On Space-Time Code Design

D. Mihai Ionescu, *Member, IEEE*

Abstract— **It is shown that the separation between space-time codematrices can be described in terms of a metric of Euclidean type, which is defined via the singular values of difference codematrices, and arises naturally from a minimization of the pairwise error probability. Essentially, the distance between complex space-time codematrices is the Euclidean distance between the respective—demultiplexed and concatenated—transmit antenna streams, expressed in terms of the structure inherent to the multiple antenna arrangement. It is further shown that the determinant criterion can be strengthened, in a manner that not only suggests an optimum space-time codematrix structure, but also outlines the central role played by the Euclidean distance in quasistatic fading. Theorem 5—which claims that in order to optimize the product distance one must optimize the Euclidean distance—establishes a close inter-dependence between product and Euclidean distances; it thereby links the performance determining factors in quasistatic and independent fading, and rigorously establishes the relevance of combining space-time coding and modulation in fading environments. A multidimensional space-time constellation for two transmit antennas, and its coset partitioning—based on traces of differences between constellation matrices—are described. Example codes constitute the first reporting of a space-time coded modulation scheme for fading channels, whereby a space-time constellation is partitioned in cosets.**

Keywords—**Diversity, space-time coded modulation, space-time coset codes, equal eigenvalue criterion, product distance.**

I. Introduction

Space-time coding constitutes a means for combating the signal fading inherent to wireless communications, and is aimed in its most general form at enhancing both the coding gain and the level of diversity—where the latter includes transmit diversity. Motivated by the information-theoretic results of Winters [29], Foschini and Gans [12], and Telatar [27], early ideas on transmit diversity schemes—wherein, e.g., a transmit antenna sends a delayed replica of another transmit antenna's signal—have been refined by the works of Guey *et al.* [16], Tarokh *et al.* [26]. Reviews on the evolution of these ideas can be found in the introductory sections of [26], [16], [17]. Since it is advantageous to separate the problem of combating fades from that of channel equalization, the criteria for designing space-time codes are usually derived in the context of narrowband modulation and frequency nonselective fading; this isolates transmit diversity from those forms of diversity associated with the radio channel, e.g. due to multipath. Coding with spatial and temporal redundancy is accomplished by finding an efficient way to allocate symbols to different transmit antennas, while adding some type of redundancy—for implementing forward error correction—jointly across antennas. For each of the symbol streams associated with different antennas, the system can then resort to other means for combating frequency selective fading; e.g., orthogonal frequency division multiplexing (OFDM) [1].

In the single transmit antenna case, independent fading (IF)—from one complex symbol interval to another—is known to bring up diversity inherent to the code, via the minimum complex symbol Hamming distance between codewords [10], [4]. In contrast, simultaneous use of several transmit antennas allows for diversity even in flat quasistatic fading (QF). The optimality of space-time codes has been so far characterized via two criteria—Guey *et al.* [16], Tarokh *et al.* [26]—namely the determinant criterion based on the product distance [10], [4], [5], and the rank criterion. In [17], Hammons and El Gamal address binary design aspects in connection with the rank criterion.

For single transmit antenna systems, the importance of the product distance has been demonstrated in connection with fading channels by the work of Divsalar and Simon [10], who derived the first design rules. It should be remembered, nonetheless, that this quasi-distance fails to verify all of the axioms of a metric. In the transmit diversity scenario, the determinant criterion is non-constructive and thereby does not easily provide a means to implement it. Thereby, known space-time codes have either been designed based on the rank criterion, or found by some computer search [26], [15], [3], [30]. In [5], [4] good reviews of the known approaches to coding for fading channels are given for the single transmit antenna case, along with references. In [20] it was shown that an Euclidean distance arising from the singular values of difference code matrices is relevant in fading. In [21], [18], Marzetta and Hochwald present a capacity achieving structure for space-time signal constellations. The focus in [21] is on a space-time modulation scheme, whereby signaling is accomplished with waveforms; one waveform covers a number of complex symbol epochs, which amount to the relevant channel correlation time.

The space-time code design problem is approached below from a codeword separation perspective, with the goal of minimizing the code matrix pairwise error probability (PEP). Traditionally, two characterizations of codeword separation in a space-time code have been used: the product distance [10], [4] and the Hamming distance. Divsalar and Simon's results in the single transmit antenna case with interleaving show that the product distance determines the coding gain, while the minimum complex symbol Hamming distance sets the diversity level derived from coding [23]. Tarokh *et al.* arrived at a similar conclusion in the case of transmit diversity in rapid fading [26]. This paper is organized as follows. After defining the problem setting, a rigorous metric between two code elements is identified [20], and shown to be an Euclidean distance directly relevant to maximum likelihood detection. In Section III-B, it is shown that the determinant criterion [26] can be strengthened, in a manner that suggests an optimum space-time codematrix structure, and outlines the central role played by the Euclidean distance in quasistatic fading. In Section III-D, Theorem 5 establishes a close inter-dependence between product and Euclidean distances. Apart from constituting an insightful link between the performance determining factors in QF and IF, Theorem 5 establishes the relevance of, and motivation for, combining space-time coding and modulation in fading environments. In Section III-E, a multidimensional space-time constellation for two transmit antennas, and its coset partitioning are described; example codes constitute the first reporting of a space-time coded modulation scheme for fading channels, whereby a space-time constellation is partitioned in cosets via Euclidean distances.

II. The Problem Setting

Consider the problem of space-time code design for linear modulation on frequency nonselective fading channels. For a system with L transmit and M receive antennas, where the fading is uncorrelated across antennas, the total diversity level achieved is M times the diversity of a single receive antenna system. While $M = 1$ will be eventually assumed, the general notation is used below until the closed form for the received signal is obtained. Let l be the number of symbol epochs covered by a codeword (a frame in [26]). It is meaningful to regard l as the number of adjacent complex symbol epochs processed simultaneously, to some extent, in the detector. A codeword is the concatenation of all symbols sent over all of the L antennas during the corresponding l consecutive symbol epochs (starting, say, at time k); e.g., a codeword $\boldsymbol{c}$ starting at instant k is

$$\boldsymbol{c} = \left[c_k^{(1)} \dots c_k^{(L)} \dots c_{k+l-1}^{(1)} \dots c_{k+l-1}^{(L)} \right]^T = \left[\boldsymbol{c}_k^T \dots \boldsymbol{c}_{k+l-1}^T \right]^T \quad (1)$$

where c_κ^i is a complex symbol, from the complex signal constellation (with unit average energy, same for all transmit antennas), transmitted at discrete time instant κ over transmit antenna i. Alternatively, a more meaningful representation of the codeword $\boldsymbol{c}$ is via the code matrix

$$\boldsymbol{D}_{\boldsymbol{c},\,k} = \begin{bmatrix} c_k^{(1)} & c_k^{(2)} & \cdots & c_k^{(L)} \\ c_{k+1}^{(1)} & c_{k+1}^{(2)} & \cdots & c_{k+1}^{(L)} \\ \vdots & \vdots & \ddots & \vdots \\ c_{k+l-1}^{(1)} & c_{k+l-1}^{(2)} & \cdots & c_{k+l-1}^{(L)} \end{bmatrix}. \quad (2)$$

Let an arbitrary symbol on any transmit antenna be transmitted with energy E_s. Also, let the channel attenuation coefficients between transmit antenna i and receive antenna j be $\alpha_{i,j}(\cdot)$, with $E\{|\alpha_{i,j}|^2\} = 1$, $\forall i, j$; eventually, after defining the general setting, j will be constrained to 1.

Remark 1: If one is to compare the L-transmit antenna system with a single transmit antenna system, then E_s should be replaced by E_s/L, for the same total energy.

At time t, the signal received at receive antenna j is

$$x^{(j)}(t) = \textstyle\sum_{i=1}^{L} \alpha_{i,j}(t) s^{(i)}(t) \sqrt{E_s} + \eta^{(j)}(t). \quad (3)$$

In (3) it is implicitly assumed that the fading changes from one symbol epoch to another, hence the time dependence; and that $\eta^{(j)}(t)$ is zero mean complex Gaussian noise with variance $N_0/2$ per dimension. Assume, for simplicity, negligible intersymbol interference and synchronism. Symbols are sampled at $t = kT$ and the detector is presented with $x_k^{(j)} = \sum_{i=1}^{L} \alpha_{i,j}[k] c_k^{(i)} \sqrt{E_s} + \eta_k^{(j)}$. Denote

$$\gamma_k^{(i,j)} = \sqrt{E_s} \alpha_{i,j}[k], \quad (4)$$

which in the continuum case has the well-known autocorrelation function $E_s J_0(2\pi f_D^{(i,j)} \tau)$; then

$$x_k^{(j)} = \textstyle\sum_{i=1}^{L} c_k^{(i)} \gamma_k^{(i,j)} + \eta_k^{(j)}. \quad (5)$$

Finally, the single receive antenna scenario reduces (5) to

$$x_k = \textstyle\sum_{i=1}^{L} c_k^{(i)} \gamma_k^{(i)} + \eta_k. \quad (6)$$

In the sequel, both quasistatic (block) and rapid fading scenarios are discussed. The former assumes $\alpha_{i,j}(t)$ to be constant over the duration of one codeword (l complex symbol epochs), but change from one codeword to another. The rank criterion [26] determines the diversity level in QF. In rapid fading, the parameter determining diversity is a Hamming distance [26], [10]. Both are addressed below.

