# TECHNICAL RESEARCH REPORT

## *Quantization of Memoryless and Gauss-Markov Sources Over Binary Markov Channels*

*by N. Phamdo, F. Alajaji and N. Farvardin*

**T.R. 94-79**




*Sponsored by*
*the National Science Foundation*
*Engineering Research Center Program,*
*the University of Maryland,*
*Harvard University,*
*and Industry*

# Quantization of Memoryless and Gauss-Markov Sources Over Binary Markov Channels*

*Nam Phamdo*‡, *Fady Alajaji*† and *Nariman Farvardin*†

‡ Electrical Engineering Department
State University of New York
at Stony Brook
Stony Brook, NY 11794-2350

† Electrical Engineering Department
Institute for Systems Research
University of Maryland
College Park, MD 20742

**Abstract**

Joint source-channel coding for stationary memoryless and Gauss-Markov sources and binary Markov channels is considered. The channel is an additive-noise channel where the noise process is an $M$-th order Markov chain. Two joint source-channel coding schemes are considered. The first is a channel-optimized vector quantizer – optimized for both source and channel. The second scheme consists of a scalar quantizer and a maximum a posteriori detector. In this scheme, it is assumed that the scalar quantizer output has residual redundancy that can be exploited by the maximum a posteriori detector to combat the correlated channel noise. These two schemes are then compared against two schemes which use channel interleaving. Numerical results show that the proposed schemes outperform the interleaving schemes. For very noisy channels with high noise correlation, gains of 4 to 5 dB in signal-to-noise ratio are possible.

**Keywords:** Joint source-channel coding, Markov channel, vector quantization, MAP detection.

* The work of N. Phamdo is supported in part by NTT Corporation. The work of F. Alajaji and N. Farvardin is supported in part by National Science Foundation grant NSFD CD-8803012. Parts of this paper were presented at the 1994 Conference on Information Sciences and Systems, Princeton University, Princeton, NJ, March 1994.

# 1 Introduction

Source and channel coding are two problems that have traditionally been dealt with independently. This is due mainly to Shannon's source-channel separation principle [1], [2], which states that the two problems can be treated separately without loss of optimality. However, the separation principle holds only in the asymptotic case — when delay and complexity are not constrained. Recent works [3], [4], [5] have shown that, when delay and/or complexity are constrained, treating these problems jointly (i.e., joint source-channel coding) may result in improved performance over the traditional technique of tandem source-channel coding.

With the exception of [6] and [7], most of the previous work on joint source-channel coding has assumed that the channel is memoryless, disregarding the fact that real-world communication channels often have memory. In this work, we will consider two joint source-channel coding schemes for channels with memory. More specifically, the channel is assumed to be a binary stationary ergodic $M$-th order Markov channel derived from the Polya contagion urn model [8]. This is an additive-noise channel where the noise sample, $Z_i$, depends only on the *sum* of the previous $M$ noise samples $(Z_{i-1}, Z_{i-2}, \ldots, Z_{i-M})$. Memoryless sources with generalized Gaussian distributions and Gauss-Markov sources will be considered.

We first consider the design of a $k$-dimensional, rate $R$ bits/sample channel-optimized vector quantizer (COVQ) [9], [10] designed for the given source and channel. The COVQ encoder output is transmitted over the Markov channel. For each block of $k$ source samples, the COVQ encoder produces $kR$ bits for transmission. We assume that $kR$ is large enough with respect to $M$ $(kR > M)$ so that the memory in the channel can be exploited in $kR$ channel uses. Thus, by a proper design of the COVQ, we exploit the intra-block memory of the channel — but not the inter-block memory. The COVQ design algorithm is a straightforward extension of the algorithm described in [9] and [10], where the $2^{kR} \times 2^{kR}$ channel transition matrix is now given in terms of the transition probabilities of the Markov channel.

We then exploit both intra-block and inter-block memories of the channel. Here, we consider a scalar quantizer (SQ) designed for the noiseless channel. The SQ output is assumed to be redundant so that its entropy (in bits/channel use) is strictly less than the channel capacity (bits/channel use). After a proper assignment of binary indices to the SQ output, we transmit the indices directly over the channel. At the receiver, we exploit the redundancy of the SQ output and the memory of the channel through the use of a sequence maximum a posteriori (MAP) detector. This is analogous to previous works on MAP detection of a Markov source

over a memoryless channel [11], [12].

The performances of the two proposed schemes are compared against the performances of two interleaving schemes. In the interleaving systems, the Markov channel is rendered memoryless by an interleaver and de-interleaver[1]. Here, we assume that the source and channel codes are designed for the memoryless channel. Thus, the purpose of the interleaver and de-interleaver is to convert the Markov channel (with memory) into a memoryless channel. In the first interleaving scheme, we consider a COVQ designed for a memoryless channel with the same bit error rate as the Markov channel. This COVQ is then used over the interleaved channel (combination of interleaver, Markov channel and de-interleaver). This system is compared against the COVQ designed for the Markov channel. In the second interleaving system, we consider an SQ with its output transmitted over a memoryless (interleaved) channel. A sequence MAP detector, designed for the memoryless channel, is then used at the receiver. This scheme is compared against the MAP detection scheme operating directly on the Markov channel (without the use of interleaving).

The rest of this paper is organized as follows. In Section II, we present the Markov channel model. The two joint source-channel coding schemes are described in Section III. Simulation results are provided in Section IV. In Section V, comparisons between the proposed schemes and the corresponding interleaving schemes are made. Finally, the conclusions are stated in Section VI.

# 2 Channel Model

Consider a discrete channel with memory, with common input, noise and output binary alphabets and described by the following equation: $Y_i = X_i \oplus Z_i$, for $i = 1, 2, 3, \ldots$ where:

- $\oplus$ represents the addition operation modulo 2.
- The random variables $X_i$, $Z_i$ and $Y_i$ represent, respectively, the input, noise and output of the channel.
- $\{X_i\} \perp \{Z_i\}$, i.e., the input and noise sequences are independent from each other.
- The noise process $\{Z_i\}_{i=1}^{\infty}$ is a stationary mixing (hence ergodic) Markov process of order $M$. By this we mean that the noise sample, $Z_i$, depends only on the previous $M$ noise

[1]* It is assumed that the interleaver and de-interleaver are *ideal* so that the Markov channel is perfectly rendered memoryless.

samples, i.e., for $i \geq M+1$,

$$\Pr\{Z_i = e_i | Z_1 = e_1, \ldots, Z_{i-1} = e_{i-1}\} = \Pr\{Z_i = e_i | Z_{i-M} = e_{i-M}, \ldots, Z_{i-1} = e_{i-1}\}.$$

We assume that the marginal distribution of the noise process is given by

$$\Pr\{Z_i = 1\} = \epsilon = 1 - \Pr\{Z_i = 0\},$$

where $\epsilon \in [0, 1/2)$ is the channel bit error rate (BER). Furthermore, we assume that the process $\{Z_i\}$ is generated by the finite-memory contagion urn model described in [8]. According to this model, the noise sample $Z_i$ depends only on the *sum* of the previous $M$ noise samples[2]. Thus, for $i \geq M+1$,

$$\begin{aligned}\Pr\{Z_i = 1 | Z_{i-M} = e_{i-M}, \ldots, Z_{i-1} = e_{i-1}\} &= \Pr\{Z_i = 1 | \sum_{j=i-M}^{i-1} Z_j = \sum_{j=i-M}^{i-1} e_j\} \\ &= \frac{\epsilon + (\sum_{j=i-M}^{i-1} e_j)\delta}{1 + M\delta},\end{aligned}$$

where $e_j = 0$ or 1, for $j = i-M, \ldots, i-1$. The non-negative parameter $\delta$ determines the amount of correlation in $\{Z_i\}$. The correlation coefficient of the noise process is $\delta/(1+\delta)$. Note that if $\delta = 0$, the noise process $\{Z_i\}$ becomes independent and identically distributed (i.i.d.) and the resulting additive noise channel reduces to a binary symmetric channel (BSC). Finally, we note that the channel is entirely characterized by three parameters: $\epsilon$, $\delta$ and $M$.

