

Universal Space–Time Codes From Demultiplexed Trellis Codes

Cenk Köse and Richard D. Wesel, *Senior Member, IEEE*

Abstract—In broadcast scenarios or in the absence of accurate channel probability distribution information, code design for consistent channel-by-channel performance, rather than average performance over a channel distribution, may be desirable. Root and Varaiya’s compound channel theorem for linear Gaussian channels promises the existence of universal codes that operate reliably whenever the channel mutual information (MI) is above the transmitted rate. This paper presents two-dimensional trellis codes that provide such universal performance over the compound linear vector Gaussian channel when demultiplexed over two, three, and four transmit antennas. The presented trellis codes, found by exhaustive search, guarantee consistent performance on every matrix channel that supports the information transmission rate with an MI gap that is similar to the capacity gap of a well-designed additive white Gaussian noise (AWGN)-specific code on the AWGN channel. As a result of their channel-by-channel consistency, the universal trellis codes presented here also deliver comparable, or, in some cases, superior, frame-error rate and bit-error rate performance under quasi-static Rayleigh fading, as compared with trellis codes of similar complexity that are designed specifically for the quasi-static Rayleigh-fading scenario.

Index Terms—Compound multiple-input multiple-output (MIMO) channel, multiple antennas, space–time systems, trellis codes, universal codes.

I. INTRODUCTION

THE use of multiple antennas at both transmitter and receiver is crucial in order to harvest the capacity of rich propagation environments. For example, when the path gains between transmit and receive antenna pairs are independent Gaussian random variables (RVs), Foschini [1] and Telatar [2] showed that outage capacity increases linearly with the number of transmit–receive antenna pairs. As is common in the current literature, we refer to a channel resulting from the use of multiple antennas as a space–time channel.

An instance of signal transmission over a space–time channel with N_t transmit antennas and N_r receive antennas is often modeled as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \quad (1)$$

Paper approved by A. Anastasopoulos, the Editor for Wireless Communications of the IEEE Communications Society. Manuscript received November 12, 2004; revised August 19, 2005. This work was supported in part by the University of California and Conexant Systems under MICRO Grant 04-100, and in part by the National Science Foundation under Grant CCR-0209110. This paper was presented in part at the IEEE GLOBECOM, Dallas, TX, 2004.

C. Köse is with Conexant Systems, Inc., San Diego, CA 92121 USA (e-mail: cenk.kose@conexant.com).

R. D. Wesel is with the Electrical Engineering Department, University of California at Los Angeles, Los Angeles, CA 90095-1594 USA (e-mail: wesel@ee.ucla.edu).

Digital Object Identifier 10.1109/TCOMM.2006.877967

where \mathbf{H} is an $N_r \times N_t$ complex matrix of path gains, \mathbf{x} is the complex input vector, \mathbf{y} is the complex output vector, and \mathbf{w} is the complex additive white Gaussian noise (AWGN) vector with variance N_0 per dimension. For vector channels of the form (1), signal design criteria for average error-probability performance were established in [3] and [4] for the case when the path gains are characterized by complex Gaussian RVs (Rayleigh fading). Since then, the design of coded space–time diversity schemes based on the average performance criterion for quasi-static Rayleigh fading has been an active research area [5]–[10].

While good average error performance is desirable, it does not guarantee consistently good “channel-by-channel” performance (i.e., universal performance). In fact, most of the code designs proposed for the average performance criterion degrade to uncoded performance under certain rank-deficient channels. Such channels have been of interest due to their deleterious effect on average channel capacity in certain propagation environments [11]–[13]. A simple line-of-sight channel is also rank-deficient when multiple transmit and receive antennas are employed. On the other hand, space–time block-code techniques such as Alamouti repetition [5] perform extremely well on singular channels, but lose performance on unitary channels. As an alternative to these extremes, good universal performance implies good performance on singular as well as unitary channels, guaranteeing good average performance irrespective of the quasi-static distribution.

The need for consistent performance across a family of channels arises in many communication scenarios, including:

- broadcast channels [14];
- frequency-hopped transmissions (e.g., in GSM, EDGE [15], [16]) where different codewords may experience different \mathbf{H} instances;
- multicarrier transmission scenarios [17], [18] where different frequency responses are different diagonal \mathbf{H} instances.

The information-theoretic foundation for the consistent channel-by-channel performance of universal codes is presented in Root and Varaiya’s compound channel-coding theorem [19], which indicates that a *single* code with rate R b/symbol can achieve reliable transmission of information over any linear Gaussian channel \mathbf{H} that induces R b/symbol of mutual information (MI), i.e., over any channel with $\text{MI}(\mathbf{H}, E_s) \geq R$, where

$$\text{MI}(\mathbf{H}, E_s) = \log_2 \det \left(\mathbf{I} + \frac{E_s}{N_0} \mathbf{H}\mathbf{H}^\dagger \right) \quad (2)$$

is the MI (in b/symbol) between the input vector \mathbf{x} and the output vector \mathbf{y} in (1), and E_s is the transmit energy per antenna

per symbol.¹ The implication of this result is that good error performance over a set of channels does not have to come at the expense of significant performance degradation over others. A universal code delivers similar error performance over all channels with the same MI. Consider a specific code that transmits R b/symbol over N transmit antennas. Let E_s^* denote the transmit energy per antenna that this specific code requires to achieve the target bit-error rate (BER) on the (static) channel \mathbf{H} . The excess MI (EMI) requirement of that specific code on this channel is the difference between the channel MI $\text{MI}(\mathbf{H}, E_s^*)$ and the rate R

$$\text{EMI}(\mathbf{H}) = \text{MI}(\mathbf{H}, E_s^*) - R \quad \text{bits.} \quad (3)$$

Let $\mathcal{H}_R = \{\mathbf{H}: \text{MI}(\mathbf{H}) \geq R\}$ denote the set of channels that comprise the most comprehensive compound $N \times N$ channel with capacity R b/symbol. The goal of universal R b/symbol code design is to minimize, over this $N \times N$ compound channel with capacity R , the worst-case EMI $\sup_{\mathbf{H} \in \mathcal{H}_R} \text{EMI}(\mathbf{H}) - R$ at fixed target BER [or frame-error rate (FER)], latency, and decoding complexity. To this end, our study [20] on universal codes concentrated on vector-labeled universal trellis codes for the 2×2 compound channel. This paper extends the ideas in [20] to propose universal space-time trellis codes (STTCs) formed by straightforward demultiplexing of two-dimensional (2-D) linear trellis codes over two, three, and four transmit antennas. These demultiplexed trellis codes have simpler maximum-likelihood (ML) decoding than general multidimensional (vector-labeled) trellis codes, while still providing excellent performance.

