Unified Multi-Antenna Code Design for Fading Channels with Spatio-Temporal/Spectral Correlations

Mahesh K. Varanasi, *Senior Member, IEEE*, and Pranav Dayal, *Member, IEEE*

***Abstract*— A unified framework for coherent multiple antenna communications is presented leading to a general space-time code design criterion valid for arbitrary spatial and temporal fading correlations. This framework provides insights into the effect of each of spatial and temporal correlations on space-time code design. The single code design expression applies to quasi-static, fast fading and multiple block fading channels and more generally also to channels with arbitrary correlations in the temporal dimension as well as in the spatial dimension with arbitrary rank. The general coding strategy proposed consists of precoding a space-time code with its size determined by the rank of transmit correlation, the structure determined by the Cholesky factorization of the temporal correlation, and the precoding matrix obtained from the eigenvectors of the transmit correlation matrix.**

***Index Terms*— Antenna diversity, correlated fading, error analysis, MIMO-OFDM, signal design, space-time modulation, space-time codes, spatial correlation, temporal correlation, wireless communication.**

I. INTRODUCTION

EARLY works on wireless fading channels dealt with single antenna transmission techniques that involved interleaving to provide diversity via signal space methods [1]–[3]. The purpose of interleaving was to provide a channel with a white temporal fading process. Subsequently, the more practical case of imperfect interleaving leading to a non-trivial temporal fading correlation matrix was considered in [4]. Signal design with multiple transmit and receive antennas with identity spatial correlation matrix was considered in [5]. However, only the extreme cases of temporal correlation corresponding to quasi-static and fast fading were considered, and that too in a separate manner. In any case, such spatially white fading correlation assumes rich scattering which is rarely encountered in practice. Spatial correlations may arise due to insufficient antenna spacings and/or non-uniform scattering (which may occur due to a clustering of reflectors). The case of the most general spatial correlation for multiple antennas has been addressed with respect to code design in [6], [7]. A simplified practical model for spatial correlation was provided and analyzed in [8] but it did not account for any temporal correlation.

Manuscript received January 5, 2005; revised June 28, 2005 and August 2, 2005; accepted August 20, 2005. The associate editor coordinating the review of this paper and approving it for publication was T. Duman. This work was supported in part by NSF Grants CCF-0434410 and CCF-0431170. The material in this paper was presented in part at the Annual Asilomar Conference on Signals, Systems and Computers, Monterey, CA, Nov. 2004.

M. K. Varanasi is with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309-0425 USA (e-mail: varanasi@colorado.edu).

P. Dayal is with Qualcomm, Inc., San Diego, CA 92121 USA (e-mail: dayalp@dsp.colorado.edu). This work was done when P. Dayal was with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309-0425 USA.

Digital Object Identifier 10.1109/TWC.2006.04855

In this work, we formulate a unified framework for analysis and design of multiple antenna space-time communication systems under a combined spatial and temporal fading correlation scenario. Our approach is more generally applicable than previous works and more insightful as well for several reasons. Firstly, a succinct analytical formula is obtained that captures at once the diversity order performance as a function of (separable) spatial and general temporal correlation matrices. Secondly, code design rules in the spatial-correlation-only model and the temporal-correlation-only model are rigorously dealt with separately in detail, thereby improving upon previous designs even in each of these cases. Finally, we combine the two sets of design rules to propose a new code design strategy that addresses code design in channels with spatial and temporal correlations. Our view is that since the channel correlations change much more slowly as compared to the channel realizations, they can be made available to the transmitter and hence exploited (rather than be defended against) to adapt transmissions according to the current statistics.

One of the first attempts to deal with temporal correlation was the design of "smart-greedy" space-time codes of [5]. However, the smart-greedy codes designed for the quasi-static fading channel are only meant to exploit the additional temporal diversity of a fast fading channel. The effect of temporal correlation in the absence of spatial correlation has been addressed in [9]. However, only two special cases were considered in [9], namely, a full rank temporal correlation scenario and the case with the coherence time being less than the number of transmit antennas. Even the so-called "smart-robust" space-time codes of [10] are only designed for the case of a block interleaved model with a full rank temporal correlation matrix. In this paper, we propose a general code design strategy for arbitrary temporal correlation matrices. Our framework clearly shows how the code design criterion is modified as one goes from the quasi-static regime to the fast fading regime via other general temporal correlation matrices.

The effect of spatial correlation in the quasi-static fading regime has been addressed in [8]. It was stated in [8] that a space-time code that leads to full diversity under spatially white fading is sufficient to achieve the maximum diversity of a spatially correlated channel as well. However, in this paper we show that this need not be the optimal strategy. Specifically, for rank deficient spatial fading correlation, we justify the advantage of precoding a reduced dimensional space-time code over using space-time codes designed for the spatially white fading channel. The idea of precoding a reduced rank space-time code for correlated channels has been considered before. For instance, the choice of the precoding matrix with orthogonal space-time block codes [11] has been considered in [12], [13]. For non-orthogonal space-time coding, the choice of precoding matrix is considered in [14]. In these works, the choice of precoding matrices is dictated by the PEP expressions for a pair of space-time codewords. For orthogonal space-time codes, exact expressions of symbol error rates can be used to propose a choice of the optimal precoding matrix as done in [15]. However, no justification is provided in these works for the advantage of precoding reduced dimensional space-time codes over simply using space-time codes that provide full diversity under spatially white fading. Also, the dimensions to choose for the reduced diversity space-time codes is not addressed. Additionally, one should expect a better performance from using full rate space-time codes such as TAST [16] instead of orthogonal designs even for the spatially correlated fading scenario. In our formulation of code design for spatially correlated channels, we fix the precoding matrix to be the one obtained from an eigen-decomposition of the transmit correlation matrix. We also specify that the reduced dimensional space-time code (to be precoded) should be the one designed for as many antennas as the rank of the spatial correlation matrix. We then prove that the resulting class of precoded space-time codes necessarily leads to better pairwise error probabilities than the class of codes that provide full diversity on the spatially white channel.

Spatial correlation was also considered for broadband OFDM multi-antenna systems in [17], [18]. In these works, diversity order of space-frequency codes is discussed under various transmit and receive antenna correlation matrices. Again, the advantage of code design using the precoding strategy is not considered in [17], [18] and, consequently, the exact diversity order for their coding strategy was not obtained under rank deficient spatial correlation. There are also other works such as [7], [19] that derive pairwise error probability expressions under general spatio-temporal correlations. However, they do not explore the relationship between the structure of the correlation matrices and the space-time code design nor do they obtain a succinct analytical formula that captures the diversity order as a function of spatial and temporal correlations. For arbitrarily correlated channels, [19] does not advocate anything beyond using full diversity codes designed for i.i.d channel. The work in [7] suggests robust codes with respect to diversity order but that do not involve precoding. Furthermore, these codes are designed only according to the sum of ranks criterion that is applicable to the special case of the multiple-block quasi-static fading channels but without regard to any further structure in the spatio-temporal correlations. While [7] shows performance gains relative to the smart-greedy codes of [5], it is possible to obtain further significant gains over the robust codes of [7] and design codes for more general channels than those considered in [7] using the systematic code design methodology developed in this work.

In this work, we first provide code design strategies that offer a considerable improvement in the understanding of code design under separate temporal and spatial correlations. Subsequently, we combine the insights gained as a result to propose a general space-time coding strategy for the case of arbitrary spatio-temporal fading correlations. Simulation results confirm the improvement in performance obtained by following the proposed code design strategy instead of using space-time codes that are not specifically designed to utilize the structure of the fading correlation.

Throughout this paper, we consider coherent communications in which the exact channel realization is known to the receiver but unknown to the transmitter. For the case of noncoherent communication, the performance analysis under correlated fading for the special class of Unitary Space-Time Modulation (USTM) codes is presented in [20]. The design of coding schemes for noncoherent systems under nontrivial spatio-temporal fading correlations is addressed in [21].

Even though the results of this paper are presented throughout for frequency non-selective channels with temporal correlations, they are also applicable to frequency-selective channels that employ OFDM in which case the temporal degrees of freedom here are to be interpreted as spectral degrees of freedom with the length of the code T denoting the number of subcarriers and the temporal correlation matrix denoting the correlation between channel gains across subcarriers. More generally, the temporal dimension T in this paper could represent the sum of the temporal and spectral degrees of freedom (as in MIMO-OFDM systems that code across OFDM symbols). Correlations–and often rank-deficient correlations—across subcarriers are more the rule rather than the exception (even in systems that employ frequency-hopping) in such systems.

This paper is organized as follows. The system model and the associated performance analysis is presented in Section II. The special case of non-trivial temporal correlation but spatially white fading is considered in Section III. The special case of quasi-static fading but non-trivial spatial correlation is considered in Section IV. The general case of non-trivial spatio-temporal correlation is considered in Section V. Sample code constructions are discussed in Section VI and conclusions are provided in Section VII.

II. System Model and PEP Analysis

In this paper, the superscripts $*, T$ and $\dagger$ refer to complex conjugate, transpose and hermitian transpose, respectively. The vec() operator stacks the columns of the matrix argument into one vector. The symbols $\circ$ and $\otimes$ refer to the Hadamard product and the Kronecker product, respectively [22]. For an N dimensional vector $\mathbf{x}$, the matrix $\mathrm{diag}(\mathbf{x})$ refers to a diagonal matrix with $\mathbf{x}_n, 1 \leq n \leq N$, being the n-th diagonal entry. For a set of matrices $\{\mathbf{A}_n\}_{n=1}^{N}$, the matrix $\mathrm{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_N)$ refers to the block diagonal matrix with $\mathbf{A}_n, 1 \leq n \leq N$, being the n-th diagonal block.