The above channel model offers an interesting and less complex alternative to the Gilbert-Elliott model and others [13].

## 2.1 Distribution of the Noise

For an input block $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ and an output block $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$, we denote the block channel transition probability matrix $\Pr\{\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{x}\}$ by $Q(\mathbf{y}|\mathbf{x})$.

- For block length $n \leq M$, we have [8]:

[2]For $M = 1$, the model is general, i.e., it can represent any binary first-order Markov chain with positive transition probabilities.

$$Q(\mathbf{y}|\mathbf{x}) = L(n, d, \epsilon, \delta),$$

where $d = d_H(\mathbf{x}, \mathbf{y})$ is the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$ and

$$L(n, d, \epsilon, \delta) = \frac{\left[\prod_{i=0}^{d-1}(\epsilon + i\delta)\right]\left[\prod_{j=0}^{n-d-1}(1 - \epsilon + j\delta)\right]}{\left[\prod_{l=0}^{n-1}(1 + l\delta)\right]}.$$

- For $n \geq M + 1$, we have [8]:

$$\begin{aligned} Q(\mathbf{y}|\mathbf{x}) &= \Pr\{\mathbf{Z} = \mathbf{e}\} \\ &= L(M, s, \epsilon, \delta) \prod_{i=M+1}^{n} \left[\frac{\epsilon + s_i\delta}{1 + M\delta}\right]^{e_i} \left[1 - \frac{\epsilon + s_i\delta}{1 + M\delta}\right]^{1-e_i}, \end{aligned} \tag{1}$$

where $\mathbf{e} = (e_1, e_2, \ldots, e_n)$, $e_i = x_i \oplus y_i$, $s = e_1 + \cdots + e_M$ and $s_i = e_{i-1} + \cdots + e_{i-M}$.

## 2.2 Capacity of the Channel

The capacity $C$ of this channel is given by [8]:

$$C = 1 - \sum_{s=0}^{M} \binom{M}{s} L(M, s, \epsilon, \delta) h_b\left(\frac{\epsilon + s\delta}{1 + M\delta}\right)$$

where $h_b(x) = -x\log_2(x) - (1 - x)\log_2(1 - x)$ is the binary entropy function. The capacity is monotonically increasing with $\delta$ (for fixed $\epsilon, M$) and $M$ (for fixed $\epsilon, \delta$), and monotonically decreasing with $\epsilon$ (for fixed $\delta, M$).

# 3 Joint Source-Channel Coding Schemes

## 3.1 Channel Optimized Vector Quantizer (COVQ)

The ensuing formulation of COVQ follows that of [10]. Consider a real-valued stationary and ergodic source, $\mathcal{V} = \{V_i\}_{i=1}^{\infty}$. The source is to be encoded by a $k$-dimensional, $n$-bit/vector COVQ whose output is to be transmitted over the binary Markov channel. The coding system, depicted in Figure 1, consists of an encoder mapping, $\gamma$, and a decoder mapping, $\beta$. The


Figure 1: Block Diagram of a COVQ System.

encoder mapping $\gamma : \mathbb{R}^k \mapsto \{0,1\}^n$ is described in terms of a partition $\mathcal{P} = \{S_{\mathbf{x}} \subset \mathbb{R}^k : \mathbf{x} \in \{0,1\}^n\}$ of $\mathbb{R}^k$ according to

$$\gamma(\mathbf{v}) = \mathbf{x} \quad \text{if } \mathbf{v} \in S_{\mathbf{x}}, \ \ \mathbf{x} \in \{0,1\}^n,$$

where $\mathbf{v} = (v_1, v_2, \ldots, v_k)$ is a block of $k$ successive source samples. The channel takes an input sequence $\mathbf{x}$ and produces an output sequence $\mathbf{y}$. It is given in terms of the block channel transition matrix $Q(\mathbf{y}|\mathbf{x})$. Finally, the decoder mapping $\beta : \{0,1\}^n \mapsto \mathbb{R}^k$ is described in terms of a codebook $\mathcal{C} = \{\mathbf{c_y} \in \mathbb{R}^k : \mathbf{y} \in \{0,1\}^n\}$ according to

$$\beta(\mathbf{y}) = \mathbf{c_y}, \qquad \mathbf{y} \in \{0,1\}^n.$$

The encoding rate of the above system is $R = n/k$ bits/sample and its average squared-error distortion per sample is given by [10]:

$$D = \frac{1}{k} \sum_{\mathbf{x}} \int_{S_{\mathbf{x}}} f(\mathbf{v}) \left\{ \sum_{\mathbf{y}} Q(\mathbf{y}|\mathbf{x}) ||\mathbf{v} - \mathbf{c_y}||^2 \right\} d\mathbf{v}, \tag{2}$$

where $f(\mathbf{v})$ is the $k$-dimensional source probability density function (p.d.f.). For a given source, channel, $k$ and $n$, we wish to minimize $D$ by a proper choice of $\mathcal{P}$ and $\mathcal{C}$.

From (2), we see that for a fixed $\mathcal{C}$ the optimal partition $\mathcal{P}^* = \{S^*_{\mathbf{x}}\}$ is given by [10]:

$$S^*_{\mathbf{x}} = \left\{ \mathbf{v} : \sum_{\mathbf{y}} Q(\mathbf{y}|\mathbf{x}) ||\mathbf{v} - \mathbf{c_y}||^2 \right.$$

$$\left. \leq \sum_{\mathbf{y}} Q(\mathbf{y}|\tilde{\mathbf{x}}) ||\mathbf{v} - \mathbf{c_y}||^2, \forall \tilde{\mathbf{x}} \in \{0,1\}^n \right\}, \tag{3}$$

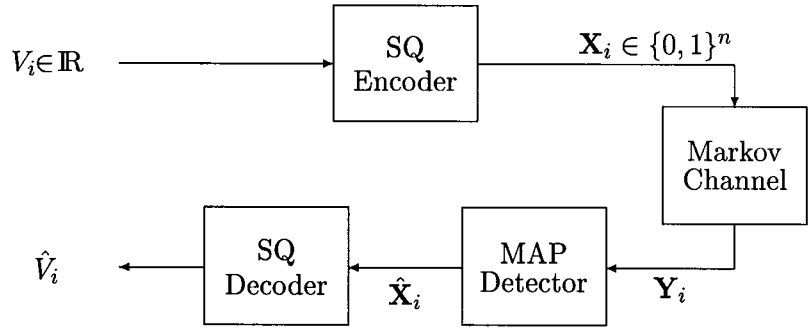


Figure 2: Block Diagram of Joint Source-Channel Coding System Using MAP Detection (SQ-MAP).

