Joint Source-Channel Decoding of Multiple Description Quantized Markov Sequences

Xiaolin Wu, Xiaohan Wang
Department of Electrical and Computer Engineering
McMaster University, Hamilton, ON, Canada, L8S 4K1
xwu@ece.mcmaster.ca/xhwang@grads.ece.mcmaster.ca

Jia Wang
Department of Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China 200240
jiawang@sjtu.edu.cn

Abstract

This paper proposes a framework for joint source-channel decoding of Markov sequences that are coded by a fixed-rate multiple description quantizer (MDQ), and transmitted via a lossy network. This framework is suited for lossy networks of primitive energy-deprived source encoders. Our technical approach is one of maximum *a posteriori* probability (MAP) sequence estimation that exploits both the source memory and the correlation between different MDQ descriptions. We solve the MAP estimation problem by computing the longest path in a weighted directed acyclic graph, at a complexity of $O(L^2NK)$, where N is the number of source symbols in the input sequence, K is the number of MDQ descriptions, and L is the number of codewords of the central quantizer. If the source sequence is Gaussian Markovian, the decoder complexity can be reduced to $O(LNK)$.

For MDQ-compressed Markov sequences impaired by both bit errors and erasure errors, the performance of joint source-channel MAP decoder can be 6dB higher than the conventional hard-decision decoder. Furthermore, the new MDQ decoding technique unifies the treatments of different subsets of the K descriptions available at the decoder, circumventing the thorny issue of requiring up to $2^K - 1$ MDQ side decoders.

1 Introduction

In many modern signal communication applications, such as those in wireless and sensor networks, the source encoders have to contend with limited power and primitive computing capability (“smart dust”) and at the same time attain resiliency in adverse network conditions. This calls for an asymmetric codec design that shifts the computational burden from encoders to decoders, considering that the decoder often situates at a central control site that has practically unbounded power supply and computing resources, and can furthermore fuse data collected from multiple channels. In this paper we study a particular coding system in the above design principle.

Assume that the source to be encoded and communicated via noisy channel(s) is a Markov sequence $\chi^{\mathcal{N}} = \chi_1, \chi_2, \cdots, \chi_{\mathcal{N}}$. Due to the constraints of battery capacity and computing power, the encoder has to do without sophisticated modeling and source coding operations and it is forced to forgo channel coding altogether. The maximum work load that the encoder can manage is to quantize (scalar or vector) $\chi^{\mathcal{N}}$ into $K \geq 2$ descriptions, and then send these descriptions through a lossy network, in fixed length code without entropy coding. To keep multiple description coding simple, multiple description scalar quantizer (MDSQ) or multiple description lattice vector quantizer (MDLVQ) should be used. We refer the reader to [1, 2] for information on MDSQ and MDLVQ, and to [3] on multiple description coding in general.

As a result of the expediency on the part of the encoder, the decoder is furnished with rich forms of statistical redundancy:

- the memory of the Markov sequence that is unexploited due to scalar coding or simple suboptimal block code (e.g., lattice VQ);
- residual source redundancy for lack of entropy coding;
- the correlation that is intentionally introduced between the K descriptions by MDQ.

The inquiry of this paper is how a decoder can best utilize these available redundancies to correct the channel errors.

Joint source-channel decoding of multiple description scalar quantized (MDSQ) Markov sequences was reported [4,5], in which the memory of the Markov source and the correlation between the descriptions of MDSQ are exploited in tandem. The main contribution of this paper is a new joint source-channel MDQ decoding technique that can simultaneously utilize the inter-description and intra-description redundancies. The new technique applies to MAP decoding of multiple description vector quantization (MDVQ) as well.

We first pose the problem as one of MAP sequence estimation, and then solve it by a graph theoretical algorithm of complexity $O(L^2NK)$, where N is the number of symbols in the input sequence, K is the number of side quantizers, and L is the number of codewords in the central quantizer. Moreover, for MDSQ-coded Gaussian Markov sequences we can reduce the algorithm complexity to $O(LNK)$, which is comparable to those of typical Vertibi-type decoding algorithms. Many authors published MAP estimation algorithms for joint source-channel decoding of single description quantized Markov sequences [6–9]. This paper is a generalization of these algorithms to joint source-channel MDQ decoding. Remarkably, it turns out that the generalized new MDQ decoding algorithm eliminates the need for $2^K - 1$ side decoders, which is a well-known operational difficulty associated with hard MDQ decoding process.