Consider coding for an N_{T} transmit and N_{R} receive antenna system where the codeword spans a frame of length T time slots. Let $\mathbf{S}$ and $\mathbf{R}$ be constant square matrices of size N_{T} and N_{R}, respectively, that represent square roots of the transmit and receive spatial correlation matrix as introduced in [8]. The fading matrix $\mathbf{H}(t)$ for the t-th time slot is given by the $N_{\mathrm{T}} \times N_{\mathrm{R}}$ matrix $\mathbf{S}\mathbf{W}(t)\mathbf{R}^*$, where the entries of $\mathbf{W}(t)$ for a given t are i.i.d zero-mean unit variance complex normal random variables. However, the T-length vector of the (m,n)-th entries of $\mathbf{W}(t)$ for $1 \le t \le T$ exhibits a correlation matrix $\mathbf{\Sigma}$. In order to have a reasonable code design strategy as proposed in this paper, we may assume that the transmitter knows $\mathbf{S}$ and $\mathbf{\Sigma}$ and that the receiver knows $\mathbf{S}$ and $\mathbf{R}$. We also focus on the coherent scenario which means that the receiver knows the exact values of $\mathbf{H}(t)$ for all t. We consider coding at the physical layer and assume that a higher-layer protocol learns the required correlation matrices using some pilot-based correlation estimation technique and provides the same to the receiver/transmitter.

If $\mathbf{X}(t)$ is the $1 \times N_{\mathrm{T}}$ vector transmitted in the t-th time slot, then the $1 \times N_{\mathrm{R}}$ received vector $Y(t)$ can be written as

$$\begin{aligned} \mathbf{Y}(t) &= \sqrt{\rho}\mathbf{X}(t)\mathbf{S}\mathbf{W}(t)\mathbf{R}^* + \mathbf{N}(t) \quad (1)\\ &= \sqrt{\rho}\tilde{\mathbf{X}}(t)\mathbf{W}(t)\mathbf{R}^* + \mathbf{N}(t), \quad (2)\end{aligned}$$

where $\mathbf{N}(t)$ is the noise vector with i.i.d zero-mean unit variance complex normal entries, ρ is a measure of the signal-to-noise ratio and the superscript ~refers to post multiplication with $\mathbf{S}$. Applying the vec operation to the above set of equations and using the fact that $\mathrm{vec}(\mathbf{A}\mathbf{B}\mathbf{C}^{\mathsf{T}}) = (\mathbf{C} \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$, we get

$$\mathrm{vec}(\mathbf{Y}(t)) = \sqrt{\rho}(\mathbf{I}_{N_{\mathrm{R}}} \otimes \tilde{\mathbf{X}}(t))(\mathbf{R}^\dagger \otimes \mathbf{I}_{N_{\mathrm{T}}})\mathrm{vec}(\mathbf{W}(t)) + \mathrm{vec}(\mathbf{N}(t)). \quad (3)$$

To combine the above for each $1 \le t \le T$, define

$$\begin{aligned} \mathbf{h} &= [\mathrm{vec}(\mathbf{W}(1))^{\mathsf{T}}, \ldots, \mathrm{vec}(\mathbf{W}(T))^{\mathsf{T}}]^{\mathsf{T}}\\ \mathbf{y} &= [\mathrm{vec}(\mathbf{Y}(1))^{\mathsf{T}}, \ldots, \mathrm{vec}(\mathbf{Y}(T))^{\mathsf{T}}]^{\mathsf{T}}\\ \mathbf{n} &= [\mathrm{vec}(\mathbf{N}(1))^{\mathsf{T}}, \ldots, \mathrm{vec}(\mathbf{N}(T))^{\mathsf{T}}]^{\mathsf{T}}\end{aligned}$$

so that

$$\begin{aligned}\mathbf{y} = \sqrt{\rho}\mathrm{diag}((\mathbf{I}_{N_{\mathrm{R}}} \otimes \tilde{\mathbf{X}}(1))(\mathbf{R}^\dagger \otimes \mathbf{I}_{N_{\mathrm{T}}}), \ldots,\\ \ldots, (\mathbf{I}_{N_{\mathrm{R}}} \otimes \tilde{\mathbf{X}}(T))(\mathbf{R}^\dagger \otimes \mathbf{I}_{N_{\mathrm{T}}}))\mathbf{h} + \mathbf{n} \quad (4)\end{aligned}$$

represents the complete input output relationship in the system. Note that $\mathbf{h}$ is a vector of zero-mean complex normal entries with a covariance matrix of $\mathbf{\Sigma} \otimes \mathbf{I}_{N_{\mathrm{R}}N_{\mathrm{T}}}$ and $\mathbf{n}$ consists of i.i.d zero-mean unit variance complex normal entries. Let s, r and σ be the ranks of $\mathbf{S}, \mathbf{R}$ and $\mathbf{\Sigma}$, respectively.

One can interpret the collection of vectors $\mathbf{X}(t)$, for $1 \le t \le T$, as a codeword of a space-time codebook. Let the i-th such codeword be represented as $\{\mathbf{X}_i(t)\}_{t=1}^T$ and for ease of presentation let us refer to it as $\{\mathbf{X}_i(t)\}$. The maximum-likelihood detection rule employed by the receiver under the assumptions made here selects the hypothesis i_{ML} given by

$$\begin{aligned} i_{ML} = \arg\min_j \\ \|\mathbf{y} - \sqrt{\rho}\mathrm{diag}((\mathbf{I}_{N_{\mathrm{R}}} \otimes \tilde{\mathbf{X}}_j(1))(\mathbf{R}^\dagger \otimes \mathbf{I}_{N_{\mathrm{T}}}), \ldots,\\ (\mathbf{I}_{N_{\mathrm{R}}} \otimes \tilde{\mathbf{X}}_j(T))(\mathbf{R}^\dagger \otimes \mathbf{I}_{N_{\mathrm{T}}}))\mathbf{h}\|^2.\end{aligned}$$

With the receiver employing the maximum-likelihood detection rule, the pairwise error probability (PEP) between two possible codewords $\{\mathbf{X}_i(t)\}$ and $\{\mathbf{X}_j(t)\}$ conditioned on a channel realization $\mathbf{h}$ is given by

$$\begin{aligned} P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}/\mathbf{h}) &= Q\left(\sqrt{\frac{\rho}{2}}\|\mathbf{G}_{ij}\mathbf{h}\|\right) \quad (5)\\ &= \frac{1}{\pi}\int_0^{\pi/2} e^{-\frac{\rho}{4\sin^2(\theta)}\|\mathbf{G}_{ij}\mathbf{h}\|^2} d\theta, \quad (6)\end{aligned}$$

where $\mathbf{G}_{ij} = \mathrm{diag}((\mathbf{I}_{N_{\mathrm{R}}} \otimes (\tilde{\mathbf{X}}_i(1) - \tilde{\mathbf{X}}_j(1)))(\mathbf{R}^\dagger \otimes \mathbf{I}_{N_{\mathrm{T}}}), \ldots, (\mathbf{I}_{N_{\mathrm{R}}} \otimes (\tilde{\mathbf{X}}_i(T) - \tilde{\mathbf{X}}_j(T)))(\mathbf{R}^\dagger \otimes \mathbf{I}_{N_{\mathrm{T}}}))$. The Craig's formula $Q(x) = \frac{1}{\pi}\int_0^{\pi/2} e^{-\frac{x^2}{2\sin^2(\theta)}} d\theta, \forall x > 0$, has been used in the second step above. The exact PEP is then obtained by taking the expectation over the distribution of $\mathbf{h}$.

$$P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}) = \frac{1}{\pi}\int_0^{\pi/2} \frac{d\theta}{\left|\mathbf{I} + \frac{\rho}{4\sin^2(\theta)}\mathbf{G}_{ij}(\mathbf{\Sigma} \otimes \mathbf{I}_{N_{\mathrm{T}}N_{\mathrm{R}}})\mathbf{G}_{ij}^\dagger\right|} \quad (7)$$

Useful bounds on the PEP can also be obtained by first simplifying the conditional PEP in (6) before taking the expectation over the distribution of $\mathbf{h}$.

Proposition 1: Define $\mathbf{\Omega}_{ij} = \mathbf{G}_{ij}(\mathbf{\Sigma} \otimes \mathbf{I}_{N_{\mathrm{T}}N_{\mathrm{R}}})\mathbf{G}_{ij}^\dagger$ and let $\delta \in (0, \frac{\pi}{2})$ be a constant. Then,

$$\begin{aligned}\left(\frac{1}{2} - \frac{\delta}{\pi}\right) \cdot \frac{1}{\left|\mathbf{I} + \frac{\rho}{4\sin^2(\delta)}\mathbf{\Omega}_{ij}\right|} \\ \le P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}) \\ \le \frac{1}{2} \cdot \frac{1}{\left|\mathbf{I} + \frac{\rho}{4}\mathbf{\Omega}_{ij}\right|}. \quad (8)\end{aligned}$$

Proof: Replacing θ by $\pi/2$ in the integrand of (6) and then taking the expectation over $\mathbf{h}$, we obtain the Chernoff bound on the pairwise error probability

$$P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}) \le \frac{1}{2} \cdot \frac{1}{\left|\mathbf{I}_{N_{\mathrm{T}}N_{\mathrm{R}}T} + \frac{\rho}{4}\mathbf{G}_{ij}(\mathbf{\Sigma} \otimes \mathbf{I}_{N_{\mathrm{T}}N_{\mathrm{R}}})\mathbf{G}_{ij}^\dagger\right|}, \quad (9)$$

which is the right-side inequality in (8). One can also lower bound the Q-function using the Craig's formula as

$$\begin{aligned} Q(x) &= \frac{1}{\pi}\int_0^{\pi/2} e^{-\frac{x^2}{2\sin^2(\theta)}} d\theta\\ &\ge \frac{1}{\pi}\int_\delta^{\pi/2} e^{-\frac{x^2}{2\sin^2(\theta)}} d\theta\\ &\ge \left(\frac{1}{2} - \frac{\delta}{\pi}\right) e^{-\frac{x^2}{2\sin^2(\delta)}} \quad (10)\end{aligned}$$

Using (10) for the conditional PEP in (6) and taking the expectation over the distribution of $\mathbf{h}$, we obtain the lower bound

$$P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}) \ge \left(\tfrac{1}{2} - \tfrac{\delta}{\pi}\right) \cdot \frac{1}{\left|\mathbf{I}_{N_{\mathrm{T}}N_{\mathrm{R}}T} + \frac{\rho}{4\sin^2(\delta)}\mathbf{G}_{ij}(\mathbf{\Sigma} \otimes \mathbf{I}_{N_{\mathrm{T}}N_{\mathrm{R}}})\mathbf{G}_{ij}^\dagger\right|}, \quad (11)$$

which is the left-side inequality in (8). ■

Let $P_e(\rho)$ denote the actual codeword error probability of the system. The diversity order, defined as

$$d = \lim_{\rho\to\infty} \frac{-\log(P_e(\rho))}{\log(\rho)}, \quad (12)$$

is a measure of performance of the system at high SNR. Define the codeword difference $\mathbf{X}_{ij} = [(\mathbf{X}_i(1) - \mathbf{X}_j(1))^\mathsf{T}, \ldots, (\mathbf{X}_i(T) - \mathbf{X}_j(T))^\mathsf{T}]^\mathsf{T}$ for any $i \neq j$. The following proposition shows the dependency of the diversity order on each of the spatial and temporal correlation matrices and the space-time codebook.

