Orthogonal Design of Unitary Constellations for Uncoded and Trellis-Coded Noncoherent Space–Time Systems

Wanlun Zhao, *Student Member*, Geert Leus, *Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—We construct unitary noncoherent space–time constellations, which can be considered as a concatenation of a training block with an orthogonal design. With a simple construction, our constellations are easy to design, enjoy full antenna diversity, allow for a simplified maximum-likelihood (ML) detector, and achieve error performance comparable to existing designs that rely on computer search. To exploit the constellation structures and improve coding gains, we further pursue a trellis-coded modulation (TCM) approach. Based on the sequence pairwise error analysis, we identify two simple parameters to quantify the asymptotic error performance, which enables us to compare among different TCM schemes or uncoded alternatives.**

***Index Terms*—Fading channel, noncoherent, performance analysis, space–time (ST), trellis-coded modulation (TCM).**

I. Introduction

Recent information-theoretic results [8], [18], [13], [20] on multiantenna communications have awakened great interest in coherent, differential, and noncoherent space–time (ST) system designs. In this work, we focus on noncoherent ST system designs.

When the channel is unknown at both the transmitter and the receiver, capacity analysis suggests a scaled unitary matrix ST signaling structure [13]. Motivated by this result, unitary ST modulation was introduced in [10]. Exact and Chernoff bound expressions for the pairwise error probability (PEP) were given, and an initial unitary ST constellation design method was described. More advanced constellation design methods were reported in [11], [1] and [14], relying on different distance metrics: [11], [1] used the chordal distance, which does not guarantee full diversity; whereas [14] adopted a distance metric derived from the asymptotic union bound (AUB), which guarantees full diversity. For large constellations, however, the AUB is generally a loose bound for the SEP. Constellations achieving the best AUB do not necessarily achieve the best performance. In [1] and [14], no extra structure was imposed on the unitary ST constellations, and the proposed design methods were based on a cumbersome computer search. In [11], on the other hand, a block circulant correlation structure was imposed on the unitary constellations. This rendered the computer search and optimization more tractable, but still rather inefficient when the constellation size is large. Pilot-based noncoherent designs became available recently [5], [6], [17]. In these approaches, channel coefficients are first estimated based on the pilot symbols. Then, coherent detection is performed using the obtained channel estimates. However, this detection approach is inherently suboptimal.

Manuscript received March 18, 2003; revised January 9, 2004. This work was supported by the ARO/CTA under Grant DAAD19-01-2-0011. The material in this correspondence was presented in part at the International Conference on Communications (ICC), Anchorage, AK, May 2003.

W. Zhao and G. B. Giannakis are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA (e-mail: wlzhao@ece.umn.edu; georgios@ece.umn.edu).

G. Leus was with the Department of Electrical Engineering, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium. He is now with the Department of Electrical Engineering, Delft University of Technology, 2628CD Delft, The Netherlands (e-mail: leus@cas.et.tudelft.nl).

Communicated by G. Caire, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2004.828154

In the first part of this work, we introduce novel unitary constellations based on the orthogonal design. They can be viewed as a special pilot-based design that enables low-complexity noncoherent maximum-likelihood (ML) detection. In Section II, we describe our channel and system models. In Section III, we present our unitary constellation in the framework of pilot-based designs, identify several useful properties that characterize the error performance and structure of this constellation, and most importantly, develop a low-complexity detector that enjoys ML optimality. We also compare the error performance of our constellations with those in [11].

It is widely accepted that ST signals can serve as their own channel codes [15], [16]. Nonetheless, due to the inherent orthogonal structure, our constellations may exhibit limited coding gains. To exploit the rich diversity and boost coding gains, we pursue a trellis-coded modulation (TCM) approach for the noncoherent ST system in the second part of our work. In Section IV, we describe the block diagram of the noncoherent ST TCM scheme, perform set partitioning for the unitary constellations, and identify two simple parameters to quantify asymptotic error performance based on the sequence PEP. Of these two parameters, the primary one is the diversity order; whereas the secondary one is the asymptotic coding gain. We also provide TCM design examples and simulations to verify our analysis. Finally, we conclude our work in Section V.

Notation: Upper (lower) bold face letters denote matrices (column vectors); $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote conjugate, transpose, and Hermitian, respectively; $\|\cdot\|$ represents the Frobenius norm; $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix; and $\boldsymbol{0}_{M\times N}$ denotes the $M \times N$ all-zero matrix. Finally, $\mathbb{N}$ and $\mathbb{C}$ stand for the natural and complex number field, respectively.

II. Channel and System Modeling

We consider a noncoherent ST communication system with M transmit and N receive antennas. The channels between the different transmit and receive antennas are assumed to be mutually independent and block Rayleigh faded. Let $\boldsymbol{H}$ be the channel coefficient matrix with (i,j)th entry, $h_{i,j}$, denoting the channel coefficient from the ith transmit to the jth receive antenna. Then, $h_{i,j} \sim \mathcal{CN}(0,1)$, $i \in \{1,\ldots,M\}$, $j \in \{1,\ldots,N\}$. The matrix $\boldsymbol{H}$ remains constant during the channel's coherence interval of length T, and varies independently from block to block. During each block interval of length T, a matrix $\boldsymbol{X} \in \mathcal{C}$ is transmitted, where $\mathcal{C}$ is a noncoherent ST constellation and comprises L scaled unitary matrices with dimension $T \times M$. Mathematically, $\boldsymbol{X}^H\boldsymbol{X} = T\,\boldsymbol{I}_M, \forall\ \boldsymbol{X} \in \mathcal{C}$. Transmitting the M columns of $\boldsymbol{X}$ simultaneously via the M transmit antennas, we obtain the received matrix $\boldsymbol{Y}$ in additive white Gaussian noise (AWGN) as

$$\boldsymbol{Y} = \sqrt{\frac{\rho}{M}}\boldsymbol{X}\boldsymbol{H} + \boldsymbol{N} \tag{1}$$

where ρ is the signal-to-noise ratio (SNR) per receive antenna and $\boldsymbol{N}$ is the AWGN matrix with independent and identically distributed (i.i.d.) entries $n_{i,j} \sim \mathcal{CN}(0,1)$, $i \in \{1,\ldots,T\}$, and $j \in \{1,\ldots,N\}$.

III. Constellation Design

In this section, we will present a simple constructive design of unitary ST constellations. After reviewing some preliminaries, the proposed design for two transmit antennas, $M = 2$, will be studied in detail. Some useful properties of the obtained constellation indicating its error performance and structure will be identified. Furthermore, a

simple ML detector will be derived, and comparisons with the systematic design of [11] will be provided. Finally, we will give an example guiding the generalization of this design to more than two transmit antennas.

A. Preliminaries

To put our work in context, we will first go over some existing results on unitary ST modulations. The following results are available in [10] and [11]. Given the data model (1) and the constellation $\mathcal{C}$, the ML estimate for $\boldsymbol{X}$ is

$$\hat{\boldsymbol{X}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{X}\in\mathcal{C}} p(\boldsymbol{Y}|\boldsymbol{X}) = \arg\max_{\boldsymbol{X}\in\mathcal{C}} \|\boldsymbol{X}^H\boldsymbol{Y}\|^2 \tag{2}$$

where $\hat{\boldsymbol{X}}_{\mathrm{ML}}$ is found by maximizing the energy of $\boldsymbol{X}^H\boldsymbol{Y}$. We can derive a union-type upper bound for the symbo error probability (SEP) as

$$P_s \le \frac{1}{L}\sum_{\boldsymbol{X},\boldsymbol{X}'\in\mathcal{C},\,\boldsymbol{X}\neq\boldsymbol{X}'} P_e(\boldsymbol{X}\to\boldsymbol{X}') \tag{3}$$

where L is the cardinality of $\mathcal{C}$ and $P_e(\boldsymbol{X}\to\boldsymbol{X}')$ is the PEP defined as

$$P_e(\boldsymbol{X}\to\boldsymbol{X}') := P(\|\boldsymbol{X}'^H\boldsymbol{Y}\|^2 > \|\boldsymbol{X}^H\boldsymbol{Y}\|^2 \mid \boldsymbol{X}).$$

