Channel-Optimized Quantization With Soft-Decision Demodulation for Space–Time Orthogonal Block-Coded Channels

Firouz Behnamfar, *Member, IEEE*, Fady Alajaji, *Senior Member, IEEE*, and Tamás Linder, *Senior Member, IEEE*

***Abstract*—We introduce three soft-decision demodulation channel-optimized vector quantizers (COVQs) to transmit analog sources over space–time orthogonal block (STOB)-coded flat Rayleigh fading channels with binary phase-shift keying (BPSK) modulation. One main objective is to judiciously utilize the soft information of the STOB-coded channel in the design of the vector quantizers while keeping a low system complexity. To meet this objective, we introduce a simple space–time decoding structure that consists of a space–time soft detector, followed by a linear combiner and a scalar uniform quantizer with resolution q. The concatenation of the space–time encoder/modulator, fading channel, and space–time receiver can be described by a binary-input, 2^q-output discrete memoryless channel (DMC). The scalar uniform quantizer is chosen so that the capacity of the equivalent DMC is maximized to fully exploit and capture the system's soft information by the DMC. We next determine the statistics of the DMC in closed form and use them to design three COVQ schemes with various degrees of knowledge of the channel noise power and fading coefficients at the transmitter and/or receiver. The performance of each quantization scheme is evaluated for memoryless Gaussian and Gauss–Markov sources and various STOB codes, and the benefits of each scheme is illustrated as a function of the antenna-diversity and soft-decision resolution q. Comparisons to traditional coding schemes, which perform separate source and channel coding operations, are also provided.**

***Index Terms*—Channel-optimized vector quantization, convolutional codes, diversity, joint source-channel coding, multiantenna fading channels, soft-decision demodulation, space–time coding, wireless communications.**

I. Introduction

SPACE–TIME orthogonal block (STOB) coding [3], [32] was recently developed to improve the error performance of wireless communication systems. Like many other error protection schemes[1] that are designed in the spirit of Shannon's separation theorem [26], space–time codes are designed to operate on uniform independent and identically distributed (i.i.d.) bit streams. However, Shannon's separation theorem does not take into consideration constraints on system complexity and delay. As real-world communication systems are constrained, systems with independent source and channel codes, known as tandem coding systems, may have inferior performance compared with systems which perform source and channel coding jointly. This issue was quantitatively studied, for example, in [19], where joint source-channel coding was shown to outperform tandem coding for systems having delay or complexity below a certain threshold. Furthermore, in a recent information theoretic study [35], it was shown that the joint source-channel coding reliability function (the error exponent of optimal joint source-channel coding) can be twice as large as the tandem coding reliability function (the error exponent of concatenated optimal source and channel coding) for a large class of discrete memoryless source and channel pairs.

Joint source-channel coding may be implemented in various ways. When the input to the space–time encoder is a nonuniform binary sequence, maximum *a posteriori* (MAP) detection may be applied to enhance detection and improve system performance. For single-input single-output (SISO) systems, joint source-channel coding via MAP detection is studied, for example, in [2], [21], [23], [25], and [29]. For STOB-coded channels, this problem is considered in [6], where a closed-form expression for the pairwise error probability (PEP) of symbols that undergo STOB coding and MAP detection is derived and significant improvements are shown over maximum-likelihood (ML) detection. Another joint source-channel coding approach is the optimization of index assignment in vector quantizers. This approach is studied in [7], [13], and [34] for SISO channels. The study in [7] considers hard decoding and applies a simulated annealing based algorithm to minimize the distortion caused by the channel noise via optimizing the index assignment.

In this paper, we consider the transmission of continuous-alphabet (analog) sources over channels with multiple transmit and multiple receive antennas. We employ channel-optimized vector quantizers (COVQs), which are other joint source-channel coders for compressing the source while rendering it robust against channel errors. COVQ design was originally studied in [8], [14] for arbitrary discrete memoryless channels. In [1] and [22], COVQ with soft-decision demodula-

[1]Throughout this paper, by "error protection scheme," we mean the wider class of techniques, which includes classical error-correction codes as well as diversity (such as space–time modulation) codes.

Manuscript received December 27, 2004; revised September 18, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Helmut Boelcskei. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Premier's Research Excellence Award (PREA) of Ontario. Parts of this paper were presented at The 2003 Canadian Workshop on Information Theory (CWIT'03), Waterloo, ON, Canada, May 2003 and The 2004 International Symposium on Information Theory and Its Applications (ISITA'04), Parma, Italy, October 2004.

F. Behnamfar and F. Alajaji are with the Department of Mathematics and Statistics and the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: firouz@mast.queensu.ca; fady@mast.queensu.ca).

T. Linder is with the Department of Mathematics and Statistics and the Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N6, Canada. He is also with the Computer and Automation Research Institute of the Hungarian Academy of Sciences, Budapest H-1111, Hungary (e-mail: linder@mast.queensu.ca).

Digital Object Identifier 10.1109/TSP.2006.879328

tion was implemented for channels with Rayleigh flat fading, and white and colored additive Gaussian noise and intersymbol interference, respectively. In [28] and [30], COVQ systems with optimal and suboptimal soft-decoding based on Hadamard matrices were introduced and studied.

We consider a multiple-input multiple-output (MIMO) system used in conjunction with STOB coding. Our proposed system adds only two blocks, with modest computational needs, to a conventional space–time coded system. It is well known that for STOB codes with perfect channel state information available at the receiver, the MIMO channel can be explicitly decomposed into a set of independent parallel SISO channels with chi-squared fading [4], [5], [27]. However, when the system has multiple receive antennas, it is not obvious how to extend hard decoding to soft-decision demodulation.[2] Therefore, an important task is the proper processing of the received signals from the different antennas. We propose to address this problem by performing space–time soft-decoding followed by linear combining at the receiver. The linear combiner has the following key advantages: 1) it has a very simple structure, 2) it allows the COVQ index transition probabilities to be determined in closed form, 3) its output is continuous, making soft-decision demodulation possible, and 4) its outputs are sufficient statistics and do not cause loss of soft information. Inspired by the work in [1] and [22], we use soft-decision demodulation as opposed to soft-decoding methods such as in [28] to exploit the soft information available at the output of the linear combiner. Our choice is motivated by several factors. First, soft-decision demodulation may be implemented via a q-bit uniform quantizer at the receiver (not to be confused with the COVQ blocks at the transmitter and the receiver), which makes the task of decoding computationally simple. In contrast, the first version of soft decoding in [28] needs the computation of trigonometric functions and matrix multiplication and the second version also requires matrix inversion. Second, we observe that the performance of our decoder converges to that of soft decoding as the resolution q of the soft-decision demodulator grows to infinity. It also requires less computational complexity (although its storage complexity may be higher).

We show that the concatenation of the space–time encoder, the MIMO channel, the space–time soft-decoder, the combiner, and the uniform scalar quantizer is equivalent to a binary-input, 2^q-output discrete memoryless channel (DMC) used kr times, where k and r are the quantizer dimension and rate, respectively. The step size of the uniform quantizer used for soft-decision demodulation is numerically selected so that the capacity of the equivalent DMC is maximized for each value of the channel signal-to-noise ratio (CSNR). This is a suboptimal criterion, but as the simulation results of [22] demonstrate, there is a substantial correlation between maximizing channel capacity and minimizing distortion.[3] We show that the transition probabilities of this equivalent DMC can be expressed in terms of the symbol PEP of the ML-decoded STOB coded channel. Hence, these probabilities can be determined using the results of [4].

We design three soft-decision demodulation COVQs for the equivalent DMC. The first COVQ is the classical COVQ which assumes that the index transition probabilities are known at both the transmitter and the receiver. The encoder and decoder codebooks are determined iteratively using the modified generalized Lloyd algorithm (GLA) [11]. As the CSNR is not always available at the transmitter, we consider the design of two fixed-encoder adaptive-decoder (FEAD) COVQs. In a FEAD COVQ, the encoder is designed for a fixed CSNR and the decoder, which can estimate the channel fading coefficients and the CSNR, adapts itself to the channel conditions. Our first FEAD COVQ uses only the knowledge of the CSNR at the receiver (as in [31]), while the second one, which we call the on-line FEAD COVQ, employs also the knowledge of the channel fading coefficients at the receiver and, as a result, it outperforms the FEAD COVQ. An important feature of FEAD COVQ is that its decoder codebooks are *computed* in terms of the transmitter parameters, and not through a training process as for classical COVQ. Therefore, this method does not need a large memory at the receiver to store a different codebook for each value of the CSNR. We demonstrate that with a proper choice of the design CSNR, the performance loss of FEAD COVQ can be significantly reduced as compared with the classical COVQ.

