$. For simplicity we assume balanced description in rate ($R_1 = R_2$) and side distortion ($d_s = d_1 = d_2$). Designing an MD quantizer is usually formulated as a Lagrangian function, with the purpose to minimize the central distortion subject to the two side distortions,

$$L(d_0, d_s, \lambda) = d_0 + \lambda 2d_s. \quad (4)$$

For two independent erasure channels with equal loss probability p , the mean distortion \bar{d} can be written as

$$\bar{d}(p) = (1-p)^2 d_0 + 2p(1-p)d_s + p^2 \sigma^2, \quad (5)$$

where σ^2 is the variance of the source, which for convenience is scaled to one. The central and side distortions, d_0 and d_s , minimizing eq. (5) for a given loss probability \hat{p} are denoted by $d_0(\hat{p})$ and $d_s(\hat{p})$, respectively. These optimal distortions can be found by,

$$\begin{aligned} \{d_0(\hat{p}), d_s(\hat{p})\} &= \arg \min_{d_0, d_s} \left(d_0 + 2 \frac{\hat{p}}{1-\hat{p}} d_s + \frac{\hat{p}^2 \sigma^2}{(1-\hat{p})^2} \right) \\ &= \arg \min_{d_0, d_s} (d_0 + 2\lambda d_s). \end{aligned} \quad (6)$$

The minimization problem in eq. (6) is equivalent to the minimization of the Lagrange function in eq. (4), when $\lambda = \frac{\hat{p}}{1-\hat{p}}$. In line with common practice, we in this paper use Lagrange-like λ notation in the weighted minimization eq. (6).

Now, for a given loss probability \hat{p} and rate, the optimum solution set $\{d_0(\hat{p}), d_s(\hat{p})\}$ in eq. (6) can be determined numerically, when using the implicit relationship between d_s and d_0 in eq. (3). It is straight forward to show that \bar{d} is a convex function of d_s . Hence, when the numerical algorithm converges to a minimum, this minimum is the global minimum.

At this point we introduce the case where a static MD quantizer is optimized at loss probability \hat{p} and applied at a loss probability p . We denote the resulting distortion $\bar{d}_{stat}(p, \hat{p})$. This distortion can be written as,

$$\bar{d}_{stat}(p, \hat{p}) = (1-p)^2 d_0(\hat{p}) + 2p(1-p)d_s(\hat{p}) + p^2 \sigma^2. \quad (7)$$

A lower bound for this distortion is obtained when $\hat{p} = p$ for every application loss probability p . This lower bound corresponds to a scenario in which the rate-distortion optimum MD quantizer optimized for loss probability p is available and used for any given value of channel loss probability p . We therefore refer to this as the dynamically optimized mean distortion $\bar{d}_{Dyn}(p)$.

III. MINIMIZING MEAN DISTORTION

We now introduce the notion of a stochastic loss probability. That is, the loss probability p is seen as a stochastic variable defined by the pdf $f(p)$. We consider the design of K static MD quantizers, indexed by the set $\mathcal{K} = \{1, 2, \dots, K\}$, where the k 'th static MD quantizer is selected by a selection rule, $k = \Phi(p)$. In this framework, when $K = 1$ we address the stochastic loss probability. Conversely, when $K > 1$ we have the ability to adapt the quantization to a time-varying loss probability. For this to work, the selected index k must be transmitted as side information along with the quantization indexes. When the rate for transmitting k is low compared to R_1 and R_2 , the adjoint side-information overhead of this system can be neglected. In this system, we denote the side distortion

for the k 'th static MD quantizer as $d_s^{(k)}$ and the central distortion as $d_0^{(k)}$.