$\mathbf{x} \in \{0,1\}^n$. Similarly, the optimal codebook $\mathcal{C}^* = \{\mathbf{c}_\mathbf{y}^*\}$ for a given partition is [10]:

$$\mathbf{c}_\mathbf{y}^* = \frac{\sum_\mathbf{x} Q(\mathbf{y}|\mathbf{x}) \int_{S_\mathbf{x}} \mathbf{v} f(\mathbf{v}) d\mathbf{v}}{\sum_\mathbf{x} Q(\mathbf{y}|\mathbf{x}) \int_{S_\mathbf{x}} f(\mathbf{v}) d\mathbf{v}}. \tag{4}$$

The COVQ design algorithm is a straightforward extension of the iterative algorithm in [10], [14]. The algorithm starts out with an initial codebook, $\mathcal{C}^{(0)}$. With this fixed, it finds the optimal partition, $\mathcal{P}^{(1)}$, using (3). With $\mathcal{P}^{(1)}$ fixed, it uses (4) to find the optimal codebook, $\mathcal{C}^{(1)}$. This procedure is repeated until the relative change in distortion is sufficiently small. Note that the average distortion, $D$, forms a monotonically non-increasing sequence. Thus, the algorithm is guaranteed to converge to a locally optimal solution (since $D \geq 0$). We will assume that $n \geq M + 1$. Therefore, the block channel transition matrix, $Q(\mathbf{y}|\mathbf{x})$, will always be given by (1).

## 3.2 MAP Detection

Next consider the system depicted in Figure 2. Here, instead of using COVQ we use a scalar quantizer (SQ). The SQ is also described by $\gamma$ and $\beta$ as above — except that $k = 1$ and $R = n$ bits/sample. Instead of optimizing the SQ for the Markov channel, we make use of the residual redundancy of the SQ to combat channel noise. This is in the spirit of the approaches in [5], [11] and [12].

The SQ in Figure 2 is designed using the Lloyd-Max formulation [15], [16] which assumes

the channel is noise-free.

Since the source, $\mathcal{V} = \{V_i\}_{i=1}^{\infty}$, is stationary and $\mathbf{X}_i$ is a function of $V_i$ for each $i$, the SQ encoder output process, $\mathcal{X} = \{\mathbf{X}_i\}_{i=1}^{\infty}$, is also stationary. Let $H_\infty(\mathcal{X})$ be the entropy rate of $\mathcal{X}$ and $H(\mathbf{X}_1)$ be the entropy of $\mathbf{X}_1$. Define

$$\begin{aligned} \rho_D &\triangleq R - H(\mathbf{X}_1), \\ \rho_M &\triangleq H(\mathbf{X}_1) - H_\infty(\mathcal{X}), \\ \rho_T &\triangleq \rho_D + \rho_M = R - H_\infty(\mathcal{X}), \end{aligned}$$

as the redundancy due to the non-uniformity of the distribution, the redundancy due to the memory and the total redundancy, respectively. We will assume that $\rho_T > 0$. In the following, we will make use of the redundancy in $\mathcal{X}$ (the amount of which is measured by $\rho_T$) to combat channel errors. This is accomplished by using a sequence MAP detector. We first consider the case where $\mathcal{V}$ is i.i.d.

If $\mathcal{V}$ is i.i.d., so is $\mathcal{X}$. Thus $\rho_M = 0$ and the only redundancy is that due to the non-uniform distribution of $\mathbf{X}_i$. In this case the sequence MAP detector is described as follows. The sequence MAP detector observes a sequence $\mathbf{y}^N = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N) \in \{0,1\}^{nN}$ and makes an estimate of the sequence $\mathbf{x}^N = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \in \{0,1\}^{nN}$ according to

$$\hat{\mathbf{x}}^N = \arg\max_{\mathbf{x}^N} \Pr\{\mathbf{X}^N = \mathbf{x}^N | \mathbf{Y}^N = \mathbf{y}^N\}.$$

It can be easily shown that (see [12]) if $n \geq M$,

$$\begin{aligned} \hat{\mathbf{x}}^N = \arg\max_{\mathbf{x}^N} \Big\{ & \log\left[Q(\mathbf{y}_1|\mathbf{x}_1)p(\mathbf{x}_1)\right] \\ & + \sum_{i=2}^{N} \log\left[\tilde{Q}(\mathbf{e}_i|\mathbf{e}_{i-1})p(\mathbf{x}_i)\right] \Big\}, \end{aligned} \tag{5}$$

where $\mathbf{e}_i = \mathbf{x}_i \oplus \mathbf{y}_i \in \{0,1\}^n$, $\tilde{Q}(\mathbf{e}_i|\mathbf{e}_{i-1}) = \Pr\{\mathbf{Z}_i = \mathbf{e}_i | \mathbf{Z}_{i-1} = \mathbf{e}_{i-1}\}$, and $p(\mathbf{x}_i) = \Pr\{\mathbf{X}_i = \mathbf{x}_i\}$. Here, $p(\mathbf{x}_i)$ is determined by integrating the marginal source p.d.f. $f(v_i)$ on the interval $S_{\mathbf{X}_i}$. Note that for $i \geq 2$,

$$\tilde{Q}(\mathbf{e}_i|\mathbf{e}_{i-1}) = \prod_{j=n(i-1)+1}^{ni} \left[\frac{\epsilon + s_j\delta}{1+M\delta}\right]^{e_j} \left[1 - \frac{\epsilon + s_j\delta}{1+M\delta}\right]^{1-e_j},$$

where $\mathbf{e}_{i-1} = (e_{n(i-2)+1}, e_{n(i-2)+2}, \ldots, e_{n(i-1)})$, $\mathbf{e}_i = (e_{n(i-1)+1}, e_{n(i-1)+2}, \ldots, e_{ni})$ and $s_j = e_{j-1} + \cdots + e_{j-M}$.

As expressed in (5), the sequence MAP detector can be implemented using a modified version of the Viterbi algorithm, where $\mathbf{x}_i$ is the state at time instant $i$. The trellis has $2^n$ states with $2^n$ branches leaving and entering each state. For a branch leaving state $\mathbf{x}_{i-1}$ and entering state $\mathbf{x}_i$, the path metric is $\log[\tilde{Q}(\mathbf{x}_i \oplus \mathbf{y}_i | \mathbf{x}_{i-1} \oplus \mathbf{y}_{i-1}) p(\mathbf{x}_i)]$.

If $\mathcal{V}$ has memory, it is in general difficult to characterize the memory of $\mathcal{X}$. However, in this case, we will make a simplifying assumption that $\mathcal{X}$ forms a first-order discrete Markov chain[3] with transition probability matrix

$$P(\mathbf{x}_i | \mathbf{x}_{i-1}) = \Pr\{\mathbf{X}_i = \mathbf{x}_i | \mathbf{X}_{i-1} = \mathbf{x}_{i-1}\}, \tag{6}$$

$\mathbf{x}_i, \mathbf{x}_{i-1} \in \{0,1\}^n$. In this case, (5) is replaced by

$$\hat{\mathbf{x}}^N = \arg\max_{\mathbf{x}^N} \Big\{ \log\left[Q(\mathbf{y}_1 | \mathbf{x}_1) p(\mathbf{x}_1)\right] + \sum_{i=2}^{N} \log\left[\tilde{Q}(\mathbf{e}_i | \mathbf{e}_{i-1}) P(\mathbf{x}_i | \mathbf{x}_{i-1})\right] \Big\}, \tag{7}$$

and the path matrix from state $\mathbf{x}_{i-1}$ to state $\mathbf{x}_i$ is $\log[\tilde{Q}(\mathbf{x}_i \oplus \mathbf{y}_i | \mathbf{x}_{i-1} \oplus \mathbf{y}_{i-1}) P(\mathbf{x}_i | \mathbf{x}_{i-1})]$. The transition probability matrix, $P(\mathbf{x}_i | \mathbf{x}_{i-1})$, is determined by measuring the relative frequency of occurences of a long training sequence (640,000 source samples). From here on, the above scheme will be referred to as SQ-MAP. We note that the complexity and delay of SQ-MAP is due mainly to the MAP detector.