The paper is structured as follows. Section 2 formulates a general framework for joint source-channel MAP decoding of MDQ-coded Markov sequences. Section 3 presents a longest path algorithm for solving the MAP MDQ decoding problem. In Section 4, a more efficient solution is developed for Gaussian Markov sequences. Simulation results are reported in Section 5. Section 6 concludes.

2 Problem Formulation

Fig. 1 schematically depicts the proposed joint source-channel MDQ decoding system. The input to the system is a finite Markov sequence $\chi^{\mathcal{N}} = \chi_1, \chi_2, \cdots, \chi_{\mathcal{N}}$. A central quantizer

$q : \mathbb{R} \to \mathcal{C}$ maps a source symbol (MDSQ) or a block of source symbols (MDVQ) to a codeword in central codebook $\mathcal{C} = \{1, 2, \cdots, L\}$, where L is the number of codecells of the central quantizer. Let the codebooks of the K side quantizers be $\mathcal{C}_k = \{1, 2, \cdots, L_k\}$, where L_k is the number of codecells of side quantizer k, $L \le \prod_{k=1}^{K} L_k$ and $L_k \le L$, $k = 1, 2, \cdots, K$. The K-description MDQ is specified by an index assignment function $\alpha_k : \mathcal{C} \to \mathcal{C}_k$ [2]. The redundancy carried by the descriptions is reflected by a rate $1 - \log_2 L / \sum_{k=1}^{K} \log_2 L_k$.

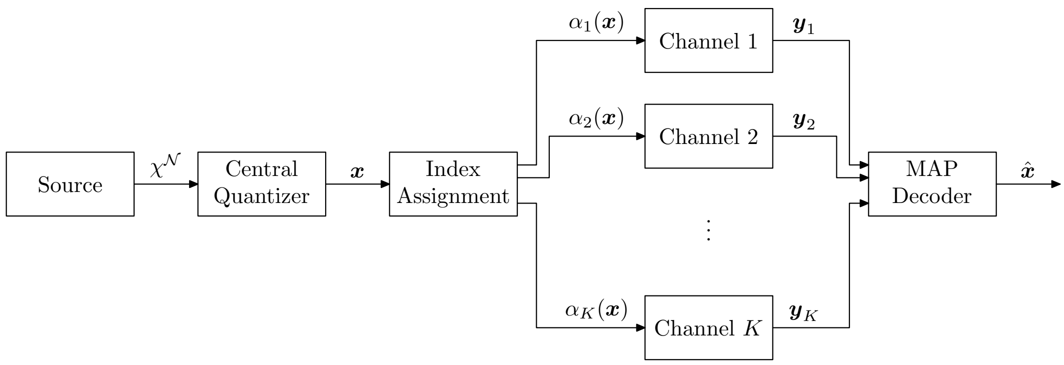

Figure 1: Block diagram of a MDSQ based communication system with a MAP decoder.

Let $\boldsymbol{x} = x_1 x_2 \cdots x_N \in \mathcal{C}^N$ be the output sequence of $\chi^{\mathcal{N}}$ produced by the central quantizer, $N = \mathcal{N}$ for MDSQ, or $N = \iota\mathcal{N}$ for MDVQ with ι being the VQ dimension. The K descriptions of MDQ, $\alpha_k(\boldsymbol{x}) \in \mathcal{C}_k^N$, $k = 1, 2, \cdots, K$, are transmitted via K noisy channels. We assume that the K noisy channels are memoryless, mutually independent, and do not introduce phase errors such as insertion or deletion of code symbols or bits. Consequently, a received description may have inversion or/and erasure errors, but it has the same number of bits as the one generated by MDQ. Denote the received code streams by $\boldsymbol{y}_k = y_{k,1} y_{k,2} \cdots y_{k,N}$, with $y_{k,n}$ being the n^{th} codeword of description k that is observed by the decoder.

