# Optimal Detection of Discrete Markov Sources Over Discrete Memoryless Channels - Applications to Combined Source-Channel Coding

*by N. Phamdo and N. Farvardin*

# Optimal Detection of Discrete Markov Sources Over Discrete Memoryless Channels — Applications to Combined Source-Channel Coding [†]

*Nam Phamdo and Nariman Farvardin*

Electrical Engineering Department
and Systems Research Center
University of Maryland
College Park, Maryland 20742

## Abstract

We consider the problem of detecting a discrete Markov source which is transmitted across a discrete memoryless channel. The detection is based upon the *maximum a posteriori* (MAP) criterion which yields the minimum probability of error for a given observation. Two formulations of this problem are considered: (i) a *sequence* MAP detection in which the objective is to determine the most probable transmitted sequence given the observed sequence and (ii) an *instantaneous* MAP detection which is to determine the most probable transmitted symbol at time $n$ given all the observations prior to and including time $n$. The solution to the first problem results in a "Viterbi-like" implementation of the MAP detector (with large delay) while the latter problem results in a recursive implementation (with no delay). For the special case of the binary symmetric Markov source and binary symmetric channel, simulation results are presented and an analysis of these two systems yields explicit critical channel bit error rates above which the MAP detectors become useful.

Applications of the MAP detection problem in a combined source-channel coding system are considered. Here it is assumed that the source is highly correlated and that the source encoder (in our case, a vector quantizer (VQ)) fails to remove all of the source redundancy. The remaining redundancy at the output of the source encoder is referred to as the "residual" redundancy. It is shown, through simulation, that the residual redundancy can be used by the MAP detectors to combat channel errors. For small block sizes, the proposed system beats Farvardin and Vaishampayan's channel-optimized VQ by wide margins. Finally, it is shown that the instantaneous MAP detector can be combined with the VQ decoder to form a minimum mean-squared error decoder. Simulation results are also given for this case.

**Index Terms:** Markov source; discrete memoryless channel; MAP detection; combined source-channel coding.

---

# I. Introduction

In his celebrated paper [1], Shannon stated that in information transmission over a noisy channel, "redundancy must be introduced in the proper way to combat the particular noise structure involved. However, any redundancy in the source will usually help if it is utilized at the receiving point. In particular, if the source already has a certain redundancy and no attempt is made to eliminate it in matching to the channel, this redundancy will help combat noise."

This statement, though made more than forty years ago, forms the foundation of the present paper. The principle assumption here is that the source to be transmitted has a certain redundancy[1] and due to certain constraints (for example, on the complexity), the transmitter makes no attempt to "match" the source to the channel. Instead, the source is transmitted directly over the channel. The problem thus is to design a receiver which fully "utilizes" the source redundancy to combat the effect of channel noise. Because of the constraint on the transmitter, the system described above inevitably suffers from a loss of optimality. Nevertheless, given this constraint, we seek to obtain the optimum receiver for a given source and channel.

First, discrete sources and channels are considered. It is hypothesized that the source is in the form of a discrete Markov chain [2] and that the channel is a discrete memoryless channel [2]. The receiver is a *maximum a posteriori* (MAP) receiver (detector) which is optimum in the sense of minimizing the probability of error [3]. Here, the redundancy between successive symbols of the Markov source is used (by the MAP detector) to provide some protection against channel errors — in a manner not unlike error correction used in convolutional coding.

The above formulation has been considered before by several authors. Bahl *et al.* [4] looked at the problem of determining the *a posteriori* probabilities of a finite-

---

[1]The term "redundancy" usually has three connotations: (i) the source has memory, (ii) the source distribution is non-uniform and (iii) the source has memory and a non-uniform distribution. In this paper, "redundancy" implies cases (i) or (iii).

state Markov source (this is a generalization of the discrete Markov source considered in this paper) observed through a discrete memoryless channel in the context of decoding linear block and convolutional codes. Devore [5] considered observing a binary Markov source through a binary symmetric channel and gave necessary and sufficient conditions for the optimality of certain detectors. More recently, Sayood and Borkenhagen [6] considered the detection of a discrete Markov source over a discrete memoryless channel in a joint source-channel DPCM image coding system. The main difference between the current paper and [6] is that the detector in [6] is suboptimum, hence, it is not a MAP detector. Furthermore, we consider two variations of this problem: (i) *sequence* MAP detection which involves a large delay and (ii) *instantaneous* MAP detection which involves no delay. In [6], only sequence detection was considered. In sequence MAP detection, the objective is to determine the most probable transmitted *sequence* given an observed sequence while the objective of instantaneous MAP detection is to determine the most probable transmitted *symbol* at a particular time given all the observations up to that time.

Upon solving these two problems, we apply their solutions to a combined source-channel coding system in which the source is no longer discrete but is continuous-valued. The channel remains discrete. In this case, the use of an explicit source encoder is unavoidable since the source entropy (which is infinite) is higher than the channel capacity (which is finite) [2]. The purpose of the source encoder is to reduce the source entropy to a level lower than the channel capacity by introducing some distortion [2]. If the source encoder is optimum in the sense that it has the lowest possible rate for a given distortion criterion, then it must remove all of the source redundancy – that is, the source encoder output must have no redundancy. However, similar to [6], we assume that the source encoder (in our case, a vector quantizer (VQ) with a small block size) is suboptimum, i.e., it fails to remove all of the source redundancy. Again, this may be attributed to some complexity constraint. The re-

dundancy which remains at the output of the source encoder is called the *residual redundancy* [6]. Here, MAP detection is applied to exploit the residual redundancy. More specifically, the VQ encoder output is modeled as a discrete Markov source and the problem reduces to that of discrete sources and channels mentioned previously. Simulation results for this system on a Gauss-Markov source are obtained and comparisons are made with Farvardin and Vaishampayan's channel-optimized VQ [7, 8]. It is then shown that the instantaneous MAP detector can be combined with the VQ decoder to form an optimum (minimum mean-squared error) decoder.

The rest of this paper is organized as follows. In Section II, preliminary notations are established. The two MAP detection problems and their solutions are given in Section III. In Section IV, the special case of the binary symmetric Markov source and binary symmetric channel is studied and simulation results are provided for this case. Section V includes a study of combined source-channel coding using MAP detection. Finally, conclusions are stated in Section VI.

## II. Preliminaries

A stationary stochastic process, $\{I_n\}_{n=1}^{\infty}$, with alphabet $\mathcal{J}_N \triangleq \{0, 1, \dots, N-1\}$ of size $N$ is said to be a discrete Markov source if $\forall\, n > 1$,

$$\Pr\{I_n = i_n | \mathbf{I}_1^{n-1} = \mathbf{i}_1^{n-1}\} = \Pr\{I_n = i_n | I_{n-1} = i_{n-1}\}, \tag{1}$$

$\forall\, i_k \in \mathcal{J}_N, k = 1, 2, \dots, n$, where

$$\mathbf{I}_1^{n-1} \triangleq (I_1, I_2, \dots, I_{n-1}), \tag{2}$$

$$\mathbf{i}_1^{n-1} \triangleq (i_1, i_2, \dots, i_{n-1}). \tag{3}$$

Because of stationarity, the conditional probability function, $\Pr\{I_n = i_n | I_{n-1} = i_{n-1}\}$, is independent of $n$. Thus, the Markov source can be completely characterized by the source transition probability matrix,

$$P(i_n | i_{n-1}) \triangleq \Pr\{I_n = i_n | I_{n-1} = i_{n-1}\}, \quad i_n, i_{n-1} \in \mathcal{J}_N, \tag{4}$$

and the initial source distribution,

$$P_1(i_1) \triangleq \Pr\{I_1 = i_1\}, \quad i_1 \in \mathcal{J}_N. \tag{5}$$

The simplest example of a discrete Markov source is the binary symmetric Markov source (BSMS) with $N = 2$, transition matrix,

$$P(i_n|i_{n-1}) = \begin{cases} p & \text{if } i_n = i_{n-1} \\ 1-p & \text{if } i_n \neq i_{n-1} \end{cases}, \quad i_n, i_{n-1} \in \mathcal{J}_2, \tag{6}$$

and initial distribution $P_1(0) = P_1(1) = 0.5$. Note that for a given BSMS, we may construct a $2^m$-ary Markov source by taking $m$ successive samples of the binary source and appropriately renaming each letter of the alphabet $\mathcal{J}_{2^m}$. We will make use of this construction in a later section.

It is assumed that the Markov source,$\{I_n\}_{n=1}^{\infty}$, is transmitted across a discrete memoryless channel (DMC) with the same alphabet, $\mathcal{J}_N$, as the source. The received sequence (at the output of the channel) is denoted by $\{J_n\}_{n=1}^{\infty}$. The channel is described by the channel transition probability matrix,

$$Q(j_n|i_n) = \Pr\{J_n = j_n|I_n = i_n\}, \quad i_n, j_n \in \mathcal{J}_N. \tag{7}$$

The simplest example of a DMC is the binary symmetric channel (BSC) with $N = 2$ and transition matrix,

$$Q(j_n|i_n) = \begin{cases} 1-\epsilon & \text{if } j_n = i_n \\ \epsilon & \text{if } j_n \neq i_n \end{cases}, \quad j_n, i_n \in \mathcal{J}_2. \tag{8}$$

Again, note that we may construct a $2^m$-ary DMC through $m$ uses of a BSC.

## III. Statement of Problems and Solutions

The two problems to be considered in this paper are as follows. Given $N, P, P_1, Q$ and an observation sequence $\mathbf{J}_1^n = \mathbf{j}_1^n$, determine (i) the most probable transmitted sequence,

$$\hat{\mathbf{i}}_1^n = \arg \max_{\mathbf{i}_1^n \in \mathcal{J}_N^n} \Pr\{\mathbf{I}_1^n = \mathbf{i}_1^n | \mathbf{J}_1^n = \mathbf{j}_1^n\}, \tag{9}$$

4

and (ii) the most probable transmitted symbol at time $n$,

$$\hat{i}_n = \arg \max_{i_n \in \mathcal{J}_N} \Pr\{I_n = i_n | \mathbf{J}_1^n = \mathbf{j}_1^n\}. \tag{10}$$

These two problems are referred to as the *sequence* MAP detection and the *instantaneous* MAP detection problems, respectively.