Throughout this paper, $x_{i,j}$ denotes a scalar entry at row i and column j of a deterministic matrix $\mathbf{X}$. The i^{th} element of a deterministic vector $\mathbf{x}$ will be denoted by x_i. Similarly, scalar random variable $X_{i,j}$ is an entry of a random matrix $\boldsymbol{X}$ and $\boldsymbol{x}$ is a random vector whose i^{th} element is X_i. Some constants are also indicated by italic capitals, but their difference from scalar random variables should be clear from the context.

The rest of this paper is organized as follows. Section II reviews STOB coding and then describes the system components and their design in detail. The three soft-decision demodulation COVQ schemes are presented in Section III. Section IV provides numerical results, and the conclusions are given in Section V.

II. System Components

The system block diagram is shown in Fig. 1. A COVQ encoder forms a vector $\boldsymbol{x}$ of dimension k from the incoming scalar source samples. It then encodes $\boldsymbol{x}$ at a rate of r bits per sample (bps) into a binary index I of length kr bits. Encoding is specified by the decision regions $\{\mathcal{S}_i\}_{i=0}^{N_e-1}$ $(N_e = 2^{kr})$, which form a partition of the k-dimensional space, using the rule that $I = i$ if $\boldsymbol{x} \in \mathcal{S}_i$. Index I is then sent over the channel and received as index J. The COVQ decoder simply uses $\boldsymbol{y}_J$, the J^{th} element in the codebook, to reconstruct $\boldsymbol{x}$ as $\hat{\boldsymbol{x}} = \boldsymbol{y}_J$. The goal in COVQ design is to minimize the expected value of $||\boldsymbol{x} - \hat{\boldsymbol{x}}||^2$ through finding the optimal partition and codebook.

Our objective is to model the concatenation of the blocks between the encoder and the decoder of the vector quantizer by a discrete channel and then design efficient vector quantization

[2]As it is detailed in (3) and (4), the space–time soft-decoded signals at different receive antennas have various noise powers. The received signals should then be properly combined according to some optimality criterion.

[3]In a related work [12], it is shown that for Gaussian channels, maximizing the channel capacity also minimizes the overall (mean-squared error) distortion.

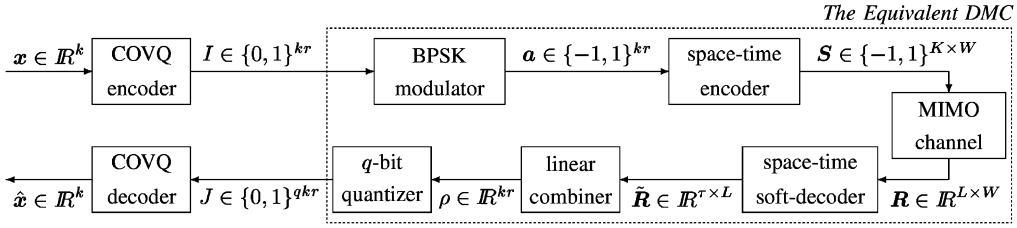

Fig. 1. System block diagram, where every τ bits in a kr-bit index I is transmitted via a space–time codeword $\mathbf{S}$, received as $\boldsymbol{R}$, and space–time soft-decoded as $\tilde{\boldsymbol{R}}$. For simplicity, we have assumed here that $\tau = kr$.

systems for this channel. The system is designed so that it judiciously incorporates the soft information of the STOB-coded channel and admits a closed-form expression for the COVQ transition probabilities. This is achieved by designing a soft-decision space–time receiver which consists of a space–time soft detector, followed by a linear combiner (whose output provides sufficient statistics to compute the conditional COVQ index probabilities), and a simple q-bit scalar uniform quantizer (whose cell size is chosen so that the capacity of the equivalent discrete channel is maximized). As a result of the simple memoryless structure of our receiver and due to the space–time code's orthogonality, we obtain that the equivalent discrete channel is indeed a binary-input 2^q-output memoryless channel whose distribution can be easily determined analytically in terms of the system parameters. In the following, we describe the system components in detail.

A. MIMO Channel

The communication system considered here employs K transmit and and L receive antennas. The channel is assumed to be Rayleigh flat fading so that the path gain from transmit antenna i to receive antenna j, denoted by $H_{j,i}$, has a unit-variance i.i.d. Rayleigh distribution. We assume that the receiver has perfect knowledge of the path gains. It is also assumed that the channel is quasi-static, meaning that the path gains remain constant during a space–time codeword transmission, but vary in an i.i.d. fashion among codeword intervals. The additive noise at receive antenna j and symbol interval t, $N_{j,t}$, is assumed to have a zero-mean unit-variance Gaussian distribution, denoted by $\mathcal{N}(0,1)$. Based on the above, for a CSNR γ_s at each receive antenna, the signal at receive antenna j at symbol interval t can be written as $R_{j,t} = \sqrt{\gamma_s/K}\sum_{i=1}^{K} H_{j,i}s_{i,t} + N_{j,t}$, for $t = 1,2,\ldots,W$, where W is the codeword length and the space–time modulated symbols $\sqrt{\gamma_s/K}\{s_{i,t}\}_{i=1}^{K}$ are simultaneously transmitted (see Section II-B). In matrix form, we have

$$\boldsymbol{R} = \sqrt{\frac{\gamma_s}{K}}\boldsymbol{H}\boldsymbol{S} + \boldsymbol{N}. \tag{1}$$

We assume that the noise, signal, and fading coefficients are statistically independent of each other.

B. Space–Time Orthogonal Block Codes

Let $\mathbf{c} = (c_1,\ldots,c_\tau)^T$ be a vector of τ consecutive constellation points and $\mathbf{S} = (\mathbf{s}_1,\ldots,\mathbf{s}_W)$ be the space–time code corresponding to it, where ^{T} indicates transposition. In the case of STOB codes, we have $W = g\tau$, where g is the coding gain and $\mathbf{S}\mathbf{S}^\dagger = g\|\mathbf{c}\|^2\mathbf{I}_K$, where $\mathbf{I}_K$ is the $K \times K$ identity matrix and † represents complex conjugate transposition. As an example, for the real code in [32, eq. (4)], $g = 1$ and $W = \tau = 4$, and for Alamouti's code [3], $g = 1$ and $W = \tau = 2$. Our system uses binary phase-shift keying (BPSK) modulation with baseband signals located at -1 and 1. It can be shown (see [4] and [16]) that when the signal constellation is real, we have

$$\boldsymbol{r}_j = \sqrt{\frac{\gamma_s}{K}}\tilde{\boldsymbol{H}}_j\mathbf{c} + \boldsymbol{n}_j \qquad j = 1,\cdots,L \tag{2}$$

where $\boldsymbol{r}_j = (R_{j,1},\ldots,R_{j,W})^T$, $\boldsymbol{n}_j = (N_{j,1},\ldots,N_{j,W})^T$, and $\tilde{\boldsymbol{H}}_j$ is a $W \times \tau$ matrix which is derived from the j^{th} row of $\boldsymbol{H}$, and it has orthogonal columns (see [16]), i.e., $\tilde{\boldsymbol{H}}_j^T\tilde{\boldsymbol{H}}_j = g\bar{H}_j\mathbf{I}_\tau$, with $\bar{H}_j = \sum_i H_{j,i}^2$. Therefore, multiplying (2) from the left by $\tilde{\boldsymbol{H}}_j^T$ yields the following at the output of the space–time soft-decoder:

$$\tilde{\boldsymbol{r}}_j \triangleq \tilde{\boldsymbol{H}}_j^T\boldsymbol{r}_j = g'\bar{H}_j\mathbf{c} + \tilde{\boldsymbol{n}}_j \tag{3}$$

where $g' = g\sqrt{\gamma_s/K}$. Note that each entry $\tilde{R}_{j,t}$ of $\tilde{\boldsymbol{r}}_j$ is associated with only *one* transmitted symbol in $\mathbf{c}$. It is not difficult to show that

$$\tilde{N}_{j,t} \sim \text{i.i.d.}\,\mathcal{N}(0, g\bar{H}_j), \qquad j = 1,\ldots,L, t = 1,\ldots,\tau. \tag{4}$$

It follows that the t^{th} space–time-coded symbol can be detected by only considering the t^{th} entry of the vectors $\tilde{\boldsymbol{r}}_j$, $1 \le j \le L$. For our application, this will imply that the bits corresponding to a COVQ index can be detected independently.

It can be shown (see, for example, [4], [5], and [20]) that the PEP of a pair of STOB-coded BPSK symbols c_i and c_j under ML decoding is given by

$$\begin{aligned}\Lambda(\delta) &\triangleq P(c_i \to c_j)\\ &= E_{\bar{H}}\left[Q(\delta\sqrt{\bar{H}})\right]\\ &= \frac{1}{2}\left(1 - \frac{\delta}{\sqrt{2+\delta^2}}\sum_{k=0}^{KL-1}\binom{2k}{k}\frac{1}{(2\delta^2+4)^k}\right)\end{aligned} \tag{5}$$

where $Q(x) = (1/\sqrt{2\pi})\int_x^\infty e^{-\lambda^2/2}d\lambda$ is the Gaussian error function, $\bar{H} = \sum_{i=1}^{K}\sum_{j=1}^{L} H_{j,i}^2$ and $\delta = \sqrt{2g\gamma_s/K}$.