The QF assumption implies $\gamma_k^{(i)} = \cdots = \gamma_{k+l-1}^{(i)} \stackrel{\text{def}}{=} \gamma^{(i)}$, $\forall i \in \{1, \dots, L\}$. In matrix form, $\boldsymbol{x} = \boldsymbol{D_c} \boldsymbol{\gamma} + \boldsymbol{\eta}$ where subscript k was dropped and the obvious notations $\boldsymbol{x} = [x_k, \dots, x_{k+l-1}]^T$, $\boldsymbol{\gamma} = \left[\gamma^{(1)}, \dots, \gamma^{(L)}\right]^T$, $\boldsymbol{\eta} = \left[\eta^{(1)}, \dots, \eta^{(L)}\right]^T$, were used. Clearly, when fading is uncorrelated across the different transmit antennas, $\gamma_i \stackrel{\text{def}}{=} \gamma^{(i)}$ are i.i.d. zero mean complex Gaussian, with variance E_s. It is known [7] that the probability $\Pr\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}$ of decoding $\boldsymbol{D_c}$ when $\boldsymbol{D_e}$ was transmitted is upper bounded by a quantity which, in the QF scenario with perfect channel state information (CSI), becomes

$$\Pr{}_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\} = \Pr\{2\,\Re\left(\boldsymbol{\eta}^\dagger \boldsymbol{D_{ec}} \gamma\right) > \|\boldsymbol{D_{ec}} \gamma\|^2\}, \quad (7)$$

(see also [16]). In (7), $\Re(\cdot)$ denotes the argument's real part, "†" means conjugated transposition, and

$$\boldsymbol{D_{ec}} = \boldsymbol{D_e} - \boldsymbol{D_c} \quad (8)$$

is the code difference matrix for codewords $\boldsymbol{e}$ and $\boldsymbol{c}$.

III. Code Design

The separation between space-time codewords is characterized in terms of a distance arising naturally from, and strongly related to, the minimization of (7). The relevant distance will be found to be Euclidean in nature, and a general criterion that minimizes (7) will be formulated.

A. A Distance Metric via Schatten Norms

Consider an $l \times L$ difference code matrix $\boldsymbol{D_{ec}}$. Via singular value decomposition (SVD) [19]

$$\boldsymbol{D_{ec}} = \boldsymbol{V}^\dagger \boldsymbol{\Sigma} \boldsymbol{W}. \quad (9)$$

In (9), $\boldsymbol{V}$, $\boldsymbol{W}$, are unitary $l \times l$ and $L \times L$ matrices, respectively; $\boldsymbol{\Sigma}$ is a $l \times L$ nonnegative matrix whose elements verify $\sigma_{ij} = 0$, $\forall i \neq j$, and $\sigma_{11} \geq \cdots \geq \sigma_{rr} > \sigma_{r+1,r+1} = \cdots = \sigma_{qq} = 0$, where $q \overset{\text{def}}{=} \min\{l, L\}$ and $r \overset{\text{def}}{=} \text{rank}(\boldsymbol{D_{ec}})$. The σ_{ii}, $i = 1, \ldots, q$, are the singular values of $\boldsymbol{D_{ec}}$; they are also the non-negative square roots of the eigenvalues of $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$ ($L \leq l$ assumed). Via (7) and (9)

$$\text{Pr}_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\} = \Pr\{2\,\Re\left[\boldsymbol{\eta'}^{\dagger}\boldsymbol{\gamma'}\right] > \boldsymbol{\gamma'}^{\dagger}\boldsymbol{\gamma'} = \|\boldsymbol{\gamma'}\|^2\} \quad (10)$$

where $\boldsymbol{\eta'} = \boldsymbol{V}\boldsymbol{\eta}$ and $\boldsymbol{\gamma'} = \boldsymbol{\Sigma}\boldsymbol{W}\boldsymbol{\gamma}$. This version of (7), as a particular case of the result from [7], was communicated in [16]. It will become clear that the above form of $\text{Pr}_1(\boldsymbol{D_c} \mapsto \boldsymbol{D_e})$ *does not reveal all of the structure desirable of a space-time code, and that* (10) *can be further processed to gain insight.*

It is immediate that $\boldsymbol{\eta'}$ is a zero mean Gaussian random vector with i.i.d. components and covariance matrix $N_0\boldsymbol{I}_l$; also, that $\boldsymbol{\gamma'}$ is a zero mean Gaussian random vector with a diagonal covariance matrix $E_s\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\dagger}$. Now consider the l^2 (or Euclidean) vector norm. When applied to matrices instead of vectors, it acts as a generalized matrix norm, called the Frobenius norm $\|\cdot\|_2$. The latter verifies [19]

$$\|\boldsymbol{A}\|_2^2 = \textstyle\sum_{i\,=\,1}^{\min\{m,\,n\}} \sigma_i^2(\boldsymbol{A})\,,\ \boldsymbol{A} \in M_{m,n}, \quad (11)$$

where $\sigma_i(\boldsymbol{A})$ are the singular values of $\boldsymbol{A} \in M_{m,n}$ and $M_{m,n}$ is the set of all $m \times n$ complex matrices; note that $\sigma_i^2(\boldsymbol{A})$ are the eigenvalues of $\boldsymbol{A}^{\dagger}\boldsymbol{A}$, denoted $\lambda_i(\boldsymbol{A}^{\dagger}\boldsymbol{A})$. The Frobenius norm is a particular case of the more general Schatten norms, which are unitarily invariant norms on $M_{m,n}$. Then

Proposition 1: Let $q = \min\{l, L\}$, and $d_E(\cdot,\cdot)$ denote the Euclidean distance. Given a space-time code $\mathcal{C}$ of block length l, for L transmit antennas, the mapping $d : \mathcal{C} \times \mathcal{C} \mapsto \mathbb{R}$ defined as $d(\boldsymbol{D_e}, \boldsymbol{D_c}) = (\sum_{i=1}^{q} \sigma_i^2(\boldsymbol{D_{ec}}))^{1/2}$ is a space-time metric on $\mathcal{C}$, and $d(\boldsymbol{D_e}, \boldsymbol{D_c}) = d_E(\boldsymbol{e}, \boldsymbol{c})$.

Proof: Let $M_{m,n}$ be the set of complex $m \times n$ matrices. Since $(-\boldsymbol{A})^{\dagger}(-\boldsymbol{A}) = \boldsymbol{A}^{\dagger}\boldsymbol{A}$, $\forall \boldsymbol{A} \in M_{l,L}$, and the singular values of $\boldsymbol{D_{ec}}$ are the nonnegative square roots of the eigenvalues of $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$, then $\boldsymbol{D_{ec}}$ and $-\boldsymbol{D_{ec}}$ share the same set of singular values, i.e. $d(\boldsymbol{D_e}, \boldsymbol{D_c}) = d(\boldsymbol{D_c}, \boldsymbol{D_e})$. Also, $d(\boldsymbol{D_e}, \boldsymbol{D_c}) = 0 \overset{a}{\iff} \|\boldsymbol{D_e} - \boldsymbol{D_c}\|_2 = 0 \Leftrightarrow \boldsymbol{D_e} = \boldsymbol{D_c} \Leftrightarrow \boldsymbol{e} = \boldsymbol{c}$, where (a) follows from (11). Finally, $\forall \boldsymbol{D_g} \in \mathcal{C}$, $d(\boldsymbol{D_e}, \boldsymbol{D_c}) = \|\boldsymbol{D_e} - \boldsymbol{D_c}\|_2 = \|\boldsymbol{D_e} - \boldsymbol{D_g} + \boldsymbol{D_g} - \boldsymbol{D_c}\|_2 \leq \|\boldsymbol{D_{eg}}\|_2 + \|\boldsymbol{D_{gc}}\|_2 = d(\boldsymbol{D_e}, \boldsymbol{D_g}) + d(\boldsymbol{D_g}, \boldsymbol{D_c})$. Therefore $d(\cdot,\cdot)$ is a metric and $d(\boldsymbol{D_e}, \boldsymbol{D_c}) = d_E(\boldsymbol{e}, \boldsymbol{c})$ via (11). □

Via Proposition 1, the relevant distance between space-time code matrices is essentially Euclidean. Thus, at least when the channel can be estimated and the fading is quasistatic, the relevant distance between the space time code matrices $\boldsymbol{D_e}$, $\boldsymbol{D_c}$ is the Euclidean distance between $\boldsymbol{e}$, $\boldsymbol{c}$, but expressed in terms of the structure inherent to the multiple antenna arrangement, via the singular values of $\boldsymbol{D_{ec}}$. Since the product distance [10], [15] is not a metric, Proposition 1 suggests that features like geometric uniformity should be still sought relative to the Euclidean distance.