We first address the sequence MAP detection problem. Using Bayes' Theorem, equation (9) can be re-written as

$$\hat{\mathbf{i}}_1^n = \arg \max_{\mathbf{i}_1^n \in \mathcal{J}_N^n} \frac{\Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n | \mathbf{I}_1^n = \mathbf{i}_1^n\} \Pr\{\mathbf{I}_1^n = \mathbf{i}_1^n\}}{\Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n\}}. \tag{11}$$

Since the term in the denominator does not depend on $\mathbf{i}_1^n$, it suffices to maximize only the numerator:

$$\hat{\mathbf{i}}_1^n = \arg \max_{\mathbf{i}_1^n \in \mathcal{J}_N^n} \Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n | \mathbf{I}_1^n = \mathbf{i}_1^n\} \Pr\{\mathbf{I}_1^n = \mathbf{i}_1^n\}. \tag{12}$$

By using the fact that the channel is memoryless and by successively applying conditional probability and the Markovian property of the source, the above can be expressed as

$$\hat{\mathbf{i}}_1^n = \arg \max_{\mathbf{i}_1^n \in \mathcal{J}_N^n} \left[ \prod_{k=1}^n Q(j_k | i_k) \right] \left[ \prod_{k=2}^n P(i_k | i_{k-1}) P_1(i_1) \right]. \tag{13}$$

Since the logarithm is monotonic, the above is equivalent to

$$\hat{\mathbf{i}}_1^n = \arg \max_{\mathbf{i}_1^n \in \mathcal{J}_N^n} \left[ \sum_{k=2}^n \log[Q(j_k | i_k) P(i_k | i_{k-1})] + \log[Q(j_1 | i_1) P_1(i_1)] \right]. \tag{14}$$

In the above form, the sequence MAP detector can be implemented straightforwardly using the well-known Viterbi algorithm [9]. The trellis has $N$ states. There are (at most) $N$ branches entering and leaving each state. The path metric of the branch leaving state $i_{k-1}$ at time $k-1$ and entering state $i_k$ at time $k$ is $\log[Q(j_k | i_k) P(i_k | i_{k-1})]$. This metric is similar to that of Sayood and Borkenhagen [6] except that they have an additional term $-\log[\Pr\{J_k = j_k | I_{k-1} = i_{k-1}\}]$ which requires more computational

5

burden than our metric. Unlike [6], the metric given here is optimum. We point out that the objective here is to minimize the sequence error probability rather than the symbol error probability. To minimize the symbol error probability for the given (large) delay, an approach similar to [4] can be taken. However, this requires large complexity and in most applications yields little or no improvement [4]. Therefore, we will not consider it.

The delay of the sequence MAP detector is quite large ($n - 1$ samples). That is, the sequence MAP detector must wait to observe the entire sequence $\mathbf{j}_1^n$ before making a decision on the value of $i_1$. In the following, we show how this delay problem[2] can be alleviated.

We now turn to the instantaneous MAP detection problem. The instantaneous MAP detector is different from the sequence MAP detector in that it makes a decision on $i_n$ as soon as $j_n$ is received. Also, the symbol error probability is minimized in this case. Using similar arguments as earlier, equation (10) is re-written as

$$
\begin{aligned}
\hat{i}_n &= \arg\max_{i_n \in \mathcal{J}_N} \Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n | I_n = i_n\} \Pr\{I_n = i_n\}, \\
&= \arg\max_{i_n \in \mathcal{J}_N} \Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n, I_n = i_n\}.
\end{aligned}
\tag{15}
$$

Let us denote by

$$
f^{(n)}(i_n) \triangleq \Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n, I_n = i_n\},
\tag{16}
$$

the objective function[3] that is to be maximized at time $n$. $f^{(n)}(i_n)$ can be expressed as the sum of the joint probabilities:

$$
\begin{aligned}
f^{(n)}(i_n) &= \sum_{\mathbf{i}_1^{n-1} \in \mathcal{J}_N^{n-1}} \Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n, I_n = i_n, \mathbf{I}_1^{n-1} = \mathbf{i}_1^{n-1}\}, \\
&= \sum_{\mathbf{i}_1^{n-1} \in \mathcal{J}_N^{n-1}} \Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n, \mathbf{I}_1^n = \mathbf{i}_1^n\}.
\end{aligned}
\tag{17}
$$

---

[2]Another way of reducing the delay is to limit the trellis path memory using some symbol release rule [10]. This will be explored in a later section.

[3]Actually, the objective function also depends on the observations $\mathbf{j}_1^n$. However, for notational simplicity we will not explicitly express this dependency.

Each term in the summation can be expanded using the definition of conditional probability,

$$f^{(n)}(i_n) = \sum_{\mathbf{i}_1^{n-1} \in \mathcal{J}_N^{n-1}} \Pr\{\mathbf{J}_1^n = \mathbf{j}_1^n | \mathbf{I}_1^n = \mathbf{i}_1^n\} \Pr\{\mathbf{I}_1^n = \mathbf{i}_1^n\}. \tag{18}$$

Note that the two terms in (18) are the same as those in (12). We thus proceed as before, yielding

$$f^{(n)}(i_n) = \sum_{\mathbf{i}_1^{n-1} \in \mathcal{J}_N^{n-1}} \left[ \prod_{k=1}^n Q(j_k|i_k) \right] \left[ \prod_{k=2}^n P(i_k|i_{k-1})P(i_1) \right]. \tag{19}$$

Taking out the common factor $Q(j_n|i_n)$ results in

$$\begin{aligned} f^{(n)}(i_n) &= Q(j_n|i_n) \sum_{i_{n-1} \in \mathcal{J}_N} P(i_n|i_{n-1}) \\ &\times \sum_{\mathbf{i}_1^{n-2} \in \mathcal{J}_N^{n-2}} \left[ \prod_{k=1}^{n-1} Q(j_k|i_k) \right] \left[ \prod_{k=2}^{n-1} P(i_k|i_{k-1})P(i_1) \right]. \end{aligned} \tag{20}$$

Comparing equations (19) and (20), we can see that the second summation in (20) is nothing but $f^{(n-1)}(i_{n-1})$, the objective function that was maximized at time $n-1$. Hence, the instantaneous MAP detector can be implemented using the following recursion,

$$\begin{aligned} f^{(1)}(i_1) &= Q(j_1|i_1)P_1(i_1), \\ f^{(n)}(i_n) &= Q(j_n|i_n) \sum_{i_{n-1} \in \mathcal{J}_N} P(i_n|i_{n-1})f^{(n-1)}(i_{n-1}), \quad n = 2, 3, \ldots, \end{aligned} \tag{21}$$

where

$$\hat{i}_n = \arg \max_{i_n \in \mathcal{J}_N} f^{(n)}(i_n), \quad \text{for } n = 1, 2, \ldots. \tag{22}$$

It is noted that (21) is a special case of a similar equation in [4]. Implementation of the instantaneous MAP detector using the recursion (21) requires $2(N^2 + N)$ words of memory, $N^2 + N$ floating point operations and $N - 1$ comparisons per unit time.

We should point out that there is an obvious gap between the sequence and the instantaneous MAP detectors. One has a very large delay while the other has

7

no delay. It would be useful to have a detector whose delay is in between that of the instantaneous and the sequence MAP detectors. We will show how this can be accomplished in the next section.

## IV. Special Case: BSMS and BSC

In this section, we consider the special case of BSMS and BSC. Simulation results for the sequence and instantaneous MAP detectors for BSMSs with $p = 0.9, 0.95$ and 0.99 are plotted in Figures 1, 2 and 3, respectively, for various values of $\epsilon$. The simulation was performed on 10,000 samples of the BSMS and the experiment was repeated 50 times. In Figures 1-3, $\hat{\epsilon} \triangleq \Pr\{\text{error}\} = \Pr\{\hat{I}_n \neq I_n\}$ and the average values of $\hat{\epsilon}$ (over the 50 experiments) are shown with the error bars indicating the minimum and maximum values of $\hat{\epsilon}$. The curve labeled "w/o MAP" indicates the probability of error when no MAP detection is performed (hence $\hat{\epsilon} = \epsilon$).

We also present in these figures the result of the instantaneous MAP detector for a $2^m$-ary Markov source (constructed from $m$ successive samples of the BSMS) and a $2^m$-ary DMC (constructed through $m$ uses of the BSC) for $m = 4$. Actually, we just transmitted the BSMS directly over the BSC and allowed the instantaneous MAP detector to operate on a block of $m$ samples at a time. Hence, we have accomplished the task of bridging the delay gap between the instantaneous and sequence MAP detectors. In this case, the delay is $m - 1$ samples (of the BSMS). Note that the performance curve of the instantaneous MAP (with $m=4$) lies in between that of the instantaneous MAP (with $m=1$) and the sequence MAP.

Observe that for small values of $\epsilon$ all four curves coincide with each other. In fact, we found that in these cases, $\hat{I}_n \equiv J_n$, i.e., the MAP detector does absolutely nothing. At a certain point on the graph, the "Inst. MAP ($m=4$)" and the "Seq. MAP" curves diverge from the others. At another point, the "Inst. MAP ($m=1$)" curve diverges from the "w/o MAP" curve. This phenomena can be explained by the

following theorems and corollaries.

**Theorem 1** *Given $p \in (\frac{1}{2}, 1), \epsilon \in (0, \frac{1}{2})$. For $n \geq 3$, $\hat{\mathbf{I}}_1^n = \mathbf{J}_1^n$ is the optimum sequence detection rule if and only if*

$$\epsilon \leq \epsilon_1 \triangleq \frac{(1-p)^2}{p^2 + (1-p)^2}. \tag{23}$$

**Proof:** We first show that if $\epsilon \leq \epsilon_1$, then $\forall\, \mathbf{j}_1^n \in \mathcal{J}_2^n$,

$$\Pr\{\mathbf{I}_1^n = \mathbf{j}_1^n | \mathbf{J}_1^n = \mathbf{j}_1^n\} \geq \Pr\{\mathbf{I}_1^n = \mathbf{i}_1^n | \mathbf{J}_1^n = \mathbf{j}_1^n\}, \quad \forall\, \mathbf{i}_1^n \in \mathcal{J}_2^n. \tag{24}$$

Define $d \triangleq d_H(\mathbf{i}_1^n, \mathbf{j}_1^n) =$ Hamming distance between the two sequences $(0 \leq d \leq n)$.
Then

$$\alpha \triangleq \frac{\Pr\{\mathbf{I}_1^n = \mathbf{j}_1^n | \mathbf{J}_1^n = \mathbf{j}_1^n\}}{\Pr\{\mathbf{I}_1^n = \mathbf{i}_1^n | \mathbf{J}_1^n = \mathbf{j}_1^n\}} = \frac{(1-\epsilon)^n \prod_{k=2}^n P(j_k | j_{k-1}) P_1(j_1)}{\epsilon^d (1-\epsilon)^{n-d} \prod_{k=2}^n P(i_k | i_{k-1}) P_1(i_1)}. \tag{25}$$