Having the source and channel statistics and knowing the design of MDQ, the MDQ decoder can perform joint source-channel decoding of $\boldsymbol{y}_k$, $k = 1, 2, \cdots, K$, to best reconstruct $\boldsymbol{x}$. In a departure from the current practice of designing multiple side decoders (up to $2^K - 1$ of them!), we develop a single unified MDQ decoder that operates the same way regardless what subset of the K descriptions is available to the decoder. Our MDQ decoder takes the approach of MAP sequence estimation, and it reconstructs, given the observed sequences $\boldsymbol{y}_k$, ($k = 1, 2, \cdots, K$, some of which may be empty), the input sequence $\boldsymbol{x}$ such that the *a posteriori* probability $P(\boldsymbol{x} | \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_K)$ is maximized. Namely, the MDQ decoder emits

$$\begin{aligned} \hat{\boldsymbol{x}} &= \arg\max_{\boldsymbol{x} \in \mathcal{C}^*} P(\boldsymbol{x} | \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_K) \\ &= \arg\max_{\boldsymbol{x} \in \mathcal{C}^*} \log P(\boldsymbol{x} | \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_K). \end{aligned} \tag{1}$$

According to the Bayes' theorem,

$$
\begin{aligned}
P(\boldsymbol{x}|\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_K) &= \frac{P(\boldsymbol{x})P(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_K|\boldsymbol{x})}{P(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_K)} \\
&\overset{(a)}{\propto} P(\boldsymbol{x})P(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_K|\boldsymbol{x}) \\
&= P(\boldsymbol{x})P(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_K|\alpha_1(\boldsymbol{x}),\alpha_2(\boldsymbol{x}),\cdots,\alpha_K(\boldsymbol{x})) \\
&\overset{(b)}{=} P(\boldsymbol{x})\prod_{k=1}^{K} P(\boldsymbol{y}_k|\alpha_k(\boldsymbol{x})) \\
&\overset{(c)}{=} \prod_{n=1}^{N}\left\{P(x_n|x_{n-1})\prod_{k=1}^{K} P_k(y_{k,n}|\alpha_k(x_n))\right\}.
\end{aligned} \tag{2}
$$

In the above derivation, step (a) is due to the fact that $\boldsymbol{y}_1$ through $\boldsymbol{y}_K$ are fixed in the objective function; step (b) is because of the mutual independency of the K channels; and step (c) is under the assumption that $\boldsymbol{x}$, the output of the central quantizer, is first-order Markovian and the channels are memoryless. This assumption certainly holds, if the original source sequence $\chi^{\mathcal{N}}$ before MDQ is first-order Markovian, and it remains a good approximation for a high-order Markov sequence $\chi^{\mathcal{N}}$ as well, if $\chi^{\mathcal{N}}$ is vector quantized.

In (2) we let $P(x_1|x_0) = P(x_1)$ as convention. $P_k(\boldsymbol{b}'|\boldsymbol{b})$ is the probability of receiving a codeword $\boldsymbol{b} = b_1 b_2 \cdots b_B$ from channel k as $\boldsymbol{b}' = b'_1 b'_2 \cdots b'_B$. Because the channel is memoryless, we have

$$P_k(\boldsymbol{b}'|\boldsymbol{b}) = \prod_{i=1}^{B} P_k(b'_i|b_i). \tag{3}$$

Specifically, if the K channels can be modeled as memoryless error-and-erasure channels (EEC), where each bit is either transmitted intact, or inverted, or erased (the erasure can be treated as the substitution with a new symbol '\$'), then $\boldsymbol{b} \in \{0,1\}^B, \boldsymbol{b}' \in \{0,1,\$\}^B$ and

$$P_k(b'_i|b_i) = \begin{cases} p_\phi, & \text{if } b'_i = \$, \\ (1-p_\phi)(1-p_c), & \text{if } b'_i = b_i, \\ (1-p_\phi)p_c, & \text{otherwise}, \end{cases} \tag{4}$$

where p_ϕ is the erasure probability and p_c is the inversion or crossover probability.