Proposition 2: (Additivity) If $\boldsymbol{D_e}, \boldsymbol{D_c} \in \mathcal{C}$, then $d^2(\boldsymbol{D_e}, \boldsymbol{D_c}) \overset{a}{=} \text{tr}\,(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}) \overset{b}{=} \text{tr}\,(\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger})$. If $\boldsymbol{c} = [\boldsymbol{c}_1^T \ldots \boldsymbol{c}_k^T \ldots]^T$, $\boldsymbol{e} = [\boldsymbol{e}_1^T \ldots \boldsymbol{e}_k^T \ldots]^T$ are branch label sequences in the trellis for a space-time code $\mathcal{C}$—i.e. $\boldsymbol{c}_\kappa = [c_\kappa^{(1)} \cdots c_\kappa^{(L)}]^T$, similarly for $\boldsymbol{e}_\kappa$—then $d^2(\boldsymbol{D_c}, \boldsymbol{D_e}) = \sum_\kappa^\infty d_E^2(\boldsymbol{c}_\kappa, \boldsymbol{e}_\kappa)$.

Proof: (a) follows from the singular values of $\boldsymbol{A}$ being the nonnegative square roots of the eigenvalues of $\boldsymbol{A}^{\dagger}\boldsymbol{A}$, $\forall \boldsymbol{A} \in M_{l,L}$, $L \leq l$, and from $\text{tr}(\boldsymbol{A})$ being the eigenvalue sum of $\boldsymbol{A}$. Since $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}} = \boldsymbol{W}^{\dagger}\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\boldsymbol{W}$ and $\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger} = \boldsymbol{V}^{\dagger}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\boldsymbol{V}$ (via (9)), $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$, $\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ are unitarily diagonalizable, with the same nonzero eigenvalues. The last part is straightforward. □

As to the PEP between $\boldsymbol{D_c}, \boldsymbol{D_e} \in \mathcal{C}$, the closed form

$$\text{Pr}_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\} = (1/2)\text{erfc}\left(\|\boldsymbol{\gamma'}\|/\sqrt{4N_0}\right), \quad (12)$$

is derived in Appendix A; it allows for exact calculation of the average quantity $E\{\text{Pr}_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}\}$, which can be obtained by noting that $\|\boldsymbol{\gamma'}\|^2$ is a quadratic form $\boldsymbol{\gamma}^{\dagger}\boldsymbol{W}^{\dagger}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{\Sigma}\boldsymbol{W}\boldsymbol{\gamma}$—in complex Gaussian variables γ_i—for which the p.d.f. is well known (see Appendix B). Then

Proposition 3: In i.i.d. L-transmit-antenna Rayleigh fading, assuming perfect CSI at the receiver, $E\{\text{Pr}_1(\boldsymbol{D_c} \mapsto \boldsymbol{D_e})\} \geq (1/2)\text{erfc}(\sqrt{d^2(\boldsymbol{D_c}, \boldsymbol{D_e})E_s/4N_0})$. Equality occurs if $\boldsymbol{\gamma'} = \boldsymbol{\Sigma}\boldsymbol{W}\boldsymbol{\gamma}$ lies on the $\mathbb{R}^{2L}$ hyperellipsoid $\|\boldsymbol{\gamma'}\|^2 = E\{\|\boldsymbol{\gamma'}\|^2\} = E_s\sum_{i=1}^{q}\sigma_i^2(\boldsymbol{D_{ec}})$. If all singular values of $\boldsymbol{D_{ec}}$ differ, then

$$E\{\text{Pr}_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}\} = \frac{1}{2}\sum_{i=1}^{L} A_i \left(1 - \sqrt{\frac{\sigma_{ii}^2\frac{E_s}{4N_0}}{1 + \sigma_{ii}^2\frac{E_s}{4N_0}}}\right) \quad (13)$$

where $A_i \overset{\text{def}}{=} \sigma_{ii}^{2(L-1)}/\prod_{\substack{k=1\\k\neq i}}^{L}(\sigma_{ii}^2 - \sigma_{kk}^2)$. If all singular values are equal, $\sigma_{ii} = \sigma$, $\forall i = 1, \ldots, L$, then

$$E\{\text{Pr}_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}\} = \left[\tfrac{1}{2}\left(1 - \mu\right)\right]^L \textstyle\sum_{k=0}^{L-1}\binom{L-1+k}{k}\left[\tfrac{1}{2}\left(1 + \mu\right)\right]^k \quad (14)$$

for $\mu \overset{\text{def}}{=} \sqrt{(\sigma^2 E_s/4N_0)/(1+\sigma^2 E_s/4N_0)}$, $\sigma = (\text{tr}(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}))/L$.

Intermediate cases are similar. As alternatives to complex residue formulae [16, eq. (29)], (13), (14) offer a perspective linked to more familiar expressions of diversity. Equality in Proposition 3 minimizes $\text{Pr}_1(\boldsymbol{D_c} \mapsto \boldsymbol{D_e})$, conditioned on $d(\boldsymbol{D_c}, \boldsymbol{D_e})$; increasing $d(\boldsymbol{D_c}, \boldsymbol{D_e})$ can further reduce $\text{Pr}_1(\boldsymbol{D_c} \mapsto \boldsymbol{D_e})$.

B. Space-Time Code Design via an Equal Eigenvalue Criterion (EEC)

The optimal structure of the matrix $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$, as well as the interaction between Euclidean and product distances, are characterized. Let $\boldsymbol{I}_L$ denote the unit $L \times L$ matrix.

Proposition 4 (EEC) In i.i.d., L transmit antenna, quasistatic Rayleigh fading with perfect CSI, $\text{Pr}_1(\boldsymbol{D_c} \mapsto \boldsymbol{D_e})$ is made as small as possible at diversity L iff, for all pairs $\boldsymbol{D_c}$, $\boldsymbol{D_e} \in \mathcal{C}$, the Euclidean squared distances $\text{tr}(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}})$ are made as large as possible, and the non-square matrices $\boldsymbol{D_{ec}}$ are, up to certain proportionality factors, semi-unitary—i.e. $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}} = \left(\text{tr}(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}})/L\right)\boldsymbol{I}_L$. Suboptimal codes (relative to $\Pr_1(\boldsymbol{D_c} \mapsto \boldsymbol{D_e})$) are characterized by matrices $\boldsymbol{D_{ec}^\dagger D_{ec}}$ whose main diagonal elements are as close as possible to each other (or $\mathrm{tr}(\boldsymbol{D_{ec}^\dagger D_{ec}})/L$), and for which the row-wise sum of the absolute values of the elements off the main diagonal is as small as possible for each row.

Proof: Suppose $l \geq L$, whereby $q = \min\{l, L\} = L$. By the determinant criterion [26], one must maximize $\prod_{i=1}^{L} \sigma_{ii}^2$, where σ_{ii}^2 are the eigenvalues $\lambda_i(\boldsymbol{D_{ec}^\dagger D_{ec}})$ of $\boldsymbol{D_{ec}^\dagger D_{ec}}$. Necessarily, $\boldsymbol{D_{ec}^\dagger D_{ec}}$ is positive definite, as all its eigenvalues $\sigma_{ii}^2 > 0$ (diversity L assumed). By Hadamard's theorem [19], maximization of the product of the eigenvalues of some square, positive definite matrix $\boldsymbol{A} = [a_{ij}]$, with a_{ii} fixed, occurs iff the matrix is diagonal. Specifically, $\prod_i \lambda_i(\boldsymbol{A}) = |\boldsymbol{A}| \leq \prod_i a_{ii}$, where $|\boldsymbol{A}|$ is the determinant of $\boldsymbol{A}$, and equality holds iff $a_{ij} = 0$ when $i \neq j$. Further maximization of $\prod_i \lambda_i(\boldsymbol{A})$ is possible by adjusting a_{ii} so as to maximize $\prod_i a_{ii}$. Via the arithmetic-mean geometric-mean (AMGM) inequality, the necessary and sufficient condition for maximizing $\prod_{i=1}^{L} \lambda_i(\boldsymbol{D_{ec}^\dagger D_{ec}})$ is $\sigma_{ii}^2 = \lambda = \mathrm{tr}(\boldsymbol{D_{ec}^\dagger D_{ec}})/L, \forall i = 1, \ldots, q$. Hence, the minimum eigenvalue product is maximized iff the Euclidean distance $\mathrm{tr}(\boldsymbol{D_{ec}^\dagger D_{ec}})$ is made as large as possible for all pairs $\boldsymbol{D_c}$, $\boldsymbol{D_e} \in \mathcal{C}$, and each $\boldsymbol{D_{ec}^\dagger D_{ec}}$ is diagonal with all diagonal elements equal to $\mathrm{tr}(\boldsymbol{D_{ec}^\dagger D_{ec}})/L = d^2(\boldsymbol{D_e}, \boldsymbol{D_c})/L = d_E^2(\boldsymbol{e}, \boldsymbol{c})/L$. Sub-optimality follows from the continuous dependence of polynomial zeros on polynomial coefficients—applied to the eigenvalues and characteristic polynomial of $\boldsymbol{D_{ec}^\dagger D_{ec}}$—and Geršgorin's theorem [19, pp. 343-364]. □