Section II summarizes our results on the worst-case minimum distance of a space-time code under linear transformations with equal MI, and derives a simple approximate criterion for universal behavior. This section also formulates the encoder rate, constellation size, and trellis complexity requirements for universal STTCs formed by demultiplexing a 2-D trellis code.

Section III presents several linear trellis codes, found by exhaustive search over their respective encoder classes. These codes provide universal performance when demultiplexed over two, three, and four transmit antennas and received with ML decoding. The performance variation of universal codes, as compared with other existing space-time codes over different channel instances, is illustrated via extensive simulations. Our discussion ends with simulation results showing that the average error performance of the proposed universal codes over quasi-static Rayleigh fading is comparable, and in some cases superior, to existing space-time codes with similar decoding complexity designed specifically for the average error-probability performance. A channel-by-channel look at the 2×2 and 3×3 quasi-static Rayleigh fading demonstrates why the proposed universal codes deliver excellent average Rayleigh fading error-probability performance, and why these universal codes provide a channel-by-channel reliability that codes designed only for good average performance may not provide. Section IV delivers the conclusions.

¹Assuming the input vector has a complex Gaussian distribution with covariance matrix $E_s \mathbf{I}_{N_t}$.

II. DESIGN GUIDELINES FOR UNIVERSAL STTCs

Let \mathbf{X} and $\hat{\mathbf{X}}$ be two different codewords of a space-time code \mathcal{C} for N_t transmit antennas, and let $\mathbf{E} = \hat{\mathbf{X}} - \mathbf{X}$ denote the codeword difference matrix. Under ML decoding, the probability that the decoder mistakes $\hat{\mathbf{X}}$ for \mathbf{X} conditioned on the perfect knowledge of the channel matrix \mathbf{H} at the receiver is given by

$$P(\mathbf{X} \rightarrow \hat{\mathbf{X}}|\mathbf{H}) = Q\left(\sqrt{\frac{d^2(\mathbf{E}|\mathbf{H})}{2N_0}}\right) \quad (4)$$

where $d^2(\mathbf{E}|\mathbf{H}) = \|\mathbf{H}\mathbf{E}\|^2 = \text{trace}(\mathbf{H}\mathbf{E}\mathbf{E}^\dagger\mathbf{H}^\dagger)$ is the squared Euclidean norm of the codeword difference matrix \mathbf{E} when transformed by the channel \mathbf{H} , and $Q(\cdot)$ is the standard Gaussian tail integral function. For a fixed channel \mathbf{H} , the minimum of the squared distance $d^2(\mathbf{E}|\mathbf{H})$ over all \mathbf{E} will be called the $d_{\min}^2(\mathbf{H})$ of the code.

With ML decoding, the universal performance of an STTC results from its ability to sustain good $d_{\min}^2(\mathbf{H})$ over a family \mathcal{H}_R of channel instances that support transmission rate R . The smallest value of $d_{\min}^2(\mathbf{H})$ over this compound channel is a function of the eigenvalues of the codeword difference matrices. For a given codeword difference matrix \mathbf{E} , let $\zeta = (\zeta_1, \dots, \zeta_{N_t})$ be the vector of eigenvalues of $\mathbf{E}\mathbf{E}^\dagger$ with the ordering

$$\zeta_1 \geq \zeta_2 \geq \dots \geq \zeta_{N_t}. \quad (5)$$

We will write $\mathbf{E} \cong \zeta$. The eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$, where \mathbf{H} is an $N_r \times N_t$ complex channel-gain matrix, will be denoted by $\lambda_i, i = 1, \dots, N_r$ with the ordering

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N_r}. \quad (6)$$

We will write $\mathbf{H} \cong \lambda$ where $\lambda = (\lambda_1, \dots, \lambda_{N_r})$. Throughout the paper, we will assume that $N_r \leq N_t$.

A. Worst-Case Distance Over the Compound Channel

Consider the $N_r \times N_t$ compound channel with capacity R b/symbol

$$\mathcal{H}_R(N_r, N_t) = \{\mathbf{H} \in \mathbb{C}^{N_r \times N_t}: \text{MI}(\mathbf{H}) > R\}. \quad (7)$$

For a particular code, we are interested in the minimum of the squared Euclidean norm, $d^2(\mathbf{E}|\mathbf{H})$, of a codeword difference matrix $\mathbf{E} \cong \zeta$ over all instances \mathbf{H} of the compound channel $\mathcal{H}_R(N_r, N_t)$. In order to find the minimum of $d^2(\mathbf{E}|\mathbf{H})$ over all $\mathbf{H} \in \mathcal{H}_R(N_r, N_t)$, we first consider the minimum over all $\mathbf{H} \cong \lambda$ for fixed λ . It is shown in [20] that for a given codeword difference matrix \mathbf{E} and a given set of channel eigenvalues $\lambda, d^2(\mathbf{E}|\mathbf{H}) = \|\mathbf{H}\mathbf{E}\|^2$ is minimized when the channel aligns its strongest modes with the weakest eigendirections of $\mathbf{E}\mathbf{E}^\dagger$, and vice versa. In particular, if $\mathbf{E} \cong \zeta$,

$$\inf\{d^2(\mathbf{E}|\mathbf{H}): \mathbf{H} \cong \lambda\} = \sum_{i=1}^{N_r} \lambda_i \zeta_{N_t+1-i} \quad (8)$$

therefore

$$\inf_{\mathbf{H} \in \mathcal{H}_R(N_r, N_t)} d^2(\mathbf{E}|\mathbf{H}) = \inf_{\boldsymbol{\lambda}: \sum_{i=1}^{N_r} \log_2(1+\lambda_i) \geq R} \sum_{i=1}^{N_r} \lambda_i \zeta_{N_t+1-i}. \quad (9)$$