If b_i is binary phase-shift keying (BPSK) modulated and transmitted through an additive white Gaussian noise (AWGN) channel, then

$$P_k(b'_i|b_i) = \frac{1}{\sqrt{\pi N_0}} e^{-(b'_i-b_i)^2/N_0}, \tag{5}$$

where N_0 is the noise power spectral density of the k^{th} channel.

The prior distribution $P(x)$ and transition probability matrix $P(x_n|x_{n-1})$ for the first-order Markov sequence $\boldsymbol{x}$ can be determined from the source distribution and the particular MDQ in question.

In the case of MDSQ, if the stationary probability density function of the source is $p_s(\chi)$ and the conditional probability density function is $p_s(\chi_n|\chi_{n-1})$, then

$$P(x) = \int_{\chi:q(\chi)=x_1} p_s(\chi)d\chi, \tag{6}$$

and

$$P(x_n|x_{n-1}) = \frac{\iint_{\substack{\chi_1:q(\chi_1)=x_n \\ \chi_2:q(\chi_2)=x_{n-1}}} p_s(\chi_1|\chi_2)p_s(\chi_2)d\chi_2 d\chi_1}{\int_{\chi:q(\chi)=x_{n-1}} p_s(\chi)d\chi}. \tag{7}$$

If MDLVQ is the source coder of the system, the transition probability matrix for $P(x_n|x_{n-1})$'s can be determined numerically either from a known close-form source distribution or from a training set.

3 Joint Source-Channel MDQ Decoding Algorithm

In this section, we devise a graph theoretical algorithm for joint source-channel MDQ decoding algorithm. Combining (1) and (2), we have

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\in\mathcal{C}^N} \sum_{n=1}^{N} \left\{ \log P(x_n|x_{n-1}) + \sum_{k=1}^{K} \log P_k(y_{k,n}|\alpha_k(x_n)) \right\}. \tag{8}$$

Because of the additivity of (8), we can structure the MAP estimation problem into the following subproblems:

$$w(n, x_n) = \max_{(x_1,x_2,\cdots,x_{n-1})\in\mathcal{C}^{n-1}} \sum_{i=1}^{n} \left[\log P(x_i|x_{i-1}) + \sum_{k=1}^{K} \log P_k(y_{k,i}|\alpha_k(x_i)) \right], \qquad x_n \in \mathcal{C}, \quad 1 \le n \le N. \tag{9}$$

Then, the solution of the optimization problem (1) is given by

$$\hat{\boldsymbol{x}} = \arg\max_{c\in\mathcal{C}} w(N, c). \tag{10}$$

The subproblems $w(\cdot,\cdot)$ can be expressed recursively as

$$\begin{aligned} w(n, x_n) &= \max_{(x_1,x_2,\cdots,x_{n-1})\in\mathcal{C}^{n-1}} \left\{ \sum_{i=1}^{n-1} \left[\log P(x_i|x_{i-1}) + \sum_{k=1}^{K} \log P_k(y_{k,i}|\alpha_k(x_i)) \right] \right. \\ &\qquad \left. + \log P(x_n|x_{n-1}) + \sum_{k=1}^{K} \log P_k(y_{k,n}|\alpha_k(x_n)) \right\} \\ &= \max_{c\in\mathcal{C}} \left\{ w(n-1, c) + \log P(x_n|c) \right\} + \sum_{k=1}^{K} \log P_k(y_{k,n}|\alpha_k(x_n)). \end{aligned} \tag{11}$$

The above recursion allows us to reduce the MAP estimation problem to one of finding the longest path in a weighted directed acyclic graph (WDAG) [8], which is given in Fig. 2. The underlying graph G has $LN + 1$ vertices, which consists of N stages with L vertices in each stage. Each stage corresponds to a codeword position in $\boldsymbol{x}$. Each vertex in a stage represents a possible codeword at the position. There is also one starting node s, corresponding to the beginning of $\boldsymbol{x}$.