To begin the derivation, the mean distortion for the system can be written as

$$\bar{d} = \int \left((1-p)^2 d_0^{(k)} + 2p(1-p)d_s^{(k)} + p^2 \right) f(p) dp, \quad (8)$$

where $k = \Phi(p)$. Our overall goal is now to find the selection rule and the K distortion sets $\{d_0^{(k)}, d_s^{(k)}\}$ that minimize eq. (8). We can without losing generalization interpret the k 'th static MD design as an MD design that is optimized to an explicit packet loss probability \hat{p}_k . This enables us to use the notation introduced in Section II and to describe the K designs by a set of loss probabilities, $\hat{\mathcal{P}} = \{\hat{p}_1, \hat{p}_2, \dots, \hat{p}_K\}$. The central distortion $d_0^{(k)}$ and side distortion $d_s^{(k)}$ can be determined from \hat{p}_k by using eq. (6). Furthermore, we assume that the set of loss probabilities $\hat{\mathcal{P}}$ is ordered, $\hat{p}_1 < \hat{p}_2 < \dots < \hat{p}_K$. Using this notation, the mean distortion can be written as

$$\bar{d} = \int \bar{d}_{stat}(p, \hat{p}_k) f(p) dp, \quad k = \Phi(p). \quad (9)$$

To reach the overall goal, we start out by finding the optimum selection rule given K static MD quantizer designs. The selection rule is to select the k 'th MD design for a given p when,

$$\begin{aligned} k &= \arg \min_{i \in \mathcal{K}} \left((1-p)^2 d_0^{(i)} + 2p(1-p)d_s^{(i)} + p^2 \right) \\ &= \arg \min_{i \in \mathcal{K}} \bar{d}_{stat}(p, \hat{p}_i). \end{aligned} \quad (10)$$

Naturally, when $p = \hat{p}_k$ the selection rule in eq. (10) will select the k 'th MD design. Conversely, when p is in between \hat{p}_k and \hat{p}_{k+1} , one of the two MD designs will be selected. For a fixed MD design the mean distortion $\bar{d}_{stat}(p, \hat{p}_k)$ is a 2'nd order polynomial in p . Hence the selection rule as a function of p in eq. (10) is to choose between K 2'nd order polynomials. This reduces the selection rule to the following. When p is in between \hat{p}_k and \hat{p}_{k+1} the k 'th MD design is selected when p is below a certain boundary point t_k and above this boundary point the $k+1$ 'th design is selected. The boundary point is where the two MD designs perform the same, as expressed by,

$$\begin{aligned} (1-t_k)^2 d_0^{(k)} + 2t_k(1-t_k)d_s^{(k)} + t_k^2 &= \\ (1-t_k)^2 d_0^{(k+1)} + 2t_k(1-t_k)d_s^{(k+1)} + t_k^2. \end{aligned} \quad (11)$$

Therefore, the nonzero root of eq. (11) defines the desired boundary point:

$$t_k = \frac{d_0^{(k)} - d_0^{(k+1)}}{d_0^{(k)} - d_0^{(k+1)} - 2(d_s^{(k)} - d_s^{(k+1)})}. \quad (12)$$

We are now in a position to describe the entire selection rule as a set of partitions $\mathcal{V} = \{V_1, V_2, \dots, V_K\}$, where the k 'th MD design is selected when p lies in the partition V_k .

The partition V_k can in turn be described by its boundary points as follows.

$$V_k = [t_{k-1}; t_k], \quad \forall k \quad (13)$$

where t_0 and t_K are the lower and upper boundaries for the support region of the pdf, respectively.

Subsequently, we determine the optimum MD quantizers for a fixed selection rule. Since the selection rule can be specified by a set of partitions we can write the mean distortion as

$$\bar{d} = \sum_{k \in \mathcal{K}} \int_{V_k} \left((1-p)^2 d_0^{(k)} + 2p(1-p)d_s^{(k)} + p^2 \right) f(p) dp. \quad (14)$$

The mean distortion in eq. (14) is a sum of independent contributions from each MD quantizer. Thus, we can minimize eq. (14) by minimizing each contribution. By minimizing the contribution from the k 'th MD design we can determine the optimum distortion set $\{d_0^{(k)}, d_s^{(k)}\}$ as

$$\{d_0^{(k)}, d_s^{(k)}\} = \arg \min_{d_0, d_s} \int_{V_k} \left((1-p)^2 d_0 + 2p(1-p)d_s + p^2 \right) f(p) dp. \quad (15)$$