The solution to the optimization problem

$$\begin{aligned} \min_{\mathbf{H}} \sum_{i=1}^{N_r} \lambda_i \zeta_{N_t+1-i} \\ \text{subject to } \text{MI}(\mathbf{H}) = \sum_{i=1}^{N_r} \log_2(1 + \lambda_i) \geq R \end{aligned}$$

has the waterpouring form with the worst-case channel

$$\mathbf{H}^o = \sum_{i=1}^{N_r} \sqrt{\lambda_i^o} \mathbf{u}_i \mathbf{n}_{N_t+1-i}^\dagger \quad (10)$$

where $\{\mathbf{u}_k\}_{k=1, \dots, N_r}$ is any basis in $\mathbb{C}^{N_r \times 1}$, and \mathbf{n}_i is the eigenvector of $\mathbf{E}\mathbf{E}^\dagger$ that corresponds to ζ_i . The worst-case channel eigenvalues are given by

$$\lambda_i^o = \begin{cases} 2^{R/N^*} \frac{G_{N^*}(\zeta)}{\zeta_{N_t+1-i}} - 1, & i = 1, \dots, N^* \\ 0, & i = N^* + 1, \dots, N_r \end{cases} \quad (11)$$

where $N^* = \max\{k: 2^{R/k} G_k(\zeta) > \zeta_{N_t+1-k}, k \leq N_t\}$ is the number of nonzero eigendirections that the worst-case channel uses and

$$G_k(\zeta) = \left(\prod_{i=N_t-k+1}^{N_t} \zeta_i \right)^{1/k}, \quad k = 1, \dots, N_t \quad (12)$$

is the geometric mean of the smallest k eigenvalues of $\mathbf{E}\mathbf{E}^\dagger$.

This exact solution, however, does not yield a simple criterion for universality except for $N_r = 1$, in which case, the worst-case channel is given by $\mathbf{h}^o = (2^R - 1)^{1/2} \mathbf{n}_{N_t}^\dagger$, where \mathbf{n}_{N_t} is the eigenvector corresponding to the weakest eigenvector of $\mathbf{E}\mathbf{E}^\dagger$.

The purpose of universal coding is to deliver R b/symbol with consistently good error probability across all instances of the $N_t \times N_t$ compound channel $\mathcal{H}_R(N_t, N_t)$. This would guarantee consistent performance for any $N_r \leq N_t$. In fact, for any $N_r \leq N_t$, $\mathcal{H}_R(N_r, N_t)$ is equivalent to the subset of $\mathcal{H}_R(N_t, N_t)$ consisting of square matrices \mathbf{G} of rank less than or equal to N_r , i.e., $\mathcal{H}_R(N_r, N_t) \equiv \{\mathbf{G} \in \mathcal{H}_R(N_t, N_t): \text{rank}(\mathbf{G}) \leq N_r\}$, where \equiv is in the sense that for $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$, $\mathbf{G} \in \mathbb{C}^{N_t \times N_t}$, $\mathbf{H} \equiv \mathbf{G}$ if and only if $\mathbf{U}\mathbf{G} = [\mathbf{H}; \mathbf{0}]$ for some unitary \mathbf{U} . For example, the subset of singular channels \mathbf{G} in $\mathcal{H}_R(N_t, N_t)$ is equivalent to the compound channel $\mathcal{H}_R(1, N_t)$ with $N_r = 1$. The equivalence $\mathbf{H} \equiv \mathbf{G}$ implies that a space-time code would have identical error-probability performance under \mathbf{G} and \mathbf{H} with ML decoding.

An approximate criterion for universality over $\mathcal{H}_R(N, N)$ is obtained by bounding the worst-case distance over the compound channel. First, an upper bound on the worst-case min-

imum distance on the compound channel comes from the minimum distance over the equal-eigenvalue ($\lambda_i = 2^{R/N} - 1, i = 1, \dots, N$) channels. For a codeword difference \mathbf{E}

$$\begin{aligned} \lambda: \sum_{i=1}^N \log_2(1+\lambda_i) \geq R \sum_{i=1}^N \lambda_i \zeta_{N+1-i} (\mathbf{E}\mathbf{E}^\dagger) \\ \leq \left(2^{\frac{R}{N}} - 1\right) \sum_{i=1}^N \zeta_{N+1-i} (\mathbf{E}\mathbf{E}^\dagger) \end{aligned} \quad (13)$$

$$= \left(2^{\frac{R}{N}} - 1\right) \Delta_E(\mathbf{E}) \quad (14)$$

where $\Delta_E(\mathbf{E}) = \text{trace}(\mathbf{E}\mathbf{E}^\dagger)$ is the squared Euclidean norm of \mathbf{E} .

Now, a lower bound is obtained as

$$\begin{aligned} \sum_{i=1}^N \lambda_i \zeta_{N+1-i} (\mathbf{E}\mathbf{E}^\dagger) \\ = \left(\sum_{j=1}^N \zeta_{N+1-j} \right) \left(\sum_{i=1}^N \frac{\zeta_{N+1-i}}{\sum_{j=1}^N \zeta_{N+1-j}} (1 + \lambda_i) - 1 \right) \\ \stackrel{(a)}{\geq} \left(\sum_{j=1}^N \zeta_{N+1-j} \right) \left(2^{\sum_{i=1}^N \frac{\zeta_{N+1-i}}{\sum_{j=1}^N \zeta_{N+1-j}} \log_2(1+\lambda_i)} - 1 \right) \\ \stackrel{(b)}{\geq} \left(\sum_{j=1}^N \zeta_{N+1-j} \right) \left(2^{\left[\frac{R \sum_{j=1}^N \zeta_{N+1-j}}{\sum_{j=1}^N \zeta_{N+1-j}} \right]} - 1 \right) \\ \stackrel{(c)}{\geq} \Delta_E(\mathbf{E}) \left(2^{R \frac{\zeta_N}{\Delta_E(\mathbf{E})}} - 1 \right). \end{aligned} \quad (15)$$

In the above set of equations, (a) follows from the concavity of the logarithm function. To see (b), let $r_i = \log_2(1 + \lambda_i)$ and $\theta_i = \zeta_i / \sum_{j=1}^N \zeta_j$. Noting that $r_i \geq r_{i+1} \geq 0$, for the minimum of $\sum_{i=1}^N \theta_{N+1-i} r_i$ subject to $\sum_{i=1}^N r_i \geq R$, one can take $\sum_{i=1}^N r_i = R$. An equivalent problem is then to maximize $R - \sum_{i=1}^N \theta_{N+1-i} r_i = \sum_{i=1}^N (1 - \theta_{N+1-i}) r_i = \sum_{i=1}^N \alpha_i r_i$ subject to $\sum_{i=1}^N r_i = R$, where $\alpha_i = 1 - \theta_{N+1-i}$. Since $\alpha_i \geq \alpha_{i+1}$, the maximum occurs at $r_1^* = R$ and $r_i^* = 0$ for $i > 1$, i.e., at $\lambda_1^* = 2^R - 1$ and $\lambda_i^* = 0$ for $i > 1$. Finally, (c) uses the shorthand notation $\Delta_E(\mathbf{E})$ for $\text{trace}(\mathbf{E}\mathbf{E}^\dagger) = \sum_{i=1}^N \zeta_i (\mathbf{E}\mathbf{E}^\dagger)$. The following lemma summarizes these results.