C. Soft-Decision Demodulation and the Equivalent DMC

Similar to other COVQ-based systems, the decoder of the system proposed in this paper is "memoryless" in the sense that its output is a function of only one transmitted index, as opposed to (a subset of) the entire stream of indices. The optimal receiver, in terms of minimizing the mean-square error (MSE), is therefore given by

$$\hat{\boldsymbol{x}} = E[\boldsymbol{x}|\boldsymbol{R}] = \sum_{i=0}^{N_e-1} E[\boldsymbol{x}|I=i]\Pr(I=i|\boldsymbol{R}) \tag{6}$$

where we have assumed for simplicity that $\tau = kr$ (a similar but slightly more complicated expression for $\hat{\boldsymbol{x}}$ holds for $\tau \neq kr$). The soft decoder in [28] builds a relationship in terms of the Hadamard transform between each centroid $E[\boldsymbol{x}|I=i]$ and the bits of its corresponding encoder output index i. The decoder then maps the received noisy bits at the decoder back into estimates of the source vectors. Here, we propose a suboptimal receiver (based on [22]) which has low complexity and also exploits (in part) the soft information available in channel output.

1) Soft-Decision Demodulation: Linear Combining and Scalar Quantization: Many communication systems employ hard-decoding in processing the received signals of space–time coded systems and so do not exploit the soft information available at the space–time soft-decoder outputs $\tilde{\boldsymbol{R}}$. Methods that exploit soft information can provide significant performance gains. In our case, in addition to using the soft information efficiently, the solution should allow the COVQ index transition probabilities to be determined in closed form, since this is required for the COVQ design and encoding phases. As we illustrate below and in Section IV, linear combining has both of the above properties.

In order to employ the signals of all receive branches, we use Bayes' law to see, from (2), that the conditional index probability in (6) is equal to

$$\begin{aligned}
&\Pr(I=i|\boldsymbol{R}) \\
&\quad = \frac{f_{\boldsymbol{R}|I}(\boldsymbol{R}|i)p(i)}{f_{\boldsymbol{R}}(\boldsymbol{R})} \\
&\quad = \frac{p(i)\prod_{j=1}^{L}\exp\left(-\left\|\boldsymbol{r}_j-\sqrt{\gamma_s/K}\tilde{\boldsymbol{H}}_j\mathbf{c}(i)\right\|^2/2\right)}{\sum_{l=0}^{N_e-1}p(l)\prod_{j=1}^{L}\exp\left(-\left\|\boldsymbol{r}_j-\sqrt{\gamma_s/K}\tilde{\boldsymbol{H}}_j\mathbf{c}(l)\right\|^2/2\right)} \\
&\quad = \frac{p(i)\exp\left(-\sqrt{\gamma_s/K}\left(\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j\right)^T\mathbf{c}(i)\right)}{\sum_{l=0}^{N_e-1}p(l)\exp\left(-\sqrt{\gamma_s/K}\left(\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j\right)^T\mathbf{c}(l)\right)}
\end{aligned} \tag{7}$$

where $f(\cdot)$ and $f(\cdot|\cdot)$ denote unconditional and conditional probability density functions (pdf's) and $\mathbf{c}(i)$ is the vector of BPSK symbols that correspond to index i. We have assumed that $\tau = kr$ in the above derivation and we have treated the general case ($\tau \neq kr$) in the Appendix. The above shows that the *sum* of the $\tilde{\boldsymbol{r}}_j$'s conveys all the information in channel output $\boldsymbol{R}$ that is required by the optimal decoder in (6). In other words, to employ the information contained in the received signals $\tilde{\boldsymbol{R}}$, it is enough (and optimal) to add the entries in each column of $\tilde{\boldsymbol{R}}$ and the combiner should therefore compute $\sum_{j=1}^{L}\tilde{R}_{j,t}$ for $t = 1, 2, \ldots, kr$.

In order to make the blocks that follow the combiner independent of the fading coefficients and the CSNR, we use a normalized version of the combiner output which is given by

$$\rho_t \triangleq \frac{1}{g'\bar{H}}\sum_{j=1}^{L}\tilde{R}_{j,t}. \tag{8}$$

With the above choice, the output of the combiner will be equal to

$$\rho_t = c_t + \nu_t \tag{9}$$

where $\nu_t \triangleq \sum_{j=1}^{L}\tilde{N}_{j,t}/(g'\bar{H})$ is i.i.d. additive Gaussian noise characterized by

$$\nu_t \sim \mathcal{N}\left(0, \frac{K}{g\gamma_s\bar{H}}\right), \qquad t = 1, \ldots, \tau. \tag{10}$$

The linear combiner output ρ_t is next fed into a "uniform" scalar quantizer which acts as the soft-decision demodulator. Let us indicate the decision levels of this quantizer by $\{u_i\}_{i=-1}^{N-1}$, where $N = 2^q$ is the number of the codewords. As ρ_t can take any real value, the quantizer should have two unbounded decision regions. The decision regions of the uniform quantizer are given by

$$u_i = \begin{cases} -\infty, & \text{if } i = -1 \\ (i+1-N/2)\Delta, & \text{if } i = 0, \cdots, N-2 \\ +\infty, & \text{if } i = N-1 \end{cases}$$

and the quantization rule $a(\cdot)$ is simply

$$a(\rho) = i, \quad \text{if } \rho \in (u_{i-1}, u_i], \qquad i = 0, \cdots, N-1.$$

The use of a nonuniform scalar quantizer may lead to improved overall system performance at the cost of a significant increase in complexity, since the determination of the nonequal step-sizes that maximize the capacity of the equivalent discrete channel will require an exhaustive numerical search. Note that when $q = 1$ (i.e., hard-decoding), the above linear combiner together with the q-bit quantizer reduce to a symbol-to-symbol ML decoder.

2) Transition Probabilities of the Equivalent DMC: For COVQ design, we need to derive the transition probabilities of the 2^{kr}-input 2^{qkr}-output discrete channel represented by the concatenation of the space–time encoder, the MIMO channel, the space–time soft-decoder, the linear combiner, and the uniform quantizer. Since the detection of bits that correspond to each quantizer index is decoupled as explained below (4), we note that the discrete channel is equivalent to a binary-input 2^q-output DMC used kr times. We shall refer to this discrete channel as the "equivalent DMC."

TABLE I
CAPACITY (IN BITS/CHANNEL USE) OF THE DMC DERIVED FROM q-BIT SOFT-DECISION DEMODULATION OF BPSK-MODULATED SPACE–TIME-CODED MIMO CHANNEL WITH $K = 2$ AND $L = 1$. Δ IS THE STEP SIZE THAT MAXIMIZES THE CAPACITY AND q IS THE NUMBER OF SOFT-DECISION DEMODULATION BITS

CSNR	$q=1$	$q=2$		$q=3$		$q=4$		$q=5$	
(dB)	C	C	Δ	C	Δ	C	Δ	C	Δ
16.0	0.9945	0.9969	0.414	0.9973	0.219	0.9974	0.183	0.9974	0.108
14.0	0.9881	0.9929	0.403	0.9937	0.214	0.9939	0.109	0.9940	0.109
12.0	0.9752	0.9842	0.394	0.9858	0.209	0.9862	0.107	0.9864	0.111
10.0	0.9506	0.9660	0.390	0.9695	0.207	0.9702	0.107	0.9705	0.113
8.0	0.9070	0.9333	0.393	0.9381	0.208	0.9392	0.108	0.9397	0.116
6.0	0.8374	0.8760	0.405	0.8833	0.214	0.8849	0.111	0.8855	0.120
4.0	0.7386	0.7891	0.430	0.7987	0.226	0.8009	0.118	0.8015	0.128
2.0	0.6164	0.6741	0.472	0.6852	0.248	0.6877	0.129	0.6884	0.139
0.0	0.4849	0.5427	0.536	0.5540	0.280	0.5565	0.146	0.5572	0.157
-2.0	0.3608	0.4142	0.627	0.4223	0.327	0.4246	0.170	0.4252	0.183
-4.0	0.2560	0.2974	0.751	0.3057	0.390	0.3076	0.203	0.3081	0.219

The required set of the transition probabilities for the binary-input 2^q-output DMC are $P(m|b)$, where b is a data bit and m takes values in the set $\{0, 1, \ldots, 2^q - 1\}$. Decision is made in favor of the m^{th} codepoint if the output of the linear combiner falls into the $(u_{m-1}, u_m]$ interval of length Δ. Using (9) and (10), we can write

$$\begin{aligned} P(m|b, \boldsymbol{H}) &= P(u_{m-1} \le c_t + \nu_t < u_m | \boldsymbol{H}) \\ &= Q\left((u_{m-1} - c_t)\delta\sqrt{\bar{H}}\right) \\ &\quad - Q\left((u_m - c_t)\delta\sqrt{\bar{H}}\right) \end{aligned} \tag{11}$$

where c_t is the BPSK signal which corresponds to b. The expectation over $\bar{\boldsymbol{H}}$ of each of the above $Q(\cdot)$ functions can be determined using (5) to yield

$$P(m|b) = \Lambda\left((u_{m-1} - c_t)\delta\right) - \Lambda\left((u_m - c_t)\delta\right) \tag{12}$$

where $\Lambda(\cdot)$ is defined in (5). Note that the DMC transition probability matrix is symmetric in the sense of [10].