C. Multidimensional space-time trellis codes (MSTTC)

Proposition 4 identifies a desirable structure for an arbitrary codematrix pair $\boldsymbol{D_c}$, $\boldsymbol{D_e}$, irrespective of how each codematrix is generated. With the different transmit antenna sequences along columns, a codematrix can be generated by supplying new rows—e.g. via a trellis—one or more at a time. In the class of space-time trellis codes (STTCs) discussed in [26], [15], [3], [30] every transition through the trellis contributes one row to $\boldsymbol{D_c}$; essentially, the space-time modulator is the L-fold Cartesian product between the individual complex constellations used on the transmit antennas. In order to increase the dimensionality of the overall space-time modulator one must create additional dimensions by using, e.g., consecutive time slots. In such a multidimensional trellis approach each trellis transition supplies an $L' \times L$ matrix—rather than a $1 \times L$ vector—and thereby determines the symbols transmitted from the L transmit antennas over L' consecutive symbol epochs. Denote the length of the shortest error event path (EEP) in the trellis by $p_{\min}$. Then, rather than enforcing Proposition 4 on all valid pairs $\boldsymbol{D_c}$, $\boldsymbol{D_e}$, one can focus on those that dominate performance, i.e. on the difference code matrices corresponding to EEPs of length(s) up to some $p' \geq p_{\min}$.

Let $L' = L$, and let L divide l. View $\boldsymbol{D_c}$, $\boldsymbol{D_e}$, $\boldsymbol{D_{ec}}$ as $(l/L) \times 1$ arrays of $L \times L$ sub-matrices with entries from the modulator constellation(s). In a multidimensional STTC (MSTTC) any code matrix is regarded as a sequence of l/L, block $L \times L$ sub-matrices, formed via a trellis whose branches span L modulator symbol epochs, and are therefore labeled with valid $L \times L$ sub-matrices. A trellis path is selected depending on both the current state and the current block of new input symbols. A MSTTC is said to be optimal up to EEPs of length $p' \geq p_{\min}$ if any difference code matrix pertaining to an EEP of length $p \leq p'$ trellis transitions (pL modulator symbols) is optimal in the sense of Proposition 4.

D. The Connection with the Product Distance

The maximization of the product distance is now linked to the determinant criterion, in order to prove that its strengthening is nontrivial. Let $\mathcal{K} \stackrel{\text{def}}{=} \left\{p | p \in \{0, \ldots, l-1\}, \sum_{i=1}^{L} |e_p^{(i)} - c_p^{(i)}|^2 \neq 0\right\}$; the product distance defined in [4, pp. 719-720] is generalized to multiple transmit antennas (codewords span complex symbol epochs $0, \ldots l-1$). Define the product distance

$$\delta^2(\boldsymbol{D_e}, \boldsymbol{D_c}) \stackrel{\text{def}}{=} \left(\prod_{m \in \mathcal{K}} \sum_{i=1}^{L} \left| e_m^{(i)} - c_m^{(i)} \right|^2 \right)^{1/d_H(\boldsymbol{D_e}, \boldsymbol{D_c})}; \quad (15)$$

$\delta^{2 d_H(\boldsymbol{D_e}, \boldsymbol{D_c})}(\boldsymbol{D_e}, \boldsymbol{D_c})$ has occurred naturally in the discussion of rapid fading by Tarokh *et al.* [26, II.D, eq. (17)], and $\delta^2(\boldsymbol{D_e}, \boldsymbol{D_c})$ acts as a coding gain in IF; the underlying assumption in [26, II.D] was that the flat fading coefficients change independently between complex symbol epochs.

Remark 2: (15) remains valid in a mixture of IF and QF—i.e. when fading is block-wise constant within individual blocks of L complex symbol epochs, and independent between such disjoint blocks in a codeword (see Appendix C); e.g., one may interleave entire blocks of L consecutive epochs. In effect, the flat fading coefficients would be independent from one block of L consecutive epochs to the next, while remaining constant within a block. However, interleaving in this manner forces all coordinates of the multidimensional space-time constellation points to be interleaved together, as blocks.

In (15), $d_H(\boldsymbol{D_e}, \boldsymbol{D_c})$ is the cardinality $|\mathcal{K}|$ of $\mathcal{K}$, or the row Hamming distance between $\boldsymbol{D_e}$, $\boldsymbol{D_c}$ (transmit antennas on columns); it equals the number of complex symbol epochs wherein $\boldsymbol{D_e}$, $\boldsymbol{D_c}$ send different complex symbols at least through one antenna, and will be called a 'row Hamming distance,' as every row spans one complex symbol epoch. When $L = 1$, (15) reduces to the expression in [4, pp. 719] where it is shown that increasing the product distance lowers the codeword PEP in IF.

Theorem 5: The Euclidean distance squared $d_E^2(\boldsymbol{e}, \boldsymbol{c})$ restricts the extent to which the product distance $\delta^2(\boldsymbol{D_e}, \boldsymbol{D_c})$ can be increased, via

$$\frac{\mathrm{tr}(\boldsymbol{D_{ec}^\dagger D_{ec}})}{d_H(\boldsymbol{D_e}, \boldsymbol{D_c})} = \frac{d_E^2(\boldsymbol{e}, \boldsymbol{c})}{d_H(\boldsymbol{D_e}, \boldsymbol{D_c})} \geq \delta^2(\boldsymbol{D_e}, \boldsymbol{D_c}), \quad (16)$$

and the optimal space-time code structure maximizes $\delta^2(\boldsymbol{D_e}, \boldsymbol{D_c})$, given $\mathrm{tr}(\boldsymbol{D_{ec}^\dagger D_{ec}})$.

Proof: Consider the $l \times l$ matrix $\boldsymbol{D_{ec} D_{ec}^\dagger}$, whose nonzero diagonal elements are $\sum_{i=1}^{L} |e_m^{(i)} - c_m^{(i)}|^2$, $m \in \mathcal{K}$, and whose trace is $\sum_{m \in \mathcal{K}} \sum_{i=1}^{L} |e_m^{(i)} - c_m^{(i)}|^2$. Since tr $(\boldsymbol{D_{ec}^\dagger D_{ec}}) =$

tr $(\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger})$, (16) follows directly from the AMGM inequality. In order to characterize the condition that achieves equality, note first that, since singular values are invariant to complex conjugated transposition, the (rank L) matrices $\boldsymbol{D_{ec}}$ and $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ have the same singular values. By the SVD theorem, the L singular values of the matrices $\boldsymbol{D_{ec}}$, $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ are, respectively, the nonnegative square roots of as many eigenvalues of $\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ and, respectively, $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$. Clearly, when $l > L$, the rank L, $l \times l$, Hermitian matrix $\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ has exactly L nonzero eigenvalues and $l - L$ zero eigenvalues, thereby being singular. While this prevents a straightforward application of Hadamard's theorem, it does imply that the nonzero eigenvalue product is the same for both $\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ and $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$, and thereby maximized under the equal eigenvalue condition of Proposition 4. On another hand, the product of the L nonzero eigenvalues of $\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ equals the sum of its (nonzero) principal $L \times L$ minors. They are positive because the Hermitian matrix is positive semidefinite. Finally, apply Hadamard's inequality to each positive definite principal minor, to infer that maximization requires of all nonsingular, nested, $L \times L$ principal sub-matrices of $\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ to be diagonal, and—via the AMGM inequality—have equal main diagonal entries. Given tr$(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}})$, all nonzero main diagonal entries in $\boldsymbol{D_{ec}}\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}$ must be equal, thus achieving equality in (16). □

By letting $l = L$, considerations similar to those used in the proof of Theorem 5 lead to the following corollary, which can characterize a mixture of QF and IF (see Remark 2, Appendix C, Corollary 7).