Lemma 1: The worst-case minimum distance of a space-time code over the compound $N \times N$ channel of capacity R b/symbol is bounded as

$$\begin{aligned} (2^{\frac{R}{N}} - 1) \Delta_E^* &\geq \inf_{\mathbf{H} \in \mathcal{H}_R(N, N)} d_{\min}^2(\mathbf{H}) \\ &\geq \min_{\mathbf{E}} \left(2^{R \frac{\zeta_N}{\Delta_E(\mathbf{E})}} - 1 \right) \Delta_E(\mathbf{E}) \end{aligned}$$

where $\Delta_E^* = \min_{\mathbf{E}} \text{trace}(\mathbf{E}\mathbf{E}^\dagger)$ is the minimum squared Euclidean distance of the code. The first inequality of *Lemma 1* implies that a universal code should have good minimum Euclidean distance. The second inequality of the lemma leads us to choose, among good minimum Euclidean-distance codes, those codes with high minimum eigenvalue $\zeta_{N_t}^* = \min_{\mathbf{E}} \zeta_{N_t}(\mathbf{E}\mathbf{E}^\dagger)$.

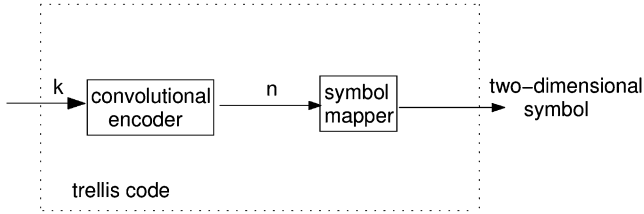


Fig. 1. A k b/symbol 2-D trellis code.

Ultimately, the universal performance of a code over the compound channel will be measured by its EMI requirement. Nevertheless, this criterion provides us with a basic rule to prune the search for universal STTCs.

B. Demultiplexing a 2-D Trellis Code

The requirements for design of universal vector-labeled STTCs for two transmit antennas were considered in [20]. In this paper, we focus on STTCs generated by straightforward demultiplexing of a 2-D trellis code over the transmit antennas. This approach results in more manageable code searches. More importantly, the resulting STTCs have the same ML decoding complexity as the demultiplexed code over a single-transmit antenna channel [6]. Consider a rate- k/n convolutional encoder with memory ν . For R b/symbol transmission over N_t transmit antennas using a 2^m -phase-shift keying/quadrature amplitude modulation (PSK/QAM) constellation, we use this encoder $l = R/k$ times (assume for simplicity that k divides R evenly). Let $(b_0, \dots, b_{n-1}), \dots, (b_{(l-1)n}, \dots, b_{ln-1})$ denote the ln codeword bits that the binary encoder would produce for lk successive input bits. If $ln = N_t m$, then we map $(b_{(i-1)m}, \dots, b_{im-1})$ onto the 2^m -PSK/QAM constellation for the i th transmit antenna, $i = 1, \dots, N_t$. If $ln > N_t m$, we puncture $ln - N_t m$ out of Rn bits, and group the remaining $N_t m$ bits similarly, keeping the index order. Fig. 1 illustrates a 2-D trellis code.

C. Constellation Size and Trellis Complexity Requirements for Demultiplexed Universal Space-Time Codes

Universal performance over the compound channel requires the following design rules for demultiplexed 2-D trellis codes.

- **Constellation size:** Universal codes are designed to maintain redundancy over any instance of the compound channel. Therefore, the constellation size should be large enough to transmit R b/symbol redundantly, even when the channel matrix has $N_t - 1$ zero rows.² Thus, the constellation size should satisfy $m > R$. For example, for 2 b/symbol transmission over $N_t = 2$ antennas, we propose rate-1/3 convolutional encoders mapping 8-PSK constellations.
- **Trellis complexity:** The effective code length (ECL) is the smallest symbol-wise Hamming weight of an error event. For a k -b/symbol linear trellis code, $\text{ECL} \leq \lfloor \nu/k \rfloor + 1$, where ν is the memory of the encoder [21]. In order to have $\zeta_{N_t}^* > 0$, the shortest error event of the single-dimensional code should be long enough to occupy at least

N_t channel symbols when demultiplexed over N_t transmit antennas. Therefore, $\lfloor \nu/k \rfloor \geq N_t(N_t - 1)$ [22]. This was noted earlier in [22]. The trellis complexity requirements can be relaxed by sacrificing one or more levels of transmit diversity. In general, for r levels of transmit diversity, the necessary trellis complexity obeys

$$\lfloor \nu/k \rfloor \geq N_t(r - 1). \quad (16)$$

With $r < N_t$ levels of transmit diversity, there is always a singular channel that supports the transmission rate, under which two codewords cannot be reliably (with error probability less than 1/2) distinguished at the receiver. However, a diversity- r code with good ζ_r^* can provide universal performance over all channels that establish at least $N_t - r + 1$ equally strong spatial eigenmodes. Moreover, as will be illustrated in Section III, missing levels of transmit diversity can be restored by an outer code.

III. UNIVERSAL SPACE-TIME CODES FROM 2-D TRELLIS CODES

The search for universal codes is complicated by the nonadditive nature of error-event eigenvalues over trellis branches. In previous work [20], for $N_t = 2$, a finite sampling of channels was used to approximate the worst-case minimum eigenvalue of the code over multiple forward trellis searches. In this paper, since we want to consider $N_t = 3$ and $N_t = 4$, as well as $N_t = 2$, a small set of test channels (including the unitary and several rank-deficient diagonal channels) are used in this paper to prune the set of codes considered by the stack-based algorithm of [23] to determine codes with the best worst-case eigenvalues.