Now, since the integration in eq. (15) is a constant,

$$\begin{aligned} \{d_0^{(k)}, d_s^{(k)}\} &= \arg \min_{d_0, d_s} \left(d_0 + 2 \frac{\int_{V_k} p(1-p)f(p) dp}{\int_{V_k} (1-p)^2 f(p) dp} d_s \right) \\ &= \arg \min_{d_0, d_s} (d_0 + 2\lambda_k d_s). \end{aligned} \quad (16)$$

This minimization is equivalent to the minimization in eq. (6), where the MD design is optimized for an explicit loss probability \hat{p} . Thus, we can interpret the k 'th static MD design as an MD design that is optimized to an explicit packet loss probability \hat{p}_k , where

$$\hat{p}_k = \frac{\int_{V_k} p(1-p)f(p) dp}{\int_{V_k} (1-p)f(p) dp}. \quad (17)$$

One method to find the best joint \mathcal{V} and $\hat{\mathcal{P}}$ is by iterations reminiscent to those of Lloyd [11]. To do this, we define centroids by eq. (17) and do nearest neighbor partitioning using eq. (12) and (13). To terminate the iterations, we introduce a small positive number δ in a stop criterion. Next follows the outline of our proposed minimum-mean algorithm:

- (1) Initial guess of $\hat{\mathcal{P}}_0$, set counter $l \leftarrow 1$ and $E[\bar{d}(p, \hat{\mathcal{P}}_0)] \leftarrow \infty$.
- (2) Update \mathcal{V}_l using eq. (12) and eq. (13).
- (3) Calculate new $\hat{\mathcal{P}}_l$ from eq. (17).
- (4) If $\frac{E[\bar{d}(p, \hat{\mathcal{P}}_l)] - E[\bar{d}(p, \hat{\mathcal{P}}_{l-1})]}{E[\bar{d}(p, \hat{\mathcal{P}}_l)]} < \delta$ then stop, else $l \leftarrow l + 1$ and return to step (2).

Convergence of this algorithm can be shown using the same arguments as for the original Lloyd algorithm. One

question that raises naturally is if this two-step method always converges to the global minimum. We have not been able to show if this is indeed the case. However, experiments conducted with several different pdf's and several different initializations indicate, that for a given pdf the point of convergence is unchanged with the initialization. This indicates that the procedure does not stall at local minima.

IV. MINIMIZING MAXIMUM DISTORTION

When we apply the coding system at a given loss probability p , the distortion will be larger than that of a rate-distortion optimum system for this particular loss probability. This sub-optimality can be quantified as,

$$\Delta d(p, \hat{p}) = \bar{d}_{Stat}(p, \hat{p}) - \bar{d}_{Dyn}(p). \quad (18)$$

At design time we do not know which loss probability the system will be applied at. It is therefore of interest to minimize the sub-optimality, i.e., minimize the maximum gap in distortion between system performance and rate-distortion optimum performance over all p . The maximum distortion difference for partition V_k is given by,

$$\Delta d_{max,k} = \max_{p \in V_k} \Delta d(p, \hat{p}_k), \quad (19)$$

and the maximum distortion difference for all partitions can therefore be defined as,

$$\Delta D_{max} = \max_{k \in \mathcal{K}} \Delta d_{max,k}. \quad (20)$$

We assume that the pdf is fully supported in the interval $[t_0; t_K]$ and we note that, when minimizing the maximum distortion, only the support of the pdf influence the result. In the following, we outline a procedure to minimize eq. (20).

First, we observe that since the distortion gap in eq. (18) is not a convex function there is no simple solution in general. However, we argue that for $p \neq \hat{p}$ the distortion must be a positive and increasing function of $|p - \hat{p}|$. Hence for a given partition k the distortion gap $\Delta d(p, \hat{p}_k)$ will be maximum in its boundaries t_k and t_{k+1} , and will be denoted $\Delta d_{max,k}^{(1)} = \Delta d(t_{k-1}, \hat{p}_k)$ and $\Delta d_{max,k}^{(2)} = \Delta d(t_k, \hat{p}_k)$. This argument holds for all intervals except for the interval between the last centroid \hat{p}_K and t_K . In this interval the distortion gap is still positive, but for both $p = \hat{p}_K$ and $p = t_K$ the distortion gap equals zero. Thus a full examination is needed in this interval,

$$\Delta d_{max,K}^{(2)} = \max_{p \in (\hat{p}_K; t_K]} \Delta d(p, \hat{p}_K). \quad (21)$$