Corollary 6: If $\boldsymbol{D_e}$, $\boldsymbol{D_c}$ are $L \times L$ codematrices, and $\boldsymbol{D_{ec}}$ is full rank, then

$$|\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}|^{1/L} \leq \left(\prod_{m=1}^{L}\sum_{i=1}^{L}|e_m^{(i)} - c_m^{(i)}|^2\right)^{1/L} = \delta^2(\boldsymbol{D_e}, \boldsymbol{D_c}). \quad (17)$$

Equality is achieved iff $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$ is diagonal, in particular when the equal eigenvalue condition is met.

By Theorem 5, any attempt to increase $\delta^2(\boldsymbol{D_e}, \boldsymbol{D_c})$—e.g. by adjusting $\sum_{i=1}^{L}|e_m^{(i)} - c_m^{(i)}|^2$ while preserving $d_H(\boldsymbol{D_e}, \boldsymbol{D_c})$—is as successful as the squared Euclidean distance tr$(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}})$ permits. Theorem 5 is transparent to the number of complex symbol epochs covered by one trellis transition; also, $d_H(\boldsymbol{D_e}, \boldsymbol{D_c})$ is a key parameter—which should be increased—as it determines the diversity level when interleaving is used. Thereby, the code's performances in quasistatic and fast fading are closely related.

As trellis coded modulation (TCM) attempts to increase the minimum Euclidean distance (per trellis transition), Theorem 5 implies that TCM techniques remain relevant to the design of STTCs. In multidimensional space-time TCM (Section III-C), a multidimensional space-time constellation is partitioned via Euclidean distances (traces of branch label differences, see Section III-E.1), and a trellis selects the relevant $L \times L$ codematrices (symbols transmitted by L antennas in L consecutive symbol epochs). Such structure should be enforced at least on the codeword pairs yielding the shortest EEPs.

E. Discussion and Examples

Alamouti's scheme [2] for $L = 2$ transmit antennas uses the Hurwitz-Radon (HR) transform [25]; like the full-rate space-time block codes of [25], it does obey the structure outlined above, and offers a simple means to implement Proposition 4: concatenate any encoder, a mapper from encoded symbols to constellation points, and a HR transform. However, such a code would have a subunitary rate, i.e., with 4PSK modulators on the individual antennas, the code would send less than 2 bits/s/Hertz.

In order to illustrate the effects of the eigenvalue spread the expectation $E\{\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}\}$ is plotted vs. E_s/N_0 in Figure 1, for different sets of eigenvalues of $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$, in the two cases $L = 2$, $L = 4$. Note that tr$(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}})$ equals 4 in both cases, and that increasing tr$(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}})$ does not influence the *relative* spacing between curves. The effect of eigenvalue imbalance is visible and illustrates the importance of allocating tr$(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}})$ equally among the eigenvalues of $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$. While the values indicated in Figure 1 for squared singular values are chosen only for illustrative purposes, they do allow one to verify that the separation between curves, in dB, is as predicted by the ratio between the geometric and arithmetic means of the eigenvalues of $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$, i.e. $10\log_{10}[(\prod_{i=1}^{L}\sigma_{ii}^2)^{1/L}/(\frac{1}{L}\sum_{i=1}^{L}\sigma_{ii}^2)]$; e.g., the curve marked with a '△' is asymptotically $|10\log_{10}[(0.08 \cdot 3.92)^{1/2}/(4/2)]| \approx 5.5$dB away from the equal eigenvalue curve for $L = 2$. Asymptotically with the signal-to-noise ratio (SNR), all curves pertaining to one value of L are parallel—i.e. they exhibit the same diversity.

Two examples of STTCs for 4PSK and $L = 2$ transmit antennas, designed along the lines discussed in Section III-C, follow. The constructions—along with the multidimensional constellation of Table I—guarantee that the code matrix pairs corresponding to the shortest EEPs in

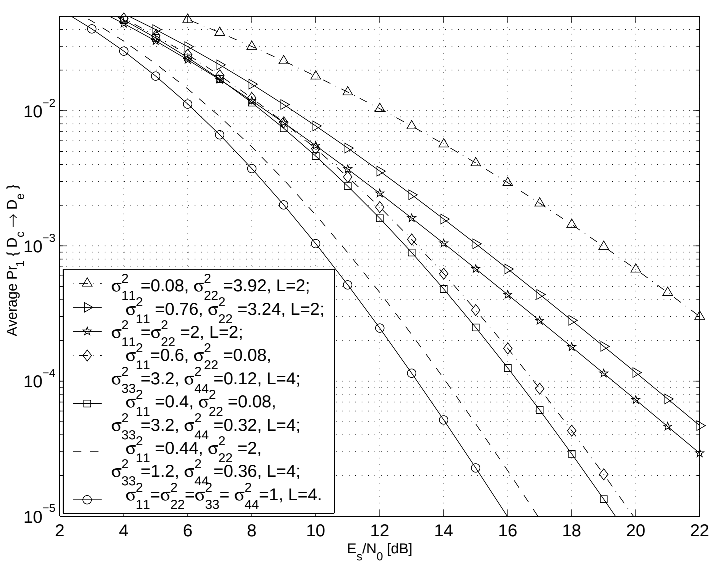

Fig. 1. Effect—on PEP—of singular value imbalance, and quantitative illustration of the coding gain. For $L = 2$, $L = 4$ transmit antennas, the average $\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}$ is plotted vs. E_s/N_0, for different sets of eigenvalues of $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}$; tr$(\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D_{ec}}) = 4$.

the trellis verify the equal singular value (ESV) condition. While an exhaustive search was not conducted to maximize Euclidean distances between such pairs, or minimize their number, the constructions are a convenient way to enforce the ESV structure via properties of orthogonal matrices.

E.1 An 8-state multidimensional space-time trellis code

The trellis diagram of an 8-state multidimensional code constructed via the approach discussed in Section III-C is shown in Figure 2. This code verifies the ESV condition for code difference matrices up to EEPs of length $p' = 2$ transitions (see Section III-C). This is accomplished by the use of an augmented set of Alamouti matrices, as follows. For L=2 and 4PSK, the 16 orthogonal complex matrices discussed by Alamouti [2] do exhibit the aforementioned (optimal) ESV structure for all pairwise differences. In order to achieve the desired spectral efficiency of 2 b/s/Hz, one must have enough 2×2 constituent matrices in the multidimensional constellation; this requires augmenting the optimal matric set—e.g., by a reflection of itself, see the two halves of Table I. Note that, in effect, some pairs of codematrices from the augmented set will not obey the ESV structure. Nevertheless, each half of Table I does form an orthogonal set of 16 2×2 Alamouti matrices of the form $\begin{bmatrix} a & b \\ b^* & -a^* \end{bmatrix}$ or $\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}$. The design is to guarantee that all difference code matrices pertaining to an EEP of length $k \leq 2$ transitions ($2k$ modulator symbols) have ESVs. This is easily achieved by a trellis labeling which uses, e.g., the first half of Table I for transitions starting in even states, and the second half of Table I for transitions starting in odd states; since each half is a complete Alamouti set, the ESV structure is inherited by differences between the labels of parallel transitions (EEPs of length 1), and between sequences of labels corresponding to EEPs of length 2.

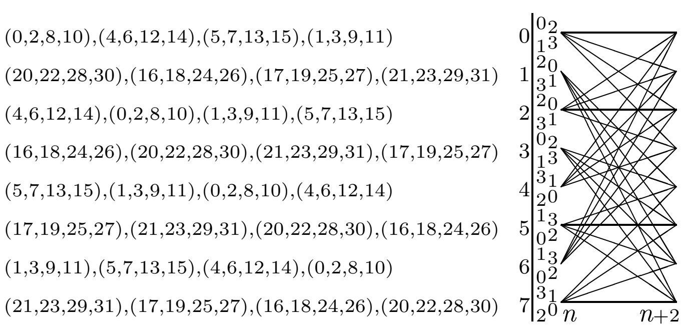

Fig. 2. Trellis diagram of 8-state STTC code; two *coded*, two *uncoded* bits. A connection between two states represents a cluster of four parallel transitions. For each transition, the branch label represents the subscript index of a matrix C_i, $i = 0, \ldots, 31$, from Table I; each C_i reflects the complex symbols transmitted during two consecutive 4PSK symbol epochs from both transmit antennas. One pair of parentheses encloses labels of parallel transitions (within respective cluster) ordered by increasing decimal value ($0 \ldots 3$, left to right) of corresponding *uncoded* bit pair. The decimal values of different pairs of *coded* bits—corresponding to respective clusters—are shown next to each node.

The branch labels (listed on the left hand side) are grouped in 4-tuples—corresponding to groups of 4 parallel transitions from each state—and represent subscript indices of the matrices C_i, $i = 0, \ldots, 31$, from Table I ; e.g., from state 0 with two coded, and two uncoded, input bits that have base-10 representations of 3 and 1, respectively, the code transitions to state 2, sending C_7.