The terms $P_1(j_1)$ and $P_1(i_1)$ cancel because they are both $\frac{1}{2}$. Now define,

$$\mathcal{K} \triangleq \{2, 3, \ldots, n\},$$

$$\mathcal{A} \triangleq \{k \in \mathcal{K} : i_k = j_k, i_{k-1} = j_{k-1}\},$$

$$\mathcal{B} \triangleq \{k \in \mathcal{K} : k \notin \mathcal{A}\}. \tag{26}$$

Hence

$$\alpha = \left(\frac{1-\epsilon}{\epsilon}\right)^d \frac{\prod_{k \in \mathcal{A}} P(j_k | j_{k-1}) \prod_{k \in \mathcal{B}} P(j_k | j_{k-1})}{\prod_{k \in \mathcal{A}} P(i_k | i_{k-1}) \prod_{k \in \mathcal{B}} P(i_k | i_{k-1})}. \tag{27}$$

By definition of $\mathcal{A}$, the products over $\mathcal{A}$ are equal and they cancel out,

$$\alpha = \left(\frac{1-\epsilon}{\epsilon}\right)^d \prod_{k \in \mathcal{B}} \left[\frac{P(j_k | j_{k-1})}{P(i_k | i_{k-1})}\right]. \tag{28}$$

Note that each term inside the product is either $\frac{p}{1-p}$, $1$ or $\frac{1-p}{p}$. The smallest of these is $\frac{1-p}{p}$ (since $p > \frac{1}{2}$). Thus

$$\alpha \geq \left(\frac{1-\epsilon}{\epsilon}\right)^d \left(\frac{1-p}{p}\right)^{|\mathcal{B}|}. \tag{29}$$

9

A moment of reflection would reveal that $|\mathcal{B}| \leq 2d$ (equality is achieved if the set $\{k \in \mathcal{K} \cup \{1\} : i_k \neq j_k\}$ does not contain either any pair of consecutive integers, $\{1\}$ or $\{n\}$). Hence

$$\alpha \geq \left(\frac{1-\epsilon}{\epsilon}\right)^d \left(\frac{1-p}{p}\right)^{2d} \geq 1. \tag{30}$$

The last inequality follows from our assumption on $\epsilon$ ($\leq \epsilon_1$).

To prove the converse, we note that $\exists$ sequences $\mathbf{i}_1^n, \mathbf{j}_1^n \in \mathcal{J}_2^n$ such that if $\epsilon > \epsilon_1$,

$$\alpha = \left(\frac{1-\epsilon}{\epsilon}\right)^d \left(\frac{1-p}{p}\right)^{2d} < 1. \tag{31}$$

Here we use the fact that $n \geq 3$.$\square$

For instantaneous MAP detection, we have the following theorem and corollaries; the proof of the theorem is provided in [5].

**Theorem 2** *Given* $p \in (\frac{1}{2}, 1), \epsilon \in (0, \frac{1}{2})$. *For* $n \geq 1$, $\hat{I}_n = J_n$ *is the optimum instantaneous detection rule if and only if*

$$(1 - 2\epsilon)^2 \geq 4p - 3. \tag{32}$$

**Corollary 1** *Given* $p \in (\frac{1}{2}, \frac{3}{4}]$. *For* $n \geq 1$, $\hat{I}_n = J_n$ *is the optimum instantaneous detection rule* $\forall \epsilon \in (0, \frac{1}{2})$.

**Corollary 2** *Given* $p \in [\frac{3}{4}, 1), \epsilon \in (0, \frac{1}{2})$. *For* $n \geq 1$, $\hat{I}_n = J_n$ *is the optimum instantaneous detection rule if and only if*

$$\epsilon \leq \epsilon_2 \triangleq \frac{1}{2} - \sqrt{p - \frac{3}{4}}. \tag{33}$$

The above theorems[4] state that under certain conditions, the simple detection rule $\hat{I}_n = J_n$ is optimum. Hence, any detection algorithm would be useless under

---

[4]Theorem 1 can be extended to a class of non-binary Markov sources. This class consists of discrete Markov sources with alphabet size $N = 2^m$, $p =$ the probability of returning to the same state and $q = (1 - p)/(N - 1) =$ the probability of going to any other state $(p > q)$. The channel is obtained through $m$ uses of a BSC. In this case, the critical BER is $\epsilon_1 = (1-p)^2/[(N-1)^2 p^2 + (1-p)^2]$. The proof of this is similar to that of Theorem 1 and is omitted.

these conditions. For a given BSMS with parameter $p$, $\epsilon_1$ is defined to be the *critical channel BER* for the sequence MAP detector. If $p > 0.75$, then $\epsilon_2$ is defined to be the critical BER for the instantaneous MAP detector. Sequence (instantaneous) MAP detection is useful if and only if $\epsilon > \epsilon_1$ ($\epsilon > \epsilon_2$). The values of $\epsilon_1$ and $\epsilon_2$ are indicated in each Figures 1-3. It can be seen that the simulation results in these figures agree with Theorem 1 and Corollary 2.

Although the critical BER for the "Inst. MAP ($m$=4)" curve has not been derived, it seems from Figures 1-3 that this value is close to $\epsilon_1$ (not $\epsilon_2$). This is due mainly to the $m - 1$ samples of delay in this case. Also, notice the change in the slope of this curve at $\epsilon = \epsilon_2$.

## V. Combined Source-Channel Coding

We now apply the solutions of the MAP detection problems to a combined source-channel coding system. The system to be considered is depicted in Figure 4. The source code is a VQ with blocklength $k$ and rate $R = \frac{1}{k}\log_2 N$ bits/sample (or bits/dimension), where $N$ is the size of the VQ codebook. The VQ is memoryless in the sense that it operates on a block of $k$ source samples independent of every other samples. The source, $\{\mathbf{X}_n\}_{n=1}^{\infty}$, is assumed to be correlated from block to block and, therefore, the VQ encoder output, $\{I_n\}_{n=1}^{\infty}$, is also correlated. This correlation is called the *residual* redundancy [6]. We model the residual redundancy by assuming that the sequence, $\{I_n\}_{n=1}^{\infty}$, forms a discrete Markov source. As before, $\{I_n\}_{n=1}^{\infty}$ is transmitted across a DMC. At the receiver, either sequence or instantaneous MAP detection is performed. The MAP detector output, $\{\hat{I}_n\}_{n=1}^{\infty}$, is then provided to the VQ decoder to obtain the reconstructed vector, $\tilde{\mathbf{x}}_n = \mathbf{c}_{i_n}$, where $\mathcal{C} = \{\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{N-1}\}$ is the VQ codebook. The system in Figure 4 is similar to that of Sayood and Borkenhagen [6] except that in [6], they used DPCM rather than VQ for the source code.

11

Though the purpose of the VQ encoder is to remove all of the source redundancy, it fails to do so mainly because of the limitation on the complexity and hence on the blocklength $k$ [11]. Since, by assumption, the source is highly correlated and since $k$ is small, the VQ encoder output is also correlated from one block to the next. This inherent redundancy in the data is exploited by the MAP detector to combat channel errors. The VQ encoder can be thought of as a combined source-channel encoder. It acts like a source encoder because it reduces the source entropy. It also functions like a channel encoder since its output contains redundancy. Thus, in Figure 4, there is no explicit channel encoder (though the MAP detector can be thought of as a channel decoder). Note that the system in Figure 4 does not require an increase in bandwidth like those of traditional channel coding techniques.

Simulation results for a first-order Gauss-Markov source, i.e., an autoregressive (AR(1)) source, with correlation coefficient $\rho = 0.9$ are given in Figures 5-9. The transition probability matrix, $P(\cdot|\cdot)$, was estimated using a long training sequence (1.2 million source samples). The initial distribution, $P_1(\cdot)$, was chosen to be the equilibrium distribution [2]. The simulation was performed for 50,000 vectors[5] from outside the training sequence and the experiment was repeated 50 times using 50 different realizations of the channel. The average squared-error distortion (averaged over the 50 trials) are shown (in dB). The variation between the minimum and maximum average distortion is quite small and is not shown in these figures. For $k = 1$ and $R = 1.0$ (Figure 5), the estimated transition matrix for $\{I_n\}_{n=1}^{\infty}$ is close to that of a BSMS with $p = 0.856$; the corresponding values of $\epsilon_1$ and $\epsilon_2$ are shown in Figure 5. Note the dip in the performance curve at the critical BERs. For comparison, we also present the results of Farvardin and Vaishampayan's channel-optimized VQ (COVQ)

---

[5]Due to the large complexity in the case of sequence MAP detection, we have limited the trellis path memory to $K = 50N$ vectors. The symbol release rule is as follows. The best path at time instant $n$ is traced backward $K$ time units. The symbol on this path is then released. We found that there was no loss of optimality for this large value of $K$.

[7, 8][6]. In many cases, both the sequence and instantaneous MAP detection systems beat the COVQ by wide margins. For completeness, these results are tabulated in Tables 1 and 2.

Next, we consider a method of improving the performance of the system in Figure 4. Observe that, similar to hard-decision convolutional decoding [10], there is an irreversible loss of information due to the "hard decision" made by the MAP detector. For example, the instantaneous MAP detector provides the most probable index $\hat{i}_n$ to the VQ decoder; but it does not tell the decoder how much confidence it has that this choice is correct. This information is embedded in the objective function $f^{(n)}(i_n)$. The remedy, therefore, is to combine the MAP detector and the VQ decoder into a single *minimum mean-squared error* (MMSE) decoder. In this case, the MMSE decoder can be thought of as a combined source-channel decoder. We consider only the instantaneous MMSE decoding problem; the sequence MMSE problem is left open.

From classical estimation theory [3], the MMSE estimate of $\mathbf{X}_n$ given $\mathbf{J}_1^n = \mathbf{j}_1^n$ is

$$\tilde{\mathbf{x}}_n = E[\mathbf{X}_n | \mathbf{J}_1^n = \mathbf{j}_1^n]. \tag{34}$$

Using conditional expectation, (34) is equivalent to

$$\tilde{\mathbf{x}}_n = \sum_{i_n \in \mathcal{J}_N} E[\mathbf{X}_n | I_n = i_n, \mathbf{J}_1^n = \mathbf{j}_1^n] \Pr\{I_n = i_n | \mathbf{J}_1^n = \mathbf{j}_1^n\}. \tag{35}$$

Note that the expectation in (35) is just the centroid of the encoding region $i_n$, denoted by $\bar{\mathbf{c}}_{i_n}$ (for LBG-VQ [12], $\bar{\mathbf{c}}_{i_n} = \mathbf{c}_{i_n}$ is a codevector). Furthermore, the second term in the summation in (35) is, up to a scale factor, equal to $f^{(n)}(i_n)$. Thus equation (35) can be re-written as,

$$\tilde{\mathbf{x}}_n = \frac{1}{\beta} \sum_{i_n \in \mathcal{J}_N} \bar{\mathbf{c}}_{i_n} f^{(n)}(i_n), \tag{36}$$

where

$$\beta = \sum_{i_n \in \mathcal{J}_N} f^{(n)}(i_n), \tag{37}$$

---

[6]The results taken from [7] are for a memoryless Gaussian source. However, since $k = 1$, we assume that the same results hold for a Gauss-Markov source.

is the scale factor. Implementation of the MMSE decoder requires only an additional $kN + N$ multiply-add per vector over the instantaneous MAP system in Figure 4.