Our procedure for minimizing the overall distortion gap ΔD_{max} is to move one centroid at a time and update the boundary points t_k and t_{k+1} by using eq. (12), until the difference between the two candidates are under a given threshold, $|\Delta d_{max,k}^{(1)} - \Delta d_{max,k}^{(2)}| < \delta$. This procedure is repeated until this criterion is satisfied for all partitions.

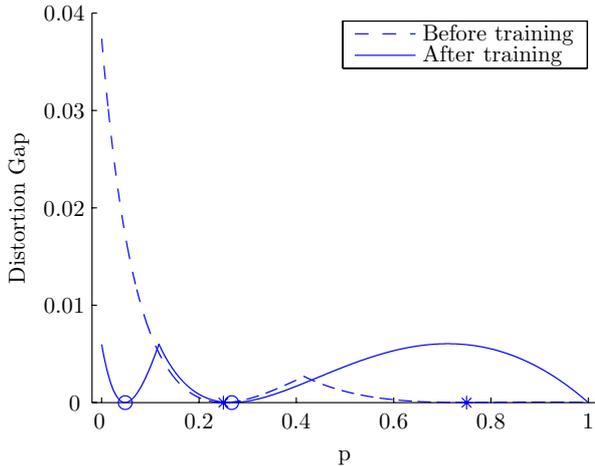


Fig. 1. Distortion gap $\Delta d(p, \hat{\mathcal{P}})$, for rate 1. The centroids before training are marked with stars and centroids after training are marked with circles.

The resulting algorithm is summarized as follows.

- (1) Make an initial guess of $\hat{\mathcal{P}}_0$ and calculate boundary points by using eq. (12), set partition counter $k \leftarrow 1$ and iteration counter $l \leftarrow 1$
- (2) Move centroid \hat{p}_k and update boundary points t_k and t_{k+1} until $|\Delta d_{max,k}^{(1)} - \Delta d_{max,k}^{(2)}| < \delta$.
- (3) If $k \neq K$ then update $k \leftarrow k + 1$ and goto step (2). Else set $k \leftarrow 1$ and $l \leftarrow l + 1$, and terminate if $\hat{\mathcal{P}}_l = \hat{\mathcal{P}}_{l+1}$.

An example of the distortion gap before and after this procedure as a function of p for two partitions is shown in Fig. 1. In this example, we see how the maximum distortion gap is decreased from $37 \cdot 10^{-3}$ to $6 \cdot 10^{-3}$ by the use of the minimize-maximum approach. We have not been able to show that the algorithm always converges to the global minimum. But again, by making random initializations of the algorithm, we have observed that the algorithm converges to the same solution, which is thus likely to be the global minimum. Since it is only the support of the pdf that influence the solution, the observation holds for all pdf's. In the following we analyze the performance gain in ΔD_{max} and compare benefits and drawbacks of this procedure with the minimize-mean procedure.

V. NUMERICAL RESULTS FOR RATE-DISTORTION OPTIMUM QUANTIZERS

Based on the methods described in the previous sections, we have carried out several numerical experiments. In these experiments, $f(p)$ is assumed uniform and a full pdf support in the interval $[0; 1]$. The two procedures are initialized with K equally distributed centroids $\hat{\mathcal{P}}$.