Proposition 2: The diversity order is given by

$$d = \min_{i \neq j} \operatorname{rank}(\mathbf{Z}_{ij}) \tag{13}$$

with

$$\mathbf{Z}_{ij} = ((\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \circ \boldsymbol{\Sigma}) \otimes (\mathbf{R}^\dagger\mathbf{R}) \tag{14}$$

Furthermore, the coding gain is given by the minimum product of the non-zero eigenvalues of $\mathbf{Z}_{ij}$ among all $i \neq j$ such that $\operatorname{rank}(\mathbf{Z}_{ij})$ is the least possible.

Proof: Using Proposition 1, we obtain that the diversity order of the PEP for any $i \neq j$ is given by

$$\lim_{\rho \to \infty} \frac{-\log(P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}))}{\log \rho} = \operatorname{rank}(\mathbf{G}_{ij}(\boldsymbol{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T}N_\mathrm{R}})\mathbf{G}_{ij}^\dagger) \tag{15}$$

We now simplify the expression for $\mathbf{G}_{ij}(\boldsymbol{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T}N_\mathrm{R}})\mathbf{G}_{ij}^\dagger$. Explicitly multiplying the matrices in the product $\mathbf{G}_{ij}(\boldsymbol{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T}N_\mathrm{R}})\mathbf{G}_{ij}^\dagger$, we obtain a block partitioned matrix whose (r,c)-th block is given by

$$\begin{aligned}
&\boldsymbol{\Sigma}_{r,c}(\mathbf{I}_{N_\mathrm{R}} \otimes (\tilde{\mathbf{X}}_i(r) - \tilde{\mathbf{X}}_j(r)))(\mathbf{R}^\dagger \otimes \mathbf{I}_{N_\mathrm{T}}) \times \\
&\qquad (\mathbf{R} \otimes \mathbf{I}_{N_\mathrm{T}})(\mathbf{I}_{N_\mathrm{R}} \otimes (\tilde{\mathbf{X}}_i(c) - \tilde{\mathbf{X}}_j(c))^\dagger) && (16)\\
&= \boldsymbol{\Sigma}_{r,c}(\mathbf{R}^\dagger\mathbf{R} \otimes (\tilde{\mathbf{X}}_i(r) - \tilde{\mathbf{X}}_j(r))(\tilde{\mathbf{X}}_i(c) - \tilde{\mathbf{X}}_j(c))^\dagger) && (17)\\
&= \boldsymbol{\Sigma}_{r,c}(\tilde{\mathbf{X}}_i(r) - \tilde{\mathbf{X}}_j(r))(\tilde{\mathbf{X}}_i(c) - \tilde{\mathbf{X}}_j(c))^\dagger)\mathbf{R}^\dagger\mathbf{R}, && (18)
\end{aligned}$$

where the first equality follows from the fact that $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD})$ and the second because $(\tilde{\mathbf{X}}_i(r) - \tilde{\mathbf{X}}_j(r))(\tilde{\mathbf{X}}_i(c) - \tilde{\mathbf{X}}_j(c))^\dagger)$ is simply a scalar. Using the symbol $\circ$ that represents component-wise multiplication or the Hadamard product, we see that $\mathbf{G}_{ij}(\boldsymbol{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T}N_\mathrm{R}})\mathbf{G}_{ij}^\dagger$ can be written as

$$\left(\left\{\begin{bmatrix} \tilde{\mathbf{X}}_i(1) - \tilde{\mathbf{X}}_j(1) \\ \vdots \\ \tilde{\mathbf{X}}_i(T) - \tilde{\mathbf{X}}_j(T) \end{bmatrix} \times \left[(\tilde{\mathbf{X}}_i(1) - \tilde{\mathbf{X}}_j(1))^\dagger, \ldots, (\tilde{\mathbf{X}}_i(T) - \tilde{\mathbf{X}}_j(T))^\dagger\right]\right\} \circ \boldsymbol{\Sigma}\right) \otimes (\mathbf{R}^\dagger\mathbf{R}).$$

so that we have

$$\mathbf{G}_{ij}(\boldsymbol{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T}N_\mathrm{R}})\mathbf{G}_{ij}^\dagger = ((\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \circ \boldsymbol{\Sigma}) \otimes (\mathbf{R}^\dagger\mathbf{R}). \tag{19}$$

The actual codeword error probability can be upper bounded using the PEPs through the union bound and also lower bounded by considering the PEP between any one given pair of codewords. Hence, we can conclude that the diversity order of the actual codeword error probability is the minimum rank of $\mathbf{Z}_{ij} = ((\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \circ \boldsymbol{\Sigma}) \otimes (\mathbf{R}^\dagger\mathbf{R})$ among all distinct codewords $\{\mathbf{X}_i(t)\}$ and $\{\mathbf{X}_j(t)\}$. Also, the coding gain, defined as the term corresponding to the worst case PEP at high SNR, is as described in the proposition. ■

The code design principle, therefore, is to maximize the minimum rank of $\mathbf{Z}_{ij}$ over all $i \neq j$. The obtained expression for diversity order and coding gain can be shown to coincide with the well known special cases.

Corollary 1: For quasi-static fading ($\boldsymbol{\Sigma} = \mathbf{1}\mathbf{1}^\mathsf{T}$), the diversity order is given by

$$d = \min_{i \neq j} \operatorname{rank}(\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \otimes (\mathbf{R}^\dagger\mathbf{R}) \tag{20}$$

The above corollary recovers the result of the quasi-static spatially correlated channel in [8]. For any two square matrices $\mathbf{A}$ and $\mathbf{B}$, the eigenvalues of $\mathbf{A} \otimes \mathbf{B}$ are given by the product of the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$. This also implies that $\operatorname{rank}(\mathbf{A} \otimes \mathbf{B}) = \operatorname{rank}(\mathbf{A})\operatorname{rank}(\mathbf{B})$. These facts lead to the following corollary for the spatially white fading channel.

Corollary 2: For spatially white fading, i. e., $\mathbf{S} = \mathbf{I}_{N_\mathrm{T}}$ and $\mathbf{R} = \mathbf{I}_{N_\mathrm{R}}$, the diversity order is given by

$$d = N_\mathrm{R} \times \min_{i \neq j} \operatorname{rank}((\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger) \circ \boldsymbol{\Sigma}) \tag{21}$$

The coding gain is the minimum product of the non-zero eigenvalues of $((\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger) \circ \boldsymbol{\Sigma})$ over all $i \neq j$, such that the rank of $((\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger) \circ \boldsymbol{\Sigma})$ is the least possible. Further special cases are as follows.

1) Quasi-static fading, $\boldsymbol{\Sigma} = \mathbf{1}\mathbf{1}^\mathsf{T}$:

$$d = N_\mathrm{R} \times \min_{i \neq j} \operatorname{rank}(\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger) \tag{22}$$

Code design becomes maximization of the minimum rank of $\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger$ and the coding gain is the maximization of the determinant of $\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger$ among all $i \neq j$ such that the rank of $\mathbf{X}_{ij}$ is the least possible. This is referred to as the classical rank and determinant criterion.

2) Fast fading, $\boldsymbol{\Sigma} = \mathbf{I}$:

$$d = N_\mathrm{R} \times \min_{i \neq j} \operatorname{rank} \left(\operatorname{diag}(\|\mathbf{X}_i(1) - \mathbf{X}_j(1)\|^2, \ldots, \|\mathbf{X}_i(T) - \mathbf{X}_j(T)\|^2)\right)$$

Code design criterion reduces to maximizing the number of rows differing between $\mathbf{X}_i$ and $\mathbf{X}_j$ or equivalently the number of non-zero rows of $\mathbf{X}_{ij}$. The coding gain is the minimum product of the Euclidean norms of the rows of $\mathbf{X}_{ij}$ over all $i \neq j$ such that $\mathbf{X}_{ij}$ has the least number of non-zero rows. This is referred to as the row Hamming distance criterion.

3) M independent fading blocks of coherence time T', $\boldsymbol{\Sigma} = \mathbf{I}_M \otimes \mathbf{1}_{T'}\mathbf{1}_{T'}^\mathsf{T}$:

$$d = N_\mathrm{R} \times \min_{i \neq j} \sum_{m=1}^{M} \operatorname{rank}(\mathbf{X}_{ij}(m)\mathbf{X}_{ij}(m)^\dagger), \tag{23}$$

where $\{\mathbf{X}_{ij}(m)\}_{m=1}^{M}$ are the M sub-matrices of $\mathbf{X}_{ij}$ obtained by taking the consecutive T' rows of $\mathbf{X}_{ij}$. This is referred to as the *sum of ranks* criterion.

The code design guidelines described for these special cases for spatially white fading have been obtained earlier in several works [5], [16]. The general expression for $\mathbf{Z}_{ij}$ given in Proposition 2 must, however, be used to design codebooks for any given spatio-temporal correlation matrices. We first consider the cases for temporal and spatial correlations separately, and then finally address the issue of code design for arbitrary spatio-temporal correlation structures.