For our k-dimensional COVQ with rate r shown in Fig. 1, we denote the natural binary representation of the index of a decision region by $\{b_1, b_2, \ldots, b_{kr}\}$ and that of a codevector by $\{m_1, m_2, \ldots, m_{kr}\}$. The COVQ index transition probabilities can hence be computed as

$$P_{J|I}(j|i) = \prod_{t=1}^{kr} P(m_t|b_t) \tag{13}$$

with $P(m_t|b_t)$ given in (12).

D. Step Size of the Uniform Quantizer at the Decoder

The final design parameter of the system is Δ, the step size of the uniform quantizer at the receiver. For a given soft-decision resolution q and CSNR γ_s, we numerically select the $\Delta = \Delta(q)$ which maximizes the capacity of the equivalent DMC. This is a suboptimal criterion since our ultimate goal is minimizing the MSE, not maximizing the capacity, but as the simulation results of [22] demonstrate, there is a strong correlation between having a high channel capacity and reduced MSE distortion.

For a given soft-decision resolution q and CSNR γ_s, we determine the step size Δ which maximizes the capacity of the DMC by maximizing the mutual information between the DMC input and output. Because the channel transition probability matrix is symmetric, a uniform input distribution achieves channel capacity [10]. Note that the step-size does not depend on the rate or dimension of the COVQ (used to quantize the source), and is only a function of q and the CSNR. As a typical set of results, we list, in Table I, the capacity of the equivalent DMC versus the "optimal" step-size of the uniform quantizer for Alamouti's [3] dual transmit single receive setup. Similar results can be derived for systems with a different number of transmit antennas, receive antennas, or space–time codes. As shown in Fig. 2, when the step-size is very small (close to zero) or very large, soft-decision demodulation does not significantly increase channel capacity. We also note that if the CSNR is high, soft-decision demodulation is not very beneficial in terms of improving channel capacity. However, for moderate and low CSNRs, and with the optimal choice of Δ, soft-decision demodulation significantly increases channel capacity. For example, at CSNR $= -2$ dB in

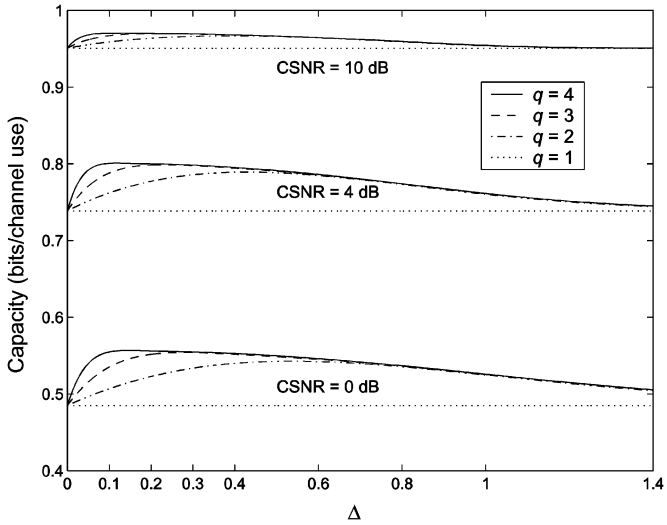

Fig. 2. DMC capacity versus the step size of the uniform quantizer, $K = 2$, $L = 1$.

Table I, there is a 15% benefit in using soft-decision demodulation with $q = 5$ bits. Also note that the channel capacity increases less than 1% from $q = 3$ to $q = 5$, even for severe channel conditions. This shows that typically $q = 3$ achieves most of the capacity gain offered by soft-decision demodulation. Finally, we note, as expected, that as $q \to \infty$, the value of Δ, which maximizes the capacity of the DMC, goes to zero.

III. Quantization With Soft-Decision Demodulation

A. Soft-Decision Demodulation COVQ

The transition probability given in (13) can be used in the modified generalized GLA algorithm [11] to design a soft-decision demodulation COVQ for space–time-coded MIMO channels as explained below. Every input k-tuple is encoded at a rate of r bits per sample (bps). Therefore, the input space is partitioned into $N_e = 2^{kr}$ subsets. As we use BPSK modulation, a vector of kr real-valued signals is received for every transmitted index. This vector is soft-decision decoded at a rate of q bits per dimension. Therefore, each k-dimensional source vector is decoded to one of the $N_d = 2^{qkr}$ codevectors. The input space partitioning and the codebook are optimized based on two necessary conditions for optimality using training data $\{\boldsymbol{x}_m, m = 0, \ldots, M-1\}$ as follows.

- *The nearest neighbor condition*: For a fixed codebook and $i = 0, \cdots, N_e - 1$, the optimal partition $\mathcal{P}^* = \{\mathcal{S}_i^*\}$ is

$$\mathcal{S}_i^* = \left\{ \boldsymbol{x} : \sum_{j=0}^{N_d-1} P_{J|I}(j|i) d(\boldsymbol{x}, \boldsymbol{y}_j) \le \sum_{j=0}^{N_d-1} P_{J|I}(j|l) d(\boldsymbol{x}, \boldsymbol{y}_j), \forall l \right\} \tag{14}$$

  where $\boldsymbol{x}$ is a training vector, $\{\boldsymbol{y}_j, j = 0, \ldots, N_d - 1\}$ is the codebook, $d(\boldsymbol{x}, \boldsymbol{y})$ is the squared Euclidean distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, and ties are broken according to a preset rule.

- *The centroid condition*: Given a partition $\mathcal{P} = \{\mathcal{S}_i, i = 0, \ldots, N_e - 1\}$, the optimal codebook $\mathcal{C}^* = \{\boldsymbol{y}_j^*\}$ is

$$\boldsymbol{y}_j^* = \frac{\sum_{i=0}^{N_e-1} P_{J|I}(j|i) \sum_{l:\boldsymbol{x}_l \in \mathcal{S}_i} \boldsymbol{x}_l}{\sum_{i=0}^{N_e-1} P_{J|I}(j|i) |\mathcal{S}_i|}, \qquad j = 0, \cdots, N_d - 1. \tag{15}$$

  where $|\mathcal{S}_i|$ is the number of the training vectors in $\mathcal{S}_i$.

Note that the above steps do not increase the overall distortion. As distortion is bounded from below, convergence to a local minimum is therefore guaranteed via an iterative execution of the above two steps.

Training consists of using the above conditions iteratively to update the codebook until the average training distortion given by

$$D = \frac{1}{kM} \sum_{m=0}^{M-1} \sum_{j=0}^{N_d-1} P_{J|I}\left(j|\mathcal{I}(\boldsymbol{x}_m)\right) d(\boldsymbol{x}_m, \boldsymbol{y}_j)$$

converges, where $\mathcal{I}(\boldsymbol{x})$ is the index of the partition cell to which $\boldsymbol{x}$ belongs.

B. Fixed-Encoder Adaptive-Decoder Soft-Decision Demodulation COVQ

Equations (5) and (11)–(13) show that the CSNR should be known at both the transmitter and the receiver to compute the COVQ index transition probabilities. The CSNR can be estimated at the receiver and then fed back to the transmitter. As the feedback path may not always be available, it is of particular interest to consider the case where no information about the channel state is available to the transmitter. In [31], an FEAD COVQ is proposed which addresses this issue for a different setup involving hybrid digital–analog SISO transmission systems. In the following, we show how to design a FEAD COVQ for the soft-decision decoded STOB-coded channel. In addition to having multiple antennas, our FEAD COVQ differs from the one in [31] in that we have a digital channel (instead of a hybrid digital-analog channel). The block diagram of the FEAD COVQ looks the same as in Fig. 1. The key difference is that here the encoder partition matches a "design CSNR" and the receiver codebook is adapted to the actual CSNR. The encoder uses a design codebook $\{\boldsymbol{z}_i\}_{i=0}^{N_e-1}$ to find its partition via

$$\mathcal{S}_i^* = \left\{ \boldsymbol{x} : \sum_{j=0}^{N_e-1} P_{J|I}(j|i) d(\boldsymbol{x}, \boldsymbol{z}_j) \le \sum_{j=0}^{N_e-1} P_{J|I}(j|l) d(\boldsymbol{x}, \boldsymbol{z}_j), \forall l \right\}.$$

Because we assume complete lack of information at the transmitter, the encoder codebook $\{\boldsymbol{z}_i\}$ is designed assuming the channel is hard decoded (i.e., that $q = 1$).