The matrices $\boldsymbol{C}_i$, $i = 0, \ldots, 31$, are a subset from the four-fold Cartesian product of the 4PSK constellation with itself (hence the redundancy). The entries of $\boldsymbol{C}_i$, $i = 0, \ldots, 31$, represent indices of complex points from the 4PSK constellation $s_m = (1/\sqrt{2} + j/\sqrt{2}) \exp(jm\pi/2)$, $m = 0, \ldots, 3$. Each $\boldsymbol{C}_i$ defines the 4PSK symbols to be sent over the $L = 2$ transmit antennas, during two consecutive symbol epochs. Consequently, each trellis branch covers two consecutive 4PSK symbol epochs; while this constitutes a similarity with multiple trellis coded modulation (MTCM) of multiplicity two [9], the space-time trellis in Figure 2 can not be reduced to a concatenation between a binary MTCM trellis of multiplicity two and a 2×2 orthogonal block code. Finally, complexity is judged in terms of the product between the number of states and the number of transitions emerging from each state, normalized to one complex symbol epoch; e.g., since each transition in Figure 2 spans two symbol epochs, and 16 transitions (including parallel ones) emerge from any state, complexity is $16 \times 8/2$=64—equal to that of Tarokh's 16-state STTC. The 32 matrix set in Table I is partitioned in eight cosets

TABLE I

2×2 MATRICES $\boldsymbol{C}_0, \ldots, \boldsymbol{C}_{31}$, FORM MULTIDIMENSIONAL SPACE-TIME CONSTELLATION, WITH SUBSCRIPT INDICES USED IN TRELLIS BRANCH LABELS. ENTRIES OF $\boldsymbol{C}_i$, i=0,...,31, REPRESENT INDICES OF COMPLEX POINTS FROM A 4PSK CONSTELLATION. EACH $\boldsymbol{C}_i$ DEFINES 4PSK SYMBOLS TO BE SENT OVER $L = 2$ TRANSMIT ANTENNAS, DURING TWO CONSECUTIVE COMPLEX SYMBOL EPOCHS.

$\boldsymbol{C}_0 \ldots \boldsymbol{C}_7$	$\boldsymbol{C}_8 \ldots \boldsymbol{C}_{15}$	$\boldsymbol{C}_{16} \ldots \boldsymbol{C}_{23}$	$\boldsymbol{C}_{24} \ldots \boldsymbol{C}_{31}$
$\begin{bmatrix} 0 & 3 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 2 & 3 \\ 0 & 3 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 0 & 3 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 1 & 3 \end{bmatrix}$	$\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$
$\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$	$\begin{bmatrix} 0 & 3 \\ 2 & 3 \end{bmatrix}$	$\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}$
$\begin{bmatrix} 0 & 0 \\ 3 & 1 \end{bmatrix}$	$\begin{bmatrix} 2 & 0 \\ 3 & 3 \end{bmatrix}$	$\begin{bmatrix} 0 & 2 \\ 3 & 3 \end{bmatrix}$	$\begin{bmatrix} 2 & 2 \\ 3 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$	$\begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix}$
$\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$	$\begin{bmatrix} 3 & 0 \\ 1 & 0 \end{bmatrix}$
$\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$	$\begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$	$\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$	$\begin{bmatrix} 3 & 3 \\ 2 & 0 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}$	$\begin{bmatrix} 3 & 0 \\ 3 & 2 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$	$\begin{bmatrix} 3 & 2 \\ 3 & 0 \end{bmatrix}$

by partitioning each half—a complete Alamouti set—into four cosets, via Euclidean distances between multidimensional points (traces of Gram matrices $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D}_{\boldsymbol{ec}}$ of the 2×2 difference matrices). By construction,

- Transitions starting from, or ending in, a state are labeled with matrices in the same half of Table I.
- Set partitioning maximizes the minimum Euclidean dis-

tance between branches sharing a state.

- $\boldsymbol{D}_{\boldsymbol{ec}}^{\dagger}\boldsymbol{D}_{\boldsymbol{ec}}$ has equal eigenvalues $\forall \boldsymbol{D}_{\boldsymbol{ec}}$ pertaining to EEPs of length $p \leq 2$ (4 complex 4PSK symbols).
- The row Hamming distance between labels of any two parallel transitions is 2 (diversity 2 in IF [26]).

Figure 3 compares the average frame error probability curve for this space-time code, against those of Alamouti's scheme and of two other trellis space-time codes from [26], [6]—all sending 2 bits/sec/Hz; Viterbi's algorithm was used in the decoder. Because each trellis transition covers two symbol epochs and the row Hamming distance between any two parallel transitions is two, the parallel transitions in the trellis of Figure 2 do not reduce the diversity level in IF (vs. QF, see Figure 3). However, a comparison on equal complexity grounds would pair the MSTTC in Figure 2 with, e.g., the 16-state STTC from [26]; the latter has diversity three in IF [13], and thereby outperforms the former in IF.

The parallel concatenation from [8], based on the four-state, systematic and recursive constituent code presented therein—which, like the space-time code from Figure 2, is non-binary—is easily verified to perform slightly worse after eight iterations [8, Fig. 4] than the example code above (see Figure 3). Note that the definition of complexity used in [8] is different from the one used above; when the former is observed, the turbo code from [8, Fig. 4] after eight iterations is twice as complex as the 16-state space-time code from [26], and more than five times as complex than the example code from Figure 2.

E.2 A 32-state multidimensional space-time trellis code

While the largest diversity in IF reported in the literature [13] is three, the 2b/s/Hz, 32-state code in Figure 4 avoids parallel transitions, has a minimum row Hamming distance of four, hence diversity four in IF (see Figure 3); while it clearly outperforms the best 32-state code from [13], note again that a comparison on equal complexity grounds would pair the MSTTC from Figure 4 with a 64-state code in the class discussed in [26], [6], [3], [13]—but none has been explicated, or characterized in IF.

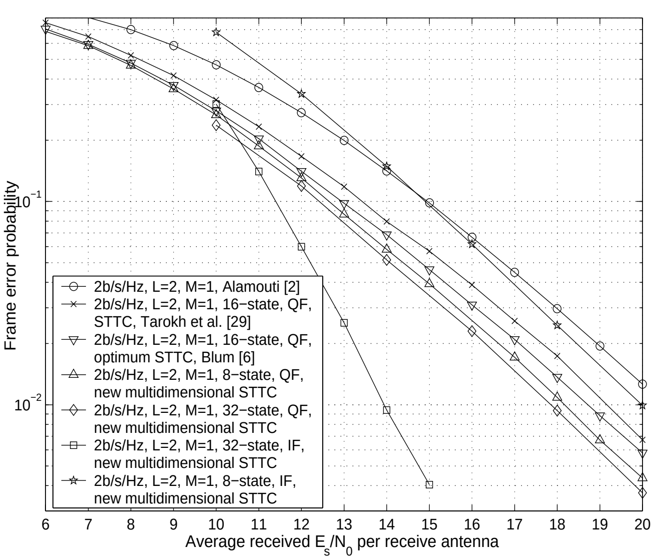

Fig. 3. Simulated frame error probability curves of several space-time transmission schemes for $L = 2$ transmit, and $M = 1$ receive, antennas, in quasistatic/independent fading; 130 symbols/frame.