III. Spatially White Fading with Non-Trivial Temporal Correlation

For spatially white fading ($\mathbf{S} = \mathbf{I}_{N_\text{T}}$ and $\mathbf{R} = \mathbf{I}_{N_\text{R}}$), the diversity order result of Corollary 2 implies that we need to maximize the minimum rank of $(\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger) \circ \mathbf{\Sigma}$ among all codeword pairs. Some preliminary observations can be made using the following proposition that specifies when codes designed for quasi-static and fast fading channels can be leveraged for other temporal correlation matrices.

Proposition 3: Suppose $\mathbf{S} = \mathbf{I}_{N_\text{T}}$ and $\mathbf{R} = \mathbf{I}_{N_\text{R}}$.

1) If $T \leq N_\text{T}$, then for any $\mathbf{\Sigma}$, a code that has the maximum diversity order of TN_R under quasi-static fading will also have the same maximum diversity order of TN_R over the channel with temporal correlation $\mathbf{\Sigma}$. Moreover, the coding gain over the correlated channel is lower bounded by that for the quasi-static channel.
2) If $\mathbf{\Sigma}$ has full rank, then for any T and N_T, a code that has the maximum diversity order of TN_R under fast fading will also have the same maximum diversity order of TN_R over the channel with temporal correlation $\mathbf{\Sigma}$. Moreover, the coding gain over the correlated channel is upper bounded by that for the fast fading channel.

Proof: We invoke the following inequality known as the Oppenheim inequality for any two positive semidefinite matrices $\mathbf{A}$ and $\mathbf{B}$ of size n [22].

$$\det(\mathbf{A}) \cdot \prod_{i=1}^{n} \mathbf{B}_{ii} \leq \det(\mathbf{A} \circ \mathbf{B}). \tag{24}$$

For the first part of the proposition, set $\mathbf{B} = \mathbf{\Sigma}$ and $\mathbf{A} = \mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger$ in (24) and note that the diagonal entries of $\mathbf{\Sigma}$ are all 1. One can then see that ensuring $T \leq N_\text{T}$ and full rank for $\mathbf{X}_{ij}$ guarantees that $(\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger) \circ \mathbf{\Sigma}$ has rank T and, therefore, that the diversity order is TN_R. Moreover, the coding gain expressions for the quasi-static and the temporally correlated channels are given by the minimum over $i \neq j$ of the left and right hand sides of (24), respectively.

For the second part of the proposition, if $\mathbf{\Sigma}$ is known to have full rank, then setting $\mathbf{A} = \mathbf{\Sigma}$ and $\mathbf{B} = \mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger$ in (24) shows that a diversity order of TN_R can be achieved if all rows of $\mathbf{X}_{ij}$ are non-zero. This coincides with the row hamming distance criterion for the fast fading channel. Also, we have the Hadamard inequality,

$$\det(\mathbf{A} \circ \mathbf{B}) \leq \prod_{i=1}^{n} (\mathbf{A} \circ \mathbf{B})_{ii}, \tag{25}$$

which implies that the coding gain for the correlated channel is upper bounded by the coding gain for the fast fading channel. ∎

The diversity order results under the special cases of Proposition 3 have also been noted in [9]. Proposition 3, however, cannot be used for code design when $T > N_\text{T}$ and $\mathbf{\Sigma}$ is not full rank. Therefore, we need a design strategy that utilizes the structure of the correlation matrix $\mathbf{\Sigma}$ and that also works for any arbitrary $\mathbf{\Sigma}$ and T. We now provide insights into such a design method for arbitrary correlation matrices.

Using the Cholesky decomposition [22], one can write $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^\dagger$, where $\mathbf{L}$ is a lower triangular matrix. If $\sigma < T$, then $\mathbf{L}$ need not be unique. We impose a particular structure on $\mathbf{L}$ so that the last $T - \sigma$ columns of $\mathbf{L}$ are zero. Also, if p_k denotes the row index of the first non-zero entry of the k-th column $\mathbf{L}_k$ of $\mathbf{L}$, then we require that $1 = p_1 < p_2 \cdots < p_\sigma \leq T$. Define the block lower triangular matrix

$$\widehat{\mathbf{X}}_{ij} = [\text{diag}(\mathbf{L}_1)\mathbf{X}_{ij}, \ldots, \text{diag}(\mathbf{L}_\sigma)\mathbf{X}_{ij}].$$

The diversity order and coding gain results are first stated in relation to the matrix $\widehat{\mathbf{X}}_{ij}$.

Proposition 4: The diversity order of the temporally correlated channel is given by

$$d = N_\text{R} \times \min_{i \neq j} \text{rank}(\widehat{\mathbf{X}}_{ij}). \tag{26}$$

Also, the coding gain is the minimum product of the non-zero eigenvalues of $\widehat{\mathbf{X}}_{ij}\widehat{\mathbf{X}}_{ij}^\dagger$ over all $i \neq j$ such that the rank of $\widehat{\mathbf{X}}_{ij}$ is the least possible.

Proof: Using the result of Corollary 2, we only need to prove that

$$(\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger) \circ \mathbf{\Sigma} = \widehat{\mathbf{X}}_{ij}\widehat{\mathbf{X}}_{ij}^\dagger. \tag{27}$$

This follows from a more general result on the Hadamard products. If $\mathbf{A}, \mathbf{B}$ are complex matrices of size $T \times m$ and $T \times n$, respectively, and $\mathbf{B}_k$ is the k-th column of $\mathbf{B}$, $1 \leq k \leq n$, then we show that

$$\mathbf{A}\mathbf{A}^\dagger \circ \mathbf{B}\mathbf{B}^\dagger = \mathbf{C}\mathbf{C}^\dagger, \tag{28}$$

where

$$\mathbf{C} = \begin{bmatrix} \text{diag}(\mathbf{B}_1)\mathbf{A} & \text{diag}(\mathbf{B}_2)\mathbf{A} & \ldots & \text{diag}(\mathbf{B}_n)\mathbf{A} \end{bmatrix}. \tag{29}$$

For any two column vectors $\mathbf{a}$ and $\mathbf{b}$, it is easily verified that

$$(\mathbf{a}\mathbf{a}^\dagger \circ \mathbf{b}\mathbf{b}^\dagger) = (\mathbf{a} \circ \mathbf{b})(\mathbf{a} \circ \mathbf{b})^\dagger. \tag{30}$$

Also, the Hadamard product satisfies the distributive law. Therefore, denoting the i-th column of $\mathbf{A}$ by $\mathbf{A}_i$, we have

$$\begin{aligned}
\mathbf{A}\mathbf{A}^\dagger \circ \mathbf{B}\mathbf{B}^\dagger &= \left(\sum_i \mathbf{A}_i\mathbf{A}_i^\dagger\right) \circ \left(\sum_j \mathbf{B}_j\ \mathbf{B}_j^\dagger\right) \\
&= \sum_i \sum_j \mathbf{A}_i\mathbf{A}_i^\dagger \circ \mathbf{B}_j\mathbf{B}_j^\dagger \\
&= \sum_i \sum_j (\mathbf{A}_i \circ \mathbf{B}_j)(\mathbf{A}_i \circ \mathbf{B}_j)^\dagger \\
&= \sum_i \sum_j (\text{diag}(\mathbf{B}_j)\mathbf{A}_i)(\text{diag}(\mathbf{B}_j)\mathbf{A}_i)^\dagger \\
&= \sum_j \text{diag}(\mathbf{B}_j) \left(\sum_i \mathbf{A}_i\mathbf{A}_i^\dagger\right) \text{diag}(\mathbf{B}_j)^\dagger \\
&= \sum_j \text{diag}(\mathbf{B}_j)\mathbf{A}\mathbf{A}^\dagger\text{diag}(\mathbf{B}_j)^\dagger \\
&= \mathbf{C}\mathbf{C}^\dagger,
\end{aligned}$$

where $\mathbf{C}$ is as defined above.

The required result of (27) now follows from (28) by setting $\mathbf{A} = \mathbf{X}_{ij}$ and $\mathbf{B} = \mathbf{L}$. ∎

Corollary 3: The maximum diversity order under spatially white correlation is upper bounded as

$$d \leq N_\text{R} \times \min(\sigma N_\text{T}, T). \tag{31}$$

Our goal is to utilize the lower triangular nature of $\mathbf{L}$ for obtaining a criterion on the codeword difference $\mathbf{X}_{ij}$ that would lead to a lower bound on the rank of $\widehat{\mathbf{X}}_{ij}$. The proposed code design strategy is described in the next proposition.

Proposition 5: For $1 \leq k \leq \sigma - 1$, let $D_k \subset [p_k, p_{k+1} - 1]$ denote the index set corresponding to the row indices between p_k and $p_{k+1} - 1$ such that the entries of $\mathbf{L}_k$ on those rows are non-zero. Also, let $D_\sigma \subset [p_\sigma, T]$ be the index set corresponding to the row indices such that the entries of $\mathbf{L}_\sigma$ on those rows are non-zero. Thus,

$$\mathbf{L}_k(j) \neq 0, \forall j \in D_k, 1 \leq k \leq \sigma.$$

Let $d_k = |D_k|, 1 \leq k \leq \sigma$. Design the space-time codebook such that for each signal pair $\{\mathbf{X}_i(t)\}$ and $\{\mathbf{X}_j(t)\}$, the submatrix of $\mathbf{X}_{ij}$ obtained by taking the rows corresponding to the index set D_k has full rank $\min(d_k, N_{\text{T}})$ for each $k \in [1, \sigma]$. Then, the diversity order of the temporally correlated channel is lower bounded as

$$d \geq N_{\text{R}} \times \sum_{k=1}^{\sigma} \min(d_k, N_{\text{T}}), \tag{32}$$

Proof: The rank of any block lower triangular matrix is lower bounded by the sum of the ranks of the diagonal blocks [22]. When the conditions imposed on the structure of $\mathbf{X}_{ij}$ are met for each $i \neq j$, each term in the summation of (32) gives the rank of a diagonal block of the block lower triangular matrix $\widehat{\mathbf{X}}_{ij}$. ■

The proposed code design procedure for any given $\mathbf{\Sigma}$, therefore, is to ensure that conditions of Proposition 5 are met for each codeword pair. As described at the end of this section, this lower bound on the diversity order could be further improved by exploring all possible permutations of stacking the entries of $\mathbf{W}(t)$ into $\mathbf{h}$.