The closed-form PEP expression is known when the M singular values of $\boldsymbol{X}^H\boldsymbol{X}'/T$ are equal. The Chernoff upper bound for the PEP can be derived as follows:

$$P_e(\boldsymbol{X}\to\boldsymbol{X}') \le \frac{1}{2}\prod_{m=1}^{M}\left[\frac{1}{1+\frac{(\rho T/M)^2\left(1-d_m^2(\boldsymbol{X},\boldsymbol{X}')\right)}{4(1+\rho T/M)}}\right]^N$$

where $1 \ge d_1(\boldsymbol{X},\boldsymbol{X}') \ge \cdots \ge d_M(\boldsymbol{X},\boldsymbol{X}') \ge 0$ are the M singular values of $\boldsymbol{X}^H\boldsymbol{X}'/T$. For high SNR, a distance metric, the diversity product distance, can be extracted from this bound to guide constellation designs

$$\begin{aligned}\delta_{DP}(\boldsymbol{X},\boldsymbol{X}') &= \left[\prod_{m=1}^{M}(1-d_m^2(\boldsymbol{X},\boldsymbol{X}'))\right]^{1/2M}\\ &= \left[\prod_{m=1}^{M}\sin(\theta_m(\boldsymbol{X},\boldsymbol{X}'))\right]^{1/M}\end{aligned} \tag{4}$$

where $\theta_m(\boldsymbol{X},\boldsymbol{X}')$ is the mth principal angle between the two subspaces spanned by the columns of $\boldsymbol{X}$ and $\boldsymbol{X}'$. A related distance metric is the subspace distance defined in [9, p. 603]

$$\begin{aligned}\delta_S(\boldsymbol{X},\boldsymbol{X}') &= \min_{m=1,\ldots,M}(1-d_m^2(\boldsymbol{X},\boldsymbol{X}'))^{1/2}\\ &= \min_{m=1,\ldots,M}\sin(\theta_m(\boldsymbol{X},\boldsymbol{X}')).\end{aligned} \tag{5}$$

Finally, the diversity product ζ of a constellation $\mathcal{C}$ is defined as

$$\zeta = \min_{\boldsymbol{X},\boldsymbol{X}'\in\mathcal{C},\,\boldsymbol{X}\neq\boldsymbol{X}'}\delta_{DP}(\boldsymbol{X},\boldsymbol{X}').$$

When $\zeta > 0$, $\mathcal{C}$ enjoys full antenna diversity MN.

B. Constellation Design for $M = 2$

We now introduce a novel class of unitary ST constellations that can be designed in an algebraic fashion, without requiring any computer search. As pointed out in [12], to guarantee full antenna diversity, a unitary ST constellation must satisfy $T \ge 2M$. We will concentrate on the case $T = 2M$. Due to its importance, we will describe and analyze the constellation for two transmit antennas $M = 2$. Specifically, our goal here is to construct a unitary ST constellation $\mathcal{C}_a$ with L matrices of dimension 4×2. Our design is inspired by the well-known Alamouti scheme [2], which is the unique complex orthogonal design for two transmit antennas that achieves full diversity at 1 symbol per channel use in coherent ST systems [15].

The design procedure for $\mathcal{C}_a$ with $L = Q^2$, $Q \in \mathbb{N}$, is as follows. We first construct a single matrix $\boldsymbol{X} \in \mathcal{C}_a$ by concatenating a known unitary matrix $\boldsymbol{T}$ with the 2×2 complex orthogonal design $\boldsymbol{O}_{2\times 2}$. Intuitively, the matrix $\boldsymbol{T}$ can be considered as a training matrix; whereas $\boldsymbol{O}_{2\times 2}$ carries information symbols. The matrices $\boldsymbol{T}$ and $\boldsymbol{O}_{2\times 2}$ are

$$\boldsymbol{T} = \begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix} \quad\text{and}\quad \boldsymbol{O}_{2\times 2} = \begin{bmatrix}s_0 & s_1\\ -s_1^* & s_0^*\end{bmatrix}$$

where $s_0, s_1 \in \mathbb{C}$ are the two information symbols. In the coherent case, $\boldsymbol{O}_{2\times 2}$ achieves 1 symbol per channel use. To guarantee the unitarity of $\boldsymbol{X} := [\boldsymbol{T}^T \quad \boldsymbol{O}_{2\times 2}{}^T]^T$, we confine s_0 and s_1 to be Q-PSK symbols. Specifically, defining $\boldsymbol{X}_{k,l}$ as

$$\boldsymbol{X}_{k,l} := \begin{bmatrix}1 & -1 & e^{j\frac{2\pi}{Q}k} & -e^{-j\frac{2\pi}{Q}l}\\ 1 & 1 & e^{j\frac{2\pi}{Q}l} & e^{-j\frac{2\pi}{Q}k}\end{bmatrix}^T$$

where $(k,l) \in S\times S$ and $S = \{0,\ldots,Q-1\}$, a unitary ST constellation of size $L = Q^2$ can be designed as $\mathcal{C}_a = \{\boldsymbol{X}_{k,l} \mid (k,l) \in S\times S\}$. It can be easily verified that $\boldsymbol{X}_{k,l}^H\boldsymbol{X}_{k,l} = 4\boldsymbol{I}_2, \forall\,(k,l) \in S\times S$. The uncoded bit rate is $(2/T)\log_2(Q) = (1/2)\log_2(Q)$.

This constellation can be considered as an example of the more general pilot-based designs. Nonetheless, the constellation's inherent orthogonality and enforced unitarity make it a special class. On the one hand, these properties may harm the spectral efficiency; on the other hand, they enable low-complexity ML detection, unlike other pilot-based designs.

C. Constellation Properties

We have illustrated a simple construction of unitary ST constellations for two transmit antennas. We discuss some important properties of these constellations next.

Property 1: For a design $\mathcal{C}_a$ and any $\boldsymbol{X}_{k,l}, \boldsymbol{X}_{k',l'} \in \mathcal{C}_a$, it can be shown that

$$0 \le d_1(\boldsymbol{X}_{k,l},\boldsymbol{X}_{k',l'}) = d_2(\boldsymbol{X}_{k,l},\boldsymbol{X}_{k',l'}) \le 1.$$

Proof: For any $\mathcal{C}_a$ and any $\boldsymbol{X}_{k,l}, \boldsymbol{X}_{k',l'} \in \mathcal{C}_a$, it is easy to verify that

$$\begin{aligned}&\frac{1}{T^2}\left(\boldsymbol{X}_{k,l}^H\boldsymbol{X}_{k',l'}\right)^H\left(\boldsymbol{X}_{k,l}^H\boldsymbol{X}_{k',l'}\right)\\ &\qquad = \frac{1}{4}\left[2+\cos\frac{2\pi}{Q}(k-k')+\cos\frac{2\pi}{Q}(l-l')\right]\boldsymbol{I}_2.\end{aligned}$$

Therefore, the matrix $\boldsymbol{X}_{k,l}^H\boldsymbol{X}_{k',l'}/T$ clearly has two equal singular values between 0 and 1, i.e.,

$$0 \le d_1(\boldsymbol{X}_{k,l},\boldsymbol{X}_{k',l'}) = d_2(\boldsymbol{X}_{k,l},\boldsymbol{X}_{k',l'}) \le 1.$$

If

$$d_{k,l;k',l'} = d_1(\boldsymbol{X}_{k,l},\boldsymbol{X}_{k',l'}) = d_2(\boldsymbol{X}_{k,l},\boldsymbol{X}_{k',l'})$$

then

$$d_{k,l;k',l'} = \frac{1}{2}\sqrt{2+\cos\frac{2\pi}{Q}(k-k')+\cos\frac{2\pi}{Q}(l-l')}. \tag{6}$$

It is clear that $d_{k,l;k',l'}$ achieves its maximum 1 when (k',l') is matched to (k,l). In general, it may not achieve its possible minimum 0. □

Property 2: The constellation $\mathcal{C}_a$ enjoys full antenna diversity MN and the corresponding diversity product is $\zeta = \sin(\pi/Q)/\sqrt{2}$.

Proof: The maximum singular value $d_{\max}$ related to any two matrices $\boldsymbol{X}_{k,l}, \boldsymbol{X}_{k',l'} \in \mathcal{C}_a$ with $\boldsymbol{X}_{k,l} \neq \boldsymbol{X}_{k',l'}$ can be calculated from (6) as

$$d_{\max} = \max_{\substack{0\le k,l,k',l'<Q \\ (k,l)\neq(k',l')}} d_{k,l;k',l'} = \frac{1}{2}\sqrt{3+\cos\frac{2\pi}{Q}}.$$

Hence, the diversity product for this design is $\zeta = \sin(\pi/Q)/\sqrt{2}$. □

Property 3: The constellation $\mathcal{C}_a$ is *geometrically uniform* with respect to the subspace distance.