In some special circumstances, the output of the MAP detector will always be identical to its input. In such cases, we say that the MAP detector is useless. As an example, when $\mathcal{X}$ is i.i.d. and $M = n = 1$, it is shown in [17] that the MAP detector is useless if

$$\left[\frac{1 - \epsilon + \delta}{\epsilon + \delta}\right] \left[\frac{1-p}{p}\right] \geq 1, \tag{8}$$

where $p = \Pr\{X = 0\} \in (1/2, 1]$. If (8) does not hold, then the sequence MAP detector will be useful for sufficiently large $N$ [17]. Detailed analyses of the sequence MAP detector are given in [17] for the case of $M = n = 1$.

[3]Note that this assumption is not accurate in general — even in the case where $\mathcal{V}$ is a first-order Gauss-Markov source.

In this paper, we are mainly interested in cases where $M = 1$ and $n > 1$. In these cases, little is known about the usefulness of the MAP detector. However, an important factor contributing to the performance of the MAP detector is how the binary codewords are assigned to the SQ quantization levels. This issue will be discussed in the following section.

# 4 Numerical Results

## 4.1 Memoryless Sources

In this subsection, we will assume that the source is i.i.d. with distribution given by

$$f(v) = \frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)} \exp\{-[\eta(\alpha,\sigma)|v|]^{\alpha}\}, \tag{9}$$

where $\eta(\alpha,\sigma) = \sigma^{-1}[\Gamma(3/\alpha)/\Gamma(1/\alpha)]^{-1/2}$, $\alpha > 0$ is the exponential rate of decay and $\sigma^2$ is distribution variance. Note that for $\alpha = 2$ the above is the Gaussian p.d.f. For $\alpha = 1$, it is the Laplacian p.d.f. Any i.i.d. source with distribution given by (9) is referred to as a generalized Gaussian source.

Numerical results for COVQ over binary Markov channels with $\delta = 10$ and $M = 1$ and generalized Gaussian sources with shape parameter $\alpha = 0.5$, 1 and 2 are presented in Tables 1, 2 and 3, respectively. Signal-to-noise ratio (SNR) performances are given in dB for rates $R = 2$, 3 and 4 bits/sample and channel BER $\epsilon = 0.0$, 0.005, 0.01, 0.05 and 0.1. Also provided in Tables 1-3 are the optimal performances theoretically attainable (OPTA) obtained by evaluating $D(RC)$, where $D(\cdot)$ is the distortion-rate function of the source for the squared-error distortion measure and $C$ is the channel capacity in bits per channel use.

The COVQ results were obtained from 500,000 training vectors. A vector quantization codebook (optimized for the noiseless channel) with codewords assigned by a simulated annealing algorithm (described in [18]) is chosen as the initial codebook for the COVQ with $\epsilon$=0.005. The final codebook for $\epsilon$=0.005 is chosen as the initial codebook for $\epsilon$=0.01, and so on.

Simulation results for SQ-MAP are given in Tables 4-6. The simulations were run 100 times, with $N$=1000 source samples used in each run. The average distortion, averaged over the 100 runs, is given in dB. The SQ's used in the simulations were *symmetric* Lloyd-Max scalar quantizers. As mentioned earlier, how the quantization levels are mapped to binary codewords is an important consideration. We have examined two codeword assignments: the natural binary code (NBC) and the folded binary code (FBC). An example of these two codes

is illustrated in Figure 3. Note that the least significant bit (LSB) is the leftmost bit. Also, the FBC sign bit is the LSB. From our observations, FBC consistently outperforms NBC. FBC was used in the SQ-MAP results in Tables 4, 5 and 6.

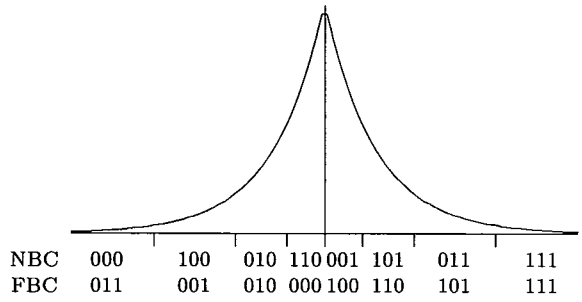


Figure 3: NBC and FBC Codeword Assignments for an 8-Level Lloyd-Max Scalar Quantizer; Generalized Gaussian Source with Shape Parameter $\alpha = 1$.

Note that, when $M$=1, $\tilde{Q}(\mathbf{e}_i|\mathbf{e}_{i-1})$ depends only on $\mathbf{e}_i$ and $e_{n(i-1)}$ (most significant bit (MSB) of $\mathbf{e}_{i-1}$). Thus, for fixed $\mathbf{y}^N$, the path metric from state $\mathbf{x}_{i-1}$ to state $\mathbf{x}_i$ depends only on $\mathbf{x}_i$ and $x_{n(i-1)}$ (MSB of $\mathbf{x}_{i-1}$). Therefore, the MSB of the binary codeword plays an important role in the Viterbi search. Now note that, because of symmetry, the MSB of NBC is 0 or 1 with equal probability. Hence, the MSB of NBC has zero redundancy. FBC, on the other hand, has the property that the MSB is much more likely to be 0 than 1. Hence, the MSB of FBC has high redundancy. Therefore, it is easier to determine whether $e_{n(i-1)}$=0 or 1 with FBC than with NBC. We believe that this is the reason for the superiority of FBC over NBC in the SQ-MAP scheme. Also, note that the performance of the MAP detector tends to increase as the amount of residual redundancy increases. The redundancies of symmetric Lloyd-Max scalar quantizers are tabulated in Table 7. We next compare COVQ and SQ-MAP.

The COVQ system is a (locally) optimal system that efficiently exploits the intra-block memory. Both encoder and decoder of this system are optimal in the sense of minimizing the mean squared error. However, this system does not make any use of the inter-block memory. On the other hand, the SQ-MAP system, which exploits both memories, consists of a sub-optimal encoder and a MAP decoder that minimizes the error probability but *not* the mean squared error. For fixed $M$, the effect of the intra-block memory of the channel becomes more dominant as $kR$ increases. Therefore, for large blocks of $kR$ bits ($kR >> M$), the COVQ system outperforms the SQ-MAP system (e.g. for $k = 1$, $R = 4$ in Tables 1-6).

So far, we have only considered the case where $M = 1$ and $\delta = 10$. In Tables 8 and 9, we provide COVQ and SQ-MAP results for source shape parameter $\alpha = 0.5$, $\delta = 10.0$, rate $R = 4$ bits/sample and $M = 0$, 1, 2 ,3 and 4. Note that for almost all cases, the performances of both schemes increase as $M$ increases. This is essentially due to the fact that as the memory $M$ increases, both intra-block and inter-block memories increase; the MAP detector exploits this increase in combating channel errors. Similarly, the COVQ scheme exploits the increase in intra-block memory as long as $kR$ is sufficiently larger than $M$. In Table 10, we provide results for parameters: $\alpha = 0.5$, $M = 1$, $R = 4$ bits/sample, and $\delta = 0, 1, 2, 5, 10$. In general, the performances increase as $\delta$, and hence, channel capacity, increases. However, there are some instances where the SNR decreases when $\delta$ goes from zero to one. For COVQ, this may be due to the poor choice of initial codebook used in the design of the COVQ. For SQ-MAP, we have observed that the bit and symbol error probabilities of the sequence MAP detector actually decreases as $\delta$ increases. However, this does not directly translate to an increase in SNR.