We use a pair (n, x), $1 \le n \le N$, $x \in \mathcal{C}$ to label a node in G. From node $(n-1, b)$ to node (n, a), $a, b \in \mathcal{C}$, there is a directed edge, whose weight is

$$\log P(a|b) + \sum_{k=1}^{K} \log P_k(y_{k,n}|\alpha_k(a)).$$

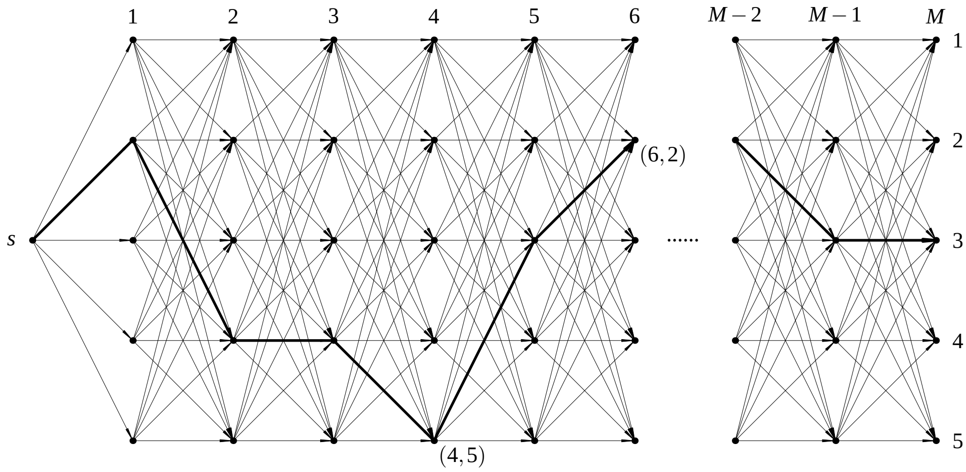

Figure 2: Graph G constructed for the joint source-channel MDQ decoding ($L = 5$).

From the starting node s to each node $(1, a)$, there is an edge whose weight is

$$\log P(a) + \sum_{k=1}^{K} \log P_k(y_{k,1}|\alpha_k(a)).$$

In graph G, the solution of the subproblem $w(n, a)$ is the weight of the longest path from the starting node s to node (n, a), which can be calculated recursively using dynamic programming. The MAP decoding problem is then converted into finding the longest path in graph G from the starting node s to nodes $(N, c), c \in \mathcal{C}$. By tracing back step by step to the starting node s, the MDQ decoder can reconstruct the input sequence $\boldsymbol{x}$ to $\hat{\boldsymbol{x}}$, the optimal result as defined in (1).

Now we analyze the complexity of the proposed algorithm. The dynamic programming algorithm proceeds from the starting node s to the nodes (N, c), through all LN nodes in G. The value of $w(n, a)$ can be evaluated in $O(L)$ time, according to (11). The quantities $\log P(a|b)$ and $\log P_k(y_{k,n}|\alpha_k(a))$ can be precomputed and stored in lookup tables so that they will be available to the dynamic programming algorithm in $O(1)$ time. Hence the term $\sum_{k=1}^{K} \log P_k(y_{k,n}|\alpha_k(a))$ in (11) can be computed in $O(K)$ time. Therefore, the total time complexity of this algorithm is $O(L^2NK)$. To reconstruct the input sequence, the selection in (11) should be recorded at each node, which results in a space complexity of $O(LN)$.

4 Monotonicity based Complexity Reduction

In [8] we proposed a monotonicity-based fast algorithm for the problem of MAP estimation of Markov sequences, in which the problem is transformed to one of matrix search and the monotonicity of the optimization objective function can be exploited to reduce the computational complexity dramatically. In this section, we show that the same technique can be applied to joint source-channel decoding of MDSQ-compressed Markov sequences under the same condition.

A two-dimensional matrix $A = A(a, b)$ is said to be *totally monotone* with respect to row maxima if the following relation holds:

$$A(a, b) \leq A(a, b') \Rightarrow A(a', b) \leq A(a', b'), \;\; a < a', b < b' \qquad (12)$$

If an $n \times n$ matrix A is totally monotone, then the row maxima of A can be found in $O(n)$ time [10]. A sufficient condition for (12) is

$$A(a, b') + A(a', b) \leq A(a, b) + A(a', b'), \ a < a', b < b', \tag{13}$$

which is also known as the Monge condition [10].