The problem is then to adapt the decoder codebook according to the actual CSNR value, while the encoder codebook remains

fixed. Let us denote the encoder index by I and the decoder index by J. The average distortion per dimension is given by

$$\begin{aligned} D &= \frac{1}{k} E\left[\|\boldsymbol{x}-\boldsymbol{y}\|^2\right] \\ &= \frac{1}{k}\sum_{j=0}^{N_d-1} E\left[\|\boldsymbol{x}-\boldsymbol{y}\|^2 | J=j\right] P_J(j) \\ &= \frac{1}{k}\sum_{j=0}^{N_d-1} E\left[\|\boldsymbol{x}-\boldsymbol{y}_j\|^2 | J=j\right] P_J(j). \end{aligned} \tag{16}$$

The goal in quantizer design is to derive the $\boldsymbol{y}_j$, which minimize (16). This gives the optimal $\boldsymbol{y}_j$ in the MMSE sense as

$$\begin{aligned} \boldsymbol{y}_j^* &= E[\boldsymbol{x}|J=j] \\ &= \sum_{i=0}^{N_e-1} E[\boldsymbol{x}|I=i, J=j] P_{I|J}(i|j) \\ &= \sum_{i=0}^{N_e-1} \boldsymbol{m}_i P_{I|J}(i|j), \qquad j=0,\cdots,N_d-1 \end{aligned} \tag{17}$$

where $\boldsymbol{m}_i = \int_{\mathcal{S}_i} \boldsymbol{x} p(\boldsymbol{x}) d\boldsymbol{x}$ denotes the mean of the samples in $\mathcal{S}_i$ and (17) follows because $E[\boldsymbol{x}|I=i, J=j] = E[\boldsymbol{x}|I=i] = \boldsymbol{m}_i$. Note that the decoder can simply determine $P_{I|J}(i|j)$ as

$$P_{I|J}(i|j) = \frac{P_{J|I}(j|i) P_I(i)}{\sum_{k=0}^{N_e-1} P_{J|I}(j|k) P_I(k)}, \tag{18}$$

where $P_I(i) = P(I=i) = P(\boldsymbol{x} \in \mathcal{S}_i)$. Also note that $P_I(i)$ and $\boldsymbol{m}_i$ can be approximated in the training phase of the encoder as follows:

$$P_I(i) \approx \frac{|\mathcal{S}_i|}{M}, \quad \boldsymbol{m}_i \approx \frac{1}{|\mathcal{S}_i|} \sum_{l:\boldsymbol{x}_l \in \mathcal{S}_i} \boldsymbol{x}_l. \tag{19}$$

From (17), we observe that unlike COVQ, given the encoder means $\{\boldsymbol{m}_i\}_{i=0}^{N_e-1}$, the FEAD-COVQ does not require a training phase to determine $\{\boldsymbol{y}_j\}$. In other words, the decoder codebook is *computed* from the channel transition probabilities and the encoder means via (17) and (18).

C. Online FEAD Soft-Decision Demodulation COVQ

One of the assumptions made in Section II-A is that the receiver has perfect knowledge of the channel fading coefficients without error. In light of this assumption, we observe in (9) that the output of the linear combiner has an identical form to the output of an additive white Gaussian noise (AWGN) channel with the variance of the additive noise ν_t known at the receiver and given by (10). This motivates us to apply soft-decision demodulation directly on ρ_t using the step-size Δ derived in [22, sec. II] for AWGN channels. The channel transition probabilities, given the path gains matrix $\boldsymbol{H}$, are given by (11). We expect the online FEAD COVQ to outperform the FEAD COVQ because it exploits the knowledge of the channel state information in the COVQ decoding phase.

Note that the above derivation is valid if the channel fading coefficients remain constant during the transmission time of an index.

D. Asymptotic Optimality of the Proposed Decoding Method

It was shown in Section II-C-1) that using the "adder" structure for the combiner does not cause suboptimality (i.e., an optimal soft decoder for COVQ over STOB coded channels could also use the same structure). As explained in the two previous sections, the linear combiner output ρ_t is quantized via a uniform quantizer with step size Δ. The two sources of suboptimality in our decoding method are therefore the uniform scalar quantizer itself and the way its step size is chosen. For finite q, Δ is not chosen in the optimal way, which would be minimizing the overall distortion, because the relationship between Δ and the overall distortion seems to be complicated. The suboptimal nature of the uniform quantizer becomes negligible as q grows. In fact, assuming that q grows without bound and the quantizer resolution becomes finer (i.e., $\lim_{q\to\infty} \Delta(q) = 0$), one can show that the proposed system converges to the optimal soft decoding (as opposed to soft-decision demodulation) COVQ as follows.

For simplicity, we assume that $kr = \tau$ (a similar argument holds for $kr \neq \tau$). Let $l \in \{0, 1, \ldots, 2^{qkr} - 1\}$ be an index input to the proposed COVQ decoder. Recall that l is formed by the concatenation of $m_1, m_2, \ldots, m_{kr}$ integers (each of which is represented by q bits) that result from the soft-decision demodulation of $\rho_1, \rho_2, \ldots, \rho_{kr}$. Given a source with a smooth pdf $f_{\boldsymbol{X}}(\cdot)$, the l^{th} centroid of the proposed COVQ with index transition probabilities $P(l|i)$ is given by (20)–(22), shown at the bottom of the next page, which is the optimal decoder in (6) and where we have made the simplifying assumption that in the limit as $q \to \infty$, the joint pdf $f(\rho_1, \rho_2, \ldots, \rho_{kr}|i)$ is constant in the kr-dimensional cube of size Δ^{kr} and we have used (8) to re-write (21) as (22).[4] Using the above approach and (18), one can verify that the FEAD COVQ is also optimal when $q \to \infty$.

IV. Numerical Results and Discussion

A. Implementation Issues

We consider the transmission of zero-mean unit-variance i.i.d. Gaussian and Gauss–Markov sources over MIMO channels. 500 000 training vectors and 850 000 test vectors are employed. Each test is performed five times, and the average signal-to-distortion ratio (SDR) in decibels is reported. MIMO systems with K transmit and L receive antennas are referred to as $(K-L)$ systems. Alamouti's code [3] is used for the dual transmit-antenna systems. The real (rate 1) code of [32, eq. (4)] is employed for the quad-transmit system to maximize throughput and because our constellation is real.

Several training strategies were examined, and the best one in terms of having consistent results and high training SDR was used as follows. For any given COVQ rate r, dimension k, and

[4]The argument (20)–(22) can be made a rigorous proof under some regularity conditions on the pdf $f_{\boldsymbol{X}}(\cdot)$. For example, it suffices to assume that $f_{\boldsymbol{X}}(\cdot)$ is continuous, differentiable, and has a light tail.

number of soft-decision demodulation bits q, we first train a k-dimensional rate-qr VQ with the split algorithm [11]. We next use the simulated annealing algorithm [7], which aims to minimize the average end-to-end distortion for a given VQ codebook through optimizing the assignment of indexes of the VQ codevectors. It can be shown that the cost function to be minimized equals

$$\sum_{i=0}^{N_e-1} P_I(i) \sum_{j=0}^{N_d-1} P_{J|I}(j|i) \left\langle \boldsymbol{y}_{b(j)}, \left(\boldsymbol{y}_{b(j)} - 2\boldsymbol{m}_i \right) \right\rangle.$$

where $\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \sum_i f_i g_i$ is the standard inner product and $b : \{0, \cdots, N_d - 1\} \to \{0, \cdots, N_d - 1\}$ is the one-to-one mapping function to be optimized. Simulated annealing is used only at the highest CSNR. We then use an approach similar to the one in [9]; namely, we use the modified generalized Lloyd algorithm to derive the COVQ codebooks starting from the highest CSNR to the lowest, and vice versa. This method is referred to as the "decrease–increase" (DI) method. Another way to obtain the codebooks could be starting from the highest CSNR down to the lowest, which we refer to as the "decreasing" method.

Table II compares the results of the DI and descending methods for various STOB coded systems. It is observed that the DI method is mostly beneficial at low CSNR values. This is because at low CSNR some encoder cells are empty. Empirical results show that these cells are optimized more efficiently through the CSNR-increasing loop.