Simulation results for the instantaneous MMSE decoder are given in Figures 5-9 and in Tables 1 and 2. Note that, in most cases, the result for the instantaneous MMSE decoder is close to that of the sequence MAP decoder. However, when the channel is very noisy, the MMSE decoder is considerably better. Of course, the delay of the instantaneous MMSE decoder is much less than that of the sequence MAP decoder.

Finally, we combine the ideas presented in this paper with that of COVQ [8]. Since these two ideas are, loosely speaking, independent, combining them should lead to a system that is better than one which uses only one of those two ideas. The system considered consists of a COVQ encoder and a MMSE decoder. Simulation results for this system are given in Table 3 for a few selected cases. Here, $\epsilon_d$ is the design BER of the COVQ encoder and $\epsilon_a$ is the actual channel BER (the MMSE decoder was designed for $\epsilon_a$). It can be seen that for all positive $\epsilon_d$ considered the proposed system beats the COVQ. However, the best result does not occur for $\epsilon_d = \epsilon_a$ but for $\epsilon_d$ slightly smaller than $\epsilon_a$. One possible explanation for this is the following. Suppose that instead of using the MMSE decoder, we use the system in Figure 4 with COVQ replacing VQ. Since the effective channel (between $I_n$ and $\hat{I}_n$) will be less noisy than the actual channel (between $I_n$ and $J_n$), the COVQ should be design for a channel with a lower level of noise than the actual channel. Of course, a possible improvement to the present system is to jointly design the COVQ encoder and the MMSE decoder using some training sequence approach. We have not tried this in the present work.

## VI. Conclusions

We have presented two formulations for MAP detection of discrete Markov sources over discrete memoryless channels. The first, sequence MAP detection, involves large

14

delay and results in a "Viterbi-like" implementation. The second, instantaneous MAP detection, requires no delay and results in a recursive implementation. For the special case of the binary symmetric Markov source and binary symmetric channel, we have shown that the two MAP detectors are useful for only a certain range of channel bit error rates. The MAP detection problem was applied to a combined source-channel coding system in which the source code is a small-block-size VQ which does not remove all of the source redundancy. The MAP detectors were applied to utilize the residual redundancy at the output of the VQ encoder. Improvements of the combined source-channel coding system were proposed in the form of instantaneous MMSE decoding and combined COVQ encoding with MMSE decoding. For the Gauss-Markov source with correlation coefficient 0.9, the proposed system beats the COVQ with the same (small) blocklength (hence same delay).

Future research directions may include alternate methods for estimating the transition probabilities of the Markov source. This is particularly important for large alphabet sizes ($N \geq 128$) since training-data approaches are impractical for these cases. Further studies on channel mismatch and source mismatch (because of the inaccuracy in estimating the transition matrix of the Markov source) are needed. Extensions of Theorems 1 and 2 to general $N$-ary Markov sources would be of interest. Methods of reducing the MAP detector's complexity and applications to real-world data, such as speech or image, should also be investigated.

Finally, we mention that the systems proposed in this paper are particularly useful in applications, such as remote telemetry, in which there is a hard constraint on the transmitter complexity.

# References

[1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.,* Vol. 27, pp. 379-423 and 623-656, 1948.

[2] R. E. Blahut, *Principles and Practice of Information Theory.* Reading, Massachusetts: Addison-Wesley, 1987.

[3] H. V. Poor, *An Introduction to Signal Detection and Estimation.* New York: Springer-Verlag, 1988.

[4] L. R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Trans. Inform. Theory,* Vol. 20, pp. 284-287, Mar. 1974.

[5] J. L. Devore, "A Note on the Observation of a Markov Source Through a Noisy Channel," *IEEE Trans. Inform. Theory,* Vol. 20, pp. 762-764, Nov. 1974.

[6] K. Sayood and J. C. Borkenhagen, "Use of Residual Redundancy in the Design of Joint Source/Channel Coders," *IEEE Trans. Commun.,* Vol. 39, pp. 838-846, Jun. 1991.

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

[8] N. Farvardin and V. Vaishampayan, "On the Performance and Complexity of Channel-Optimized Vector Quantizers," *IEEE Trans. Inform. Theory,* Vol. 37, pp. 155-160, Jan. 1991.

[9] G. D. Forney, Jr., "The Viterbi Algorithm," *Proc. IEEE,* vol. 61, pp. 268-278, Mar. 1973.

[10] S. Lin and D. J. Costello, Jr., *Error Control Coding.* Englewood Cliffs, New Jersey: Prentice-Hall, 1983.

[11] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Massachusetts: Kluwer Academic Press, 1991.

[12] Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantization Design," *IEEE Trans. Commun.,* Vol. 28, pp. 84-95, Dec. 1980.

| $\log_{10}\epsilon$ | $k=1; R=1.0$ VQ | VQ+ Inst. MAP | VQ+ Seq. MAP | COVQ [7] | VQ+ Inst. MMSE | $k=1; R=2.0$ VQ | VQ+ Inst. MAP | VQ+ Seq. MAP | COVQ [7] | VQ+ Inst. MMSE | $k=1; R=3.0$ VQ | VQ+ Inst. MAP | VQ+ Seq. MAP | COVQ [7] | VQ+ Inst. MMSE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $-\infty$ | 4.38 | 4.38 | 4.38 | 4.40 | 4.38 | 9.27 | 9.27 | 9.27 | 9.30 | 9.27 | 14.53 | 14.53 | 14.53 | 14.62 | 14.53 |
| -3.0 | 4.35 | 4.35 | 4.35 | – | 4.37 | 9.09 | 9.11 | 9.19 | – | 9.18 | 13.75 | 14.08 | 14.30 | – | 14.22 |
| -2.9 | 4.35 | 4.35 | 4.35 | – | 4.37 | 9.05 | 9.07 | 9.17 | – | 9.16 | 13.57 | 13.97 | 14.25 | – | 14.15 |
| -2.8 | 4.34 | 4.34 | 4.34 | – | 4.36 | 8.99 | 9.02 | 9.14 | – | 9.13 | 13.36 | 13.87 | 14.18 | – | 14.07 |
| -2.7 | 4.32 | 4.32 | 4.32 | – | 4.36 | 8.92 | 8.95 | 9.11 | – | 9.10 | 13.09 | 13.77 | 14.11 | – | 13.98 |
| -2.6 | 4.31 | 4.31 | 4.31 | – | 4.35 | 8.84 | 8.87 | 9.07 | – | 9.07 | 12.79 | 13.60 | 14.04 | – | 13.86 |
| -2.5 | 4.29 | 4.29 | 4.29 | – | 4.34 | 8.73 | 8.78 | 9.02 | – | 9.03 | 12.44 | 13.41 | 13.95 | – | 13.72 |
| -2.4 | 4.27 | 4.27 | 4.27 | – | 4.33 | 8.60 | 8.77 | 8.96 | – | 8.98 | 12.03 | 13.18 | 13.83 | – | 13.56 |
| -2.3 | 4.24 | 4.24 | 4.24 | 4.25 | 4.32 | 8.44 | 8.67 | 8.92 | 8.52 | 8.92 | 11.56 | 12.91 | 13.69 | 12.04 | 13.38 |
| -2.2 | 4.20 | 4.20 | 4.20 | – | 4.30 | 8.25 | 8.54 | 8.84 | – | 8.85 | 11.04 | 12.57 | 13.51 | – | 13.15 |
| -2.1 | 4.15 | 4.15 | 4.15 | – | 4.27 | 8.03 | 8.51 | 8.77 | – | 8.77 | 10.46 | 12.20 | 13.29 | – | 12.89 |
| -2.0 | 4.09 | 4.09 | 4.09 | 4.11 | 4.25 | 7.75 | 8.36 | 8.67 | 7.88 | 8.67 | 9.83 | 11.86 | 13.03 | 10.50 | 12.60 |
| -1.9 | 4.02 | 4.02 | 4.02 | – | 4.21 | 7.44 | 8.22 | 8.56 | – | 8.54 | 9.14 | 11.49 | 12.72 | – | 12.25 |
| -1.8 | 3.93 | 3.93 | 3.93 | – | 4.16 | 7.07 | 8.02 | 8.41 | – | 8.39 | 8.41 | 10.96 | 12.43 | – | 11.86 |
| -1.7 | 3.82 | 3.82 | 3.82 | – | 4.11 | 6.64 | 7.77 | 8.23 | – | 8.21 | 7.64 | 10.38 | 12.07 | – | 11.43 |
| -1.6 | 3.69 | 4.68 | 3.68 | – | 4.04 | 6.16 | 7.48 | 8.01 | – | 8.00 | 6.84 | 9.95 | 11.63 | – | 10.93 |
| -1.5 | 3.52 | 3.52 | 3.97 | – | 3.96 | 5.63 | 7.11 | 7.73 | – | 7.73 | 6.00 | 9.39 | 11.14 | – | 10.40 |
| -1.4 | 3.32 | 3.32 | 3.90 | – | 3.86 | 5.04 | 6.69 | 7.39 | – | 7.42 | 5.14 | 8.68 | 10.67 | – | 9.80 |
| -1.3 | 3.09 | 3.08 | 3.82 | 3.15 | 3.73 | 4.40 | 6.19 | 6.97 | 5.20 | 7.05 | 4.26 | 7.90 | 10.03 | 6.47 | 9.13 |
| -1.2 | 2.80 | 2.80 | 3.70 | – | 3.59 | 3.71 | 5.61 | 6.56 | – | 6.61 | 3.37 | 7.07 | 9.29 | – | 8.40 |
| -1.1 | 2.47 | 2.48 | 3.53 | – | 3.42 | 2.98 | 4.95 | 6.26 | – | 6.11 | 2.47 | 6.18 | 8.40 | – | 7.60 |
| -1.0 | 2.09 | 2.09 | 3.29 | 2.27 | 3.20 | 2.22 | 4.24 | 5.75 | 3.63 | 5.54 | 1.56 | 5.31 | 7.34 | 4.67 | 6.73 |
| -0.9 | 1.65 | 1.65 | 2.96 | – | 2.94 | 1.43 | 3.43 | 5.02 | – | 4.87 | 0.65 | 4.35 | 6.23 | – | 5.79 |
| -0.8 | 1.16 | 1.16 | 2.92 | – | 2.61 | 0.61 | 2.57 | 4.09 | – | 4.12 | -0.25 | 3.30 | 5.04 | – | 4.79 |
| -0.7 | 0.61 | 1.48 | 2.44 | – | 2.20 | -0.21 | 1.96 | 2.95 | – | 3.30 | -1.15 | 2.26 | 3.50 | – | 3.73 |
| -0.6 | -0.00 | 0.93 | 1.96 | – | 1.69 | -1.05 | 1.18 | 1.70 | – | 2.40 | -2.02 | 1.23 | 1.40 | – | 2.64 |
| -0.5 | -0.67 | 0.05 | 1.07 | – | 1.06 | -1.88 | 0.27 | -0.01 | – | 1.44 | -2.87 | 0.26 | -2.21 | – | 1.56 |