Proof: The notion of geometric uniformity was introduced in [7]. Instead of the conventional Euclidean distance, we adopt the subspace distance as our distance metric. An important consequence of this uniformity is that all $\boldsymbol{X}_{k,l} \in \mathcal{C}_a$ share the same distance profile. Our proof proceeds in two steps. First, for any two matrices belonging to $\mathcal{C}_a$, we identify a mapping from $\mathcal{C}_a$ onto $\mathcal{C}_a$ that maps one matrix to the other. Second, we show that this mapping is isometric with respect to (w.r.t.) the subspace distance. For any $\boldsymbol{X}_{k,l}, \boldsymbol{X}_{k',l'} \in \mathcal{C}_a$, we can always find some integers p and q, such that $k' = (k+p)_{\bmod Q}$ and $l' = (l+q)_{\bmod Q}$. Let us then define the mappings

$$\mathcal{M}_1(x) = (x+p)_{\bmod Q} \quad \text{and} \quad \mathcal{M}_2(x) = (x+q)_{\bmod Q}.$$

By definition, $\mathcal{M}_1(\cdot)$ maps k to k', and $\mathcal{M}_2(\cdot)$ maps l to l'. It is clear that both $\mathcal{M}_1(\cdot)$ and $\mathcal{M}_2(\cdot)$ map S onto S. Defining the mapping $\mathcal{M}$ as $\mathcal{M}(\boldsymbol{X}_{k,l}) = \boldsymbol{X}_{\mathcal{M}_1(k),\mathcal{M}_2(l)}$, $\forall\, \boldsymbol{X}_{k,l} \in \mathcal{C}_a$, we obtain

$$\mathcal{M}(\boldsymbol{X}_{k,l}) = \boldsymbol{X}_{k',l'} \quad \text{and} \quad \mathcal{M}(\mathcal{C}_a) = \mathcal{C}_a.$$

Hence, it remains to show that the mapping $\mathcal{M}$ is isometric w.r.t. the distance metric $\delta_S(\cdot,\cdot)$. Let $\mathcal{M}(\boldsymbol{X}_{k_1,l_1}) = \boldsymbol{X}_{k_1',l_1'}$ and $\mathcal{M}(\boldsymbol{X}_{k_2,l_2}) = \boldsymbol{X}_{k_2',l_2'}$. To show

$$\delta_S(\boldsymbol{X}_{k_1,l_1}, \boldsymbol{X}_{k_2,l_2}) = \delta_S(\boldsymbol{X}_{k_1',l_1'}, \boldsymbol{X}_{k_2',l_2'})$$

it is sufficient to prove that for any $\boldsymbol{X}_{k_1,l_1}, \boldsymbol{X}_{k_2,l_2} \in \mathcal{C}_a$, the matrix $\boldsymbol{X}_{k_1,l_1}^H \boldsymbol{X}_{k_2,l_2}/T$ has the same common singular value as $\boldsymbol{X}_{k_1',l_1'}^H \boldsymbol{X}_{k_2',l_2'}/T$. From the definition of $\mathcal{M}$, we can deduce that

$$(k_1-k_2)_{\bmod Q} = (k_1'-k_2')_{\bmod Q}$$
$$(l_1-l_2)_{\bmod Q} = (l_1'-l_2')_{\bmod Q}.$$

Then, it follows from (6) that the common singular value $d_{k_1,l_1;\,k_2,l_2}$ of $\boldsymbol{X}_{k_1,l_1}^H \boldsymbol{X}_{k_2,l_2}/T$ is equal to the common singular value $d_{k_1',l_1';\,k_2',l_2'}$ of $\boldsymbol{X}_{k_1',l_1'}^H \boldsymbol{X}_{k_2',l_2'}/T$. □

We discuss the consequences of these properties next. As mentioned in Section III-A, a closed-form expression exists for $P_e(\boldsymbol{X} \to \boldsymbol{X}')$, when $\boldsymbol{X}^H\boldsymbol{X}'/T$ has equal singular values. Property 1 enables such an exact PEP analysis. The importance of Property 2 for error performance is apparent, since it ensures full diversity. Following Property 3, we can simplify (3) as follows:

$$P_s \le \sum_{\boldsymbol{X}'\in\mathcal{C},\ \boldsymbol{X}'\neq\boldsymbol{X}} P_e(\boldsymbol{X}\to\boldsymbol{X}').$$

Furthermore, combining $\mathcal{C}_a$ with uniform trellises, we obtain geometrically uniform codes for noncoherent ST systems, which will be developed in Section IV.

D. Simplified ML Detection

For large constellations, a major drawback of unitary ST modulations is their relatively high ML detection complexity. However, our constellations based on the orthogonal design allow for an ML detector with much less complexity, especially for large constellation sizes. In this section, we will derive this new simplified ML detector, and compare it to the original ML detector of [10].

Based on the ML detector for unitary constellations in (2), we consider the ML detection when $\mathcal{C}_a$ is employed. For any $\boldsymbol{X}_{k,l} \in \mathcal{C}_a$, it is possible to decompose $\|\boldsymbol{X}_{k,l}^H\boldsymbol{Y}\|^2$ into three parts

$$\begin{aligned}\left\|\boldsymbol{X}_{k,l}^H\boldsymbol{Y}\right\|^2 &= \operatorname{tr}\left\{\boldsymbol{Y}^H\boldsymbol{X}_{k,l}\boldsymbol{X}_{k,l}^H\boldsymbol{Y}\right\} \\ &= 2\|\boldsymbol{Y}\|^2 + \operatorname{tr}\left\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{A}_k\right\} + \operatorname{tr}\left\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{B}_l\right\} \qquad (7)\end{aligned}$$

where

$$\boldsymbol{A}_k = \begin{bmatrix} 0 & e^{-jk\frac{2\pi}{K}} & 0 & e^{jk\frac{2\pi}{K}} \\ e^{jk\frac{2\pi}{K}} & 0 & -e^{jk\frac{2\pi}{K}} & 0 \\ 0 & -e^{-jk\frac{2\pi}{K}} & 0 & e^{jk\frac{2\pi}{K}} \\ e^{-jk\frac{2\pi}{K}} & 0 & e^{-jk\frac{2\pi}{K}} & 0 \end{bmatrix}$$

$$\boldsymbol{B}_l = \begin{bmatrix} 0 & e^{-jl\frac{2\pi}{K}} & 0 & -e^{jl\frac{2\pi}{K}} \\ e^{jl\frac{2\pi}{K}} & 0 & e^{jl\frac{2\pi}{K}} & 0 \\ 0 & e^{-jl\frac{2\pi}{K}} & 0 & e^{jl\frac{2\pi}{K}} \\ -e^{-jl\frac{2\pi}{K}} & 0 & e^{-jl\frac{2\pi}{K}} & 0 \end{bmatrix}.$$

Notice how the second term is only related to the first index k, while the third term is only related to the second index l. Based on this decomposition, the new simplified ML detector can now be expressed as

$$\hat{\boldsymbol{X}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{X}_{k,l}\in\mathcal{C}_a} \left\|\boldsymbol{X}_{k,l}^H\boldsymbol{Y}\right\|^2 = \boldsymbol{X}_{\hat{k}_{\mathrm{ML}},\hat{l}_{\mathrm{ML}}}$$

where $\hat{k}_{\mathrm{ML}}$ and $\hat{l}_{\mathrm{ML}}$ are computed as follows:

$$\hat{k}_{\mathrm{ML}} = \arg\max_{k\in S} \operatorname{tr}\left\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{A}_k\right\}$$
$$\hat{l}_{\mathrm{ML}} = \arg\max_{l\in S} \operatorname{tr}\left\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{B}_l\right\}.$$

The original ML decoder needs to calculate $\|\boldsymbol{X}_{k,l}^H\boldsymbol{Y}\|^2$, Q^2 times. However, our simplified approach only needs to calculate $\operatorname{tr}\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{A}_k\}$, Q times, and $\operatorname{tr}\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{B}_l\}$, Q times. For large-size constellations, the detection complexity is considerably reduced.

There are two ways to treat the two indexes. One is to treat them as a single symbol. Alternatively, we can treat them as two independent symbols. In the latter case, two questions arise: Will each index achieve full diversity? What is the diversity product for each index? To address these questions, we first focus on the index k. The ML detector for k is

$$\hat{k}_{\mathrm{ML}} = \arg\max_{k\in S}\ \operatorname{tr}\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{A}_k\}.$$

To show that k achieves full diversity, we consider the PEP between k and k'

$$\begin{aligned}P_e(k\to k') &= P\left(\operatorname{tr}\left\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{A}_k\right\} < \operatorname{tr}\left\{\boldsymbol{Y}\boldsymbol{Y}^H\boldsymbol{A}_{k'}\right\} \mid k\right) \\ &= P\left(\left\|\boldsymbol{X}_{k,l}^H\boldsymbol{Y}\right\| < \|\boldsymbol{X}_{k',l}^H\boldsymbol{Y}\| \mid k,l\right) \\ &= P_e(\boldsymbol{X}_{k,l}\to\boldsymbol{X}_{k',l}).\end{aligned}$$

Since the maximum singular value related to any two matrices $\boldsymbol{X}_{k,l}, \boldsymbol{X}_{k',l'} \in \mathcal{C}_a$ with $\boldsymbol{X}_{k,l} \neq \boldsymbol{X}_{k',l'}$ is the same as the one related to any two matrices $\boldsymbol{X}_{k,l}, \boldsymbol{X}_{k',l} \in \mathcal{C}_a$ with $\boldsymbol{X}_{k,l} \neq \boldsymbol{X}_{k',l}$ (see (6)), it is clear that k achieves full diversity and enjoys the same diversity product as $\boldsymbol{X}_{k,l}$. The same result holds for l.