## 4.2 Gauss-Markov Sources

In this subsection, we consider a first-order Gauss-Markov source which is described by the recursion

$$V_i = \phi V_{i-1} + U_i,$$

where $\phi \in (-1, 1)$ is called the correlation parameter of the process and $\{U_i\}$ is an i.i.d. sequence of Gaussian random variables. Results for COVQ and SQ-MAP are given in Tables 11 and 12, respectively, for $\phi = 0.9$. The SQ-MAP results here are consistently better than the results for the i.i.d. Gaussian source (Table 6). This is expected since there is an additional redundancy due to memory. The amounts of residual redundancy of symmetric Lloyd-Max scalar quantizers are listed in Table 7.

# 5 Comparisons with Interleaving

The traditional technique for handling a channel with memory is to use interleaving. In the following, we consider two channel interleaving schemes and compare their performances against COVQ and SQ-MAP. The reasoning for making such comparisons is the following. Suppose we are given a channel with memory. Suppose further that we know exactly how the channel memory is characterized (say by the Markov condition). Then how much improvement

in our system does this knowledge provide us? If we know nothing about the channel memory, the best approach is to use interleaving to render the channel memoryless and then design a system for the memoryless channel. On the other hand, if we know exactly how the channel is characterized, then we may be better off designing our system 'optimally' for this channel. In the following, we examine how much the quantization system can be improved with knowledge of the channel memory characteristics.

The first interleaving scheme, COVQ-IL, consists of a COVQ optimized for a BSC and an interleaver. It is assumed that the interleaving length is sufficiently large so that the combination of interleaver, Markov channel and de-interleaver is equivalent to a BSC. The SNR performances of this scheme are given in Tables 1, 2, 3, 8, 10 and 11. COVQ-IL is compared against COVQ (optimized for the Markov channel). Observe that in almost all cases COVQ outperforms COVQ-IL. For low values of the channel BER $\epsilon$, COVQ-IL sometimes slightly beats COVQ-IL (e.g., Table 1, $\epsilon = 0.005$, $k = 1$ and $R = 3$). This may be due to the fact that the index assignment scheme (simulated annealing), used to choose the initial codebook for the COVQ algorithm, operates under the assumption that the channel is memoryless. This assignment, may therefore perform poorly over the Markov channel. When COVQ beats COVQ-IL, the largest gain is 5.47 dB which occurs in Table 11 for $R = 4$ bits/sample, $k = 2$ and $\epsilon = 0.1$. In general, 4 to 5 dB gain is possible for large values of $\epsilon$ and high noise correlation ($\delta = 10$). The gain of COVQ over COVQ-IL is due to the fact that COVQ exploits the noise memory whereas COVQ-IL does not.

The second interleaving scheme, SQ-IL-MAP, consists of a symmetric SQ designed by the Lloyd-Max formulation, an interleaver/de-interleaver combination and a sequence MAP detector. The SQ binary codewords are assigned by FBC. The argument here is that FBC is a good codeword assignment for BSC [19] and the purpose of the interleaver/de-interleaver is to convert the Markov channel into a BSC. The MAP detector is designed for the BSC. The SNR results of SQ-IL-MAP are also provided in Tables 4, 5, 6, 9, 10 and 12. This scheme is compared against SQ-MAP. Note that SQ-MAP beats SQ-IL-MAP in most of the cases. The largest gain is 7.5 dB which occurs in Table 9 for $M = 4$ and $\epsilon = 0.05$. For comparison purposes, we also provide in Tables 4, 5, 6, 9 and 12 the results of the second interleaving scheme without MAP detection (denoted as SQ-IL). Note that for memoryless sources, the MAP detector offers no improvement in the interleaving scheme for small values of $\epsilon$. Since the interleaver renders the channel memoryless and the source is also memoryless, the sequence MAP detector is actually a memoryless MAP detector. That is, each observation $\mathbf{Y}_i$ is decoded

independently of every other observation. For such a MAP detector, it can be easily shown that MAP detection is useless whenever

$$\epsilon \leq \frac{p_{\min}}{p_{\min} + p_{\max}},$$

where $p_{\min} = \min_{\mathbf{x} \in \{0,1\}^n} \Pr\{\mathbf{X}_i = \mathbf{x}\}$ and $p_{\max} = \max_{\mathbf{x} \in \{0,1\}^n} \Pr\{\mathbf{X}_i = \mathbf{x}\}$. The above is only a sufficient condition and it is independent of the binary codeword assignment.

Finally, we note that the two interleaving schemes have large encoding and decoding delays (due to the interleaver and de-interleaver). The COVQ scheme only have a block delay of $k-1$ samples. The SQ-MAP scheme has the MAP detector delay.

# 6 Conclusions

We considered joint source-channel coding for generalized Gaussian and Gauss-Markov sources and binary Markov channels. Two schemes were considered, COVQ and SQ-MAP. COVQ outperforms SQ-MAP when $kR$ is large. COVQ has a high encoding complexity and a small decoding complexity. SQ-MAP, on the other hand, has a small encoding complexity and a large decoding complexity. These schemes were compared against two interleaving schemes. The performance gain is as much as 5 dB when the channel is very noisy with high noise correlation. This may correspond to the behavior of land mobile radio channels during deep fades.

# References


[1] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, Vol. 27, pp. 379–423 and 623–656, 1948.

[2] C. E. Shannon, "Coding Theorems for a Discrete Source with a Fidelity Criterion," *IRE Nat. Conv. Rec.*, pp. 142–163, Mar. 1959.

[3] E. Ayanoglu and R. M. Gray, "The Design of Joint Source and Channel Trellis Waveform Coders," *IEEE Transactions on Information Theory*, Vol. 33, pp. 855–865, Nov. 1987.

[4] N. Phamdo, N. Farvardin, and T. Moriya, "A Unified Approach to Tree-Structured and Multi-Stage Vector Quantization for Noisy Channels," *IEEE Transactions on Information Theory*, Vol. 39, pp. 835–850, May 1993.

[5] N. Phamdo and N. Farvardin, "Quantization Over Discrete Noisy Channels Using Rate-One Convolutional Codes," submitted to *IEEE Transactions on Information Theory*, Aug. 1993.

[6] H. S. Wang, *Finite-State Modeling, Capacity, and Joint Source/Channel Coding for Time-Varying Channels.* PhD thesis, Rutgers University, 1992.

[7] A. C. Hung and T. H.-Y. Meng, "Adaptive Channel Optimization of Vector Quantized Data," *Data Compression Conference*, 1993.

[8] F. Alajaji and T. Fuja, "A Communication Channel Modeled on Contagion," *IEEE Transactions on Information Theory*, to appear, November 1994.

[9] H. Kumazawa, M. Kasahara, and T. Namekawa, "A Construction of Vector Quantizers for Noisy Channels," *Electronics and Engineering in Japan*, Vol. 67-B, pp. 39–47, Jan. 1984.

[10] N. Farvardin and V. Vaishampayan, "On the Performance and Complexity of Channel-Optimized Vector Quantizers," *IEEE Transactions on Information Theory*, Vol. 37, pp. 155–160, Jan. 1991.

[11] K. Sayood and J. C. Borkenhagen, "Use of Residual Redundancy in the Design of Joint Source/Channel Coders," *IEEE Transactions on Communications*, Vol. 39, pp. 838–846, June 1991.