To exploit the total monotonicity for complexity reduction, we need to convert our recursion formula into a matrix search form [8]. We rewrite (11) as

$$w(n, a) = \max_{b \in \mathcal{C}} \left\{ w(n-1, b) + \log P(a|b) + \sum_{k=1}^{K} \log P_k(y_{k,n}|\alpha_k(a)) \right\}. \tag{14}$$

Then for each $1 \leq n \leq N$, we define an $L \times L$ matrix A_n such that

$$A_n(a, b) = w(n-1, b) + \log P(a|b) + \sum_{k=1}^{K} \log P_k(y_{k,n}|\alpha_k(a)). \tag{15}$$

To apply the fast algorithm to the joint source-channel MDSQ decoding problem, we check if matrix A_n satisfies the total monotonicity. Substituting A_n in (15) for A in (13), we have

$$\log P(a|b') + \log P(a'|b) \leq \log P(a|b) + \log P(a'|b'), \ a < a', b < b', \tag{16}$$

which is a sufficient condition for A_n to satisfy the total monotonicity and therefore, for the fast algorithm to be applicable. This condition, which depends only on the source statistics not the channels, is exactly the same as the one derived in [8]. It was shown by [8] that (16) holds if the source is Gaussian Markov, which includes a large family of signals studied in practice and theory.

Finally, we conclude that the time complexity of MAP decoding of MDSQ can be reduced to $O(LNK)$ for Gaussian Markov sequences. The linear dependency of the MAP decoding algorithm in the sequence length N and source codebook size L makes it comparable to the complexity of typical Vertibi-type decoders.

5 Simulation Results

We implemented the proposed MAP-based MDQ decoding algorithm and tested it on three first-order, zero-mean, unit-variance Gaussian Markov sequences with the correlation coefficient ρ being 0.1, 0.5 and 0.9 respectively. Two different two-description scalar quantizers (2DSQ) were used in our experiments, which are uniform and have the index assignment matrices shown in Fig. 3. One of them has $L = 15$ central codecells, and the other $L = 21$ codecells. For both 2DSQ's, the two side quantizers each has $L_1 = L_2 = 8$ codecells. The 2DSQ with two diagonals in its index assignment matrix has a stronger correlation between the two descriptions than the 2DSQ of three diagonals, i.e., the former has higher degree of redundancy than the latter.

For each description k, $k = 1, 2$, the codeword index $\alpha_k(x)$ is transmitted in fixed length code of three bits. The channel is simulated to be error-and-erasure channel with erasure probability p_ϕ and inversion probability p_c varying. The new MDQ decoding algorithm is compared with 1) MAP decoder for single description scalar quantization, and 2) conventional hard-decision MDQ decoder. The competing scalar quantizer (SQ) is uniform and

1							
2	3						
	4	5					
		6	7				
			8	9			
				10	11		
					12	13	
						14	15

(a) L=15

1	2						
3	4	6					
	5	7	8				
		9	10	11			
				12	14		
				13	15	16	
					17	18	20
						19	21

(b) L=21

Figure 3: The index assignments for two two-description scalar quantizers as proposed by [1].

has a fixed rate of six bits per sample so that it matches the 2DSQ's in rate and codecell structure. The system performance measure is the signal-to-noise ratio (SNR).

The simulation results are plotted in Fig. 4, 5 and 6. Over all values of ρ, p_c and p_ϕ, the joint source-channel MAP MDQ decoder outperforms the conventional hard-decision MDQ decoder, regardless the level of correlation between the two side descriptions. Not surprisingly, the performance gap between the two approaches increases as the amount of memory in the Markov source (ρ) increases. This is because the hard-decision MDQ decoder cannot benefit from the residual source redundancy in $\boldsymbol{x}$. The gap also increases as the erasure error probability p_ϕ increases, indicating that the MAP MDQ decoder can make a better use of inter-description correlation. Also, as expected, the MAP SQ decoder achieves higher SNR than the MAP MDQ decoder when the channel quality is very good, but the former loses to the latter as the channel condition deteriorates. This is when the redundancy of MDQ starts to pay off. More interestingly, we notice that joint source-channel MAP decoding of MDQ is advantageous even when the source memory is weak (see the curves for $\rho = 0.1$).