E. Comparison With Systematic Design

We will compare our unitary constellations with those from the systematic design of [11]. For systematic designs, we follow the computer search procedures described in [11], but adopt either the diversity product distance or the subspace distance to replace the original chordal distance. These new distance metrics guarantee full diversity. We have designed three sets of constellations with cardinality $L = 16$, $L = 32$, and $L = 64$. There are three constellations in each set, which are obtained from our constructive design, the systematic design with diversity product distance, and the systematic design with subspace distance,

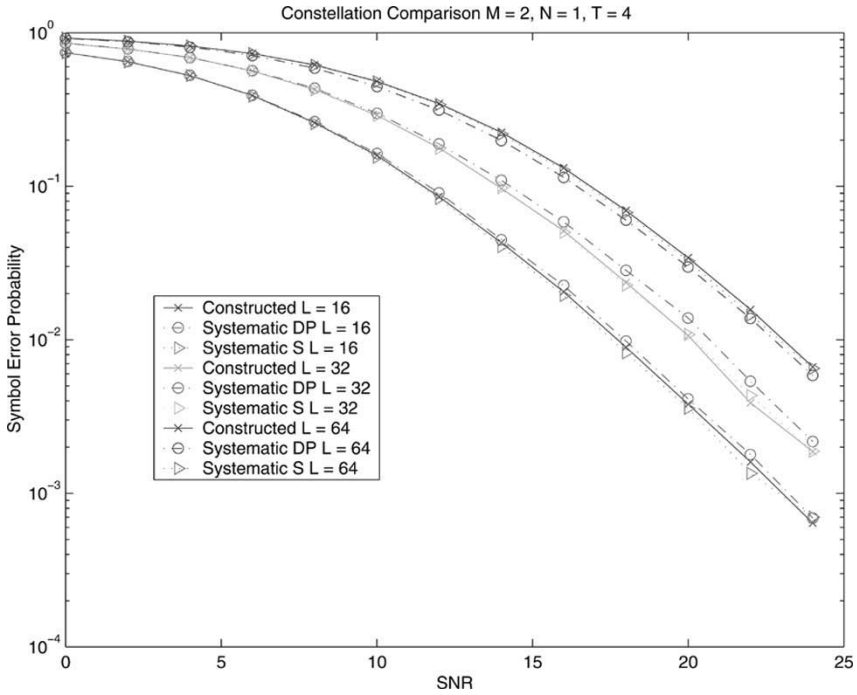

Fig. 1. Constellation performance comparison based on Monte Carlo simulations.

respectively. The Monte Carlo simulation results of SEP for these constellations are depicted in Fig. 1. We can observe that the constellations in each set enjoy similar SEP performance. Nonetheless, our constructive design enables a low-complexity ML detection.

F. Extensions for $M > 2$

To enable low-complexity ML detection, we have constructed our constellations based on the orthogonal designs. Orthogonality limits the spectral efficiency of noncoherent signaling. Hence, this design is not suitable for high spectral efficiency with large M. Having this in mind, we only generalize our constellation construction to $M = 3$ and $M = 4$, since these constellations may still be of practical interest.

We first design unitary constellations for four transmit antennas. Each matrix in the constellation has dimension $T \times M = 8 \times 4$. The construction is similar to the one described in Section III-B. For simplicity, we will focus on a constellation size of $L = Q^3$, $Q \in \mathbb{N}$, and derive only the diversity product and the simplified ML receiver. For $M = 4$, we employ the training matrix $\boldsymbol{T}_{4\times 4}$ and the rate-$3/4$ orthogonal design $\boldsymbol{O}_{4\times 4}$ as our building blocks, where

$$\boldsymbol{T}_{4\times 4} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix}$$

and

$$\boldsymbol{O}_{4\times 4} = \begin{bmatrix} s_0 & s_1 & s_2 & 0 \\ -s_1^* & s_0^* & 0 & -s_2 \\ -s_2^* & 0 & s_0^* & s_1 \\ 0 & s_2^* & -s_1^* & s_0 \end{bmatrix}.$$

The design $\boldsymbol{O}_{4\times 4}$ is available in [19]. Enforcing s_0, s_1, s_2 in $\boldsymbol{O}_{4\times 4}$ to be Q-PSK symbols, we obtain one matrix $\boldsymbol{X}_{k,l,m} \in \mathcal{C}_a$ as follows:

$$\boldsymbol{X}_{k,l,m} = \begin{bmatrix} 1 & -1 & -1 & 0 & r^k & -r^{-l} & -r^{-m} & 0 \\ 1 & 1 & 0 & 1 & r^l & r^{-k} & 0 & r^{-m} \\ 1 & 0 & 1 & -1 & r^m & 0 & r^{-k} & -r^{-l} \\ 0 & -1 & 1 & 1 & 0 & -r^m & r^l & r^k \end{bmatrix}$$

where $(k, l, m) \in S \times S \times S$, $S = \{0, \ldots, Q-1\}$, and $r = e^{j\frac{2\pi}{Q}}$. The resulting unitary ST constellation is

$$\mathcal{C}_a = \{\boldsymbol{X}_{k,l,m} \mid (k, l, m) \in S \times S \times S\}.$$

For $\mathcal{C}_a$ and any matrices $\boldsymbol{X}_{k,l,m}, \boldsymbol{X}_{k',l',m'} \in \mathcal{C}_a$, it can be derived that $\boldsymbol{X}_{k,l,m}^H \boldsymbol{X}_{k',l',m'}/T$ has four equal singular values. Letting $d_{k,l,m;k',l',m'}$ be the common singular value, we have

$$d_{k,l,m;k',l',m'} = \left\{ \frac{1}{2} + \frac{1}{6}\left[\cos\frac{2\pi}{Q}(k-k') + \cos\frac{2\pi}{Q}(l-l') + \cos\frac{2\pi}{Q}(m-m') \right] \right\}^{1/2}.$$

Hence, the diversity product is $\zeta = \sin(\pi/Q)/\sqrt{3}$. The simplified ML receiver is

$$\hat{k}_{\mathrm{ML}} = \arg\max_{k\in S} \left\| \boldsymbol{X}_{k,0,0}^H \boldsymbol{Y} \right\|^2$$

$$\hat{l}_{\mathrm{ML}} = \arg\max_{l\in S} \left\| \boldsymbol{X}_{0,l,0}^H \boldsymbol{Y} \right\|^2$$

$$\hat{m}_{\mathrm{ML}} = \arg\max_{m\in S} \left\| \boldsymbol{X}_{0,0,m}^H \boldsymbol{Y} \right\|^2.$$

As in the coherent case, the constellation for three transmit antennas can be constructed by simply deleting one specific column from the design of $M = 4$.

IV. Trellis-Coded Modulation (TCM)

Even though ST signals can serve as their own channel codes [16], [15], the inherent orthogonal structure in our unitary constellation may limit its coding gain. To improve the coding gain and exploit the rich diversity in fast fading, we pursue a TCM approach. In [14], unitary ST modulation was combined with convolutional coding to boost error performance, but no constellation expansion and set partitioning concepts were involved due to the lack of structure in the computer searched constellations. On the other hand, our constructive design exhibits uniformity and enables simple set partitioning.

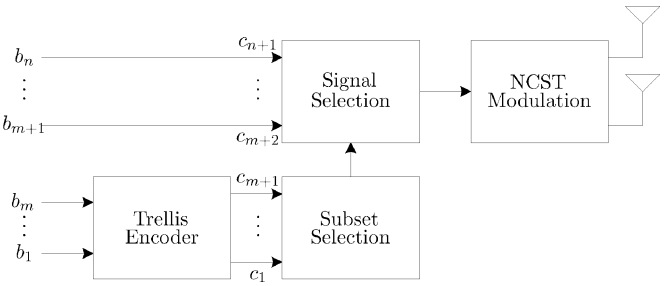

Fig. 2. TCM system model.

A. System Diagram and Trellis Decision Metric

The block diagram of the proposed TCM system is shown in Fig. 2. Assuming the uncoded system has a constellation of size 2^n, the TCM system can be described as follows. From the n-bit input vector $\boldsymbol{b} := [b_1 \quad b_2 \quad \cdots \quad b_n]^T$, m bits are fed to a rate $R = m/(m+1)$ convolutional encoder. The $m+1$ output bits are used to choose a subset of the noncoherent constellation; whereas the remaining $n-m$ bits determine a signal within this subset. Hence, one redundant bit is added every n input bits, which amounts to expanding the constellation size to 2^{n+1}. If $m < n$, parallel transitions are allowed in the trellis; If $m = n$, no parallel transitions exist.