Besides the worst-case minimum eigenvalue $\zeta_{N_t}^*$, three other code parameters will be useful for interpreting the performance of space-time codes in different scenarios. These parameters are the minimum squared Euclidean distance of the code, $\Delta_E^* = \min_{\mathbf{E}} \sum_{i=1}^{N_r} \zeta_i(\mathbf{E}\mathbf{E}^\dagger)$, the diversity order of the code, $\Delta_H^* = \min_{\mathbf{E}} \text{rank}(\mathbf{E})$, and the minimum product distance $\Delta_P^* = \min_{\mathbf{E}} \prod_{i=1}^{\Delta_H^*} \zeta_i(\mathbf{E}\mathbf{E}^\dagger)$. The worst-case i th eigenvalue over all $\mathbf{E}\mathbf{E}^\dagger$ is denoted by $\zeta_i^* = \min_{\mathbf{E}} \zeta_i(\mathbf{E}\mathbf{E}^\dagger)$.

All codes presented in this section are found by exhaustive search over their class of encoders to maximize the worst-case minimum-eigenvalue ζ_r^* under a transmit diversity constraint $\Delta_H^* = r$, while sacrificing no more than 20% of the maximum squared Euclidean distance Δ_E^* achievable within the same class. In cases where $\zeta_r^* > 0$ cannot be attained without less than 20% loss in squared Euclidean distance, the search was carried out without the squared Euclidean distance constraint. Ties are resolved by looking at the average bit multiplicity of worst-case eigenvalues.

A. Universal Codes for $N_t = 2$ Transmit Antennas

Table I lists demultiplexed universal trellis codes and comparison codes for $R = 1, 2, 3$ b/symbol transmission over two, three, and four transmit antennas. We consider the $R = 2$ b/symbol case in detail.

Example 1: $N_t = 2, R = 2$ b/symbol: Table I lists rate-1/3 + 8 PSK trellis codes for 16, 32, and 64 states (codes

²The channel matrix cannot be the zero matrix which has zero MI.

TABLE I

R B/SYMBOL TRELLIS CODES FOR N_t TRANSMIT ANTENNAS. A RATE- k/n CONVOLUTIONAL ENCODER OUTPUTS $N R$ CODE BITS WHICH ARE MAPPED ON $N_t \times 2^m$ -PSK/QAM. 16 QAM, QPSK: GRAY LABELING, 4PSK: NATURAL LABELING. 8 PSK: GRAY LABELING 0,2,3,1,5,7,6,4 IN OCTAL AROUND THE CIRCLE

Code	N_t	R	k/n	ν	g_i	Const.	Δ_H^*	$\eta_{\Delta_H}^*$	Δ_P^*	Δ_E^*
Multiplexed universal codes found by exhaustive search										
1	2	1	1/4	4	03 22 36 04	QPSK	2	$\eta_2^* = 5.53$	80	20
2	2	2	1/3	4	31 05 35	8PSK	2	$\eta_2^* = 0.67$	16.32	12.6
3	2	2	1/3	5	71 31 61	8PSK	2	$\eta_2^* = 0.89$	≤ 25.02	13.4
4	2	2	1/3	6	155 056 145	8PSK	2	$\eta_2^* = 0.70$	17.7	17.2
5	2	3	1/3	3	06 16 13	16QAM	2	$\eta_2^* = 0.40$	1.6	4
6	3	1	1/6	3	15 03 07 05 04 11	QPSK	3	$\eta_3^* = 2.44$	248	26
7	3	2	1/5	4	31 06 02 31 02 ¹	8PSK	3	$\eta_3^* = 0.14$	4.69	7.57
8	3	2	1/3	5	62 55 47	QPSK	3	$\eta_3^* = 0.24$	32	22
9	3	3	1/2	5	75 62	QPSK	2	$\eta_2^* = 2.0$	15.5	12.0
10	4	2	1/4	4	33 23 26 06	QPSK	2	$\eta_2^* = 4.60$	78.6	24
11	4	2	1/4	5	75 71 67 53	QPSK	3	$\eta_3^* = 0.29$	49.1	36
Multiplexed maximal- Δ_E^* codes										
12	2	2	1/3	4	31 27	QPSK	2	$\eta_2^* = 1.25$	16	14
13	2	2	1/2	5	53 75	QPSK	2	$\eta_2^* = 0.46$	16	8
14	2	2	1/2	6	155 117	QPSK	1	$\eta_1^* \leq 10$	18	24
15	2	1	1/4	4	25 27 33 37	QPSK	2	$\eta_2^* = 5.53$	24	24
16	3	3	1/2	5	65 57	QPSK	2	$\eta_2^* \leq 1.43$	19.7	16
17	4	2	1/4	4	25 27 33 37	QPSK	2	$\eta_2^* = 2.62$	74.14	30
Universal STTCM [20]										
18	2	2	2/6	5	[20]	8PSK	2	$\eta_2^* = 0.72$	13.0	12.0
19	2	2	2/6	6	[20]	8PSK	2	$\eta_2^* = 0.80$	13.8	16.0
Trellis codes for periodic-erasures [21]										
20	2	2	1/3	6	173 062 115	8PSK	2	$\eta_2^* = 0.56$	32.3	17.6
21	3	3	1/4	6	117 155-145 137	16QAM	2	$\eta_2^* = 0.28$	5.03	14.4
STTCM for average Rayleigh performance										
22	2	2	2/4	4	from [6]	4PSK	2	$\eta_2^* = 0.52$	8	16
23	2	2	2/4	5	from [6]	4PSK	2	$\eta_2^* = 1.44$	32	16
24	2	2	2/4	6	from [6]	4PSK	2	$\eta_2^* = 1.04$	28	18
25	3	2	2/6	4	from [6]	4PSK	2	$\eta_2^* = 0.91$	13	12
26	3	2	2/6	5	from [6]	4PSK	3	$\eta_3^* = 0.156$	8	14
27	4	2	2/8	5	from [8]	4PSK	3	$\eta_3^* = 0.144$	34.6	36
TCM + Alamouti STB										
28	2	2	2/3	6	from [24],[5]	4PSK	2	$\eta_2^* = 3$	6	9
Super-orthogonal code										
28	2	2	2/4	5	from [9]	4PSK	2	$\eta_2^* = 0.4$	8	16

#2, #3, and #4, respectively). These codes deliver universal performance over the 2×2 compound channel when they are demultiplexed over two transmit antennas. Fig. 2(a) displays the simulated BER performance of code #4 as well as two other 64-state transmit-diversity schemes over the 2×2 compound channel as a function of EMI. With perfect channel state information at the receiver, the performance of the ML decoder on unitary channels \mathbf{H} does not depend on the particular instance of \mathbf{H} . However, on singular channels, the performance of the code varies with the direction of the nonvanishing eigenvector of the channel matrix, except for the case of an orthogonal transmit design, e.g., Alamouti repetition [5]. For orthogonal transmit designs, the ML performance of the code depends only on the sum of the eigenvalues of the channel matrix.