Table 1: SNR (in dB) Performances of Combined Source-Channel Coding Schemes Using MAP Detection for AR(1) Source with $\rho = 0.9$; $k$ = Dimension; $R$ = Rate (Bits/Sample); $\epsilon$ = Channel Bit Error Rate.

| $\log_{10}\epsilon$ | k = 2; R = 1.0 | | | | | k = 4; R = 1.0 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | VQ | VQ+ Inst. MAP | VQ+ Seq. MAP | COVQ [8] | VQ+ Inst. MMSE | VQ | VQ+ Inst. MAP | VQ+ Seq. MAP | COVQ [8] | VQ+ Inst. MMSE |
| $-\infty$ | 7.92 | 7.92 | 7.92 | 7.87 | 7.92 | 10.19 | 10.19 | 10.19 | 10.18 | 10.19 |
| -3.0 | 7.80 | 7.86 | 7.87 | – | 7.88 | 9.94 | 10.01 | 10.06 | – | 10.04 |
| -2.9 | 7.77 | 7.84 | 7.86 | – | 7.86 | 9.88 | 9.96 | 10.03 | – | 10.00 |
| -2.8 | 7.74 | 7.82 | 7.84 | – | 7.85 | 9.81 | 9.91 | 10.00 | – | 9.96 |
| -2.7 | 7.69 | 7.79 | 7.82 | – | 7.83 | 9.71 | 9.83 | 9.95 | – | 9.90 |
| -2.6 | 7.63 | 7.76 | 7.81 | – | 7.81 | 9.60 | 9.76 | 9.90 | – | 9.83 |
| -2.5 | 7.56 | 7.72 | 7.78 | – | 7.78 | 9.46 | 9.66 | 9.83 | – | 9.76 |
| -2.4 | 7.48 | 7.67 | 7.75 | – | 7.75 | 9.29 | 9.54 | 9.75 | – | 9.66 |
| -2.3 | 7.37 | 7.61 | 7.70 | 7.31 | 7.71 | 9.08 | 9.40 | 9.64 | 9.15 | 9.54 |
| -2.2 | 7.23 | 7.53 | 7.65 | – | 7.66 | 8.84 | 9.22 | 9.51 | – | 9.39 |
| -2.1 | 7.07 | 7.43 | 7.58 | – | 7.59 | 8.55 | 9.01 | 9.37 | – | 9.22 |
| -2.0 | 6.88 | 7.30 | 7.49 | 6.83 | 7.51 | 8.21 | 8.76 | 9.18 | 8.37 | 9.01 |
| -1.9 | 6.65 | 7.15 | 7.39 | – | 7.42 | 7.82 | 8.46 | 8.97 | – | 8.76 |
| -1.8 | 6.37 | 6.97 | 7.25 | – | 7.30 | 7.37 | 8.11 | 8.72 | – | 8.48 |
| -1.7 | 6.05 | 6.74 | 7.19 | – | 7.16 | 6.87 | 7.70 | 8.43 | – | 8.14 |
| -1.6 | 5.67 | 6.47 | 7.06 | – | 6.99 | 6.31 | 7.24 | 8.08 | – | 7.77 |
| -1.5 | 5.24 | 6.14 | 6.86 | – | 6.77 | 5.69 | 6.72 | 7.66 | – | 7.34 |
| -1.4 | 4.75 | 5.92 | 6.62 | – | 6.52 | 5.02 | 6.14 | 7.18 | – | 6.85 |
| -1.3 | 4.21 | 5.65 | 6.31 | 4.37 | 6.21 | 4.31 | 5.51 | 6.62 | 6.23 | 6.31 |
| -1.2 | 3.61 | 5.17 | 5.95 | – | 5.85 | 3.56 | 4.82 | 6.00 | – | 5.72 |
| -1.1 | 2.97 | 4.63 | 5.50 | – | 5.42 | 2.77 | 4.06 | 5.29 | – | 5.08 |
| -1.0 | 2.27 | 4.01 | 5.04 | 2.76 | 4.91 | 1.95 | 3.27 | 4.49 | 4.65 | 4.39 |
| -0.9 | 1.54 | 3.30 | 4.44 | – | 4.30 | 1.11 | 2.42 | 3.55 | – | 3.66 |
| -0.8 | 0.77 | 2.50 | 3.82 | – | 3.60 | 0.25 | 1.56 | 2.49 | – | 2.90 |
| -0.7 | -0.03 | 1.67 | 2.82 | – | 2.83 | -0.62 | 0.69 | 1.28 | – | 2.14 |
| -0.6 | -0.85 | 0.79 | 1.42 | – | 1.98 | -1.49 | -0.15 | -0.16 | – | 1.41 |
| -0.5 | -1.68 | -0.13 | -0.23 | – | 1.11 | -2.37 | -0.78 | -2.34 | – | 0.74 |

Table 2: SNR (in dB) Performances of Combined Source-Channel Coding Schemes Using MAP Detection and MMSE Decoding for AR(1) Source with $\rho = 0.9$; $k =$ Dimension; $R =$ Rate (Bits/Sample); $\epsilon =$ Channel Bit Error Rate.

| $\log_{10} \epsilon_a$ | $\log_{10} \epsilon_d$ | $k = 1; R = 3.0$ | | $k = 4; R = 1.0$ | |
|---|---|---|---|---|---|
| | | COVQ+ MMSE | COVQ [7] | COVQ+ MMSE | COVQ [8] |
| | $-\infty$ | 9.14 | – | 6.31 | – |
| | -1.9 | 9.62 | – | 6.87 | – |
| | -1.8 | 9.62 | – | 6.90 | – |
| | -1.7 | 9.63 | – | 6.96 | – |
| -1.3 | -1.6 | 9.63 | – | 6.86 | – |
| | -1.5 | 9.61 | – | 6.87 | – |
| | -1.4 | 9.58 | – | 6.82 | – |
| | -1.3 | 9.54 | 6.47 | 6.52 | 6.23 |
| | $-\infty$ | 6.73 | – | 4.39 | – |
| | -1.6 | 7.45 | – | 5.30 | – |
| | -1.5 | 7.47 | – | 5.33 | – |
| | -1.4 | 7.48 | – | 5.35 | – |
| -1.0 | -1.3 | 7.49 | – | 5.23 | – |
| | -1.2 | 7.12 | – | 5.18 | – |
| | -1.1 | 7.03 | – | 5.11 | – |
| | -1.0 | 6.23 | 4.67 | 5.02 | 4.65 |

Table 3: SNR (in dB) Performances of Combined Source-Channel Coding Scheme Using COVQ Encoding and MMSE Decoding for AR(1) Source with $\rho = 0.9$; $k =$ Dimension; $R =$ Rate (Bits/Sample); $\epsilon_d =$ Design Bit Error Rate of COVQ Encoder; $\epsilon_a =$ Actual Channel Bit Error Rate.
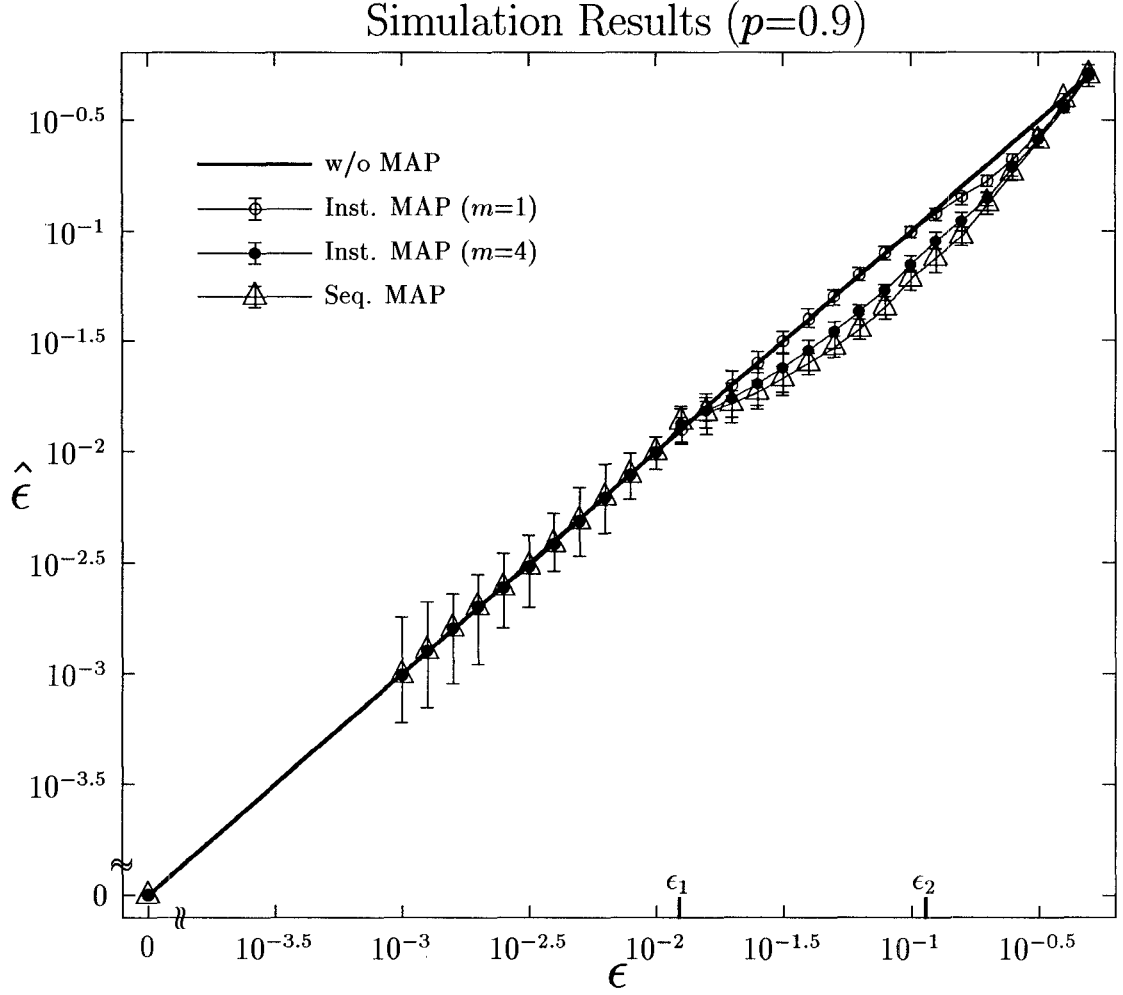
Figure 1: Performances of Instantaneous and Sequence MAP Detectors for Binary Markov Source with $p$=0.9; $\epsilon$=Channel Bit Error Rate; $\hat{\epsilon}$ = Pr{error}; $m$=Block Size of Instantaneous MAP Detector; Error Bars Indicate the Minimum and Maximum Values of $\hat{\epsilon}$ in the Simulation.