In Table I the mean distortion is given for different rates and K when using the two algorithms described in Sections III and IV. In Fig. 2 and Fig. 3 the maximum

TABLE I
COMPARISON OF MEAN DISTORTION FOR THE MINIMIZED-MEAN PROCEDURE SNR_{mean} AND MINIMIZED-MAXIMUM PROCEDURE SNR_{max}

	K	SNR_{mean}	SNR_{max}	Difference
Rate 1	1	3.3961 dB	3.3266 dB	$6.94 \cdot 10^{-2}$ dB
	2	3.4263 dB	3.4104 dB	$1.58 \cdot 10^{-2}$ dB
	4	3.4364 dB	3.4328 dB	$3.55 \cdot 10^{-3}$ dB
	8	3.4392 dB	3.4384 dB	$7.81 \cdot 10^{-4}$ dB
Rate 4	1	4.7473 dB	4.7428 dB	$4.48 \cdot 10^{-3}$ dB
	2	4.7480 dB	4.7466 dB	$1.42 \cdot 10^{-3}$ dB
	4	4.7483 dB	4.7479 dB	$3.99 \cdot 10^{-4}$ dB
	8	4.7484 dB	4.7483 dB	$9.86 \cdot 10^{-5}$ dB

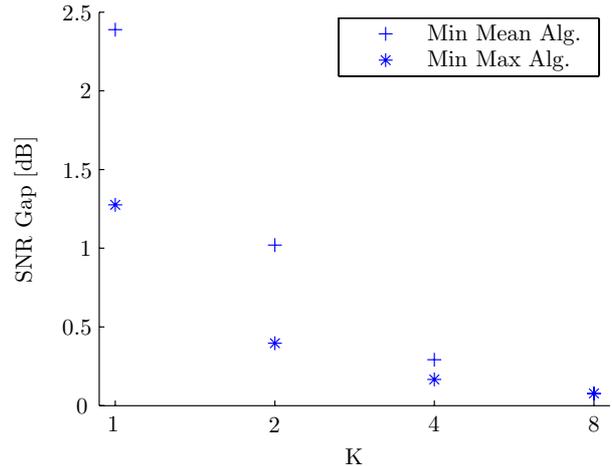


Fig. 2. Maximum distortion gap for rate 1.0.

distortion gap is analyzed for both the minimum-mean procedure and the minimum-maximum procedure.

VI. NUMERICAL RESULTS FOR REAL QUANTIZERS

In the two procedures described in Section III and Section IV a new MD quantizer could principally be designed each time a centroid is moved in the inner loop of the procedure. However, designing an MD quantizer following any of the procedures in [2], [6], [8] would have a prohibitive complexity even for moderately large vector dimensions and/or codebook sizes. We therefore suggest another approach, as explained in the following.

In an MD quantizer design the goal is to get arbitrary close to the theoretical lower-bound. We therefore use the rate-distortion lower bound as a performance model for the quantizer. Using this model allows us to directly use the two procedures, described in Section III and Section IV.

The outcome from any one of these procedures is a set of centroids $\hat{\mathcal{P}}$, and from these centroids we now design K quantizers by using principally any known design procedure for given loss probabilities, such as the ones suggested in [2], [6], [8]. By this use of a rate-distortion performance model to obtain the loss-probability centroids, the computational complexity of the overall design procedure becomes feasible.

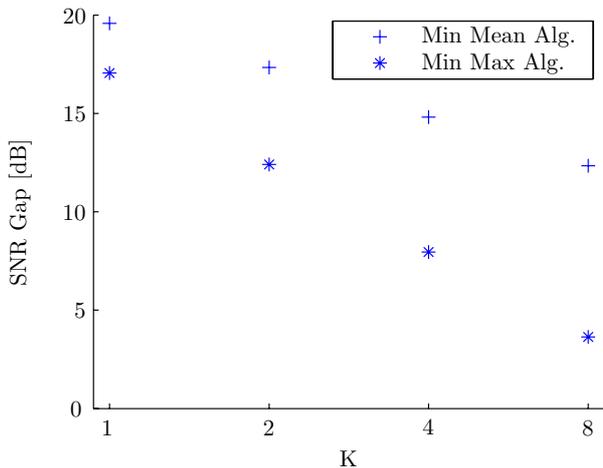


Fig. 3. Maximum distortion gap for rate 4.0.

TABLE II

THE MEAN DISTORTION AND SNR GAP FOR THE IMPLEMENTED MDVQ, FOR RATE 1.0 AND DIMENSION 4.