TABLE II
COMPARISON BETWEEN THE TRAINING SDR (IN DECIBELS) OF TWO COVQ TRAINING METHODS FOR A UNIT-VARIANCE GAUSS–MARKOV SOURCE ($\rho = 0.9$) CHANNEL-OPTIMIZED VECTOR QUANTIZED AT RATE 1.0 bps. QUANTIZATION DIMENSION IS 2. THREE MIMO SYSTEMS ARE CONSIDERED WITH $(K - L) = (2 - 1)$, (4-1), AND (2-2)

CSNR	(2-1)		(4-1)		(2-2)	
in dB	DI	decreasing	DI	decreasing	DI	decreasing
10	7.385	7.368	7.809	7.795	7.899	7.876
8	6.886	6.845	7.551	7.536	7.867	7.839
6	6.119	6.026	6.949	6.908	7.712	7.691
4	5.167	4.968	5.948	5.832	7.288	7.272
2	4.178	3.830	4.781	4.508	6.477	6.416
0	3.315	2.783	3.714	3.245	5.353	5.179
-1	2.916	2.329	3.266	2.700	4.772	4.508
-2	2.546	1.930	2.830	2.224	4.217	3.853
-3	2.202	1.578	2.436	1.818	3.713	3.245

B. COVQ for Various MIMO Channels

Fig. 3 plots the SDR curves of various COVQ-based space–time coded systems as a function of the CSNR. Even at the low COVQ dimension and rate considered here, the gain of using MIMO channels over the SISO channel is obvious. For example, at SDR $=$ 5 dB, the (2-1) system outperforms the SISO system by 6 dB (for hard decoding) and is outperformed by the (2-2) system by 4.3 dB. This figure also demonstrates the effectiveness of our linear combiner. Note that as the signal power collected by the (2-2) system is twice that of the (4-1) system, the former has a better performance, although the diversity gain of both of the systems is the same and equals $KL = 4$.

We observe in Table I that soft-decision demodulation becomes less beneficial as the CSNR grows. Increasing the number of transmit or receive antennas results in a CSNR gain due to space diversity. Therefore, we expect that increasing the number of transmit or receive antennas would leave little room for further enhancement through soft-decision demodulation. It follows from (10) that between two systems with the same diversity gain, coding gain g, and CSNR γ_s, the one with fewer transmit antennas has a higher SNR at its linear combiner

$$\lim_{q\to\infty} \boldsymbol{y}_l = \lim_{q\to\infty} \frac{\sum_{i=0}^{N_e-1} P(l|i) \int_{\mathcal{S}_i} \boldsymbol{x} f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}}{\sum_{i=0}^{N_e-1} P(l|i) \int_{\mathcal{S}_i} f_{\boldsymbol{X}}(\boldsymbol{x})} \tag{20}$$

$$= \lim_{q\to\infty} \frac{\sum_i P\left(\rho_1 \in [u_{m_1-1}, u_{m_1}), \ldots, \rho_{kr} \in [u_{m_{kr}-1}, u_{m_{kr}}) | i\right) E(\boldsymbol{x}|i)}{\sum_i P\left(\rho_1 \in [u_{m_1-1}, u_{m_1}), \ldots, \rho_{kr} \in [u_{m_{kr}-1}, u_{m_{kr}}) | i\right) p(i)}$$

$$= \lim_{q\to\infty} \frac{\sum_i f(\rho_1, \rho_2, \ldots, \rho_{kr}|i) \left(\Delta(q)\right)^{kr} E(\boldsymbol{x}|i)}{\sum_i f(\rho_1, \rho_2, \ldots, \rho_{kr}|i) \left(\Delta(q)\right)^{kr} p(i)}$$

$$= \frac{\sum_i f\left(\sum_j \tilde{\boldsymbol{r}}_j | i\right) E(\boldsymbol{x}|i)}{\sum_i f\left(\sum_j \tilde{\boldsymbol{r}}_j | i\right) p(i)} \tag{21}$$

$$= \frac{\sum_i f(\boldsymbol{R}|i) E(\boldsymbol{x}|i)}{\sum_i f(\boldsymbol{R}|i) p(i)}$$

$$= E[\boldsymbol{x}|\boldsymbol{R}]. \tag{22}$$

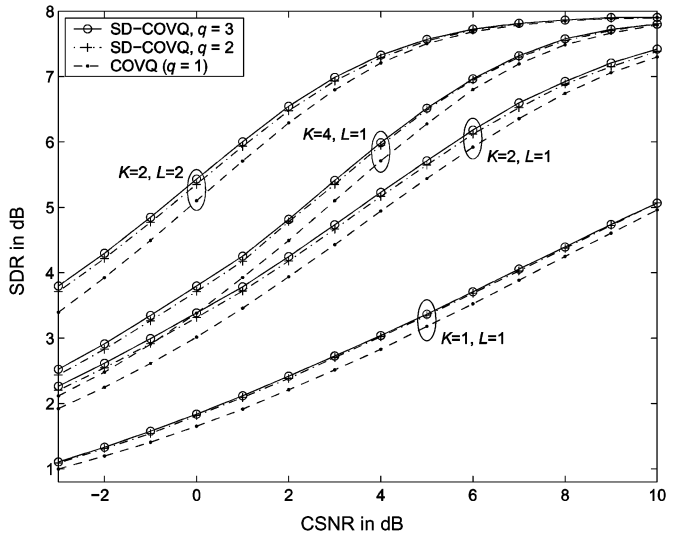

Fig. 3. Simulated SDR in decibels for a zero-mean unit-variance Gauss–Markov source ($\rho = 0.9$) vector quantized at rate 1 bps and soft-decision decoded with q bits. Quantization dimension is 2.

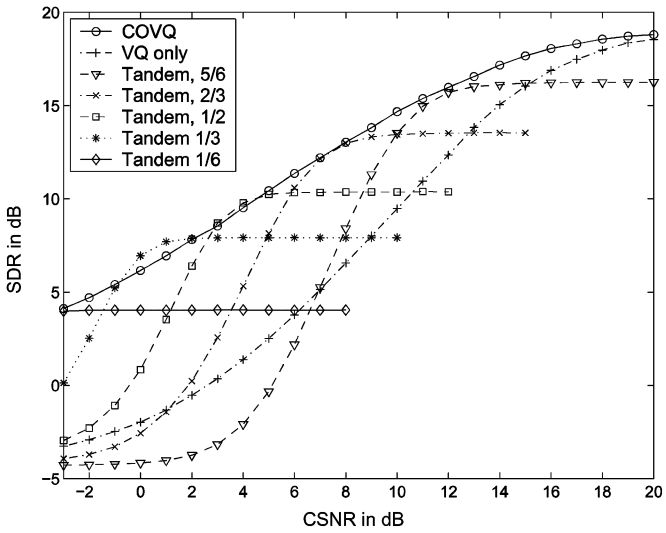

Fig. 4. Jointly designed versus tandem coding schemes for a zero-mean unit-variance Gauss–Markov source ($\rho = 0.9$). Quantization dimension is 2, and the overall rate is 3.0 bps. $K = 2$ and $L = 1$.

output. In other words, given two systems with the same diversity gain, the one with more transmit antennas obtains a larger soft-decision demodulation gain. This result can be stated more intuitively: systems with more receive branches collect more signal power. Hence, the SNR at their linear combiner would be higher, making soft-decision demodulation less effective. This observation is supported by the simulations of Fig. 3: at CSNR $= 4$ dB, the soft-decision demodulation gain in SDR is 0.29 dB for both of the (2-1) and (4-1) systems; this gain reduces to 0.12 dB for the (2-2) system.

C. COVQ Versus Tandem (Separate) Coding

We next compare our COVQ-based system with traditional tandem coding schemes which use separate source coding and channel coding blocks with VQ and convolutional coding (CC), respectively. We consider, in Fig. 4, a (2-1) system using Alamouti's code and quantization with dimension $k = 2$. The overall rate is 3.0 bps, and hence there are six choices for the (VQ, CC) code rates, namely, (0.5, 1/6), (1.0, 1/3), (1.5, 1/2), (2.0, 2/3), (2.5, 5/6), and (3.0, 0). The first four convolutional codes have 64 states and are nonsystematic with free distances of 27 [18], 14, 10, and 5 [17]. For rate 5/6, we use a rate-compatible punctured convolutional (RCPC) code with a rate-1/2 mother code.[5] Fig. 4 shows a typical behavior: the jointly designed COVQ outperforms the substantially more complex tandem systems almost everywhere. Further tests with i.i.d. sources yield even more supportive results towards COVQ.

Note that if one aims to design an unequal error protection (UEP) joint source-channel coder with the above separate coders (i.e., select the best tandem coder at each CSNR), one needs to design an algorithm to allocate the source and channel code rates for each given CSNR, thus increasing the complexity of the UEP system. COVQ does not have this problem since error protection in COVQ is built-in.

[5]The generator polynomials of the rate-1/6, 1/3, 1/2, and 2/3 convolutional codes are given by (754, 644, 564, 564, 714, 574), (574, 664, 744), (634, 564), and (3, 4, 5; 4 3 7). They are the strongest codes given in [17], [24] for the given number of states. The generator polynomials are defined as in [17]. The generator polynomials of the mother code for the rate-5/6 code are (554, 744) and its puncturing matrix is given by $\binom{11010}{10101}$ [24].