The Viterbi algorithm (VA) is employed in the ML receiver. We derive the decision metric for the VA next. A sequence $\tilde{\boldsymbol{X}} := \{\boldsymbol{X}_{k_t,l_t}\}_{t=0}^{\infty}$ of matrices drawn from $\mathcal{C}_a$ is transmitted. The received sequence is denoted by $\tilde{\boldsymbol{Y}} := \{\boldsymbol{Y}_t\}_{t=0}^{\infty}$, where $\boldsymbol{Y}_t = \boldsymbol{X}_{k_t,l_t}\boldsymbol{H}_t + \boldsymbol{N}_t$ and $\boldsymbol{H}_t$, $\boldsymbol{N}_t$ are the channel and noise matrix for the tth coherence interval, respectively. The conditional probability density function (pdf) of $\boldsymbol{Y}_t$ is

$$p(\boldsymbol{Y}_t \mid \boldsymbol{X}_{k_t,l_t}) = \frac{\exp\left(-\mathrm{tr}\left\{\left[\boldsymbol{I}_T + \rho/M\, \boldsymbol{X}_{k_t,l_t}\boldsymbol{X}_{k_t,l_t}^H\right]^{-1}\boldsymbol{Y}_t\boldsymbol{Y}_t^H\right\}\right)}{\pi^{TN}\det^N\left[\boldsymbol{I}_T + \rho/M\, \boldsymbol{X}_{k_t,l_t}\boldsymbol{X}_{k_t,l_t}^H\right]}.$$

Due to the independence of the $\boldsymbol{H}_t$'s and the $\boldsymbol{N}_t$'s, the conditional pdf of the received sequence is

$$p(\tilde{\boldsymbol{Y}} \mid \tilde{\boldsymbol{X}}) = \prod_{t=0}^{\infty} p(\boldsymbol{Y}_t \mid \boldsymbol{X}_{k_t,l_t}).$$

The ML sequence estimate for $\tilde{\boldsymbol{X}}$ generated by the VA is

$$\tilde{\boldsymbol{X}}_{\mathrm{ML}} = \arg\max_{\tilde{\boldsymbol{X}}} \sum_{t=0}^{\infty} \| \boldsymbol{X}_{k_t,l_t}^H \boldsymbol{Y}_t \|^2.$$

Based on $\tilde{\boldsymbol{X}} = \{\boldsymbol{X}_{k_t,l_t}\}_{t=0}^{\infty}$, we define two sequences $\tilde{\boldsymbol{k}} = \{k_t\}_{t=0}^{\infty}$ and $\tilde{\boldsymbol{l}} = \{l_t\}_{t=0}^{\infty}$. It follows from the constellation's orthogonality that the ML sequence estimates for $\tilde{\boldsymbol{k}}$ and $\tilde{\boldsymbol{l}}$ are

$$\tilde{\boldsymbol{k}}_{\mathrm{ML}} = \arg\max_{\tilde{\boldsymbol{k}}} \sum_{t=0}^{\infty} \mathrm{tr}\left\{\boldsymbol{Y}_t\boldsymbol{Y}_t^H\boldsymbol{A}_{k_t}\right\}$$

$$\tilde{\boldsymbol{l}}_{\mathrm{ML}} = \arg\max_{\tilde{\boldsymbol{k}}} \sum_{t=0}^{\infty} \mathrm{tr}\left\{\boldsymbol{Y}_t\boldsymbol{Y}_t^H\boldsymbol{B}_{l_t}\right\}.$$

B. Constellation Expansion and Set Partitioning

Due to the orthogonal structure and the employment of phase-shift keying (PSK) symbols as entries of our constructive design, the constellation expansion and set partitioning of our noncoherent ST constellation are similar to those of PSK. The only difference lies in the distance metric. We employ the subspace distance in the set partitioning instead of the Euclidean distance as in the conventional TCM. We consider the subspace distance between $\boldsymbol{X}_{k,l}$ and $\boldsymbol{X}_{k',l'}$ next. Plugging the two equal singular values of $\boldsymbol{X}_{k,l}^H\boldsymbol{X}_{k',l'}/T$ from (6) into (5), we obtain

$$\delta_S(\boldsymbol{X}_{k,l},\boldsymbol{X}_{k',l'}) = \frac{1}{2} - \frac{1}{4}\cos\left[\frac{2\pi}{Q}(k-k')\right] - \frac{1}{4}\cos\left[\frac{2\pi}{Q}(l-l')\right]. \tag{8}$$

Hence, the subspace distance depends only on $(k-k')_{\mathrm{mod}\ Q}$ and $(l-l')_{\mathrm{mod}\ Q}$. Examples of the considered constellation expansion and set partitioning will be given in Section IV-D.

C. Performance Analysis

In this subsection, we analyze the asymptotic error performance of the proposed noncoherent ST TCM scheme. Our analysis is based on a simple approximate bounding technique for the asymptotic error probability (AEP) using the asymptotic sequence PEP. A surprising result is that the AEP expressions can be derived for certain TCM systems, while they are not available for uncoded noncoherent ST systems.

1) Asymptotic Sequence PEP: A simple method to examine the error performance of a TCM scheme is through the sequence PEP. The PEP between two sequences

$$\tilde{\boldsymbol{X}} = \{\boldsymbol{X}_{k_t,l_t}\}_{t=0}^{\infty} \quad \text{and} \quad \tilde{\boldsymbol{X}}' = \{\boldsymbol{X}_{k'_t,l'_t}\}_{t=0}^{\infty}$$

is defined as

$$P_e(\tilde{\boldsymbol{X}} \to \tilde{\boldsymbol{X}}') = P\left(\sum_{t=0}^{\infty}\left\|\boldsymbol{X}_{k'_t,l'_t}^H\boldsymbol{Y}_t\right\|^2 > \sum_{t=0}^{\infty}\left\|\boldsymbol{X}_{k_t,l_t}^H\boldsymbol{Y}_t\right\|^2 \mid \tilde{\boldsymbol{X}}\right).$$

The asymptotic sequence PEP, denoted as $P_e^a(\tilde{\boldsymbol{X}} \to \tilde{\boldsymbol{X}}')$, can be well approximated by

$$P_e^a(\tilde{\boldsymbol{X}} \to \tilde{\boldsymbol{X}}') \approx \delta_P(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}')^{-1} \times \left[\binom{2MN-1}{MN}\left(\frac{\rho T}{M}\right)^{-MN}\right]^{\delta_H(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}')} \tag{9}$$

where $\delta_H(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}')$ is the Hamming distance between $\tilde{\boldsymbol{X}}$ and $\tilde{\boldsymbol{X}}'$ and $\delta_P(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}')$ is the product distance between $\tilde{\boldsymbol{X}}$ and $\tilde{\boldsymbol{X}}'$ defined as

$$\delta_P(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}') = \prod_{\substack{t=0 \\ \boldsymbol{X}_{k_t,l_t}\neq\boldsymbol{X}_{k'_t,l'_t}}}^{\infty}\left(1 - d^2_{k_t,l_t;k'_t,l'_t}\right)^{MN}.$$

Note that the above approximation is asymptotically accurate. In a log-log plot, the asymptotic sequence PEP coincides with the true sequence PEP curve for high SNR. Based on the discussions in Section III-D, we conclude that similar asymptotic sequence PEP expressions can be derived for index sequences $\tilde{\boldsymbol{k}}$ and $\tilde{\boldsymbol{l}}$.

We define two sequences in a trellis as *close* sequences if they achieve the minimum Hamming distance. Furthermore, we define two sequences as *the closest* sequences if they are close and they achieve the minimum product distance in all close-sequence pairs. The diversity order of the sequence error probability is dominated by the diversity order of the close sequence PEP. Letting δ_H be the Hamming distance between two close sequences, it follows that this diversity order is $\delta_H MN$. If parallel transitions exist in a trellis, then $\delta_H = 1$ and the diversity order is limited by MN. For this reason, we mainly consider TCM systems without parallel transitions.

2) AEP Bounds: We study the asymptotic error performance of noncoherent ST TCM by generalizing the approach in [3]. We will focus on uniform trellises. Combining our uniform unitary constellations with these uniform trellises, we obtain geometrically uniform codes [7]. As a direct consequence of the geometric uniformity, all $\tilde{\boldsymbol{X}}$ sequences in a trellis share the same product distance profile. Hence, performance analysis is simplified by considering the error patterns on one arbitrary sequence. We will derive approximate upper and lower bounds on the AEP that are accurate for high SNR. For certain special

trellises, we can even determine the AEP expression in closed form. Note that based on the AEP and trellis mapping, we can derive a corresponding expression for the asymptotic SEP (ASEP) or the asymptotic bit error probability (ABEP).