[12] N. Phamdo and N. Farvardin, "Optimal Detection of Discrete Markov Sources Over Discrete Memoryless Channels — Applications to Combined Source-Channel Coding," *IEEE Transactions on Information Theory*, Vol. 40, pp. 186-193, Jan. 1994.

[13] L. Kanal and A. Sastry, "Models for Channels with Memory and Their Applications to Error Control," *Proc. IEEE*, vol. 66, pp. 724-744, July 1978.

[14] Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantizer Design," *IEEE Transactions on Communications*, Vol. 28, pp. 84–94, Jan. 1980.

[15] S. P. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory,* Vol. 28, pp. 129–137, Mar. 1982.

[16] J. Max, "Quantizing for Minimum Distortion," *IRE Transactions on Information Theory,* Vol. IT-6, pp. 7-12, Mar. 1960.

[17] F. Alajaji, N. Phamdo, N. Farvardin and T. Fuja, "Detection of Binary Sources Over Discrete Channels with Additive Markov Noise," submitted to *IEEE Transactions on Information Theory*, April 1994.

[18] N. Farvardin, "A Study of Vector Quantization for Noisy Channels," *IEEE Transactions on Information Theory*, Vol. 36, pp. 799–809, July 1990.

[19] N. Farvardin and V. Vaishampayan, "Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source-Channel Coding," *IEEE Transactions on Information Theory*, Vol. 33, pp. 827–838, Nov. 1987.

| $R$ | $k$ | System | $\epsilon$=0.000 | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|---|---|
| 2 | 1 | COVQ | 5.77 | 5.10 | 4.66 | 3.07 | 2.88 |
| | | COVQ-IL | 5.77 | 5.17 | 4.71 | 2.77 | 1.69 |
| | 2 | COVQ | 8.86 | 7.33 | 7.10 | 6.04 | 4.94 |
| | | COVQ-IL | 8.86 | 7.62 | 6.82 | 4.06 | 2.48 |
| | 3 | COVQ | 10.25 | 9.05 | 8.52 | 7.36 | 6.25 |
| | | COVQ-IL | 10.25 | 8.77 | 7.83 | 4.85 | 3.13 |
| | $\infty$ | OPTA | 15.64 | 15.54 | 15.46 | 14.89 | 14.30 |
| 3 | 1 | COVQ | 10.43 | 7.80 | 7.75 | 6.52 | 5.23 |
| | | COVQ-IL | 10.43 | 8.35 | 7.23 | 4.05 | 2.32 |
| | 2 | COVQ | 14.14 | 11.58 | 10.97 | 9.17 | 7.64 |
| | | COVQ-IL | 14.14 | 10.99 | 9.65 | 5.73 | 3.75 |
| | 3 | COVQ | 16.01 | 13.59 | 12.72 | 10.35 | 9.11 |
| | | COVQ-IL | 16.01 | 11.63 | 10.50 | 6.78 | 4.43 |
| | $\infty$ | OPTA | 21.74 | 21.59 | 21.46 | 20.63 | 19.76 |
| 4 | 1 | COVQ | 15.75 | 11.26 | 10.39 | 8.73 | 7.22 |
| | | COVQ-IL | 15.75 | 10.81 | 9.06 | 5.10 | 3.41 |
| | 2 | COVQ | 20.02 | 15.21 | 14.43 | 11.58 | 9.88 |
| | | COVQ-IL | 20.02 | 13.75 | 12.11 | 7.48 | 4.94 |
| | $\infty$ | OPTA | 27.79 | 27.59 | 27.42 | 26.31 | 25.18 |

Table 1: SNR (in dB) Performances of COVQ and COVQ-IL Operating Over a Markov Channel with $\delta = 10$ and $M = 1$; Generalized Gaussian Source with Shape Parameter $\alpha = 0.5$; $R$ = Rate (Bits/Sample); $k$ = Vector Dimension; $\epsilon$ = Channel Bit Error Rate; In the Interleaved System, COVQ is Designed for Memoryless Channels; OPTA = Optimal Performance Theoretically Attainable.

| $R$ | $k$ | System | $\epsilon$=0.000 | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|---|---|
| 2 | 1 | COVQ | 7.55 | 6.99 | 6.54 | 4.53 | 4.52 |
| | | COVQ-IL | 7.55 | 6.95 | 6.45 | 4.01 | 2.51 |
| | 2 | COVQ | 8.83 | 8.09 | 7.57 | 6.70 | 5.86 |
| | | COVQ-IL | 8.83 | 8.03 | 7.41 | 4.75 | 3.31 |
| | 3 | COVQ | 9.48 | 8.71 | 8.16 | 7.27 | 6.32 |
| | | COVQ-IL | 9.48 | 8.50 | 7.80 | 5.13 | 3.59 |
| | $\infty$ | OPTA | 12.66 | 12.57 | 12.49 | 11.93 | 11.37 |
| 3 | 1 | COVQ | 12.64 | 10.50 | 9.45 | 8.27 | 7.17 |
| | | COVQ-IL | 12.64 | 10.49 | 9.17 | 5.21 | 3.62 |
| | 2 | COVQ | 14.25 | 11.88 | 10.97 | 10.00 | 8.64 |
| | | COVQ-IL | 14.25 | 11.67 | 10.28 | 6.60 | 4.47 |
| | 3 | COVQ | 15.16 | 13.01 | 12.43 | 10.68 | 9.48 |
| | | COVQ-IL | 15.16 | 11.52 | 10.67 | 7.08 | 4.84 |
| | $\infty$ | OPTA | 18.69 | 18.54 | 18.42 | 17.59 | 16.74 |
| 4 | 1 | COVQ | 18.08 | 13.57 | 13.19 | 10.54 | 8.61 |
| | | COVQ-IL | 18.08 | 12.76 | 11.03 | 6.82 | 4.79 |
| | 2 | COVQ | 20.09 | 15.38 | 15.09 | 12.27 | 10.68 |
| | | COVQ-IL | 20.09 | 14.41 | 12.92 | 8.33 | 5.71 |
| | $\infty$ | OPTA | 24.74 | 24.51 | 24.35 | 23.24 | 22.10 |

Table 2: SNR (in dB) Performances of COVQ and COVQ-IL Operating Over a Markov Channel with $\delta = 10$ and $M = 1$; Generalized Gaussian Source with Shape Parameter $\alpha = 1$; $R$ = Rate (Bits/Sample); $k$ = Vector Dimension; $\epsilon$ = Channel Bit Error Rate; In the Interleaved System, COVQ is Designed for Memoryless Channels; OPTA = Optimal Performance Theoretically Attainable.