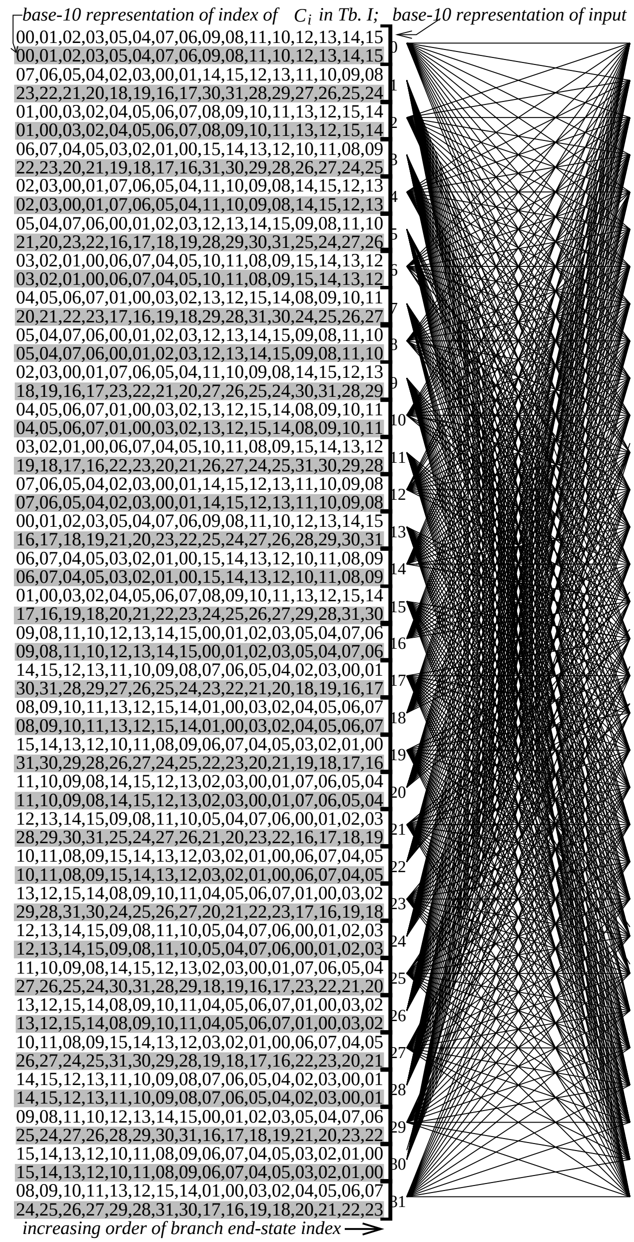

Fig. 4. Trellis diagram of 32-state STTC; all four bits *coded*. A transition *output* label (bottom row, next to each state) represents the subscript index of a matrix C_i, $i = 0, \ldots, 31$ (Table I), which reflects the complex symbols sent during two consecutive 4PSK symbol epochs from both transmit antennas. A transition *input* label (top row, next to each state) corresponds to output label below it, and is the base-10 representation of input bit 4-tuple.

IV. Conclusions

The Euclidean distance is shown to directly determine the extent to which the product distance can be increased. The determinant criterion is tightened. Theorem 5 shows that in order to increase the product distance one must increase the Euclidean distance, and that performances in

QF and IF are closely related. The relevance of combining space-time coding and modulation was established. The example codes illustrate how one can combine modulation and space-time coding in fading channels, by partitioning a space-time constellation in cosets.

V. Acknowledgment

The author wishes to thank K. K. Mukkavilli for writing the first version of the simulation code, the reviewers for their comments, and the Associated Editor for his availability and helpful suggestions.

Appendices

I. Appendix A: Proof of Proposition 3

$\Re(\cdot)$, $\Im(\cdot)$ denote real and imaginary parts, respectively; $\boldsymbol{\gamma}' = \boldsymbol{\Sigma W \gamma}$, and the assumption is QF with perfect CSI. Consider the random variable (r.v.) $2\,\Re[\boldsymbol{\eta}'^{\dagger}\boldsymbol{\gamma}'] = \sum_{i=1}^{l}(\eta_i'^{*}\gamma_i' + \eta_i'\gamma_i'^{*}) = \sum_{i=1}^{l}\xi_i$, where $\xi_i = 2\,\Re(\eta_i'\gamma_i'^{*}) = 2[\Re(\eta_i')\Re(\gamma_i') + \Im(\eta_i')\Im(\gamma_i')]$. Since $\Re(\eta_i')$, $\Im(\eta_i')$ are Gaussian, uncorrelated, with zero mean and variance $N_0/2$, they are independent and ξ_i is zero mean Gaussian, with variance $2N_0|\gamma_i'|^2$. Then $\xi_i = \sqrt{2N_0}|\gamma_i'|\zeta_i$, $\zeta_i \sim \mathcal{N}(0,1)$. Since η_i',η_j' are independent if $i \neq j$, so are ζ_i,ζ_j, $i \neq j$. Let

$$u_1 \stackrel{\text{def}}{=} \sum_{i=1}^{l}\sqrt{2N_0}|\gamma_i'|\zeta_i = 2\,\Re[\boldsymbol{\eta}'^{\dagger}\boldsymbol{\gamma}'],\ u_k = \zeta_k,\ k = 2,\ldots,l$$

Clearly, the p.d.f. of the r.v. U_1 (whose observations are u_1) is the integral over $(-\infty,\infty)^{l-1}$ with respect to u_k, $k = 2,\ldots,l$, of the joint p.d.f. of U_k, $k = 1,\ldots,l$. The modulus of the transformation Jacobian is $1/(\sqrt{2N_0}|\gamma_1'|)$ and, since the ζ_i are jointly Gaussian, one finds by induction that

$$p_{U_1}(u_1) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} du_2\cdots du_l \frac{1}{\sqrt{2N_0}|\gamma_1'|}\frac{1}{(2\pi)^{l/2}} \times \exp[-\frac{1}{2}(\frac{[u_1 - \sqrt{2N_0}|\gamma_2'|u_2 - \cdots - \sqrt{2N_0}|\gamma_l'|u_l]^2}{2N_0|\gamma_1'|^2} + u_2^2 + \cdots + u_l^2)];$$

$$p_{U_1}(u_1) = \frac{1}{\sqrt{2N_0}\sqrt{2\pi}\|\boldsymbol{\gamma}'\|}\exp\left[-u_1^2/(4N_0\|\boldsymbol{\gamma}'\|^2)\right] \quad (18)$$

Since the probability of the normal r.v. U_1 exceeding $\|\boldsymbol{\gamma}'\|^2$ is the Q-function, $Q(\cdot) = (1/2)\mathrm{erfc}(\cdot/\sqrt{2})$,

$$\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\} = \frac{1}{2}\mathrm{erfc}\left(\frac{\|\boldsymbol{\gamma}'\|}{\sqrt{4N_0}}\right) = Q\left(\sqrt{2\frac{\|\boldsymbol{\gamma}'\|^2}{4N_0}}\right) \quad (19)$$
$$\sim \frac{\sqrt{N_0}}{\sqrt{\pi}\|\boldsymbol{\gamma}'\|}\exp\left(-\frac{\|\boldsymbol{\gamma}'\|^2}{4N_0}\right) \stackrel{\text{def}}{=} f(\|\boldsymbol{\gamma}'\|).$$

via (10); (19) is an asymptotic expansion [11, pp. 1-17] of $\mathrm{erfc}(\cdot)$ to one term, as $\|\boldsymbol{\gamma}'\|/\sqrt{4N_0} \to \infty$ since $\mathrm{erfc}(\cdot)$ admits the asymptotic expansion $\mathrm{erfc}(x) \sim \exp(-x^2)/(x\sqrt{\pi})$ [22, pp. 18, 37]. This approximation is asymptotically tight as $\|\boldsymbol{\gamma}'\|/\sqrt{4N_0} \to \infty$. Thereby, $f(\|\boldsymbol{\gamma}'\|)$ upper bounds $\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}$ asymptotically tight. Moreover, from the perspective of the desired minimization of the PEP, the upper bound $f(\|\boldsymbol{\gamma}'\|)$ approximates $\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}$ pessimistically (from above) for small $\|\boldsymbol{\gamma}'\|/\sqrt{4N_0}$. Taking the expectation in (19) preserves the asymptotic expansion:

$$E\{\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}\} \sim E\{f(\|\boldsymbol{\gamma}'\|)\}. \quad (20)$$

Further, the right hand side of (20) is an asymptotically tight upper bound to $E\{\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}\}$. From (19), via Jensen's inequality applied to the convex $\mathrm{erfc}(\cdot)$, we have $E\{\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}\} \geq Q\left(\sqrt{E\{\|\boldsymbol{\gamma}'\|^2\}/2N_0}\right)$, where the expectation is with respect to $\boldsymbol{\gamma}'$ and equality occurs iff $\|\boldsymbol{\gamma}'\|^2 = E\{\|\boldsymbol{\gamma}'\|^2\}$. Note that this essentially restricts the envelopes $|\gamma_i'|$, $i = 1,\ldots,L$, to be on a hyperellipsoid. When fading is uncorrelated across antennas, $\boldsymbol{\gamma}'$ is a zero mean Gaussian random vector with a diagonal covariance matrix $E_s\boldsymbol{\Sigma\Sigma}^T$, $E\{\|\boldsymbol{\gamma}'\|^2\} = E_s\sum_i\sigma_{ii}^2 = E_s d^2(\boldsymbol{D_e},\boldsymbol{D_c})$ and the proof is completed by

$$E\{\Pr_1\{\boldsymbol{D_c} \mapsto \boldsymbol{D_e}\}\} \geq Q\left(\sqrt{E_s d^2(\boldsymbol{D_e},\boldsymbol{D_c})/2N_0}\right). \quad (21)$$

II. Appendix B: Exact Expression for $\Pr_1\{\boldsymbol{D_c}\mapsto\boldsymbol{D_e}\}$

$\|\boldsymbol{\gamma}\| = \sqrt{E_s}\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}' = \boldsymbol{\Sigma W\gamma}$, where $\boldsymbol{\alpha}$ is the complex vector of channel coefficients, yield

$$\|\boldsymbol{\gamma}'\|^2 = \boldsymbol{\gamma}'^{\dagger}\boldsymbol{\gamma}' = E_s\boldsymbol{\alpha}^{\dagger}\boldsymbol{W}^{\dagger}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{\Sigma W\alpha} = E_s f(\boldsymbol{\alpha}); \quad (22)$$