The coding gain with the proposed strategy under spatially white correlation corresponds to the product of the absolute determinant squared of the diagonal blocks of $\widehat{\mathbf{X}}_{ij}$, but only if these diagonal blocks are square. Even if the diagonal blocks of $\widehat{\mathbf{X}}_{ij}$ are not square but the matrix $\widehat{\mathbf{X}}_{ij}$ is block diagonal, then the coding gain can be written as the product of the determinants of the Hermitian form of the matrices on the diagonal of $\widehat{\mathbf{X}}_{ij}$. Besides these special cases, it is difficult to obtain a coding gain expression that only depends on the signal matrices.

It is easy to verify that the proposed strategy subsumes the code design criterion for the three cases corresponding to quasi-static fading ($\mathbf{\Sigma} = \mathbf{1}_T \mathbf{1}_T^{\mathsf{T}}, \mathbf{L} = [\mathbf{1}_T \quad \mathbf{0}_{T \times T-1}]$), fast fading ($\mathbf{\Sigma} = \mathbf{I}_T, \mathbf{L} = \mathbf{I}_T$) and multiple block i.i.d fading ($\mathbf{\Sigma} = \mathbf{I}_M \otimes \mathbf{1}_{T'} \mathbf{1}_{T'}^{\mathsf{T}}, \mathbf{L} = [\mathbf{I}_M \otimes \mathbf{1}_{T'} \quad \mathbf{0}_{T \times T-M}]$) . Moreover, this general strategy guarantees a certain diversity order depending on the structure of $\mathbf{\Sigma}$ even if the system parameters do not allow us to make any conclusions using Proposition 3.

Since the guaranteed diversity order using the proposed strategy depends on the properties of the Cholesky decomposition of $\mathbf{\Sigma}$, we can exploit the order of grouping of the entries of $\mathbf{W}(t), 1 \leq t \leq T$, into $\mathbf{h}$ to obtain the most favorable Cholesky decomposition. Specifically, let $\pi(.)$ be a permutation that sends $t, 1 \leq t \leq T$, to $\pi(t)$ and let $\mathbf{P}$ be the corresponding permutation matrix of size T so that $\mathbf{P}(t, \pi(t)) = 1$. We can then rewrite (4) as

$$\check{\mathbf{y}} = \sqrt{\rho}\text{diag}((\mathbf{I}_{N_{\text{R}}} \otimes \mathbf{X}(\pi(1))), \ldots, (\mathbf{I}_{N_{\text{R}}} \otimes \mathbf{X}(\pi(T))))\check{\mathbf{h}} + \check{\mathbf{n}},$$

where

$$\begin{aligned} \check{\mathbf{y}} &= (\mathbf{P}^{\mathsf{T}} \otimes \mathbf{I}_{N_{\text{T}} N_{\text{R}}})\mathbf{y} \\ \check{\mathbf{h}} &= (\mathbf{P}^{\mathsf{T}} \otimes \mathbf{I}_{N_{\text{T}} N_{\text{R}}})\mathbf{h} \\ \check{\mathbf{n}} &= (\mathbf{P}^{\mathsf{T}} \otimes \mathbf{I}_{N_{\text{T}} N_{\text{R}}})\mathbf{n} \end{aligned}$$

The vector $\check{\mathbf{h}}$ is now complex Gaussian with the covariance matrix $\check{\mathbf{\Sigma}} \otimes \mathbf{I}_{N_{\text{T}} N_{\text{R}}}$ where $\check{\mathbf{\Sigma}} = \mathbf{P}^{\mathsf{T}} \mathbf{\Sigma} \mathbf{P}$. The vector $\check{\mathbf{n}}$ consists of i.i.d zero mean unit variance complex normal entries. Let us now define $\check{\mathbf{X}}_{ij} = [(\mathbf{X}_i(\pi(1)) - \mathbf{X}_j(\pi(1)))^{\mathsf{T}}, \ldots, (\mathbf{X}_i(\pi(T)) - \mathbf{X}_j(\pi(T)))^{\mathsf{T}}]^{\mathsf{T}}$. Following the same design criterion as implied in Proposition 5, but with $\mathbf{\Sigma}$ replaced by $\check{\mathbf{\Sigma}}$ and imposing the rank constraints on $\check{\mathbf{X}}_{ij}$ instead of $\mathbf{X}_{ij}$, one can guarantee a diversity order of at least

$$\sum_{k=1}^{\sigma} \min(\check{d}_k, N_{\text{T}}), \tag{33}$$

where $\check{d}_k$ are the corresponding values for the Cholesky factor $\check{\mathbf{L}}$ of $\check{\mathbf{\Sigma}}$. One could, therefore, choose the permutation matrix $\mathbf{P}$ that leads to the largest value of this guaranteed diversity order. This would only require an off-line search among all $T!$ possible permutation matrices for a given correlation matrix $\mathbf{\Sigma}$. Also, note that the optimal matrix $\mathbf{P}$ is only needed to design the codebook and is not needed by the receiver to perform the actual decoding.

While an exhaustive search can provide the best permutation for our strategy, it is possible to choose the matrix $\mathbf{P}$ for a given Cholesky factor $\mathbf{L}$ to improve the lower bound on the diversity order. Since $\check{\mathbf{\Sigma}} = \mathbf{P}^{\mathsf{T}} \mathbf{\Sigma} \mathbf{P} = (\mathbf{P}^{\mathsf{T}} \mathbf{L})(\mathbf{P}^{\mathsf{T}} \mathbf{L})^{\dagger}$, we can restrict ourselves to a permutation $\mathbf{P}$ that preserves the required lower triangular structure for $\mathbf{P}^{\mathsf{T}} \mathbf{L}$. The permutation $\mathbf{P}^{\mathsf{T}}$ can be used to rearrange the rows of $\mathbf{L}$ as follows. For each $k \in [2, \sigma]$, those rows corresponding to $[p_k, p_{k+1} - 1] \backslash D_k$ with only the first column being non-zero should be moved up above row index p_2. If any such rows exist, then this permutation would effectively increase the value of d_1. Then, for each $k \in [3, \sigma]$, those rows corresponding $[p_k, p_{k+1} - 1] \backslash D_k$ in the original $\mathbf{L}$ matrix with the second column non-zero and higher columns zero should also be moved up to lie between row indices p_2 and p_3. Again, if such rows exist, then this permutation effectively increases the value of d_2. Proceeding in this fashion, it is possible to arrive at a permutation $\mathbf{P}$ such that the lower triangular matrix $\mathbf{P}^{\mathsf{T}} \mathbf{L}$ has equal or higher values of the d_k's than the original matrix $\mathbf{L}$. Thus, this particular permutation leads to a higher value of the guaranteed diversity order with the proposed strategy.

An example to illustrate this permutation is shown in Fig. 1 for $N_{\text{T}} = 4, N_{\text{R}} = 1, T = 9$ and $\sigma = 3$. Only the non-zero columns of the original $\mathbf{L}$ matrix and the effect of the proposed permutations on them are shown. The values of (d_1, d_2, d_3) in this example improve from $(2, 3, 1)$ to $(4, 4, 1)$ through the proposed row permutations. This effectively improves the lower bound on diversity order from 6 to the maximum achievable diversity order 9.

$$\underset{\text{Diversity order 6}}{\overset{(d_1,d_2,d_3)=(2,3,1)}{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0.7071 & 0.7071 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0.7071 & 0.7071 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0.7071 & 0.7071 & 0 \end{bmatrix}}} \Longrightarrow \overset{(d_1,d_2,d_3)=(4,3,1)}{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0.7071 & 0.7071 & 0 \\ 0 & 1 & 0 \\ 0.7071 & 0.7071 & 0 \\ 0 & 0 & 1 \\ 0.7071 & 0.7071 & 0 \end{bmatrix}} \Longrightarrow \underset{\text{Diversity order 9}}{\overset{(d_1,d_2,d_3)=(4,4,1)}{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0.7071 & 0.7071 & 0 \\ 0 & 1 & 0 \\ 0.7071 & 0.7071 & 0 \\ 0.7071 & 0.7071 & 0 \\ 0 & 0 & 1 \end{bmatrix}}}$$

Fig. 1. Row Permutations to increase the d_k's one at a time for a sample $\mathbf{L}$ matrix (Only the first non-zero columns of $\mathbf{L}$ are shown). $N_{\mathrm{T}} = 4, N_{\mathrm{R}} = 1, T = 9, \sigma = 3$.

IV. Spatial Correlation with Quasi-Static Fading

In the quasi-static scenario with spatial correlation, we have from Corollary 1 that the diversity order is the minimum rank of $\mathbf{Z}_{ij} = (\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \otimes (\mathbf{R}^\dagger\mathbf{R})$. The rank of $\mathbf{Z}_{ij}$ is the product of the ranks of $\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger$ and $\mathbf{R}^\dagger\mathbf{R}$ and is therefore upper bounded by $\min(T, \mathrm{rank}(\mathbf{S}))\mathrm{rank}(\mathbf{R})$. As noted in [8], choosing a space-time code that leads to full diversity N_{T} under spatially white fading will lead to the maximum diversity of rank($\mathbf{S}$)rank($\mathbf{R}$) under arbitrary spatial correlation. This was also the reason space-time codes designed for the spatially white channel were used for simulations on spatially correlated channels in [23]. If the space-time code is rank deficient under spatially white fading, then, as observed in [8], the exact diversity order under arbitrary spatial correlation can not be determined easily.