| $R$ | $k$ | System | $\epsilon$=0.000 | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|---|---|
| 2 | 1 | COVQ | 9.29 | 8.64 | 8.11 | 5.72 | 5.92 |
| | | COVQ-IL | 9.29 | 8.52 | 7.87 | 4.86 | 3.05 |
| | 2 | COVQ | 9.51 | 8.97 | 8.41 | 7.08 | 6.62 |
| | | COVQ-IL | 9.51 | 8.70 | 8.06 | 5.44 | 3.86 |
| | 3 | COVQ | 9.87 | 9.15 | 8.56 | 7.74 | 7.10 |
| | | COVQ-IL | 9.87 | 8.94 | 8.29 | 5.70 | 3.99 |
| | $\infty$ | OPTA | 12.04 | 11.95 | 11.86 | 11.31 | 10.74 |
| 3 | 1 | COVQ | 14.60 | 12.39 | 11.17 | 9.29 | 7.47 |
| | | COVQ-IL | 14.60 | 12.04 | 10.50 | 6.47 | 4.67 |
| | 2 | COVQ | 15.21 | 12.81 | 11.89 | 10.59 | 9.42 |
| | | COVQ-IL | 15.21 | 12.46 | 11.15 | 7.36 | 5.15 |
| | 3 | COVQ | 15.66 | 13.55 | 12.90 | 11.35 | 10.05 |
| | | COVQ-IL | 15.66 | 12.01 | 11.40 | 7.67 | 5.37 |
| | $\infty$ | OPTA | 18.06 | 17.92 | 17.80 | 16.96 | 16.11 |
| 4 | 1 | COVQ | 20.17 | 15.67 | 14.93 | 11.24 | 9.13 |
| | | COVQ-IL | 20.17 | 14.15 | 12.30 | 7.81 | 5.60 |
| | 2 | COVQ | 21.06 | 16.70 | 16.11 | 13.28 | 11.52 |
| | | COVQ-IL | 21.06 | 15.28 | 13.70 | 9.06 | 6.40 |
| | $\infty$ | OPTA | 24.08 | 23.89 | 23.73 | 22.61 | 21.48 |

Table 3: SNR (in dB) Performances of COVQ and COVQ-IL Operating Over a Markov Channel with $\delta = 10$ and $M = 1$; Generalized Gaussian Source with Shape Parameter $\alpha = 2$; $R$ = Rate (Bits/Sample); $k$ = Vector Dimension; $\epsilon$ = Channel Bit Error Rate; In the Interleaved System, COVQ is Designed for Memoryless Channels; OPTA = Optimal Performance Theoretically Attainable.

| $R$ | System | $\epsilon$=0.000 | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|---|
| 2 | SQ-MAP | 5.60 | 5.54 | 5.34 | 4.41 | 3.43 |
| | SQ-IL-MAP | 5.60 | 5.00 | 4.47 | 1.65 | 0.88 |
| | SQ-IL | 5.60 | 5.00 | 4.47 | 1.65 | -0.40 |
| 3 | SQ-MAP | 10.35 | 9.27 | 8.37 | 5.67 | 4.03 |
| | SQ-IL-MAP | 10.35 | 7.91 | 6.35 | 1.81 | 1.02 |
| | SQ-IL | 10.35 | 7.91 | 6.35 | 0.85 | -2.13 |
| 4 | SQ-MAP | 15.69 | 11.15 | 9.46 | 4.72 | 2.73 |
| | SQ-IL-MAP | 15.69 | 9.02 | 6.47 | 1.66 | 1.03 |
| | SQ-IL | 15.69 | 9.02 | 6.47 | -0.39 | -3.71 |

Table 4: SNR (in dB) Performances of MAP Detection Schemes for a Markov Channel with $\delta = 10$ and $M = 1$; Generalized Gaussian Source with Shape Parameter $\alpha = 0.5$; $R$ = Rate (Bits/Sample); $\epsilon$ = Channel Bit Error Rate; In SQ-IL-MAP, MAP Detector is Designed for Memoryless Channels.

| $R$ | System | $\epsilon$=0.000 | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|---|
| 2 | SQ-MAP | 7.54 | 7.11 | 6.80 | 4.92 | 3.68 |
| | SQ-IL-MAP | 7.54 | 6.93 | 6.40 | 3.56 | 1.53 |
| | SQ-IL | 7.54 | 6.93 | 6.40 | 3.56 | 1.53 |
| 3 | SQ-MAP | 12.64 | 10.78 | 9.75 | 6.01 | 3.98 |
| | SQ-IL-MAP | 12.64 | 10.36 | 8.87 | 3.61 | 0.83 |
| | SQ-IL | 12.64 | 10.36 | 8.87 | 3.61 | 0.83 |
| 4 | SQ-MAP | 18.13 | 12.90 | 10.86 | 5.73 | 3.43 |
| | SQ-IL-MAP | 18.13 | 12.14 | 9.71 | 3.71 | 1.11 |
| | SQ-IL | 18.13 | 12.14 | 9.71 | 3.15 | 0.09 |

Table 5: SNR (in dB) Performances of MAP Detection Schemes for a Markov Channel with $\delta = 10$ and $M = 1$; Generalized Gaussian Source with Shape Parameter $\alpha = 1$; $R$ = Rate (Bits/Sample); $\epsilon$ = Channel Bit Error Rate; In SQ-IL-MAP, MAP Detector is Designed for Memoryless Channels.

| $R$ | System | $\epsilon$=0.000 | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|---|
| 2 | SQ-MAP | 9.27 | 8.58 | 8.10 | 5.21 | 3.18 |
| | SQ-IL-MAP | 9.27 | 8.50 | 7.84 | 4.58 | 2.41 |
| | SQ-IL | 9.27 | 8.50 | 7.84 | 4.58 | 2.41 |
| 3 | SQ-MAP | 14.62 | 12.20 | 10.91 | 5.77 | 3.35 |
| | SQ-IL-MAP | 14.62 | 11.99 | 10.36 | 4.94 | 2.19 |
| | SQ-IL | 14.62 | 11.99 | 10.36 | 4.94 | 2.19 |
| 4 | SQ-MAP | 20.15 | 13.90 | 11.89 | 5.58 | 2.93 |
| | SQ-IL-MAP | 20.15 | 13.83 | 11.35 | 4.84 | 2.39 |
| | SQ-IL | 20.15 | 13.83 | 11.35 | 4.84 | 1.89 |

Table 6: SNR (in dB) Performances of MAP Detection Schemes for a Markov Channel with $\delta = 10$ and $M = 1$; Generalized Gaussian Source with Shape Parameter $\alpha = 2$; $R$ = Rate (Bits/Sample); $\epsilon$ = Channel Bit Error Rate; In SQ-IL-MAP, MAP Detector is Designed for Memoryless Channels.

| Source | $R$ | $\rho_D$ | $\rho_M$ | $\rho_T = \rho_D + \rho_M$ |
|---|---|---|---|---|
| GG ($\alpha$=0.5) | 2 | 0.57 | 0.0 | 0.57 |
| | 3 | 0.88 | 0.0 | 0.88 |
| | 4 | 1.05 | 0.0 | 1.05 |
| GG ($\alpha$=1) | 2 | 0.28 | 0.0 | 0.28 |
| | 3 | 0.42 | 0.0 | 0.42 |
| | 4 | 0.50 | 0.0 | 0.50 |
| GG ($\alpha$=2) | 2 | 0.09 | 0.0 | 0.09 |
| | 3 | 0.18 | 0.0 | 0.18 |
| | 4 | 0.23 | 0.0 | 0.23 |
| GM ($\phi$=0.9) | 2 | 0.09 | 0.77 | 0.86 |
| | 3 | 0.18 | 1.03 | 1.20 |
| | 4 | 0.23 | 1.14 | 1.38 |

Table 7: Residual Redundancy (in Bits/Sample) of Symmetric Lloyd-Max Scalar Quantizer; GG = Generalized Gaussian Source; $\alpha$ = Shape Parameter; GM = Gauss-Markov Source; $\phi$ = Correlation Coefficient; $R$ = Rate of Scalar Quantizer in Bits/Sample; $\rho_D$ = Redundancy Due to Non-Uniform Distribution; $\rho_M$ = Redundancy Due to Memory; $\rho_T$ = Total Redundancy. ($\rho_M$ for GM Source is Estimated from Training Data Using the Markov Chain Assumption.)