At $\text{BER} = 10^{-5}$, code #4 requires no more than 0.88 b of EMI per transmit antenna on singular channels, and requires 0.93 b of EMI per transmit antenna on unitary channels. The rate-1/2 64-state maximal-free-distance convolutional code [25] with quaternary (Q)PSK modulation (code #14 in Table I) requires 0.84 b of EMI on unitary channels at $\text{BER} = 10^{-5}$.

At $\text{BER} = 10^{-5}$, code #4 handles *every* instance of the 2×2 compound channel within 0.1 b of EMI per transmit antenna of the best rate-1/2+ QPSK AWGN trellis code of similar complexity on the AWGN channel. The performance of the code over singular 2×2 channels *is* the performance of the code over 2×1 channels.

Another scheme that delivers 2 b/symbol over two transmit antennas consists of a good AWGN-trellis-coded modulation (TCM) followed by Alamouti repetition [5] (Code #28 in Table I). The 64-state rate-2/3 Ungerboeck TCM [24] achieves $\text{BER} = 10^{-5}$ at $\text{SNR} = 8.8$ dB on the AWGN channel. On the compound 2×2 channel, the EMI requirement of this scheme is a function of the sum of the channel eigenvalues [20]. On singular channels, this concatenated scheme requires only 0.55 b of EMI per antenna at $\text{BER} = 10^{-5}$, whereas on unitary channels, the EMI is requirement is 1.26 b per transmit antenna. Among the codes examined, code #4 has the most consistent channel-by-channel performance, and the space-time block (STB) + TCM (code #28) has the least consistent channel-by-channel performance.

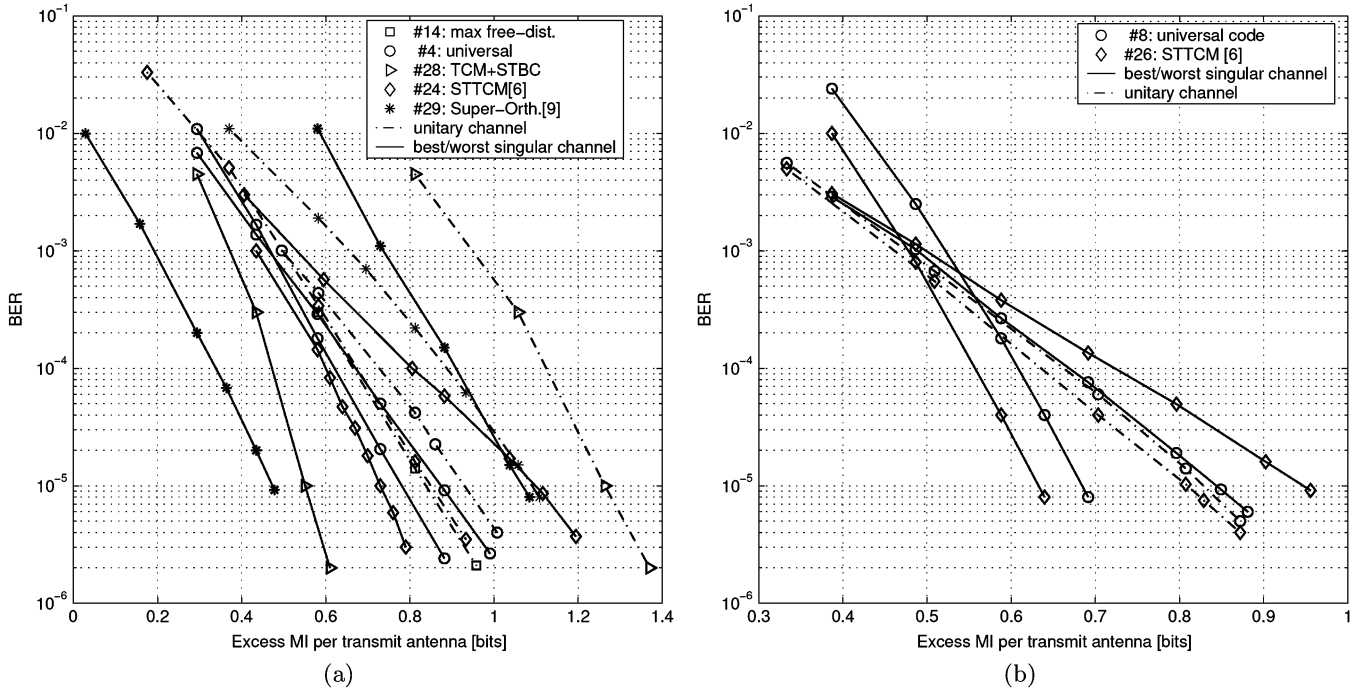


Fig. 2. (a) Channel-by-channel performance of the universal 64-state rate-1/3 + 8 PSK code (#4) over the 2×2 compound channel. Best- and worst-case (for $\text{BER} = 10^{-5}$) singular channels are identified via extensive simulation. For comparison, the AWGN performance of the 64-state maximal free-distance rate-1/2 convolutional encoder (with QPSK) (code #14), as well as the compound channel performances of the 64-state rate-2/3 + 8 PSK Ungerboeck-TCM + Alamouti block signaling (code #28), the 64-state STTCM (code #24) of Aktas and the 32-state super-orthogonal code (#29) of Siwamogsatham are provided. (b) Channel-by-channel performance of the universal $R = 2$ b/symbol QPSK code (#8) over the 3×3 compound channel. Best- and worst-case (for $\text{BER} = 10^{-5}$) singular channels are identified via extensive simulation. For comparison, the performance of the 32-state space-time code of [6] (code #26) is provided. Code #26 was proposed for average Rayleigh-fading performance. Each frame consists of 127 data symbols and 3 symbols for trellis termination.

B. Universal Codes for $N_t = 3$ and $N_t = 4$ Transmit Antennas

Example 2: $N_t = 3, R = 2$ b/symbol: For $R = 2$ b/symbol transmission over $N_t = 3$ transmit antennas, our search focused on rate-1/5 convolutional encoders. Two information bits produce ten codeword bits, one of which is punctured, and the remaining nine are used to map 3×8 PSK constellations. An exhaustive search over 16-state encoders and all 1-out-of-10 puncturing patterns resulted in code #7 in Table I (b_9 is punctured). Code #7 has $\Delta_{H,\min} = 3$ and $\zeta_3^* = 0.14$, whereas the 16-state 4-PSK STTC of [6] (code #25 in Table I) has $\Delta_{H,\min} = 2$ and $\zeta_2^* = 0.91$. Code #7 achieves $\text{BER} = 10^{-5}$ with a worst-case EMI of 1.3 b/antenna on the singular channels. Code #25 cannot perform reliably under certain singular channels, due to its limited diversity.