6 Conclusions

We proposed a framework for optimal (in MAP sense) joint source-channel decoding of Markov sequences compressed by fixed-rate MDQ. This framework allows both inter-description and intra-description correlations to be exploited for correcting bit errors as well as erasure errors. It is suitable for lossy communications involving low-power inexpensive encoders.

Efficient algorithms were developed to perform MAP decoding of MDQ at a cost comparable to that of Vertibi decoding. Operationally, the new MDQ decoding technique unifies the treatments of different subsets of descriptions available at the decoder, overcoming the difficulty of having a large number of side decoders that hinders the design of a good hard-decision MDQ decoder.

References

[1] V. A. Vaishampayan, “Design of multiple description scalar quantizers,” *IEEE Trans. on Information Theory*, vol. 39, no. 3, pp. 821–834, May 1993.

[2] S. D. Servetto, V. A. Vaishampayan, and N. J. A. Sloane, “Multiple description lattice vector quantization,” in *Data Compression Conference*, 1999, pp. 13–22.

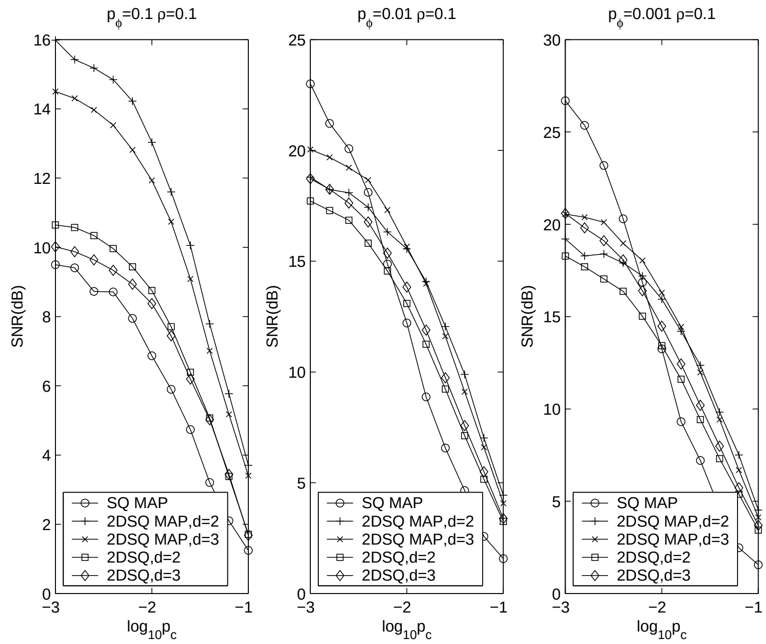

Figure 4: SNR performances of different MDQ decoders ($\rho = 0.1$), where d is the number of diagonals in the 2DSQ index assignment matrix.