In this section, we present a space-time coding strategy that guarantees the maximum diversity order of $\min(T, \mathrm{rank}(\mathbf{S})) \cdot \mathrm{rank}(\mathbf{R})$ under spatially correlated fading even though it may not achieve full diversity for spatially white fading. We prove that, whenever $T \geq N_{\mathrm{T}}$, the new strategy is better in terms of coding gain compared to the suggestion in [8] of using a space-time code that provides full diversity under spatially white fading. The new strategy is inspired from the structure of the optimum input distribution obtained from a mutual information point of view [24], [25]. Suppose the eigenvalue decomposition of $\mathbf{S}\mathbf{S}^\dagger$ is given by $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger$, where $\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$ and $\mathbf{\Lambda} = \mathrm{diag}(\mathbf{\Lambda}_s, \mathbf{0})$ with $\mathbf{\Lambda}_s$ being a square positive diagonal matrix of size s. Then, the idea is to use a space-time codebook of the form $\begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \mathbf{U}^\dagger$, where $\mathbf{C}$ is a codeword from a $T \times s$ space-time codebook. Thus, we precode a smaller size space-time codebook using the eigenvectors obtained from the transmit correlation matrix.

Consider the expression for the exact PEP obtained in (7). Since the eigenvalues of $\mathbf{A} \otimes \mathbf{B}$ are the product of the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$, we obtain that

$$\begin{aligned} &P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}) \\ &= \frac{1}{\pi}\int_0^{\pi/2} \frac{d\theta}{\left|\mathbf{I} + \frac{\rho}{4\sin^2(\theta)}\mathbf{G}_{ij}^\dagger\mathbf{G}_{ij}\right|} \\ &= \frac{1}{\pi}\int_0^{\pi/2} \prod_{n=1}^{r}\prod_{m=1}^{v} \left(1 + \frac{\rho}{4\sin^2(\theta)}\lambda_m(\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \times \right. \\ &\qquad \left. \lambda_n(\mathbf{R}^\dagger\mathbf{R})\right)^{-1} d\theta, \end{aligned} \tag{34}$$

where $\lambda_m(.)$ and $\lambda_n(.)$ represent the m-th and n-th non-zero eigenvalues of the matrix argument and v is the rank of $\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger$.

In the next proposition, we compare two classes of space-time codes both achieving the diversity order of rank($\mathbf{S}$)rank($\mathbf{R}$) and show that one of these classes of codes is better in terms of code design through pairwise error probabilities. We consider the non-trivial case of $s = \mathrm{rank}(\mathbf{S}) < N_{\mathrm{T}}$.

Proposition 6: Let $T \geq N_{\mathrm{T}}$. Let $\mathcal{C}$ be a space-time codebook consisting of $T \times N_{\mathrm{T}}$ sized matrices. Let $\mathrm{rank}(\mathbf{C}_i - \mathbf{C}_j) = N_{\mathrm{T}}, \forall \mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, i \neq j$. Consider another space-time codebook $\mathcal{C}'$ consisting of $T \times s$ sized matrices obtained by retaining only the first s columns of the space-time codewords in $\mathcal{C}$. Consider two space-time codebooks $\mathcal{X} = \mathcal{C}\mathbf{U}^\dagger$ and $\mathcal{X}' = \alpha[\ \mathcal{C}' \quad \mathbf{0}_{T \times N_{\mathrm{T}} - s}\ ]\mathbf{U}^\dagger$, where the constant $\alpha > 1$ is given by

$$\alpha = \left(\frac{\mathrm{E}_{\mathbf{C}\in\mathcal{C}}[\|\mathbf{C}\|^2]}{\mathrm{E}_{\mathbf{C}'\in\mathcal{C}'}[\|\mathbf{C}'\|^2]}\right)^{1/2}, \tag{35}$$

so that both $\mathcal{X}$ and $\mathcal{X}'$ exhibit the same average energy. Let $\mathcal{X}$ and $\mathcal{X}'$ be used on the spatially correlated channel. Then, both $\mathcal{X}$ and $\mathcal{X}'$ exhibit a diversity order of $s \cdot r$. However, for each pair of codewords $\mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, i \neq j$, that lead to pairs of codewords in $\mathcal{X}$ and $\mathcal{X}'$, the pairwise error probability

$$P(\mathbf{X}_i' \to \mathbf{X}_j') \leq P(\mathbf{X}_i \to \mathbf{X}_j) \tag{36}$$

at any given SNR ρ.

Proof: The codeword difference $(\mathbf{C}_i' - \mathbf{C}_j')$ has rank s for each $\mathbf{C}_i', \mathbf{C}_j' \in \mathcal{C}', i \neq j$, because each codeword difference in $\mathcal{C}$ has rank N_{T} and $T \geq N_{\mathrm{T}}$. Hence, the rank of $\mathbf{Z}_{ij} = (\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \otimes (\mathbf{R}^\dagger\mathbf{R})$ is equal to $s \cdot r$ when the codeword pair $(\mathbf{X}_i, \mathbf{X}_j)$ is drawn from either of $\mathcal{X}$ or $\mathcal{X}'$. Let $\mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, i \neq j$ and let $\mathbf{C}_i'$ and $\mathbf{C}_j'$ be the corresponding codewords in $\mathcal{C}'$, $\mathbf{X}_i$ and $\mathbf{X}_j$ the corresponding codewords in $\mathcal{X}$ and $\mathbf{X}_i'$ and $\mathbf{X}_j'$ the corresponding codewords in $\mathcal{X}'$. Let $\mathbf{X}_{ij} = \mathbf{X}_i - \mathbf{X}_j, \mathbf{X}_{ij}' = \mathbf{X}_i' - \mathbf{X}_j', \mathbf{C}_{ij} = \mathbf{C}_i - \mathbf{C}_j, \mathbf{C}_{ij}' = \mathbf{C}_i' - \mathbf{C}_j'$. Looking at the PEP expression in (34), we only need to compare the eigenvalues of $\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}$ and $\mathbf{X}'_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}'_{ij}$. We have that

$$\begin{aligned} \lambda_m(\mathbf{X}'_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}'_{ij}) &= \lambda_m(\alpha[\mathbf{C}'_{ij}\mathbf{0}]\Lambda\alpha[\mathbf{C}'_{ij}\mathbf{0}]^\dagger) && (37)\\ &= \alpha^2\lambda_m(\mathbf{C}'_{ij}\Lambda_s\mathbf{C}'^{\dagger}_{ij}) && (38)\\ &= \alpha^2\lambda_m(\mathbf{C}_{ij}\Lambda\mathbf{C}^\dagger_{ij}) && (39)\\ &= \alpha^2\lambda_m(\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}^\dagger_{ij}). && (40) \end{aligned}$$

Since $\alpha > 1$, we see that $\lambda_m(\mathbf{X}'_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}'_{ij}) > \lambda_m(\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}^\dagger_{ij})$ and so from (34), we get that

$$P(\mathbf{X}'_i \to \mathbf{X}'_j) \le P(\mathbf{X}_i \to \mathbf{X}_j). \quad (41)$$

■

We now show that for spatially correlated channels with $s < N_\text{T}$, precoding a reduced size space-time codebook with $\mathbf{U}^\dagger$ is a better strategy than simply using a space-time code that leads to full diversity under spatially white fading.

Proposition 7: Let $T \ge N_\text{T}$ and let $\mathcal{X}$ be a space-time code with $\text{rank}(\mathbf{X}_i - \mathbf{X}_j) = N_\text{T}, \forall \mathbf{X}_i, \mathbf{X}_j \in \mathcal{X}, i \ne j$. Then, there exists a $T \times s$ space-time code $\mathcal{C}'$ such that the new space-time codebook $\mathcal{X}' \equiv [\mathcal{C}'\mathbf{0}]\mathbf{U}^\dagger$ exhibits diversity $s \cdot r$ and the PEP for each pair of codewords in $\mathcal{X}'$ is smaller than the corresponding PEP in $\mathcal{X}$.

Proof: One can write $\mathcal{X} \equiv (\mathcal{X}\mathbf{U})\mathbf{U}^\dagger \equiv \mathcal{C}\mathbf{U}^\dagger$, where $\mathcal{C} \equiv \mathcal{X}\mathbf{U}$ is a code that meets the conditions of Proposition 6. Thus, there exists a code $\mathcal{X}'$ of the form $[\mathcal{C}'\mathbf{0}]\mathbf{U}^\dagger$, where $\mathcal{C}'$ is a space-time codebook of size $T \times s$, with the PEP for each codeword pair in $\mathcal{X}'$ smaller than the corresponding PEP in $\mathcal{X}$. ■

Using Proposition 7, we can conclude that a space-time code that leads to full diversity $N_\text{T}N_\text{R}$ under spatially white fading will provide full diversity order of $\text{rank}(\mathbf{S})\text{rank}(\mathbf{R})$ for arbitrary spatial correlations also but it is not optimum in terms of coding gain and PEP minimization when $\text{rank}(\mathbf{S}) < N_\text{T}$. The strategy of precoding a reduced dimensional space-time code, i. e., $[\mathcal{C}'\mathbf{0}]\mathbf{U}^\dagger$ as described in Proposition 7, also leads to the maximum diversity of $\text{rank}(\mathbf{S})\text{rank}(\mathbf{R})$ and the optimum selection of such a space-time code will necessarily lead to smaller PEPs. Note that one can expand the class of precoded space-time codes to be of the form $[\mathcal{C}\mathbf{0}]\mathbf{U}^\dagger$, where $\mathcal{C}$ is any code of size $T \times s$ with $\text{rank}(\mathbf{C}_i - \mathbf{C}_j) = s, \forall \mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, i \ne j$, and not necessarily restrict to codes such as $\mathcal{C}'$ described in the proof of Proposition 7. By including more general space-time codes for use in the precoding strategy, we not only explore the possibility of an improvement in performance but also exploit the low decoding complexities associated with some space-time codes. Specific choices for the the space-time codebook $\mathcal{C}$ are discussed next.