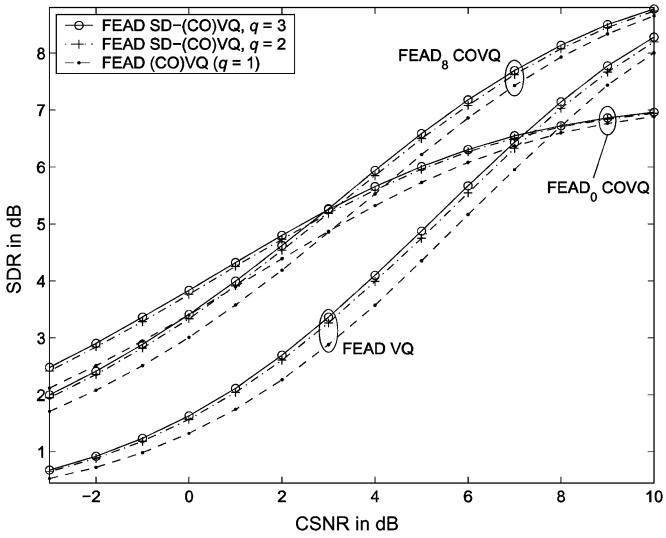

Fig. 5. Simulated SDR in dB for an i.i.d. $\mathcal{N}(0,1)$ source vector quantized at rate 2.0 bps and soft-decision demodulated with q bits. Quantization dimension is 2. $K = 2$ and $L = 1$.

D. COVQ, FEAD COVQ, and Online FEAD COVQ

Fig. 5 demonstrates the performance of a (2-1) system quantizing an i.i.d. $\mathcal{N}(0,1)$ source with various rate-2 bps FEAD COVQs with dimension $k = 2$. A FEAD$_\gamma$ COVQ is one whose design CSNR is γ dB. The FEAD VQ assumes that the channel is noiseless; hence it has a lower computational complexity at the encoder. The figure shows that such an assumption will lead to a significant SDR loss at low to medium CSNRs. FEAD$_0$ COVQ also suffers from a considerable performance degradation at high CSNRs. It seems that assuming a midrange CSNR

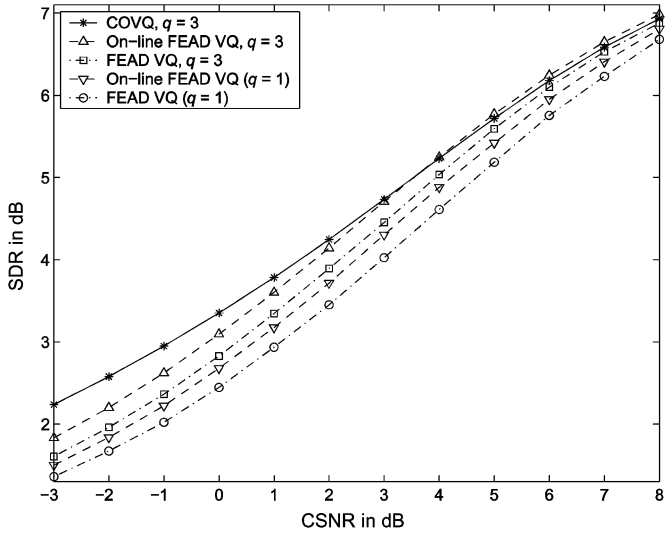

Fig. 6. Comparison among the COVQ, FEAD VQ, and online FEAD VQ for a $\mathcal{N}(0,1)$ Gauss–Markov source vector quantized at rate 1.0 bps and hard decoded. Quantization dimension is 2, $K = 2$, and $L = 1$.

of 8 dB (for the given MIMO system) will lead to reasonable performance everywhere.

The three quantizers presented in this paper are compared in Fig. 6, where a unit-variance Gauss–Markov source is quantized with dimension 2 and rate 1 bps and sent over a (2-1) system. The performance of the online and FEAD VQs (which are designed for a noiseless channel) become closer as the CSNR grows because the channel mismatch of the VQs decreases. The online FEAD VQ maintains its gain over the FEAD VQ when soft-decision demodulation is employed. The online FEAD VQ encoder assumes the channel is noiseless (CSNR $\rightarrow \infty$). Nevertheless, at high CSNR, it slightly outperforms the COVQ, which is designed for the exact CSNR. This is due to the exploitation of the knowledge of the channel fading coefficients in the online VQ decoding phase.

V. Conclusion

We presented three soft-decision demodulation COVQ-based systems for communicating analog sources over STOB coded multi-antenna channels. The proposed systems depend on whether the actual CSNR is available to the transmitter and whether the COVQ decoder is aware of the fading coefficients. The soft information of the channel is utilized through space–time soft decoding, linear combining, and scalar uniform quantization. Simple design methods were proposed for the linear combiner and the uniform quantizer. It was shown that using three soft-decision demodulation bits can achieve almost all of the gain available through soft-decision demodulation. This gain is very significant, specially when transmit diversity is employed and/or when the source is correlated. For a dual transmit antenna system and at SDR $= 5$ dB, using a second receive antenna results in 4.3-dB CSNR gain over a single-receive antenna system for a unit-variance Gauss–Markov source. For the COVQ dimension and rates considered here, the use of only two soft-decision demodulation bits results in typically 0.9-dB gain in CSNR over hard decoding. The COVQ-based system was shown to outperform tandem systems that use separate source and channel coding blocks.

Appendix
Derivation of (7) in General

For simplicity, it was assumed in (7) that each block of kr bits, which forms a COVQ encoder index, is mapped to one space–time codeword, i.e., $kr = \tau$. Here, we consider the general case. We first give the derivation for $\tau > kr$. Let i be a kr-bit COVQ encoder index. We have (23) and (24), shown at the bottom of the page, where $\mathbf{c}(i,i')$ is the vector of BPSK symbols that correspond to the concatenation of the binary forms of i and i' and we have assumed in (23) that the COVQ indexes

$$
\begin{aligned}
P(i|\boldsymbol{R}) &= \sum_{i'=0}^{2^{\tau-kr}-1} P(i,i'|\boldsymbol{R}) \\
&= \sum_{i'} \frac{p(i,i')\exp\left(-\sqrt{\gamma_s/K}\left(\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j\right)^T \mathbf{c}(i,i')\right)}{\sum_{l=0}^{2^{kr+i'}-1} p(l,i')\exp\left(-\sqrt{\gamma_s/K}\left(\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j\right)^T \mathbf{c}(l,i')\right)} \\
&= \sum_{i'} \frac{p(i)\exp\left(-\sqrt{\gamma_s/K}\left(\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j\right)^T \mathbf{c}(i,i')\right)}{2^{i'}\sum_{l=0}^{N_e-1} p(l)\exp\left(-\sqrt{\gamma_s/K}\left(\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j\right)^T \mathbf{c}(l,i')\right)} \qquad (23) \\
&= \frac{p(i)\exp\left(-\sqrt{\gamma_s/K}\sum_{t=1}^{kr} c_t(i)\sum_{j=1}^{L}\tilde{R}_{j,t}\right)}{\sum_{l=0}^{N_e-1} p(l)\exp\left(-\sqrt{\gamma_s/K}\sum_{t=1}^{kr} c_t(l)\sum_{j=1}^{L}\tilde{R}_{j,t}\right)}, \qquad (24)
\end{aligned}
$$

are independent (the memoryless assumption) and (24) follows from (23) because, from (8)

$$\left(\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j\right)^T \mathbf{c}(i,i') = \sum_{t=1}^{\tau} g'\bar{H}\rho_t c_t(i,i') = g'\bar{H}\sum_{t=1}^{kr}\rho_t c_t(i) + g'\bar{H}\sum_{t=kr+1}^{\tau}\rho_t c_t(i')$$

hence the terms pertaining to i' in $\exp(-\sqrt{\gamma_s/K}\,(\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j)^T\mathbf{c}(l,i'))$ form a multiplicative factor and cancel out in (23). Notice that (24) also depends only on the *sum* of the $\tilde{R}_{j,t}$ over all j.

Furthermore, for the case $\tau \le kr$, we can write $kr = n\tau + m$, where $n \ge 1$ and $0 \le m < \tau$ are integers. It is straightforward to verify that in this case $P(i|\boldsymbol{R})$ can be written in terms of the product of n terms as in (7) (one term per space–time codeword) and one term as in (24).

ACKNOWLEDGMENT

The authors gratefully acknowledge the useful comments of the reviewers. Special thanks are due to a reviewer who pointed out that in Section II-C-1, $\sum_{j=1}^{L}\tilde{\boldsymbol{r}}_j$ provides sufficient statistics to compute $P(I=i|\boldsymbol{R})$ and who guided us to study the asymptotic relationship between our suboptimal decoder and the optimal decoder of (6).