| $M$ | $k$ | System | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|---|
| 0 | 1 | COVQ-IL | 10.81 | 9.06 | 5.10 | 3.41 |
| | 2 | COVQ-IL | 13.75 | 12.11 | 7.48 | 4.94 |
| | $\infty$ | OPTA | 26.69 | 25.84 | 20.85 | 16.41 |
| 1 | 1 | COVQ | 11.26 | 10.39 | 8.73 | 7.22 |
| | 2 | COVQ | 15.21 | 14.43 | 11.58 | 9.88 |
| | $\infty$ | OPTA | 27.59 | 27.42 | 26.31 | 25.18 |
| 2 | 1 | COVQ | 11.50 | 10.48 | 7.90 | 6.89 |
| | 2 | COVQ | 15.46 | 14.49 | 11.56 | 9.91 |
| | $\infty$ | OPTA | 27.66 | 27.54 | 26.73 | 25.88 |
| 3 | 1 | COVQ | 11.57 | 10.56 | 7.89 | 6.56 |
| | 2 | COVQ | 15.63 | 14.56 | 11.37 | 9.75 |
| | $\infty$ | OPTA | 27.68 | 27.58 | 26.90 | 26.18 |
| 4 | 2 | COVQ | 15.67 | 14.55 | 11.21 | 9.63 |
| | $\infty$ | OPTA | 27.69 | 27.61 | 26.98 | 26.32 |

Table 8: SNR (in dB) Performances of COVQ for Different Values of $M$; $R$ = 4 Bits/Sample; $\delta$ = 10.0; Generalized Gaussian Source with Shape Parameter $\alpha$ = 0.5; $k$ = Vector Dimension; $\epsilon$ = Channel Bit Error Rate; OPTA = Optimal Performance Theoretically Attainable.

| $M$ | System | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|
| 0 | SQ-IL | 8.95 | 6.43 | -0.39 | -3.71 |
| | SQ-IL-MAP | 8.95 | 6.43 | 1.65 | 1.03 |
| 1 | SQ-MAP | 11.15 | 9.46 | 4.72 | 2.73 |
| 2 | SQ-MAP | 13.23 | 12.41 | 7.84 | 5.94 |
| 3 | SQ-MAP | 13.84 | 12.61 | 8.55 | 6.59 |
| 4 | SQ-MAP | 14.22 | 12.55 | 9.15 | 6.81 |

Table 9: SNR (in dB) Performances of SQ-MAP for Different Values of $M$; $R = 4$ Bits/Sample; $\delta = 10.0$; Generalized Gaussian Source with Shape Parameter $\alpha = 0.5$; $\epsilon$ = Channel Bit Error Rate.

| | | COVQ ($k = 2$) | | SQ-MAP | |
|---|---|---|---|---|---|
| $R$ | $\delta$ | $\epsilon = 0.01$ | $\epsilon = 0.1$ | $\epsilon = 0.01$ | $\epsilon = 0.1$ |
| 4 | 0 | 12.11 | 4.94 | 6.47 | 1.03 |
| | 1 | 11.68 | 5.06 | 4.72 | 0.52 |
| | 2 | 11.88 | 6.01 | 5.29 | 0.17 |
| | 5 | 13.18 | 8.00 | 7.30 | 1.10 |
| | 10 | 14.43 | 9.88 | 9.46 | 2.73 |

Table 10: SNR (in dB) Performances of COVQ and SQ-MAP for Different Values of $\delta$; $M = 1$; Generalized Gaussian Source with Shape Parameter $\alpha = 0.5$; $R = 4$ Bits/Sample; $\epsilon$ = Channel BER; $k$ = Vector Dimension; (For $\delta = 0$, COVQ and SQ-MAP are Equivalent to their Interleaving Counterparts).

| $R$ | $k$ | System | $\epsilon$=0.000 | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|---|---|
| 2 | 1 | COVQ | 9.26 | 8.62 | 8.08 | 5.69 | 5.90 |
| | | COVQ-IL | 9.26 | 8.49 | 7.85 | 4.84 | 3.03 |
| | 2 | COVQ | 13.46 | 11.93 | 11.02 | 10.08 | 8.97 |
| | | COVQ-IL | 13.46 | 11.52 | 10.33 | 7.26 | 5.29 |
| | 3 | COVQ | 14.94 | 13.07 | 13.14 | 11.28 | 10.30 |
| | | COVQ-IL | 14.94 | 12.54 | 11.69 | 8.49 | 6.30 |
| | $\infty$ | OPTA | 19.25 | 19.16 | 19.08 | 18.51 | 17.95 |
| 3 | 1 | COVQ | 14.57 | 12.40 | 11.20 | 9.78 | 8.56 |
| | | COVQ-IL | 14.57 | 12.00 | 10.47 | 5.61 | 4.64 |
| | 2 | COVQ | 18.95 | 15.68 | 15.70 | 12.29 | 10.57 |
| | | COVQ-IL | 18.95 | 14.69 | 13.54 | 9.27 | 6.77 |
| | 3 | COVQ | 20.57 | 17.74 | 16.62 | 14.47 | 13.00 |
| | | COVQ-IL | 20.57 | 16.05 | 14.91 | 10.71 | 8.07 |
| | $\infty$ | OPTA | 25.27 | 25.13 | 25.01 | 24.17 | 23.32 |
| 4 | 1 | COVQ | 20.13 | 15.80 | 14.50 | 11.84 | 10.64 |
| | | COVQ-IL | 20.13 | 14.05 | 11.77 | 8.21 | 5.84 |
| | 2 | COVQ | 24.74 | 19.42 | 18.69 | 15.50 | 13.77 |
| | | COVQ-IL | 24.74 | 17.71 | 16.23 | 11.29 | 8.30 |
| | $\infty$ | OPTA | 31.29 | 31.11 | 30.94 | 29.83 | 28.69 |

Table 11: SNR (in dB) Performances of COVQ and COVQ-IL Operating Over a Markov Channel with $\delta = 10$ and $M = 1$; Gauss-Markov Source with Correlation Coefficient $\phi = 0.9$; $R$ = Rate (Bits/Sample); $k$ = Vector Dimension; $\epsilon$ = Channel Bit Error Rate; In the Interleaved System, COVQ is Designed for Memoryless Channels; OPTA = Optimal Performance Theoretically Attainable.

| $R$ | System | $\epsilon$=0.005 | $\epsilon$=0.01 | $\epsilon$=0.05 | $\epsilon$=0.1 |
|---|---|---|---|---|---|
| 2 | SQ-MAP | 9.13 | 8.97 | 7.47 | 6.12 |
| | SQ-IL-MAP | 9.11 | 8.89 | 7.32 | 6.57 |
| | SQ-IL | 8.50 | 7.84 | 4.58 | 2.41 |
| 3 | SQ-MAP | 14.42 | 14.10 | 11.30 | 9.59 |
| | SQ-IL-MAP | 13.85 | 13.19 | 10.11 | 7.74 |
| | SQ-IL | 11.99 | 10.36 | 4.94 | 2.19 |
| 4 | SQ-MAP | 19.42 | 18.74 | 14.45 | 12.29 |
| | SQ-IL-MAP | 17.71 | 16.35 | 11.33 | 8.41 |
| | SQ-IL | 13.83 | 11.35 | 4.84 | 1.89 |

Table 12: SNR (in dB) Performances of MAP Detection Schemes for a Markov Channel with $\delta = 10$ and $M = 1$; Gauss-Markov Source with Correlation Coefficient $\phi = 0.9$; $R$ = Rate (Bits/Sample); $\epsilon$ = Channel Bit Error Rate; In SQ-IL-MAP, MAP Detector is Designed for Memoryless Channels.