One may consider the space-time codebook $\mathcal{C}$ to be of the form $\mathcal{T}\Upsilon$, where $\mathcal{T}$ is a $T \times s$ code that could either be a rectangular LPST code [26], a stacked extension of a code [27] or a rectangular TAST code [28]. The matrix $\Upsilon = \text{diag}(\sqrt{P_1}, \ldots, \sqrt{P_s})$ is a diagonal matrix describing the power allocation on each of the s columns of the code $\mathcal{T}$. For each choice of $\mathcal{T}$ mentioned above, the decoding at the receiver is simplified by means of the sphere-decoding algorithm [29]. If the average energy on each column of the code in $\mathcal{T}$ is the same, then one may consider the choice of P_i as obtained from a mutual information maximization point of view which can be found in [24], [25]. In the case of high SNR, it is easy to show that setting all P_i to be the same is optimal in terms of coding gain obtained from a pairwise error probability point of view. This fact is in accordance with the high SNR optimality of the equal power allocation on the s eigenmodes in terms of maximizing mutual information [30].

V. Spatio-Temporal Correlation

We now combine the insights gained in Sections III and IV to propose a coding strategy under arbitrary spatio-temporal fading correlation. As before, the diversity order expression is given by the minimum rank of $\mathbf{Z}_{ij} = ((\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}^\dagger_{ij}) \circ \mathbf{\Sigma}) \otimes (\mathbf{R}^\dagger\mathbf{R})$. To deal with the spatial correlation, we consider space-time codes with the precoding structure obtained in Section IV, namely $\mathbf{X} \equiv [\mathcal{C}\ \mathbf{0}]\mathbf{U}^\dagger$, where $\mathcal{C}$ is a $T \times s$ space-time code. Since we now also have a non-trivial temporal correlation matrix $\mathbf{\Sigma}$, the strategy proposed in Section III can be applied to choose the code $\mathcal{C}$ appropriately in order to further exploit the diversity inherent in the temporal dimension. The code design structure via the rank constraints imposed on the matrix $\mathbf{X}_{ij}$ in Proposition 5 should now be applied to the matrix difference of codewords in $\mathcal{C}$. The code design guideline and the diversity order result is summarized in the following proposition. The index sets D_k and the numbers d_k are the same as defined in Proposition 5.

Proposition 8: In the precoding strategy with $\mathcal{X} = [\mathcal{C}\ \ \mathbf{0}]\mathbf{U}^\dagger$, where $\mathcal{C}$ is a $T \times s$ space-time code, let each codeword pair $\mathbf{C}_{ij} = \mathbf{C}_i - \mathbf{C}_j, \mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, i \ne j$, be such that the submatrix of $\mathbf{C}_{ij}$ obtained by taking the rows corresponding to the index set D_k has full rank $\min(d_k, s)$ for each $k \in [1, \sigma]$. Then, the diversity order of the error probability with $\mathcal{X}$ used on a channel with general spatio-temporal correlation is lower bounded as

$$d \ge r \times \sum_{k=1}^{\sigma} \min(d_k, s) \quad (42)$$

We also have the following general result on the maximum achievable diversity order under spatio-temporal correlation.

Proposition 9: For any space-time code on a fading channel with arbitrary spatio-temporal correlations, the maximum achievable diversity order is given by

$$\text{rank}(\mathbf{R}) \times \min(T, \text{rank}(\mathbf{S}) \times \text{rank}(\mathbf{\Sigma})). \quad (43)$$

Proof: Using (27) in the proof of Proposition 4, one can write $\text{rank}((\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}^\dagger_{ij}) \circ \mathbf{\Sigma}) = \text{rank}(\widehat{\mathbf{X}}_{ij} \times (\mathbf{I}_\sigma \otimes (\mathbf{U}\sqrt{\Lambda})))$, where $\widehat{\mathbf{X}}_{ij}$ is as defined in Section III. The rank of $(\widehat{\mathbf{X}}_{ij} \times (\mathbf{I}_\sigma \otimes (\mathbf{U}\sqrt{\Lambda})))$ cannot be more than $\min(T, \sigma \cdot s)$. Hence, the maximum rank of $\mathbf{Z}_{ij} = ((\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}^\dagger_{ij}) \circ \mathbf{\Sigma}) \otimes (\mathbf{R}^\dagger\mathbf{R})$ for any $i \ne j$ is given by (43). ■

When $\mathbf{\Sigma} = \mathbf{1}\mathbf{1}^\mathsf{T}$, the maximum achievable diversity given by Proposition 9 becomes $\text{rank}(\mathbf{R}) \times \min(T, \text{rank}(\mathbf{S}))$ which was also noted in [8]. When $\mathbf{S} = \mathbf{I}$ and $\mathbf{R} = \mathbf{I}$, the maximum achievable diversity given by Proposition 9 becomes $N_\text{R} \times \min(T, N_\text{T} \times \text{rank}(\mathbf{\Sigma}))$ which was noted in [9] and also in Corollary 3. Thus, the result of Proposition 9 is a generalization of the results in both [9] and [8].

VI. CODE CONSTRUCTIONS

In this section, we provide representative examples of code constructions that meet the design guidelines proposed in this paper. Their performance is simulated for channels with spatio-temporal correlation and compared with that of codes that either do not meet the criterion proposed in this paper or are not designed to utilize the structure of the fading correlations. The first two examples consider the quasi-static fading channel with rank deficient spatial correlation. These two examples show the advantage of the precoding strategy over space-time codes that are designed for full diversity over spatially white channels. The third example describes a code designed for a channel with non-trivial temporal and spatial fading correlation. This example shows the need for choosing the space-time code for the precoding strategy appropriately. It also reiterates the advantage of the proposed code construction over simply using a space-time code designed for the quasi-static spatially white channel even though both methods lead to the same diversity order.

A. $N_T = N_R = 2, T = 2, \mathbf{\Sigma} = \mathbf{1}\mathbf{1}^T, \mathbf{R} = \mathbf{I}$

In this example, we consider three different transmit correlation structures, namely $\mathbf{S} = \mathbf{I}$ and the two correlation matrices $\mathbf{S}\mathbf{S}^\dagger = \mathbf{S}_1, \mathbf{S}_2$ where $\mathbf{S}_1$ and $\mathbf{S}_2$ are the following two correlation matrices given in [23].

$$\mathbf{S}_1 = \begin{bmatrix} 1.0000 & -0.9597 - 0.2425i \\ -0.9597 + 0.2425i & 1.0000 \end{bmatrix},$$
$$\mathbf{S}_2 = \begin{bmatrix} 1.0000 & 0.7000 + 0.7000i \\ 0.7000 - 0.7000i & 1.0000 \end{bmatrix}$$

Each of the correlation matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ have a rank of 1, equal non-zero eigenvalue but different eigenvectors. The coding scheme employed in [23] is a 2×2 full layer TAST code [16] with 4-QAM information symbols providing a rate of 4 bits per channel use (bpcu). This TAST scheme exhibits full diversity for the spatially white channel and therefore provides the maximum diversity order of 2 with the rank deficient spatial correlations $\mathbf{S}_1$ and $\mathbf{S}_2$. Our precoding scheme consists of using a code of the form

$$\begin{bmatrix} \mathbf{x}_1 & 0 \\ \mathbf{x}_2 & 0 \end{bmatrix} \mathbf{U}^\dagger,$$

where the $\mathbf{x}_i$ are independent 16-QAM information symbols and $\mathbf{U}$ is obtained from the eigendecomposition of $\mathbf{S}\mathbf{S}^\dagger$. This precoding scheme also leads to the maximum possible diversity order of 2 for spatial correlations $\mathbf{S}_1$ and $\mathbf{S}_2$. The performance of both coding schemes under each transmit correlation structure is shown in Fig. 2. It is seen that the precoding scheme is much better than the TAST scheme for both spatial correlations $\mathbf{S}_1$ or $\mathbf{S}_2$. Also, unlike the TAST scheme, the performance of the precoding scheme does not depend on the eigenvectors of the transmit correlation. Moreover, the complexity of the precoding scheme is much smaller than the TAST scheme.

B. $N_T = 4, N_R = 3, T = 4, \mathbf{\Sigma} = \mathbf{1}\mathbf{1}^T, \mathbf{R} = \mathbf{I}$

In this example, we consider a randomly generated matrix $\mathbf{S}$ of rank 2 as the transmit correlation matrix. We consider the naive application of the full diversity TAST codes of size 4×4 in this scenario. Both 2 and 3 layer TAST codes are simulated. Each of these codes provide the maximum diversity order of 6 inherent in the system. For the precoding strategy, however, we use a 4×2 space-time code that is obtained by stacking two 2×2 full diversity codes of [27]. This precoded stacked scheme also achieves the maximum diversity order of 6. The rate of each code is fixed at 12 bpcu and the performance of these codes is shown in Fig. 3. Once again, the precoding strategy leads to a higher coding gain compared to the TAST codes.

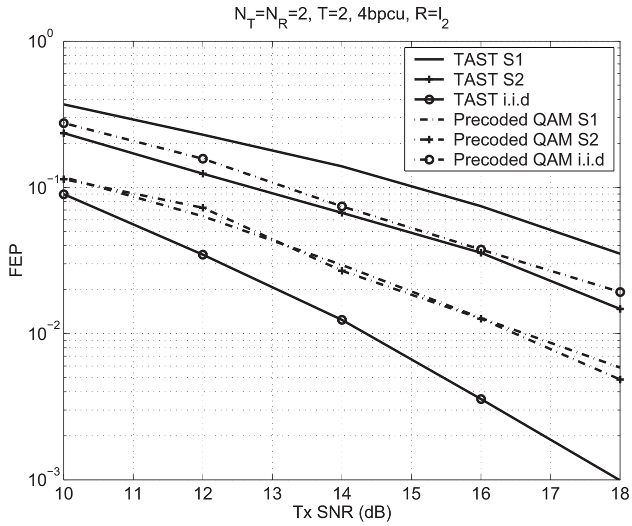


Fig. 2. Precoded QAM scheme vs. TAST scheme: Precoded QAM scheme has higher coding gain than the TAST scheme for rank deficient spatial correlations.