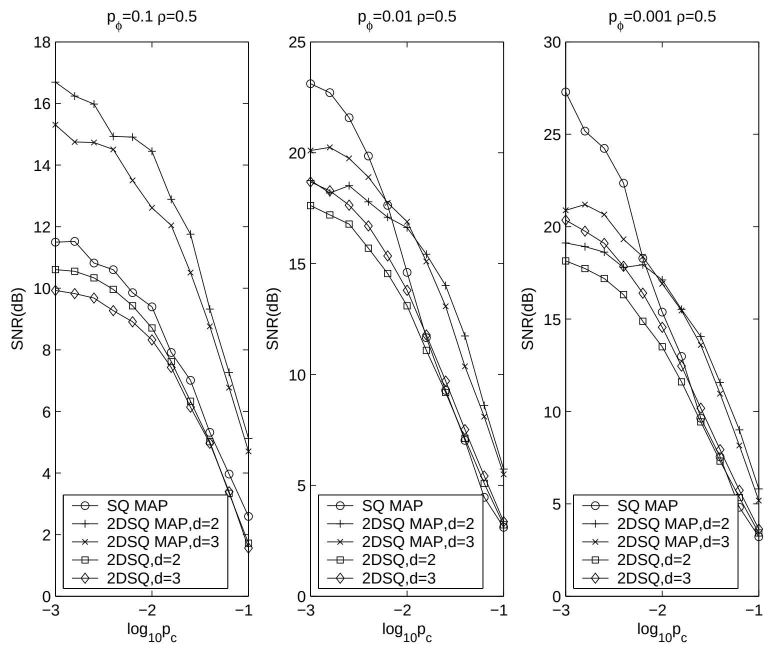

Figure 5: SNR performances of different MDQ decoders ($\rho = 0.5$)' where d is the number of diagonals in the 2DSQ index assignment matrix.

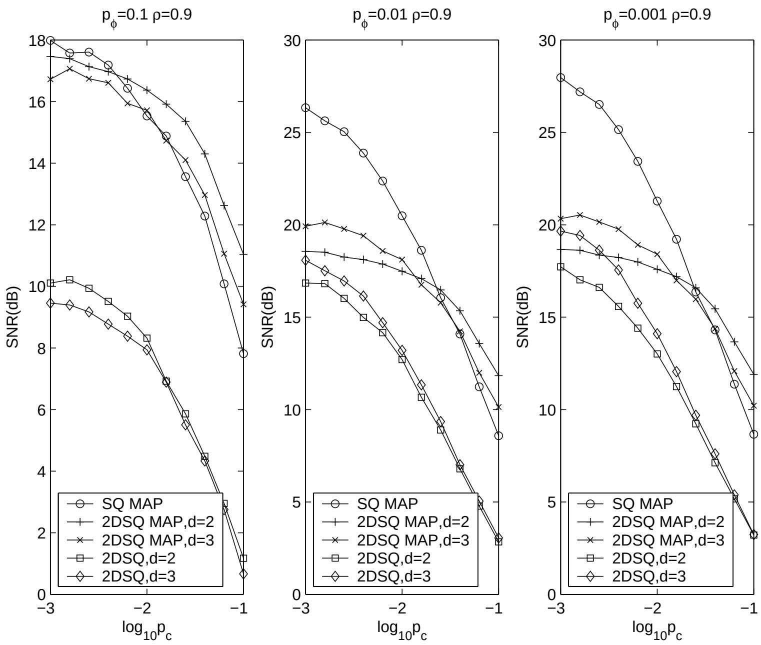

Figure 6: SNR performances of different MDQ decoders ($\rho = 0.9$), where d is the number of diagonals in the 2DSQ index assignment matrix.

[3] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, September 2001.

[4] T. Guionnet, C. Guillemot, and E. Fabre, "Soft decoding of multiple descriptions," in *IEEE International Conference on Multimedia, ICME*, vol. 2, Lausanne, Switzerland, August 26-29 2002, pp. 601–604.

[5] J. Barros, J. Hagenauer, and N. Goertz, "Turbo cross decoding of multiple descriptions," in *IEEE International Conference on Communications (ICC), 2002.*, vol. 3, 28 April-2 May 2002, pp. 1398 – 1402.

[6] M. Park and D. J. Miller, "Joint source-channel decoding for variable-length encoded data by exact and approximate MAP sequence estimation," *IEEE Trans. on Communications*, vol. 48, no. 1, pp. 1–6, January 2000.

[7] K. Sayood and J. C. Borkenhagen, "Use of residual redundancy in the design of joint source/channel coders," *IEEE Trans. on Communications*, vol. 39, no. 6, pp. 838–846, June 1991.

[8] X. Wu, S. Dumitrescu, and Z. Wang, "Monotonicity-based fast algorithms for MAP estimation of Markov sequences over noisy channels," *IEEE Trans. on Information Theory*, vol. 50, no. 7, pp. 1539–1544, July 2004.

[9] Z. Wang and X. Wu, "Length-constrained MAP decoding of variable length encoded Markov sequences," *IEEE Trans. on Communications*, (accepted).

[10] A. Aggarwal, M. Klave, S. Moran, P. Shor, and R. Wilber, "Geometric applications of a matrix-searching algorithm," *Algorithmica*, vol. 2, pp. 195–208, 1987.