Assuming that all possible sequences are equally likely and only sequences that are *close* to the transmitted one have significant probability to be the ML estimate, an asymptotically accurate approximate upper bound and a lower bound on the AEP, P^a, can be formulated as

$$P_e^a(\tilde{\boldsymbol{X}} \to \tilde{\boldsymbol{X}}_1^{'}) \leq P^a \preceq \sum_{n=1}^{N_c} P_e^a(\tilde{\boldsymbol{X}} \to \tilde{\boldsymbol{X}}_n^{'})$$

where $\preceq$ indicates that the bound is approximate, yet asymptotically accurate, $\{\tilde{\boldsymbol{X}}_n^{'}\}_{n=1}^{N_c}$ is the set of sequences close to $\tilde{\boldsymbol{X}}$, and $\tilde{\boldsymbol{X}}_1^{'}$ is a sequence that is the closest to $\tilde{\boldsymbol{X}}$. Due to Property 1, exact closed-form expressions can be calculated for the asymptotic sequence PEP. Nonetheless, we employ the asymptotically accurate approximation of the asymptotic sequence PEP in (9) for simplicity. With δ_H being the Hamming distance between two close sequences, we obtain the following asymptotically accurate approximate upper and lower bounds:

$$f(\rho,\delta_H)\delta_P(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}_1^{'})^{-1} \preceq P^a \preceq f(\rho,\delta_H)\sum_{n=1}^{N_c}\delta_P(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}_n^{'})^{-1} \quad (10)$$

where

$$f(\rho,\delta_H) = \left[\binom{2MN-1}{MN}\left(\frac{\rho T}{M}\right)^{-MN}\right]^{\delta_H}.$$

For some trellises, there is only one close sequence, i.e., $N_c = 1$. Hence, the approximate upper and lower bounds in (10) then coincide and yield the exact AEP expression. Based on the trellis mapping, we can subsequently determine expressions or approximate bounds for the ASEP P_s^a, or ABEP P_b^a.

With examples, we quantify the performance gain of one TCM system relative to another by two simple parameters. We will consider an uncoded system as a special TCM with a single-state trellis. The primary parameter is the diversity order, while the secondary parameter is the asymptotic coding gain (ACG). The diversity order of a TCM system is $\delta_H MN$. For TCM systems enjoying the same diversity order, we use ACG to further characterize the performance difference. When the ASEP expressions are known for two TCM schemes with the same minimum Hamming distance δ_H, the ACG is calculated as follows. Letting the minimum product distances be $\delta_{P,1}$ and $\delta_{P,2}$, respectively, the SEP-based ACG is $\gamma = (\delta_{P,2}/\delta_{P,1})^{1/(MN\delta_H)}$. Whenever ASEP expressions are not available, an approximate range for the ACG will be provided.

D. TCM Examples and Simulations

We have derived the ML sequence metric for a TCM receiver, illustrated the set partitioning, and developed two simple measures to quantify error performance. We are now ready to give several complete TCM examples to illustrate design procedures and verify our theoretical analysis with simulations.

1) TCM Example 1: To clarify the coding gains of TCM, we compare the error performance of a TCM system with an uncoded system. The uncoded system employs the constellation $\mathcal{C}_a = \{\boldsymbol{X}_{k,l} \mid (k,l) \in S \times S\}$ with cardinality $L = 16$, where $S = \{0,1,2,3\}$; whereas the TCM system employs the constellation $\mathcal{C}_a = \{\boldsymbol{X}_{k,l} \mid (k,l) \in S \times S\}$ with cardinality $L = 64$, where $S = \{0,1,\ldots,7\}$. Due to the orthogonal constellation structure, we can treat the indexes separately for both the uncoded and coded cases. Hence, the constellation sizes for the indexes are 4 and 8, respectively. For the coded case, we employ two identical but independent TCM schemes on the index sequences $\{\tilde{\boldsymbol{k}}_t\}_{t=0}^{\infty}$ and $\{\tilde{\boldsymbol{l}}_t\}_{t=0}^{\infty}$.

TABLE I
SET PARTITIONING FOR $\{k \mid k \in S\}$, WHERE $S = \{0,1,\ldots,7\}$

Level	Distance	Signal Sets
0	0.2706	$s_0^{(0)} = \{0,1,2,3,4,5,6,7\}$
1	0.5000	$s_0^{(1)} = \{0,2,4,6\},\ s_1^{(1)} = \{1,3,5,7\}$
2	0.7071	$s_0^{(2)} = \{0,4\},\ s_1^{(2)} = \{2,6\},$ $s_2^{(2)} = \{1,5\},\ s_3^{(2)} = \{3,7\}$

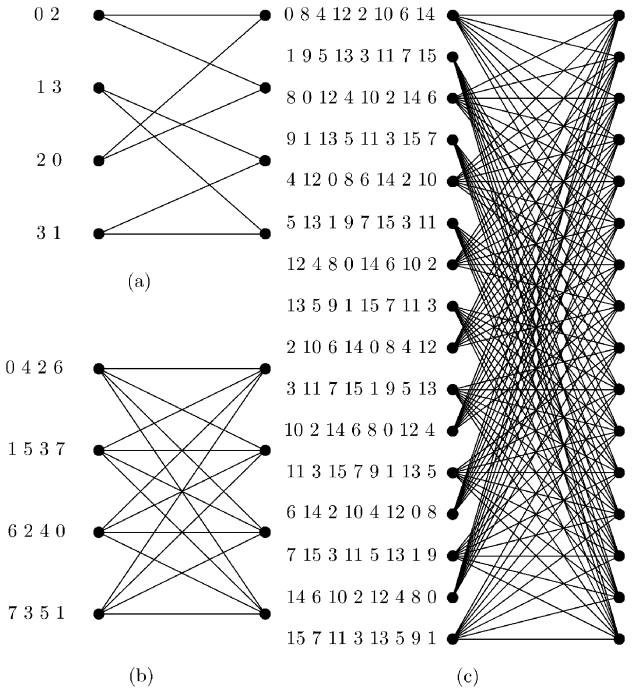

Fig. 3. Three trellis diagrams: (a) $R = 1/2$ 4-state; (b) $R = 2/3$ 4-state; (c) $R = 3/4$ 16-state.

We analyze the performance of the uncoded system first. Considering it as a one-state TCM, the minimum Hamming distance for the uncoded system is $\delta_H^u = 1$. Following (10), the approximate ASEP bounds are

$$f(\rho,1)(1-d_{0,0;1,0}^2)^{-MN} \preceq P_s^a \preceq f(\rho,1)\sum_{l=1}^{3}(1-d_{0,0;l,0}^2)^{-MN}$$

where $d_{0,0;l,0}$ is the common singular value of $\boldsymbol{X}_{0,0}^H\boldsymbol{X}_{l,0}/T$ given in (6). Plugging in these singular values, we obtain

$$f(\rho,1)\cdot 4^{MN} \preceq P_s^a \preceq f(\rho,1)\cdot[2\cdot 4^{MN} + 2^{MN}].$$

Even though the exact upper and lower bounds are available in closed form, we adopt the approximate bounds for simplicity. These bounds are sufficient to characterize the ACG.

For the TCM system, the set partitioning follows that of an 8-PSK constellation and is shown in Table I along with the intrasubset subspace distances. One input bit is fed to the rate $R = 1/2$ 4-state trellis encoder depicted in Fig. 3(a). In this trellis, the numbers next to a state correspond to the subscripts of the subsets $s_0^{(2)},\ldots,s_3^{(2)}$ shown in Table I. The 2-bit output of the trellis encoder determines a subset in $s_0^{(2)},\ldots,s_3^{(2)}$ and the remaining input bit selects a signal within thechosen subset. Due to the parallel transitions in the trellis, the minimum Hamming distance is $\delta_H^c = 1$. Since there is only one sequence close to the transmitted one, the ASEP is $P_s^a = f(\rho,1)\cdot 2^{MN}$. The

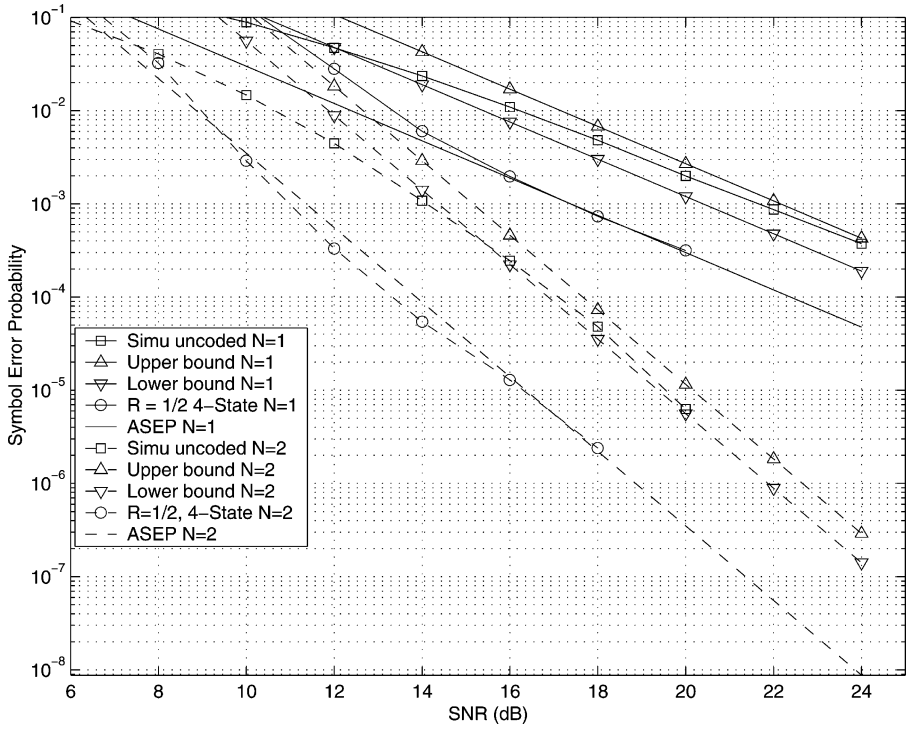

Fig. 4. SEP comparison for TCM Example 1.

approximate range of the ACG for the TCM system relative to the uncoded one is

$$\left(\frac{4^{MN}}{2^{MN}}\right)^{1/(MN)} \preceq \gamma \preceq \left(\frac{2\cdot 4^{MN}+2^{MN}}{2^{MN}}\right)^{1/(MN)}.$$

More explicitly, when $M=2$ and $N=1$, the ACG ranges from 3.01 to 4.77 dB; whereas when $M=N=2$, the ACG ranges from 3.01 to 3.79 dB. Monte Carlo simulation results for the SEP and ASEP expressions are reported in Fig. 4 for these two settings. For the TCM scheme, we can observe that the ASEP curves accurately approximate the simulated SEP at high SNR. Furthermore, the actual ACGs lie in the theoretically predicted range for both settings.