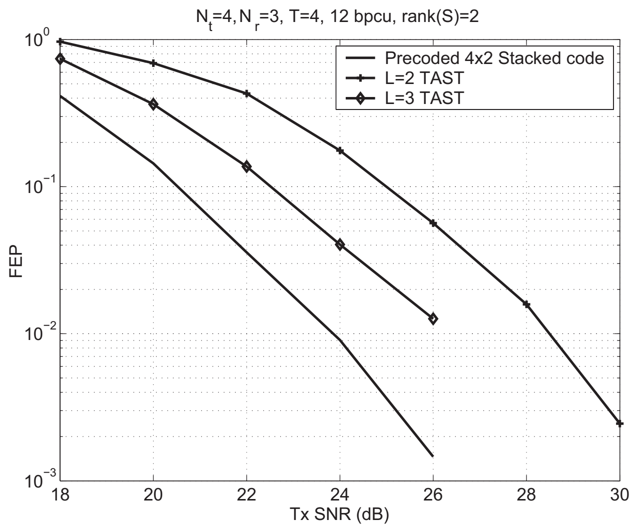


Fig. 3. Precoded Stacked code vs. TAST scheme: Precoded stacked code exhibits better performance than both two and three layer TAST schemes for rank deficient spatial correlation.

C. $N_T = 4, N_R = 2, T = 4, \mathbf{R} = \mathbf{I}$

In this example also, we consider a randomly generated matrix $\mathbf{S}$ of rank 2 as the transmit correlation matrix. The temporal correlation $\mathbf{\Sigma}$ and the corresponding $\mathbf{L}$ in this

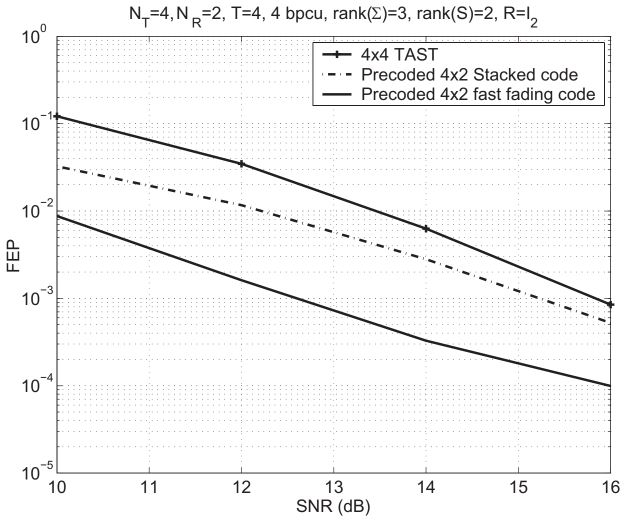


Fig. 4. Non-trivial spatial and temporal correlations: The 4×2 code used in the precoding strategy should be chosen to utilize the temporal diversity inherent in the system. With this choice, the precoding strategy is better than simply using a 4×4 TAST code.

example is set to

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \qquad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The first code we consider is a two layer four-transmit antenna TAST code of length four that was originally designed for the quasi-static spatially white channel. As described in the Introduction, many works advocate the use of such a code for channels with general spatio-temporal correlation. The second code is the precoded stacked space-time code used in Section VI-B that led to a performance improvement over the two layer TAST code on the quasi-static channel. However, in this example, due to the different structure of the temporal correlation matrix, this code provides a diversity order of only 4 and not the maximum possible value of $2\min(4, 3 * 2) = 8$ (according to Proposition 8). This can be verified by computing the diversity order of this code using Proposition 2 and the simplification of the corresponding Hadamard product as done in the proof of Proposition 9. Finally, we construct a third code as per the design rules obtained in this paper. We select first a code of length four designed for coding over *two* transmit antennas and two independent fading blocks (specifically, a TAST designed in [16] under the sum of ranks criterion), and then apply spatial precoding to it according to the rules summarized in Section V. Such a code achieves the maximum diversity order of 8. The performance of these codes is shown in Fig. 4 and it can be seen that the third code indeed achieves a very substantial coding gain over the 4×4 TAST code (which incidentally is even outperformed by the diversity-deficient second code in the SNR range considered).

VII. CONCLUSIONS

Using a new formulation of temporal correlation and simultaneous incorporation of spatial correlation, we arrive at a single code design criterion valid for arbitrary spatio-temporal fading correlations. It is found that, for rank deficient spatial correlation, precoding a smaller dimensional code leads to better coding gain than simply using a code meant for the spatially white channel even though both lead to the same diversity order. The space-time code to be precoded should be specifically designed to further exploit the diversity inherent in the temporal dimension. The proposed code design utilizes all degrees of freedom available in the spatio-temporal domain and leads to better performance with lower complexity than universal codes meant for all possible correlations.

REFERENCES

[1] D. Divsalar and M. Simon, "The design of trellis coded mpsk for fading channels: Performance criteria," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1004-1012, Sep. 1988.

[2] J. Boutros and E. Viterbo, "Signal space diversity: A power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 44, no. 4, pp. 1453-1467, July 1998.

[3] J. Boutros, E. Viterbo, C. Rastello, and J.-C. Belfiore, "Good lattice constellations for both Rayleigh fading and Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 502-518, Mar. 1996.

[4] K. Leeuwin, J. C. Belfiore, and G. K. Kaleh, "Chernoff bound of trellis-coded modulation over correlated fading channels," *IEEE Trans. Commun.*, vol. 42, no. 8, pp. 2506-2511, Aug. 1994.

[5] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.

[6] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 2383-2399, Sep. 2001.

[7] H. El Gamal, "On the robustness of space-time coding," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2417-2428, Oct. 2002.

[8] H. Bolcskei and A. J. Paulraj, "Performance of space-time codes in the presence of spatial fading correlation," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Nov. 2000, pp. 687-693.

[9] W. Su, Z. Safar, and K. J. R. Liu, "Diversity analysis of space-time modulation over time-correlated Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1832-1839, Aug. 2004.

[10] S. Siwamogsatham and M. P. Fitz, "Robust space-time codes for correlated Rayleigh fading channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2408-2416, Oct. 2002.

[11] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456-1467, July 1999.

[12] J. H. Kotecha, Z. Hong, and A. M. Sayeed, "Coding and divesity gain tradeoff in space-time codes for correlated MIMO channels," in *Proc. IEEE Global Commun. Conf.*, Dec. 2003, pp. 646-650.

[13] A. Hjorungnes, D. Gesbert, and J. Akhtar, "Precoding of space-time block coded signals for joint transmit receive correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 492-497, Mar. 2006.

[14] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 239-241, June 2002.

[15] A. Hjorungnes and D. Gesbert, "Minimum exact SER precoding of orthogonal space-time block codes for correlated MIMO channels," in *Proc. IEEE Global Commun. Conf.*, Nov. 2004, vol. 1, pp. 111-115.

[16] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1097-1119, May 2003.

[17] H. Bolcskei, M. Borgmann, and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 427-439, Apr. 2003.

[18] H. Bolcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. Wireless Communications and Networking Conf.*, Sep. 2000, pp. 1-6.

[19] M. Fozunbal, S. W. McLaughlin, and R. W. Schafer, "On space-time coding in the presence of spatio-temporal correlation," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1910-1926, Sep. 2004.

[20] D. P. Liu, Q. T. Zhang, and Q. Chen, "Structures and performance of noncoherent receivers for unitary space-time modulation on correlated fast-fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1116-1125, July 2004.

[21] S. G. Srinivasan and M. K. Varanasi, "Code design for the low snr noncoherent MIMO block Rayleigh fading channel," in *Proc. IEEE Intl. Symposium on Inform. Theory*, Sep. 2005, pp. 2218-2222.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1993.

[23] J. Wang, M. K. Simon, M. P. Fitz, and K. Yao, "On the performance of space-time codes over spatially correlated Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 877-881, June 2004.

[24] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Intl. Conf. on Communications*, June 2001, vol. 7, pp. 2266-2270.

[25] E. A. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1543-1553, Sep. 2004.

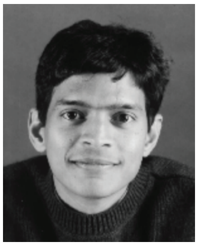
[26] P. Dayal and M. K. Varanasi, "An algebraic family of complex lattices for fading channels with application to space-time codes," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4184-4202, Dec. 2005.

[27] P. Dayal and M. K. Varanasi, "An optimal two transmit antenna space-time code and its stacked extensions," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2003, pp. 987-991.

[28] M. O. Damen, H. El Gamal, and N. C. Beaulieu, "Linear threaded algebraic space-time constellations," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2372-2388, Oct. 2003.

[29] M. O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice code decoder for space-time codes," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 161-163, May 2000.

[30] J. H. Kotecha and A. M. Sayeed, "On the capacity of correlated MIMO channels," in *Proc. IEEE Intl. Symposium on Information Theory*, June-July 2003, p. 355.



Mahesh K. Varanasi (S'87-M'89-SM'95) received the B.E. degree in electronics and communication engineering from Osmania University, Hyderabad, India, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from Rice University, Houston, TX, in 1987 and 1989, respectively. He joined the Electrical and Computer Engineering department of the University of Colorado at Boulder in 1989 as an Assistant Professor where he was later an Associate Professor during 1996-2001 and has been a Professor since 2001. His research and teaching interests are in the areas of communication theory, information theory, wireless communication and coding, detection and estimation theory, and signal processing. He has published on a variety of topics in these fields and is a highly cited researcher according to the ISI Web of Science (http://www.isihighlycited.com).

Pranav Dayal (S'03-M'05) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, India, in 2000, and the M.S. and Ph.D. degrees in electrical and computer engineering from University of Colorado, Boulder, USA, in 2003 and 2005, respectively. He joined Qualcomm, Inc., in San Diego as a Senior Engineer in August 2005 and is currently working on OFDM based wireless technologies.