K	SNR_{max}	SNR Gap
1	2.8540 dB	0.8577 dB
2	2.9646 dB	0.2103 dB
4	2.9918 dB	0.1031 dB
8	2.9901 dB	0.0983 dB

In our numerical results, we use the deterministic annealing MD design procedure [8], because this procedure can design a quantizer close to the rate-distortion bound and avoids many local minima in the training. The mean distortion result for a multiple description vector quantizer (MDVQ) is shown in Table II. For evaluation of the maximum distortion gap, the performance of the dynamically optimized MDVQ design must be found. This bound can not be found in general, but by making a full sweep of 600 MDVQ designs over the possible \hat{p} interval, the dynamically optimized lower bound is well sampled and the evaluation can be conducted. The distortion gap for the MDVQ is shown in Table II.

VII. CONCLUSIONS

In this paper, we have presented algorithms to minimize the mean distortion and to minimize the maximum of the distortion gap for K static MD quantizers for the application with stochastic or time-varying loss probabilities. Our numerical results indicate that especially minimizing the maximum distortion gap leads to significantly improved performance of the optimized bank of MD quantizers. Furthermore, this can be accomplished while sacrificing only an insignificant quantity in terms of mean distortion. This optimization procedure is relevant e.g. in voice over IP systems when the packet loss probability is slowly varying and an objective is to provide the best possible coding system for a given loss probability. Specifically, results are given in Fig. 2 and Fig. 3, where we see a significant gain

when increasing the number of MD quantizers.

The proposed procedure to first optimize the design probabilities using the rate-distortion lower bound, and subsequently apply the result on a practical MD quantizer design, enable us to only design K MD quantizers. This makes the procedure feasible with respect to computational complexity. As an example, we presented results for an implemented MD quantizer in Table II. This verifies that our proposed method is capable of significantly reducing the maximum distortion for a fully feasible MD quantizer design.

REFERENCES

- [1] L. Ozarow, "On a Source-Coding Problem with Two Channels and Three Receivers", *The Bell System Technical Journal*, vol. 59, No. 10, Dec. 1980,
- [2] Vinay A. Vaishampayan, "Design of Multiple Description Scalar Quantizers", *IEEE Transactions on Information Theory*, vol. 39, No.3, May. 1993,
- [3] Jan Østergaard, Jesper Jensen and Richard Heusdens, "Entropy Constrained Multiple Description Lattice Vector Quantization", *IEEE ICASSP*, 2004,
- [4] Vivek K. Goyal and Jelena Kovacevic, "Generalized Multiple Description Coding With Correlating Transforms", *IEEE Transactions on Information Theory*, vol. 47, no. 6, 2001,
- [5] Vivek K. Goyal and Jonathan A. Kelner and Jelena Kovacevic, "Multiple Description Vector Quantization With a Coarse Lattice", *IEEE Transactions on Information Theory*, vol. 48, 2002,
- [6] Vinay A. Vaishampayan and Jaroslaw Domaszewicz, "Design of Entropy Constrained Multiple Description Scalar Quantizers", *IEEE Transactions on Information Theory*, vol. 39, No.1, Jan. 1994,
- [7] Vinay A. Vaishampayan and N.J.A. Sloane and Sergio D. Servetto, "Multiple Description Lattice Vector Quantization with Lattice Codebooks: Design and Analysis", *IEEE Transactions on Information Theory*, vol. 47, no. 5, July 2001,
- [8] Prashant Koulgi and Shankar L. Regunathan and Kenneth Rose, "Multiple Description Quantization by Deterministic Annealing", *IEEE Transactions on Information Theory*, vol. 49, no. 8, August 2003,
- [9] Hamid Jafarkhani and Vahid Tarokh, "Multiple Description Trellis Coded Quantization", *In Proceedings of International Conference on Image Processing*, vol. 1, 1998,
- [10] Henning Schulzrinne and Stephen L. Casner and Ron Frederick and Van Jacobson, "A Transport Protocol for Real-Time Applications (RFC 1889)", *IETF*, January, 1996,
- [11] Stuart P. Lloyd, "Least Squares Quantization in PCM", *IEEE Transactions on Information Theory*, IT-28, no. 2, March 1982,