2) TCM Example 2: We compare the performance of two TCM schemes with an uncoded one. The size $L=8$ constellation employed by the uncoded system is $\mathcal{C}_a=\{\boldsymbol{X}_{k,l} \mid (k,l)\in S_1\times S_1\cup S_2\times S_2\}$, where $S_1=\{0,2\}$ and $S_2=\{1,3\}$. The uncoded system has minimum Hamming distance $\delta_H^u=1$ and diversity order MN.

For both coded systems, we use the size $L=16$ constellation $\mathcal{C}_a=\{\boldsymbol{X}_{k,l} \mid (k,l)\in S\times S\}$, where $S:=\{0,1,\ldots,3\}$. Due to (8), the set partitioning follows the set partitioning of a two-dimensional 4-PSK constellation and is shown in Table II. Two trellis encoders are considered for this constellation. The first one employs the rate $R=2/3$ 4-state trellis depicted in Fig. 3(b). The second one employs the rate $R=3/4$ 16-state trellis shown in Fig. 3(c).

In TCM scheme 1, the 3-bit trellis output is used to choose among the eight subsets, $s_0^{(3)},\ldots,s_7^{(3)}$ in Table II. The remaining bit is used to select a signal within a subset. Hence, the minimum Hamming distance is the same as for the uncoded system $\delta_H^{c1}=1$. Based on the unique close sequence, the ASEP is calculated by $P_s^a=f(\rho,1)$. Since one sequence error corresponds to one input bit error, and each code symbol corresponds to three input bits, the ABEP is $P_b^a=f(\rho,1)/3$.

In TCM system 2, the trellis encoder has a 4-bit output, which is used to choose among the 16 constellation points $s_0^{(4)},\ldots,s_{16}^{(4)}$ in Table II. There are no parallel transitions in the trellis and the minimum Hamming distance is $\delta_H^{c2}=2$. Hence, the diversity order becomes $2MN$. We determine the close paths to an arbitrary state path, say

$$0\to 0\to\cdots\to 0.$$

Its corresponding symbol sequence is $s_0^{(4)}\to s_0^{(4)}\to\cdots\to s_0^{(4)}$. For this path, we can identify four close paths from the trellis in Fig. 3(c). The first three paths differ from the all-zero state in only one position, where the trellis states are 2, 4, and 6, respectively. Correspondingly, the symbol sequences differ from $s_0^{(4)}\to\cdots\to s_0^{(4)}$ in two positions.

TABLE II
SET PARTITIONING FOR $\{(k,l) \mid (k,l)\in S\times S\}$, WHERE $S=\{0,1,2,3\}$, $S_1=\{0,2\}$, AND $S_2=\{1,3\}$

Lev.	Dist.	Signal Sets
0	0.500	$s_0^{(0)}=\{(k,l) \mid (k,l)\in S\times S\}$
1	0.707	$s_0^{(1)}=\{(k,l) \mid (k,l)\in S_1\times S_1\cup S_2\times S_2\}$ $s_1^{(1)}=\{(k,l) \mid (k,l)\in S_1\times S_2\cup S_2\times S_1\}$
2	0.707	$s_0^{(2)}=\{(k,l) \mid (k,l)\in S_1\times S_1\}$, $s_1^{(2)}=\{(k,l) \mid (k,l)\in S_2\times S_2\}$ $s_2^{(2)}=\{(k,l) \mid (k,l)\in S_1\times S_2\}$, $s_3^{(2)}=\{(k,l) \mid (k,l)\in S_2\times S_1\}$
3	1	$s_0^{(3)}=\{(0,0),(2,2)\}$, $s_4^{(3)}=\{(0,2),(2,0)\}$ $s_2^{(3)}=\{(1,1),(3,3)\}$, $s_6^{(3)}=\{(1,3),(3,1)\}$ $s_1^{(3)}=\{(0,1),(2,3)\}$, $s_5^{(3)}=\{(0,3),(2,1)\}$ $s_3^{(3)}=\{(1,0),(3,2)\}$, $s_7^{(3)}=\{(1,2),(3,0)\}$
4	–	$s_0^{(4)}=(0,0)$, $s_8^{(4)}=(2,2)$, $s_4^{(4)}=(0,2)$, $s_{12}^{(4)}=(2,0)$ $s_2^{(4)}=(1,1)$, $s_{10}^{(4)}=(3,3)$, $s_6^{(4)}=(1,3)$, $s_{14}^{(4)}=(3,1)$, $s_1^{(4)}=(0,1)$, $s_9^{(4)}=(2,3)$, $s_5^{(4)}=(0,3)$, $s_{13}^{(4)}=(2,1)$ $s_3^{(4)}=(1,0)$, $s_{11}^{(4)}=(3,2)$, $s_7^{(4)}=(1,2)$, $s_{15}^{(4)}=(3,0)$

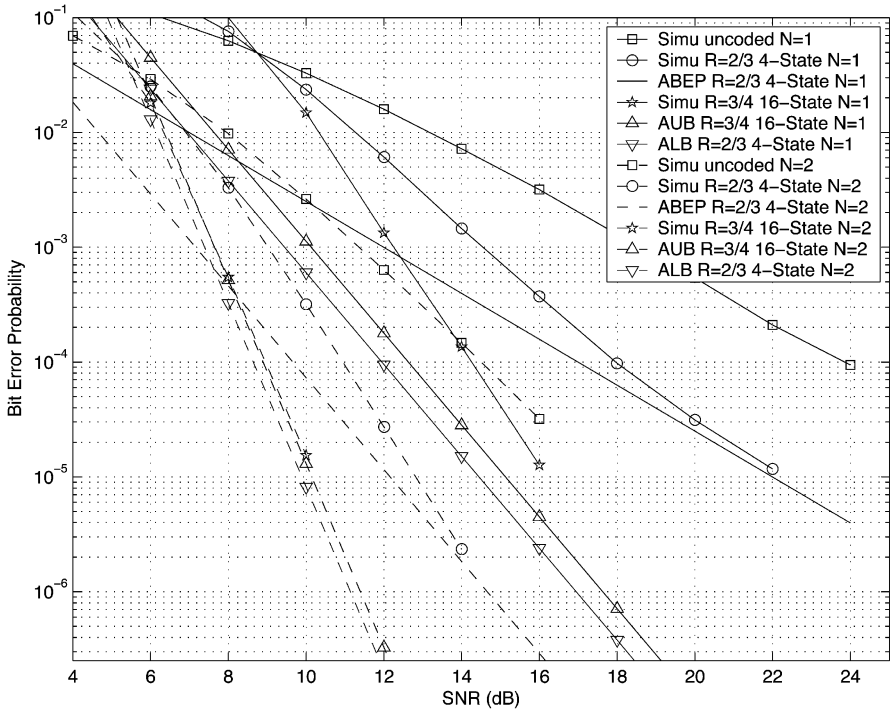

Fig. 5. BEP comparison for TCM Example 2.

They are $s_4^{(4)} \rightarrow s_8^{(4)}$, $s_2^{(4)} \rightarrow s_4^{(4)}$, and $s_6^{(4)} \rightarrow s_{12}^{(4)}$. The numbers of input bit errors involved in these events are 1, 1, and 2 bits, respectively. The remaining path achieving Hamming distance 2 diverts from the all-zero path in two positions. The state and symbol sequences are $0 \rightarrow 4 \rightarrow 2 \rightarrow 0$ and $s_2^{(4)} \rightarrow s_0^{(4)} \rightarrow s_8^{(4)}$. There are two bit errors associated with this event. Based on the bit errors for the close paths and the fact that each coded symbol corresponds to three input bits, approximate bounds for the ABEP are given by

$$2 \cdot f(\rho, 2) \cdot 4^{MN}/3 \preceq P_b^a \preceq f(\rho, 2) \cdot (4^{MN} + 2^{MN}).$$

The SNR gap between the upper and lower bound is 0.68 dB when $M = 2$, $N = 1$, and 0.25 dB when $M = N = 2$. The performance comparison in terms of the bit error probability (BEP) is shown in Fig. 5.