The Δ_E^* -constrained optimal- ζ_3^* search over 32-state encoders and all puncturing patterns resulted in an encoder that uses only four out of the eight constellation points. This encoder may simply be represented as a 32-state rate-1/3 convolutional code driving QPSK (code #8).

Fig. 2(b) displays the BER performances of code #8 and the 32-state code of [6] (code #26 in Table I) on the compound 3×3 channel. Code #8 achieves $\text{BER} = 10^{-5}$ with a worst-case EMI of 0.85 b/antenna, code #26 requires about 0.95 b of EMI in the worst case. Code #8 provides more consistent performance than code #26. Section III-D compares the average error performance of these codes under quasi-static Rayleigh fading.

At $R = 3$ b/symbol transmission over $N_t = 3$ antennas, a 3-b/symbol maximal- ζ_2^* scheme ($\Delta_H^* = 2$) was found as a result of an exhaustive search over 32-state rate-1/2 + QPSK codes (code #9). The parameters of the 32-state maximal-free-distance rate-1/2 encoder + QPSK code [25] (code #16) is provided for comparison. Code #9 is not universal over the 3×3 compound channel because it does not have full transmit diversity. However, it exhibits consistent performance over all channels with strong λ_2 presence.

C. Universal Codes by Concatenation of Algebraic Block Codes and Rank-Deficient TCMs With Good Eigenvalue Spread

When the required trellis complexity for full diversity exceeds practical limits, the concatenation of a rank-deficient trellis code with an outer code may restore the diversity order at the expense of reduced transmission rates.

Example 3: (31,21) Binary Bose-Chaudhuri-Hocquengem (BCH) Code + Code #21: We consider the (31,21) binary BCH code concatenated with the 64-state rate-1/4 + 16 QAM scheme (code #15) to deliver $R = 2.03$ b/symbol over three transmit antennas. The decoding is achieved by a Max-Log-*a posteriori* probability (APP) Bahl-Cocke-Jelinek-Raviv (BCJR) [26] followed by a single iteration Max-Log-APP decoding of the BCH code. The interleaver is a 372-b-long block interleaver hosting 12 codewords from the outer code in a 124-data-symbol packet with trellis termination. At $\text{BER} = 10^{-5}$, this scheme requires no more than 1 b of EMI on the collection of 40 channels we simulated.

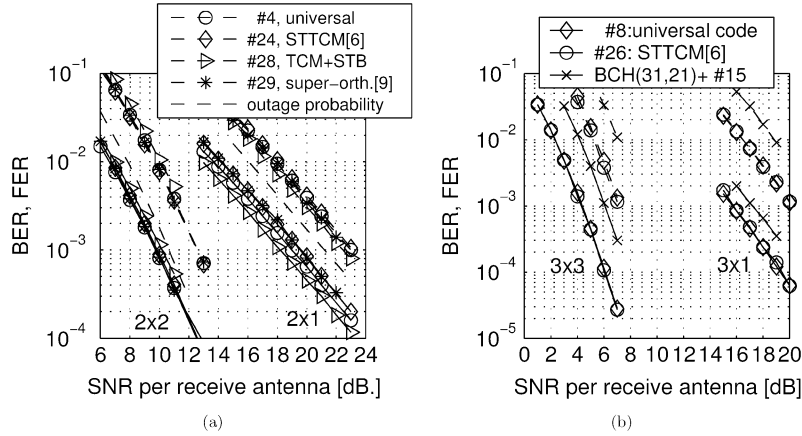


Fig. 3. (a) BER and FER performance of the 64-state universal trellis code (code #4), as compared with the 64-state code of [6], over the quasi-static Rayleigh fading channel, $N_t = 2$. Each frame consists of 127 data symbols and 3 symbols for trellis termination. ML decoding on the entire frame. (b) BER and FER performances of the 32-state universal trellis code (code #8), as compared with the 32-state code of [6] (code #26), over the 3×1 and 3×3 quasi-static Rayleigh fading channels. Each frame consists of 127 data symbols and 3 symbols for trellis termination. ML decoding on the entire frame.

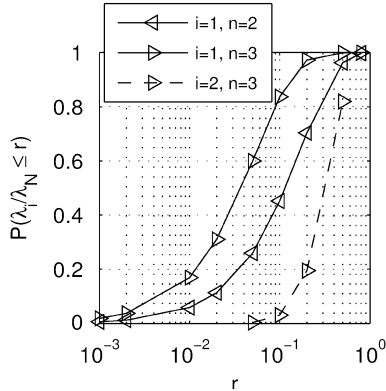


Fig. 4 Distribution of the ratio of eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$, where \mathbf{H} is an $N \times N$ matrix of independent and identically distributed complex Gaussian RVs. Top to bottom: Cumulative distribution function (CDF) of λ_3/λ_1 in 3×3 Rayleigh fading, CDF of λ_2/λ_1 in 2×2 Rayleigh fading, CDF of λ_2/λ_1 in 3×3 Rayleigh fading.

D. Universal STTCs Under Quasi-Static Rayleigh Fading

Universal STTCs deliver good average error performance under quasi-static Rayleigh fading, as long as the quasi-static period is longer than several traceback depths of the codes. Fig. 3(a) compares the FER and BER performances of code #4 and the 64-state code of [6] (code #24) over 2×1 and 2×2 quasi-static Rayleigh fading. The universal code has slightly better FER and BER over the SNR range displayed. Fig. 3(b) compares the FER and BER performances of code #8 and the 32-state code of [6] (code #25) over 3×1 and 3×3 quasi-static Rayleigh fading. The two codes have similar BER and FER performances over the range of SNRs displayed, with code #8 performing slightly worse in FER for the 3×3 scenario.

Fig. 4 displays the probability distribution of the ratio λ_N/λ_1 in $N \times N$ Rayleigh fading for $N = 2, 3$. For two transmit antennas, the probability that the eigenvalues λ_1, λ_2 are more than 10 dB apart is 0.45. For $N = 3$, this probability is 0.83. Universal code design which takes into account the performance over singular channels (through high ζ_N) results in good codes

for the average Rayleigh fading performance, while also delivering consistent channel-by-channel performance not provided by codes designed only for Rayleigh fading.