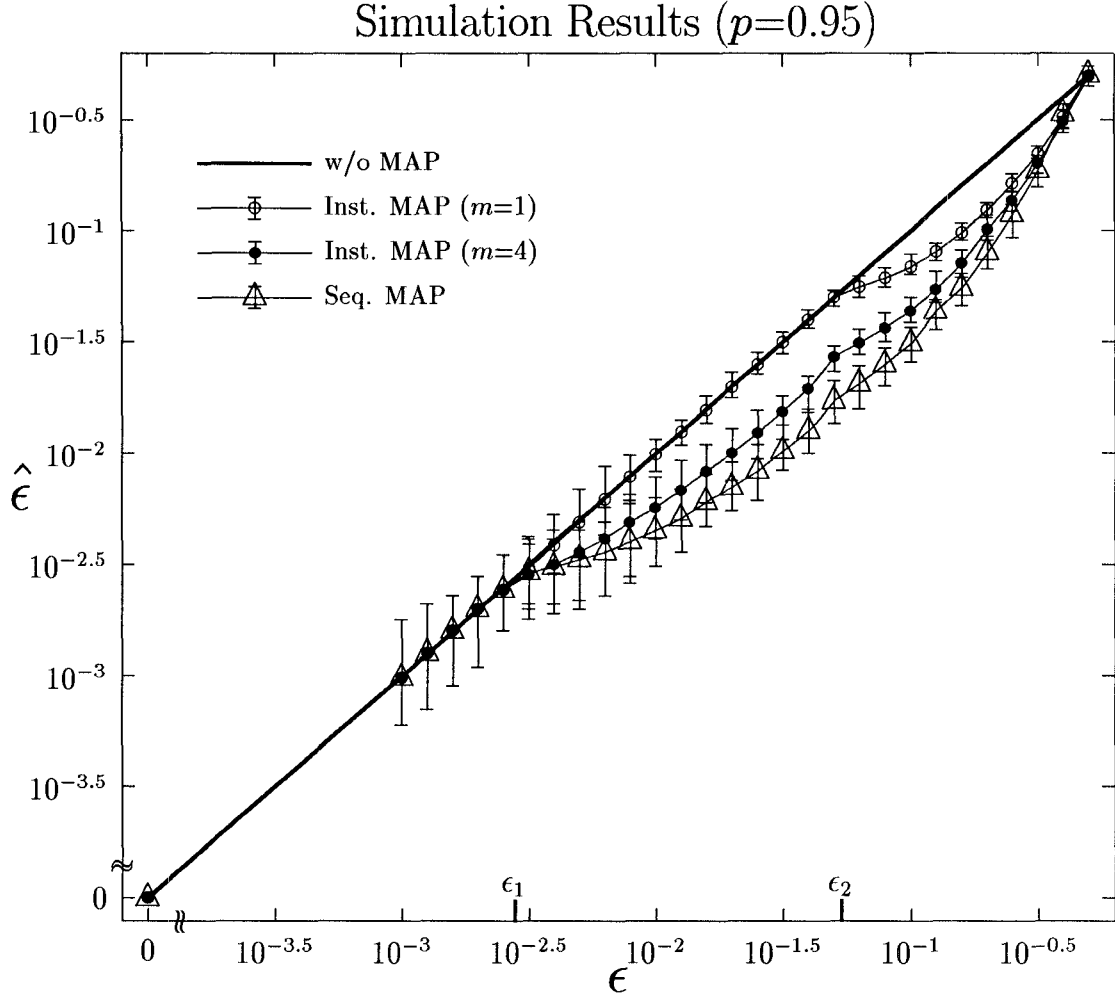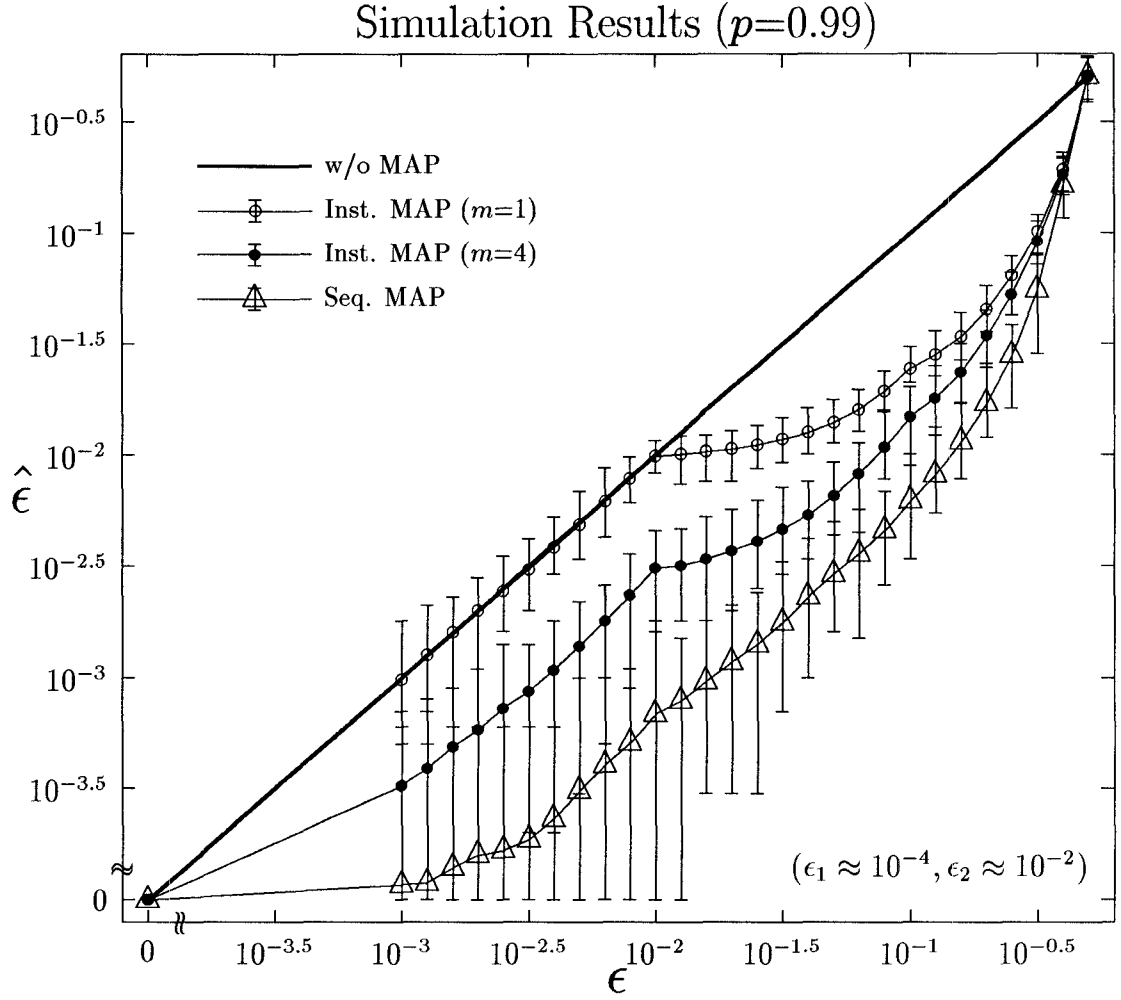Figure 2: Performances of Instantaneous and Sequence MAP Detectors for Binary Markov Source with $p$=0.95; $\epsilon$=Channel Bit Error Rate; $\hat{\epsilon} = \Pr\{\text{error}\}$; $m$=Block Size of Instantaneous MAP Detector; Error Bars Indicate the Minimum and Maximum Values of $\hat{\epsilon}$ in the Simulation.