$f(\boldsymbol{\alpha})$ is a quadratic form whose positive definite, Hermitian matrix is $\boldsymbol{F} \stackrel{\text{def}}{=} \boldsymbol{W}^{\dagger}\boldsymbol{\Sigma}^{\dagger}\boldsymbol{\Sigma W}$. From [28], [24, pp. 590-595] the characteristic function of $f(\boldsymbol{\alpha})$ can be expressed in terms of the eigenvalues ϕ_i of $\boldsymbol{R}^{*}\boldsymbol{F}$, where $\boldsymbol{R} \stackrel{\text{def}}{=} (1/2)E\left\{(\boldsymbol{\alpha} - E\{\boldsymbol{\alpha}\})^{*}(\boldsymbol{\alpha} - E\{\boldsymbol{\alpha}\})^{T}\right\}$. The characteristic function of the positive definite, Hermitian quadratic form $f(\boldsymbol{\alpha})$ is $G_f(\xi) = 1/\prod_{i=1}^{L}(1 - 2j\xi\phi_i)$, and $\boldsymbol{R} = (1/2)\boldsymbol{I}$, $\phi_i = \sigma_{ii}^2/2$, $G_f(\xi) = 1/\prod_{i=1}^{L}(1 - j\xi\sigma_{ii}^2)$. Simple singular values σ_{ii} lead to (weighted) exponential distributions; a singular value of multiplicity l produces a sum of weighted central χ^2 distributions with $2m$ degrees of freedom, $m = 1,\ldots,l$. All different and all equal singular values yield (13) and (14), respectively.

III. Appendix C: Diversity in Mixture of Independent, Quasistatic Fading

Let $M = 1$, for simplicity. Apart from the coding gain $\delta^2(\boldsymbol{D_e},\boldsymbol{D_c})$, a key parameter in flat, IF—from a complex symbol epoch to another, e.g. ideal interleaver—is the diversity level, i.e. the minimum row Hamming distance over all codematrix pairs [26]. In another possible fading scenario the channel could change independently from a block of L complex epochs to the next, while remaining essentially constant within a block. In such a mixture of IF and QF, the case when the row Hamming distance $d_H(\boldsymbol{D_e},\boldsymbol{D_c})$ still determines the diversity order, and the coding gain is still given by (15), despite of reducing by L a codeword's exposure to independent channel realizations, is characterized by

Corollary 7: Consider a MSTTC for L transmit antennas, with trellis transitions labeled by $L \times L$ complex matrices. Assume that the flat fading coefficients are constant within each block of L consecutive, complex symbol epochs determined by a trellis transition, and independent among disjoint blocks. Consider an error event of length p' trellis transitions—covering $p'L$ complex symbol epochs, and mistaking $\boldsymbol{D_e}$ for $\boldsymbol{D_c}$—such that the label difference matrices of corresponding transitions in the EEP have full rank; the pair $\boldsymbol{D_c}$, $\boldsymbol{D_e}$ experiences a diversity of $d_H(\boldsymbol{D_e}, \boldsymbol{D_c}) = p'L$, and

$$\delta^2(\boldsymbol{D_e}, \boldsymbol{D_c}) \leq \left(\prod_{m=1}^{p'L} \sum_{i=1}^{L} |e_m^{(i)} - c_m^{(i)}|^2\right)^{1/(p'L)}. \quad (23)$$

Equality in (23) is achieved iff the $L \times L$ Gram matrices of corresponding (trellis) label differences (within span of error event path) are diagonal, for which the equal eigenvalue condition is sufficient.

Proof outline: Upper-bounding the PEP of mistaking $\boldsymbol{e}$ for $\boldsymbol{c}$ starts with the squared Euclidean distance $d_E^2(\boldsymbol{e}, \boldsymbol{c}) = \sum_{t=1}^{l} |\sum_{i=1}^{L} \alpha_i(t)(e_t^{(i)} - c_t^{(i)})|^2$ [26, II.B,II.D]. Group the outer sum's terms in subsets corresponding to the disjoint blocks of epochs within which fading is constant; there are p' such nonzero terms, each containing L terms and having the form $\sum_{t=t_1+1}^{t_1+L} |\sum_{i=1}^{L} \alpha_i(e_t^{(i)} - c_t^{(i)})|^2$. Averaging over the p' independent fades can be separated into p' independent averages; each one of these is an average PEP under QF conditions, and is upper bounded by a quantity of the form $[(\prod_{i=1}^{L} \lambda_i)^{1/L} E_s/4N_0]^{-L}$, where λ_i are the nonzero eigenvalues of the full rank $L \times L$ matrices representing homologous trellis label differences along the EEP. Note that diversity is $d_H(\boldsymbol{D_e}, \boldsymbol{D_c}) = p'L$, and use Corollary 6. □

References

[1] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data rate wireless communication over wideband channels," *Proc. IEEE VTC '98*, pp. 2232-2236.

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

[3] S. Bäro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis-coded modulation," *IEEE Commun. Lett.*, vol. 4, pp. 20-22, Jan. 2000.

[4] S. Benedetto, and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*, New York: Kluwer Academic/Plenum Publishers, 1999.

[5] E. Biglieri, G. Caire, and G. Taricco, "On coding for fading channels," *International Symposium on Signals, Systems and Electronics, 1998*, pp. 101-106.

[6] R. S. Blum, "Some Analytical Tools for the Design of Space-Time Convolutional Codes," submitted to *IEEE Trans. Inform. Theory*; similar work in *Proc. Conf. on Inf. Sciences and Syst.*, Princeton Univ., Princeton, NJ, March 1999

[7] J. K. Cavers, and P. Ho, "Analysis of the error performance of trellis coded modulations in Rayleigh channels," *IEEE Trans. Commun.*, vol. 40, pp. 74-83, March 1996.

[8] D. Cui, and A. M. Haimovich, "Performance of Parallel-Concatenated Space-Time Codes," *IEEE Commun. Lett.*, vol. 5, pp. 236-238, June 2001.

[9] D. Divsalar, and M. K. Simon, "Multiple trellis coded modulation," *IEEE Trans. Commun.*, Vol. 36, pp. 410-418, Apr. 1988.

[10] D. Divsalar, and M. K. Simon, "The design of trellis coded MPSK for fading channels: Performance criteria," *IEEE Trans. Commun.*, Vol. 36, pp. 1004-1012, Sept. 1988.

[11] A. Erdélyi, *Asymptotic Expansions*, New York: Dover, 1956

[12] G. J. Foschini, and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, pp. 311-335, Mar. 1998.

[13] W. Firmanto, B. S. Vucetic, and J. Yuan, "Space-time TCM with Improved Performance on Fast Fading Channels," *IEEE Commun. Lett.*, vol. 5, pp. 154-156, April 2001.

[14] I. S. Gradshteyn, and I. M. Ryzhik, *Tables of Integrals, Series and Products*, 5th ed., San Diego: Academic Press, 1994.

[15] J. Grimm, M. P. Fitz, and J. V. Krogmeier, "Further results on space-time coding for Rayleigh fading," *Proc. Allerton Conference on Circuit and Systems Theory*, 1998

[16] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communications systems over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 527-537, Apr. 1999; *Proc. IEEE VTC'96*, pp. 136-140.

[17] A. R. Hammons, and H. El Gamal, "On the theory of space-time codes for PSK modulation," *IEEE Trans. Inform. Theory*, vol. 46, No. 2, pp. 524-542, March 2000.

[18] 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, No. 2, pp. 543-564, March 2000.

[19] R. A. Horn, and C. R. Johnson, *Matrix Analysis*, New York: Cambridge University Press, 1994.

[20] D. M. Ionescu, "New results on space-time code design criteria," *IEEE Wireless Commun. Networking Conf.*, Sept. 1999.

[21] 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.

[22] N. Nielsen, *Theorie des Integrallogarithmus und Verwandter Transzendenten*, Leipzig: Teubner, 1906.

[23] J. G. Proakis,*Digital Communications*, 3rd ed., New York: McGraw-Hill, 1995.

[24] M. Schwartz, W. R. Bennett, and S. Stein, *Communications Systems and Techniques*, Inter-University Electronics Series, vol. 4, New York: McGraw-Hill, 1966.

[25] 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.

[26] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria and code construction," *IEEE Trans. Inform. Theory*, vol. 44, No. 2, pp. 744-765, March 1998.

[27] E. I. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T Bell Labs., Tech. Rep., June 1995.

[28] G. L. Turin, "The characteristic function of Hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, pp. 199-201, June 1960.

[29] J. H. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. 5, pp. 871-878, June 1987.

[30] Q. Yan, and R. S. Blum, "Optimum Space-Time Convolutional Codes," *Wireless Commun. Networking Conf.*, Sept. 2000