IV. CONCLUSIONS

Wireless communication with multiple transmit antennas exposes the transmitted signals to a rich variety of channels. When accurate statistical characterization of the path gains is not possible, or when the code is used in a broadcast scenario, universal code design aims to deliver consistently good error-probability performance on any instance of the channel. The compound channel-coding theorem guarantees the existence of universal codes with a consistently close proximity to capacity on any instance of a continuum of space-time matrix channels. For consistent channel-by-channel performance across the rank-unconstrained $N \times N$ compound channel with ML decoding, a universal space-time code should have good minimum Euclidean distance and a good minimum eigenvalue over all codeword differences. Perhaps more importantly, universal codes should obey binary encoder rate, ECL, and constellation size requirements. Space-time codes designed for the average error performance often ignore some of these requirements.

In this paper, we presented universal STTCs formed by demultiplexing 2-D linear trellis codes over two, three, and four transmit antennas. The straightforward demultiplexing of a trellis code ensures that the ML-decoding complexity is similar to that of a single-transmit-antenna scenario. For two and three transmit antennas, trellis codes that achieve full transmit diversity with a large smallest eigenvalue and good Euclidean distance within their class were presented. For three and four transmit antennas, trellis codes that achieve two and three levels of transmit diversity with maximum smallest nonzero eigenvalue and good Euclidean distance were presented. The best- and worst-case performances over the compound channel are identified via extensive simulation. Due to their channel-by-channel consistency, these universal codes deliver excellent average error-probability performance, regardless of the statistical distribution of the channel instances.

ACKNOWLEDGMENT

The authors thank W. Zhu and M. P. Fitz of the UCLA EE Department for bringing (10) to their attention.

REFERENCES

- [1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 3, pp. 41–59, Autumn, 1996.
- [2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Euro. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.
- [3] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *Proc. IEEE Veh. Technol. Conf.*, 1996, pp. 136–140.
- [4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.
- [5] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Nov. 1998.
- [6] D. Aktas, H. El Gamal, and M. P. Fitz, "On the design and maximum-likelihood decoding of space-time trellis codes," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 854–859, Jun. 2003.
- [7] M. Tao and R. S. Cheng, "Improved design criteria and new trellis codes for space-time coded modulation in slow flat fading channels," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 313–315, Jul. 2001.
- [8] Z. Chen, B. S. Vucetic, J. Yuan, and K. L. Luo, "Space-time trellis codes for 4-PSK with three and four transmit antennas in quasi-static flat fading channels," *IEEE Commun. Lett.*, vol. 6, no. 2, pp. 67–69, Feb. 2002.
- [9] S. Siwamogsatham and M. P. Fitz, "Improved high-rate space-time codes via orthogonality and set partitioning," in *Proc. IEEE WCNC*, Mar. 2002, vol. 1, pp. 264–270.
- [10] Q. Yan and R. S. Blum, "Improved space-time convolutional codes for quasi-static slow fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 563–571, Oct. 2002.
- [11] D. Chizhik, G. J. Foschini, and R. A. Valenzuela, "Capacities of multi-element transmit and receive antennas: Correlations and keyholes," *Electron. Lett.*, vol. 36, no. 13, pp. 1099–1100, Jun. 2000.
- [12] H. Shin and J. H. Lee, "Capacities of multi-antenna fading channels: Spatial fading correlation, double scattering and keyhole," *Electron. Lett.*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.
- [13] P. Almers, F. Tufvesson, and A. F. Molisch, "Measurement of keyhole effect in a wireless multiple-input multiple-output (MIMO) channel," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 373–375, Aug. 2003.
- [14] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
- [15] H. Viswanathan and J. Balakrishnan, "Space-time signaling for high data rates in EDGE," *IEEE Trans. Veh. Technol.*, vol. 51, no. 6, pp. 1522–1533, Nov. 2002.
- [16] M. Coupechoux and V. Braun, "Space-time coding for the EDGE mobile radio system," in *Proc. IEEE Int. Conf. Pers. Wireless Commun.*, Dec. 2000, pp. 28–32.
- [17] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wide-band channels," in *Proc. IEEE Veh. Technol. Conf.*, May 1998, vol. 3, pp. 2232–2236.
- [18] Y. G. Li, J. H. Winters, and N. R. Sollenberger, "MIMO-OFDM for wireless communications: Signal detection with enhanced channel estimation," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1471–1477, Sep. 2002.
- [19] W. L. Root and P. P. Varaiya, "Capacity of classes of Gaussian channels," *SIAM J. Appl. Math.*, vol. 16, no. 6, pp. 1350–1393, 1968.
- [20] C. Köse and R. D. Wesel, "Universal space-time trellis codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2717–2727, Oct. 2003.
- [21] R. Wesel, X. Liu, and W. Shi, "Trellis codes for periodic erasures," *IEEE Trans. Commun.*, vol. 48, no. 6, pp. 938–947, Jun. 2000.
- [22] J. Grimm, M. P. Fitz, and J. V. Krogmeier, "Further results on space-time coding for Rayleigh fading," in *Proc. 36th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 1998, pp. 391–400.
- [23] D. K. Aktas and M. P. Fitz, "Computing the distance spectrum of space-time trellis codes," in *Proc. IEEE WCNC*, 2000, pp. 51–55.
- [24] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 55–67, Jan. 1982.
- [25] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [26] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.

Cenk Köse received the Ph.D. degree in electrical engineering from University of California, Los Angeles, in 2004.

He is currently with Conexant Systems, San Diego, CA. His research interests include coding and signal processing for communications with applications to wireless systems and MIMO technology.



Richard D. Wesel (S'91–M'96–SM'01) was born in Marietta, OH, in 1966. He received the S.B. and S.M. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, both in 1989, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1996.

From 1989 to 1991, he was a Member of Technical Staff with AT&T Bell Laboratories. His work at AT&T resulted in two patents. Since 1996, he has been with the University of California, Los Angeles

(UCLA), as an Assistant Professor with the Electrical Engineering Department. His research interests are in the area of communication theory, with particular interest in the topics of channel coding and distributed communication.

Dr. Wesel has received a National Science Foundation CAREER Award, an Okawa Foundation award, and the TRW Excellence in Teaching Award from the UCLA Henry Samueli School of Engineering and Applied Science. He has been an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS in the area of coding and communication theory since 1999.