Figure 3: Performances of Instantaneous and Sequence MAP Detectors for Binary Markov Source with $p$=0.99; $\epsilon$=Channel Bit Error Rate; $\hat{\epsilon} = \Pr\{\text{error}\}$; $m$=Block Size of Instantaneous MAP Detector; Error Bars Indicate the Minimum and Maximum Values of $\hat{\epsilon}$ in the Simulation.
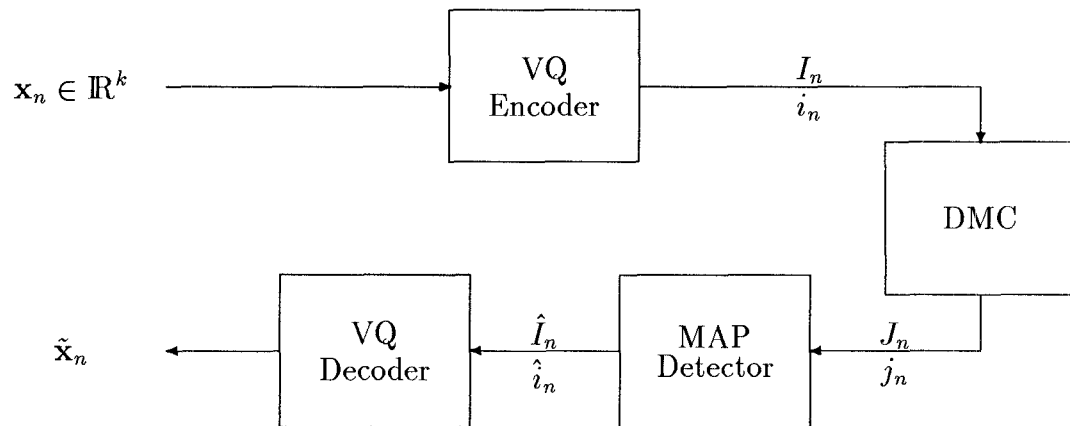
Figure 4: Block Diagram of Combined Source-Channel Coding System Using MAP Detection.
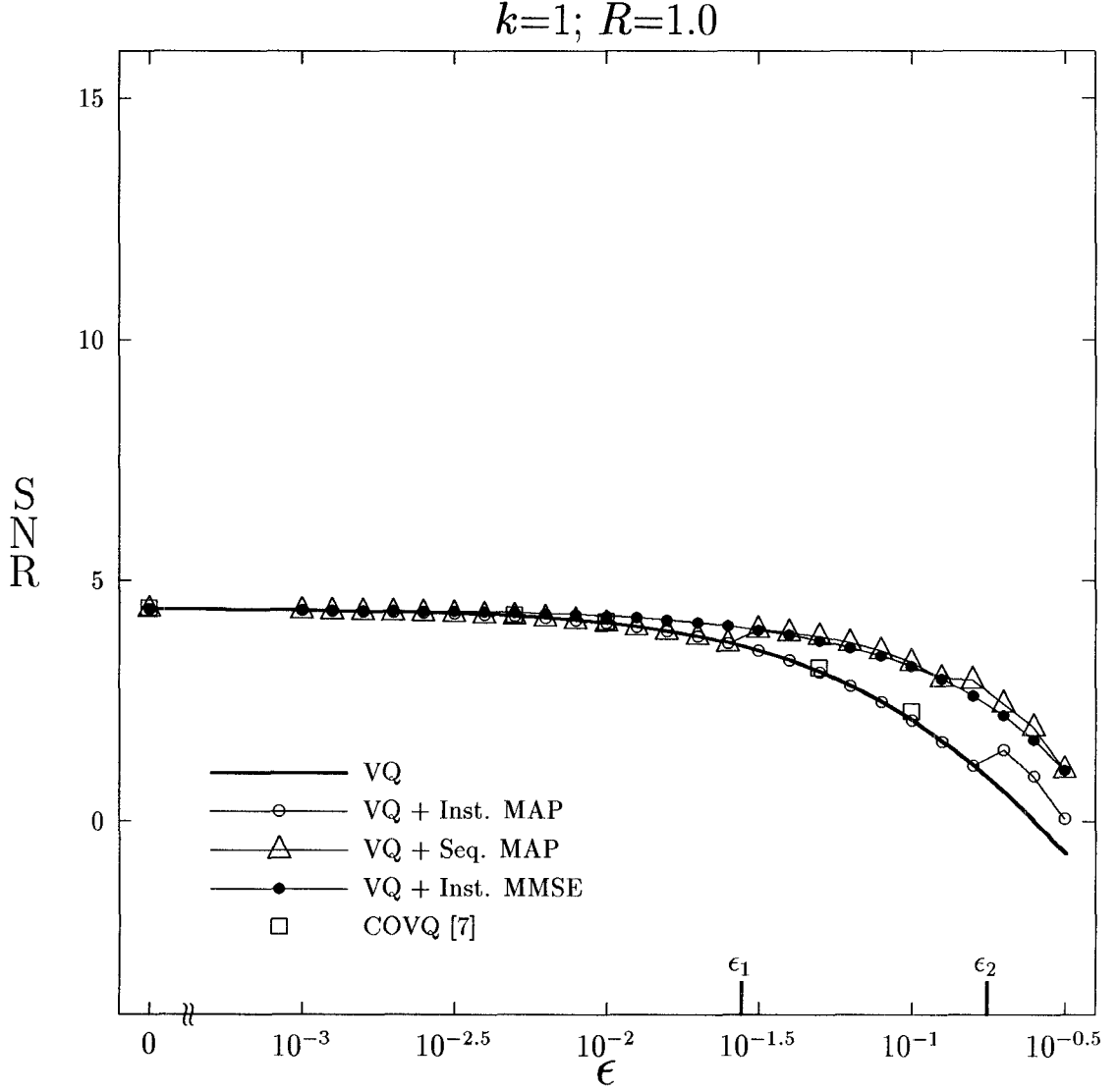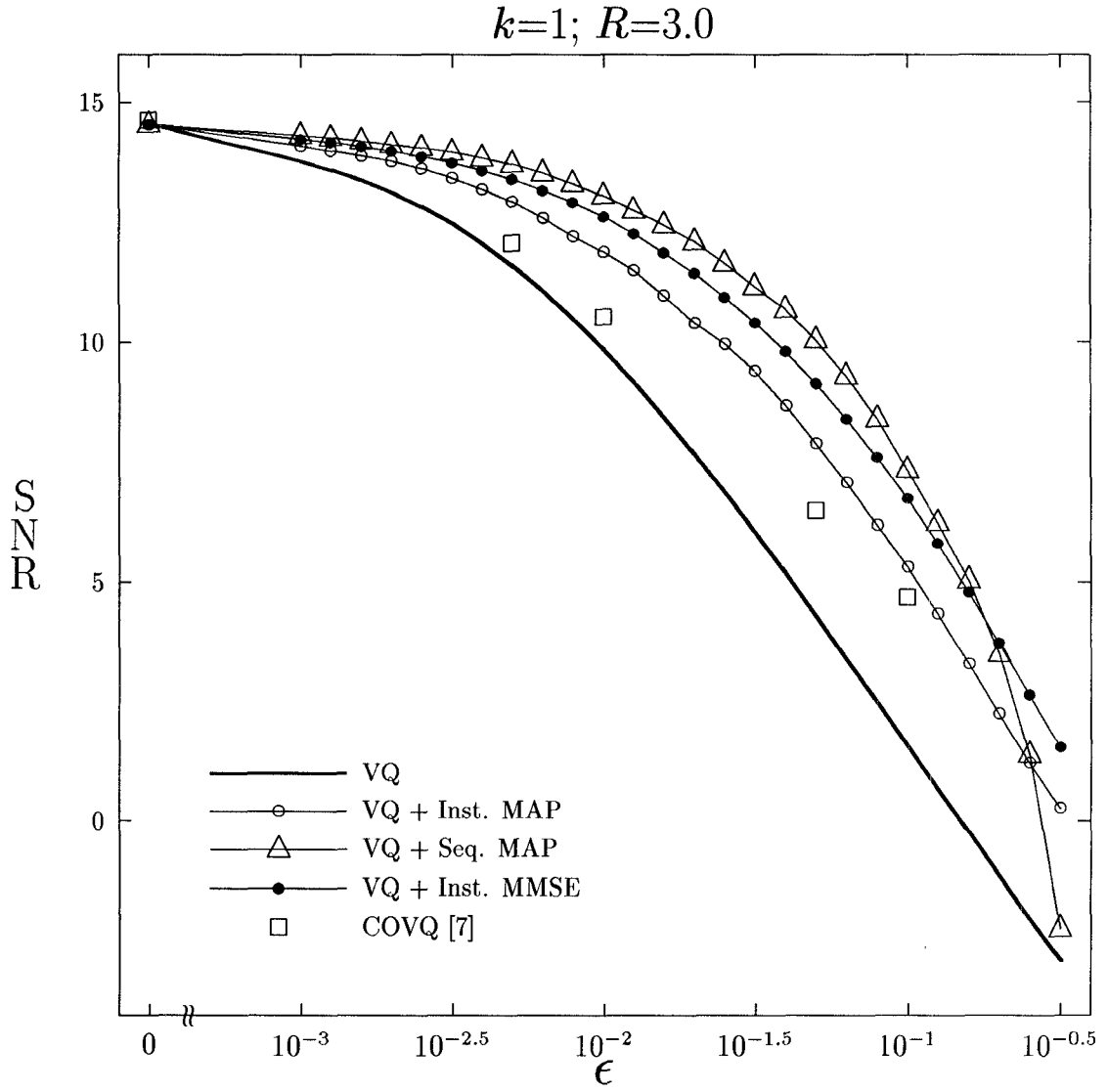
Figure 5: SNR (in dB) Performances of Combined Source-Channel Coding Schemes Using MAP Detection and MMSE Decoding for AR(1) Source with $\rho = 0.9$; $k =$ Dimension; $R =$ Rate (Bits/Sample); $\epsilon =$ Channel Bit Error Rate.