Remark 1: If one matrix in the constellation is treated as one symbol, we may achieve larger subspace distances within each constellation subset than when one matrix is treated as two independent symbols. This corresponds to doubling the dimension of the signal space. With a properly designed trellis, larger intrasubset distances mean better performance. However, the price paid for the performance gain is the considerably increased complexity.

Remark 2: When each matrix is treated as one symbol and no parallel transitions are allowed, the number of trellis states is at least equal to the cardinality of the uncoded constellation. For large constellations, the resulting trellis may be unrealistically complex. On the other hand, when each matrix is treated as two independent symbols, this problem becomes much less severe. Furthermore, the rich diversity in fast fading can be conveniently collected. For this reason, we advocate using separate TCM schemes for the two index streams.

V. Conclusion

We dealt with noncoherent ST transmissions over block fading channels and constructed simple unitary ST constellations, which can be viewed as a special pilot-based design. The novel construction's full diversity allows for low-complexity ML detection, and achieves performance comparable to existing computer searched constellations.

To further exploit the constellation structure, we also pursued a TCM approach. We designed geometrically uniform codes by combining uniform trellises with our uniform constellations w.r.t. the subspace distance. The decision metric for the trellis paths and approximate bounds on the AEP were derived. Based on the AEP and trellis mapping, we derived corresponding expressions for the ASEP or ABEP. Two simple parameters were identified to quantify the performance gain of one TCM system relative to another or to an uncoded system. The primary parameter is the diversity order; whereas the secondary one is the ACG. At the price of increased complexity, we have shown both analytically and with simulations that our TCM designs exhibit considerably improved error performance.

References

[1] D. Agrawal, T. J. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 2618–2626, Sept. 2001.

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

[3] E. Biglieri, D. Divsalar, P. McLane, and M. Simon, *Introdution to Trellis-Coded Modulation With Applications*. New York: MacMillan, 1991.

[4] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2383–2399, Sept. 2001.

[5] ——, "Training-codes for the noncoherent multi-antenna block-Rayleigh-fading channel," in *Proc. 37th Conf. Information Sciences and Systems*, Baltimore, MD, Mar. 12–14, 2003.

[6] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1097–1119, May 2003.

[7] G. D. Forney, "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241–1260, Sept. 1991.

[8] G. J. Foschini and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[9] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: John Hopkins Univ. Press, 1996, p. 599.
[10] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.
[11] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1962–1973, Sept. 2000.
[12] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.
[13] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.
[14] M. L. McCloud, M. Brehler, and M. K. Varanasi, "Signal design and convolutional coding for noncoherent space-time communication on the block-Rayleigh-fading channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1186–1194, May 2002.
[15] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.
[16] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
[17] V. Tarokh and M. Kim, "Existence and construction of noncoherent unitary space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 3112–3117, Dec. 2002.
[18] İ. E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
[19] O. Tirkkonen and A. Hottinen, "Complex space-time block codes for four tx antennas," in *Proc. IEEE GLOBECOM*, San Francisco, CA, Nov. 2000, pp. 1005–1009.
[20] L. Zheng and D. N. C. Tse, "Communicating on the grassmann manifold: A geometric approach to the noncoherent multiple antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 359–383, Feb. 2002.

Sliding-Block Decodable Encoders Between (d, k) Runlength-Limited Constraints of Equal Capacity

Navin Kashyap, *Member, IEEE,* and Paul H. Siegel, *Fellow, IEEE*

***Abstract*—We determine the pairs of (d, k)-constrained systems, $\mathcal{S}(d, k)$ and $\mathcal{S}(\hat{d}, \hat{k})$, of equal capacity, for which there exists a rate 1:1 sliding-block-decodable encoder from $\mathcal{S}(d, k)$ to $\mathcal{S}(\hat{d}, \hat{k})$. In all cases where there exists such an encoder, we explicitly describe the encoder and its corresponding sliding-block decoder.**

***Index Terms*—(d, k)-constrained systems, finite-state encoders, sliding-block decoders.**

I. Introduction

Given nonnegative integers d, k, with $d < k$, we say that a binary sequence is (d, k)-constrained if every run of zeros has length at most k and any two successive ones are separated by a run of zeros of length at least d. A (d, k)-*constrained system* is defined to be the set of all finite-length (d, k)-constrained binary sequences. The above definition can be extended to the case $k = \infty$ by not imposing an upper bound on the lengths of zero runs. In other words, a binary sequence is said to be (d, ∞)-constrained if any two successive ones are separated by at least d zeros, and a (d, ∞)-*constrained system* is defined to be the set of all finite-length (d, ∞)-constrained binary sequences. From now on, when we refer to (d, k)-constrained systems, we shall also allow k to be ∞. Note that the above definition allows finite-length (d, k)-constrained sequences to begin or end with a run of fewer than d zeros.

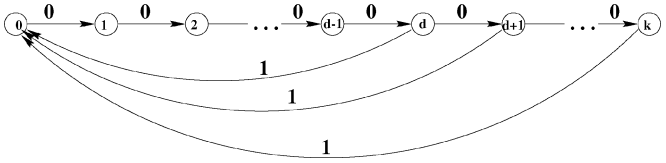

Fig. 1. Graph, $\mathcal{G}_{d,k}$, generating the (d, k)-constrained system $\mathcal{S}(d, k)$ for finite k.

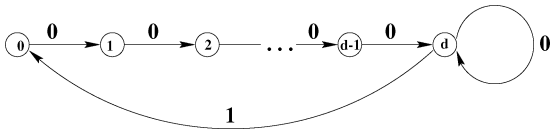

Fig. 2. Graph, $\mathcal{G}_{d,\infty}$, generating the (d, ∞)-constrained system $\mathcal{S}(d, \infty)$.

Binary sequences satisfying some (d, k) constraint are commonly used to encode information in digital and optical recording systems [1]. The parameter k is imposed to guarantee sufficient sign changes in the recorded waveform which are required to prevent clock drift during readback. The parameter d is needed to prevent intersymbol interference.

It is possible to give a convenient graphical description of (d, k)-constrained systems as follows (cf. [1], [2, Chs. 2, 3]). We define a *labeled graph*, $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{L})$, to be a finite directed graph with vertex set $\mathcal{V}$, edge set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, and edge labeling $\mathcal{L} : \mathcal{E} \to \Sigma$, where Σ is a finite alphabet. A labeled graph can be used to generate sequences of symbols from Σ by reading off the labels along paths in the graph. A *constrained system*, $\mathcal{S}$ or $\mathcal{S}(\mathcal{G})$, is the set of all finite-length[1] sequences obtained by reading off the labels along paths in a labeled graph $\mathcal{G}$. Any (d, k)-constrained system, $\mathcal{S}(d, k)$, can be generated from an appropriate labeled graph: for finite k, $\mathcal{S}(d, k)$ is the constrained system generated by the labeled graph $\mathcal{G}_{d,k}$ given in Fig. 1, while $\mathcal{S}(d, \infty)$ is generated by the labeled graph $\mathcal{G}_{d,\infty}$ shown in Fig. 2. Note that the edge labels for both these graphs come from the binary alphabet $\{0, 1\}$.

Before proceeding further, we would like to make a remark concerning the notation we shall use in this correspondence. While $\mathcal{S}(d, k)$ will be primarily used to denote the (d, k)-constrained system of finite sequences, we shall also occasionally use the same notation for the one-sided shift of infinite (d, k)-constrained sequences, or the shift space of bi-infinite (d, k)-constrained sequences. In such cases, it should be clear from the context which constrained system we mean to consider.

Manuscript received August 20, 2003; revised January 21, 2004. This work was supported by Applied Micro Circuits Corporation and by the Center for Magnetic Recording Research at the University of California, San Diego, and was performed while N. Kashyap was at the University of California, San Diego.

N. Kashyap is with the Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: nkashyap@mast.queensu.ca).

P. H. Siegel is with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA (e-mail: psiegel@ece.ucsd.edu).

Communicated by Ø. Ytrehus, Associate Editor for Coding Techniques.
Digital Object Identifier 10.1109/TIT.2004.828145

[1]Sometimes, we shall also find it necessary to consider the constrained system of *infinite* sequences $s_0 s_1 s_2 \ldots$, $s_i \in \Sigma$, of edge labels, or the constrained system of *bi-infinite* sequences $\ldots s_{-2} s_{-1} s_0 s_1 s_2 \ldots$, $s_i \in \Sigma$. In the terminology of symbolic dynamics (cf. [2]), a constrained system of bi-infinite sequences is called a *shift space*, and a constrained system of infinite sequences is called a *one-sided shift*.