REFERENCES

[1] F. Alajaji and N. Phamdo, "Soft-decision COVQ for Rayleigh-fading channels," *IEEE Commun. Let.*, vol. 2, pp. 162–164, Jun. 1998.

[2] F. Alajaji, N. Phamdo, and T. Fuja, "Channel codes that exploit the residual redundancy in CELP-encoded speech," *IEEE Trans. Speech Audio Process.*, vol. 4, pp. 325–336, Sep. 1996.

[3] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[4] G. Bauch, J. Hagenauer, and N. Seshadri, "Turbo processing in transmit antenna diversity systems," *Ann. Telecommun.*, vol. 56, pp. 455–471, Aug. 2001.

[5] F. Behnamfar, F. Alajaji, and T. Linder, "Error analysis of space–time codes for slow Rayleigh fading channels," in *Proc. IEEE Int. Symp. Information Theory*, Yokohama, Japan, Jun.–Jul. 2003, p. 12.

[6] F. Behnamfar, F. Alajaji, and T. Linder, "Performance analysis of MAP decoded space–time orthogonal block codes for non-uniform sources," in *Proc. IEEE Information Theory Workshop*, Paris, France, Mar.–Apr. 2003, pp. 46–49.

[7] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. Inf. Theory*, vol. 36, pp. 799–809, Jul. 1990.

[8] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel optimized vector quantizers," *IEEE Trans. Inf. Theory.*, vol. 37, pp. 155–160, Jan. 1991.

[9] ——, "Optimal quantizer design for noisy channels: an approach to combined source-channel coding," *IEEE Trans. Inf. Theory.*, vol. 33, pp. 827–838, Nov. 1987.

[10] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[11] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.

[12] D. Guo, S. Shamai (Shitz), and S. Verdu, "Mutual information and MMSE in Gaussian channels," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun.–Jul. 2004, p. 349.

[13] P. Knagenhjelm and E. Agrell, "The Hadamard transform—A tool for index assignment," *IEEE Trans. Inf. Theory*, vol. 42, pp. 1139–1151, Jul. 1996.

[14] H. Kumazawa, M. Kasahara, and T. Namekawa, "A construction of vector quantizers for noisy channels," *Electron. Eng. Japan*, vol. 67-B, pp. 39–47, Jan. 1984.

[15] F. H. Liu, P. Ho, and M. Cuperman, "Sequential reconstruction of vector quantized signals transmitted over Rayleigh fading channels," in *Proc. IEEE Int. Conf. Communications*, New Orleans, LA, 1994, pp. 23–27.

[16] A. Naguib, N. Seshadri, and A. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Process. Mag.*, vol. 46, pp. 76–92, May 2000.

[17] R. Johannesson and K. Zigangirov, *Fundamentals of Convolutional Coding*. New York: IEEE Press, 1999.

[18] L. H. C. Lee, *Convolutional Coding: Fundamentals and Applications*. Norwood, MA: Artech House, 1997.

[19] J. Lim and D. L. Neuhoff, "Joint and tandem source-channel coding with complexity and delay constraints," *IEEE Trans. Commun.*, vol. 51, pp. 757–766, May 2003.

[20] H. Lu, Y. Wang, P. V. Kumar, and K. Chugg, "On the performance of space–time codes," in *Proc. IEEE Information Theory Workshop*, Bangalore, India, Oct. 2002, pp. 49–52.

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

[22] N. Phamdo and F. Alajaji, "Soft-decision demodulation design for COVQ over white, colored, and ISI Gaussian channels," *IEEE Trans. Commun.*, vol. 48, pp. 1499–1506, Sep. 2000.

[23] N. Phamdo and N. Farvardin, "Optimal detection of discrete Markov sources over discrete memoryless channels-applications to combined source-channel coding," *IEEE Trans. Inf. Theory.*, vol. 40, pp. 186–193, Jan. 1994.

[24] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[25] 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.

[26] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Oct. 1948, 623-656.

[27] H. Shin and J. H. Lee, "Exact symbol error probability of orthogonal space–time block codes," in *Proc. IEEE Global Communications Conf. (IEEE GLOBECOM)*, Nov. 2002, vol. 2, pp. 1197–1201.

[28] M. Skoglund and P. Hedelin, "Hadamard-based soft-decoding for vector quantization over noisy channels," *IEEE Trans. Inf. Theory*, vol. 45, pp. 515–532, Mar. 1999.

[29] M. Skoglund, "Soft decoding for vector quantization over noisy channels with memory," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1293–1307, May 1999.

[30] ——, "Bit-estimate based decoding for vector quantization over noisy channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 48, pp. 1239–1317, Aug. 2000.

[31] M. Skoglund, N. Phamdo, and F. Alajaji, "Design and performance of VQ-based hybrid digital-analog joint source-channel codes," *IEEE Trans. Inf. Theory*, vol. 48, pp. 708–720, Mar. 2002.

[32] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1456–1467, Jul. 1999.

[33] V. Vaishampayan and N. Farvardin, "Joint design of block source codes and modulation signal sets," *IEEE Trans. Inf. Theory*, vol. 36, pp. 1230–1248, Jul. 1992.

[34] K. Zeger and A. Gersho, "Pseudo-Gray coding," *IEEE Trans. Commun.*, vol. 38, pp. 2147–2158, Dec. 1990.

[35] Y. Zhong, F. Alajaji, and L. Campbell, "When is joint source-channel coding worthwhile: an information theoretic perspective," in *Proc. 22nd Biennial Symp. Communications*, Kingston, ON, Canada, Jun. 2004, pp. 121–123.

Firouz Behnamfar (S'01–M'04) received the B.Sc. and M.Sc. degrees from Isfahan University of Technology, Isfahan, Iran, and the Ph.D. degree from Queen's University at Kingston, ON, Canada, all in electrical engineering, in 1994, 1997, and 2004, respectively.

From September 2004 to August 2006, he was a Postdoctoral Fellow and an Adjunct Assistant Professor in the Department of Mathematics and Statistics, Queen's University. He is now with Nortel Networks, Ottawa, Canada. His research interests include communication systems, wireless communications, and information theory.

Dr. Behnamfar has won a number of awards for his doctoral work, including the Outstanding Thesis in Applied Science Award from Queen's University in 2004, a postdoctoral fellowship (PDF) award, and an industrial research and development fellowship (IRDF) award, both by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Fady Alajaji (S'90–M'94–SM'00) was born in Beirut, Lebanon, on May 1, 1966. He received the B.E. degree (with distinction) from the American University of Beirut, Lebanon, and the M.Sc. and Ph.D. degrees from the University of Maryland, College Park, all in electrical engineering, in 1988, 1990 and 1994, respectively.

In 1994, he held a postdoctoral appointment at the Institute for Systems Research, University of Maryland. In 1995, he joined the Department of Mathematics and Statistics at Queen's University, Kingston, ON, Canada, where he is currently a Professor of Mathematics and Engineering. Since 1997, he has also been cross-appointed in the Department of Electrical and Computer Engineering at the same university. He held research visits to the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, in fall 2001, and to the Department of Mobile Communications, Technical University of Berlin, Germany, in winter 2004. His research interests include information theory, joint source-channel coding, error control coding, data compression, and digital communications.

Dr. Alajaji currently serves as Editor for Source and Source-Channel Coding for the IEEE TRANSACTIONS ON COMMUNICATIONS. He served as Co-Chair of the Technical Program Committee of the 2004 Biennial Symposium on Communications. In 1999, he Co-Chaired and organized the Sixth Canadian Workshop on Information Theory in Kingston, ON, Canada. He also served as Chair of the Selection Committee for the 2000 Canadian Award in Telecommunications. In 2001, he received the Premier's Research Excellence Award (PREA) from the Province of Ontario, Canada.

Tamás Linder (S'92–M'93–SM'00) was born in Budapest, Hungary, in 1964. He received the M.S. degree in electrical engineering from the Technical University of Budapest in 1988 and the Ph.D. degree in electrical engineering from the Hungarian Academy of Sciences in 1992.

He was a Postdoctoral Researcher at the University of Hawaii in 1992 and a Visiting Fulbright Scholar at the Coordinated Science Laboratory, University of Illinois, at Urbana-Champaign, from 1993 to 1994. From 1994 to 1998, he was a faculty member in the Department of Computer Science and Information Theory at the Technical University of Budapest. From 1996 to 1998, he was also a visiting research scholar in the Department of Electrical and Computer Engineering, University of California, San Diego. In 1998, he joined Queen's University, Kingston, ON, Canada, where he is currently an Associate Professor of Mathematics and Engineering in the Department of Mathematics and Statistics. His research interests include communications and information theory, source coding and vector quantization, machine learning, and statistical pattern recognition.

Dr. Linder received the Premier's Research Excellence Award of the Province of Ontario in 2002 and the Chancellor's Research Award of Queen's University in 2003. He was an Associate Editor for Source Coding of the IEEE TRANSACTIONS ON INFORMATION THEORY in 2003–2004.