24

$k=1; \ R=2.0$
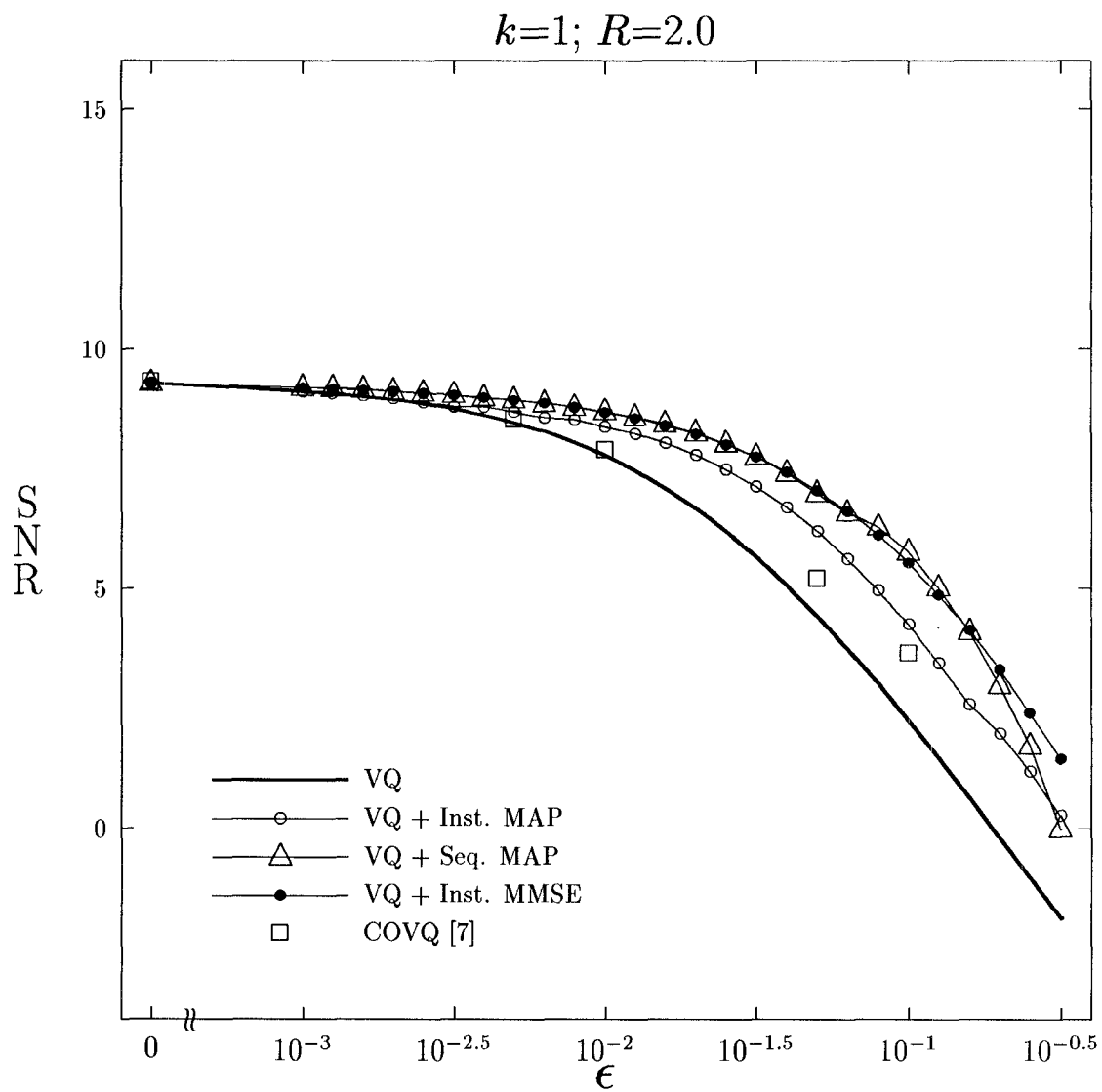
Figure 6: SNR (in dB) Performances of Combined Source-Channel Coding Schemes Using MAP Detection and MMSE Decoding for AR(1) Source with $\rho = 0.9$; $k =$ Dimension; $R =$ Rate (Bits/Sample); $\epsilon =$ Channel Bit Error Rate.
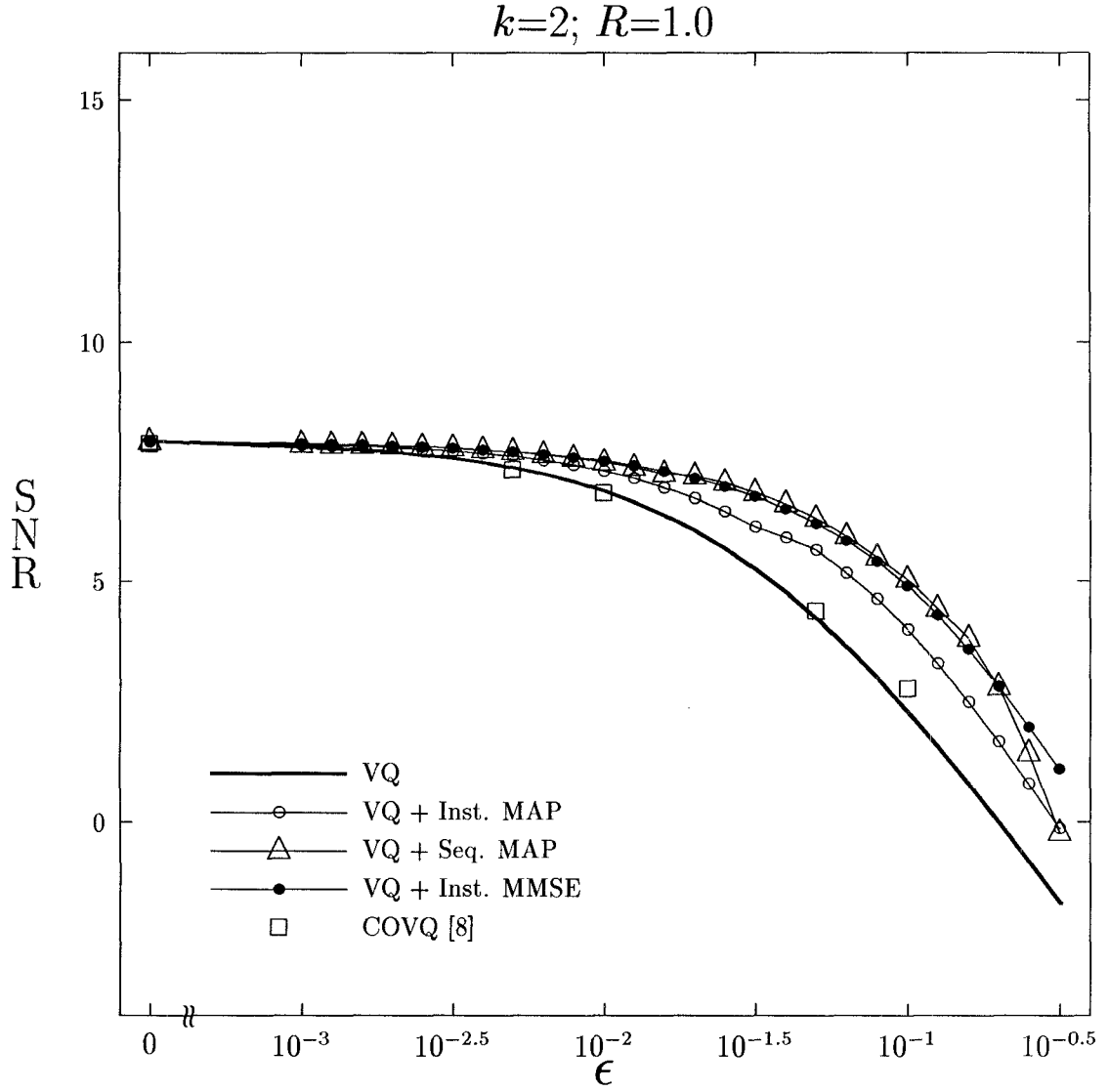
Figure 7: SNR (in dB) Performances of Combined Source-Channel Coding Schemes Using MAP Detection and MMSE Decoding for AR(1) Source with $\rho = 0.9$; $k =$ Dimension; $R =$ Rate (Bits/Sample); $\epsilon =$ Channel Bit Error Rate.
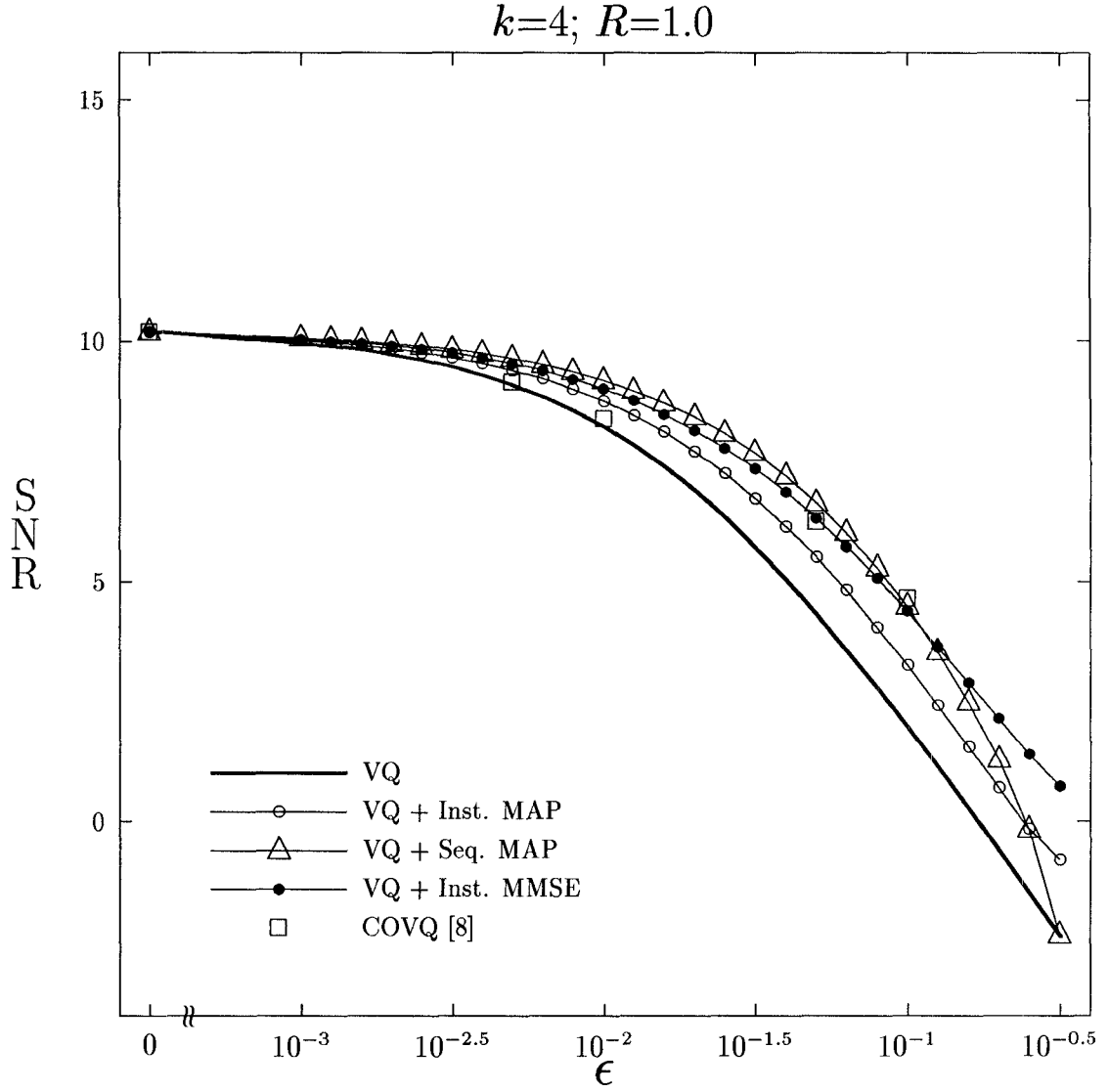
Figure 8: SNR (in dB) Performances of Combined Source-Channel Coding Schemes Using MAP Detection and MMSE Decoding for AR(1) Source with $\rho = 0.9$; $k =$ Dimension; $R = $ Rate (Bits/Sample); $\epsilon = $ Channel Bit Error Rate.

Figure 9: SNR (in dB) Performances of Combined Source-Channel Coding Schemes Using MAP Detection and MMSE Decoding for AR(1) Source with $\rho = 0.9$; $k =$ Dimension; $R =$ Rate (Bits/Sample); $\epsilon =$ Channel Bit Error Rate.