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

Pranav Dayal and Mahesh K. Varanasi
e-mail: {dayalp,varanasi}@dsp.colorado.edu
University of Colorado, Boulder, CO 80309

Abstract — **A unified framework for multiple antenna space-time coding is presented leading to a general 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 encompasses quasi-static, fast fading, multiple block fading and other general correlations in the temporal dimension as well as arbitrary rank correlations in the spatial dimension. 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.**

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 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. 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.

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. Such an analysis is more insightful than previous works because the effect of each of spatial and temporal correlations on code design can be understood more clearly. Our framework 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. Furthermore, 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 space-time code structure to achieve the maximum diversity and improved coding gain is then proposed for arbitrary spatio-temporal fading correlation.

This work was supported in part by NSF Grants CCF-0434410 and CCF-0431170.

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

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

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

where $\mathbf{N}(t)$ is the noise matrix with i.i.d zero-mean unit variance complex normal entries, ρ is a measure of the signal-to-noise ratio and the superscript $\tilde{}$ 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

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

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 $\mathbf{y} = \sqrt{\rho}\mathrm{diag}(\mathbf{R}^\dagger \otimes \tilde{\mathbf{X}}(1), \ldots, \mathbf{R}^\dagger \otimes \tilde{\mathbf{X}}(T))\mathbf{h} + \mathbf{n}$ 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. The conditional pairwise error probability $P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}/\mathbf{h})$ between two possible codewords $\{\mathbf{X}_i(t)\}$ and $\{\mathbf{X}_j(t)\}$ is

$$P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}/\mathbf{h}) = Q(\sqrt{\frac{\rho}{2}}\|\mathbf{G}_{ij}\mathbf{h}\|) \tag{4}$$

$$= \frac{1}{\pi}\int_0^{\pi/2} e^{-\frac{\rho}{4\sin^2(\theta)}\|\mathbf{G}_{ij}\mathbf{h}\|^2} d\theta, \tag{5}$$

where $\mathbf{G}_{ij} = \text{diag}(\mathbf{R}^\dagger \otimes (\tilde{\mathbf{X}}_i(1) - \tilde{\mathbf{X}}_j(1)), \ldots, \mathbf{R}^\dagger \otimes (\tilde{\mathbf{X}}_i(T) - \tilde{\mathbf{X}}_j(T)))$. 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. Replacing θ by $\pi/2$ in the integrand of (5) and averaging over the distribution of $\mathbf{h}$, we obtain the Chernoff bound on the pairwise error probability

$$P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}) \leq \frac{1}{2} \cdot \frac{1}{\left|\mathbf{I}_{N_\mathrm{T} N_\mathrm{R} T} + \frac{\rho}{4}\mathbf{G}_{ij}(\mathbf{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T} N_\mathrm{R}})\mathbf{G}_{ij}^\dagger\right|}. \tag{6}$$

For some constant $\delta \in (0, \frac{\pi}{2})$, one can also lower bound the Q-function using the Craig's formula as

$$Q(x) \geq \frac{1}{\pi}\int_\delta^{\pi/2} e^{-\frac{x^2}{2\sin^2(\theta)}} d\theta \geq (\frac{1}{2} - \frac{\delta}{\pi}) e^{-\frac{x^2}{2\sin^2(\delta)}}. \tag{7}$$

Using (7) for the conditional PEP in (4) and averaging over the distribution of $\mathbf{h}$ again, we obtain the lower bound

$$P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}) \geq \frac{\frac{1}{2} - \frac{\delta}{\pi}}{\left|\mathbf{I}_{N_\mathrm{T} N_\mathrm{R} T} + \frac{\rho}{4\sin^2(\delta)}\mathbf{G}_{ij}(\mathbf{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T} N_\mathrm{R}})\mathbf{G}_{ij}^\dagger\right|}. \tag{8}$$

Combining (6) and (8), we obtain that the diversity order of the PEP is given by

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

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

$$= \mathbf{\Sigma}_{r,c}(\mathbf{R}^\dagger\mathbf{R} \otimes (\tilde{\mathbf{X}}_i(r) - \tilde{\mathbf{X}}_j(r))(\tilde{\mathbf{X}}_i(c) - \tilde{\mathbf{X}}_j(c))^\dagger)$$
$$= \mathbf{\Sigma}_{r,c}(\tilde{\mathbf{X}}_i(r) - \tilde{\mathbf{X}}_j(r))(\tilde{\mathbf{X}}_i(c) - \tilde{\mathbf{X}}_j(c))^\dagger)\mathbf{R}^\dagger\mathbf{R},$$

where the first expression 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. Let the symbol $\circ$ represent componentwise multiplication, also known as the Hadamard product [9]. Define the codeword difference $\mathbf{X}_{ij} = [(\mathbf{X}_i(1) - \mathbf{X}_j(1))^\mathsf{T}, \ldots, (\mathbf{X}_i(T) - \mathbf{X}_j(T))^\mathsf{T}]^\mathsf{T}$. We then see that $\mathbf{G}_{ij}(\mathbf{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T} N_\mathrm{R}})\mathbf{G}_{ij}^\dagger$ can be written as

$$\mathbf{G}_{ij}(\mathbf{\Sigma} \otimes \mathbf{I}_{N_\mathrm{T} N_\mathrm{R}})\mathbf{G}_{ij}^\dagger = ((\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \circ \mathbf{\Sigma}) \otimes (\mathbf{R}^\dagger\mathbf{R}). \tag{9}$$

Using the PEP to obtain bounds on the actual codeword error probability, one can conclude that the diversity order of the actual codeword error probability is the minimum rank of $\mathbf{Z}_{ij} = ((\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^\dagger\mathbf{X}_{ij}^\dagger) \circ \mathbf{\Sigma}) \otimes (\mathbf{R}^\dagger\mathbf{R})$ among all distinct codewords $\{\mathbf{X}_i(t)\}$ and $\{\mathbf{X}_j(t)\}$. 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 rank$(\mathbf{Z}_{ij})$ is the least possible.

The general expression of $\mathbf{Z}_{ij}$ must now be used to design codebooks for arbitrary 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, $\mathbf{S} = \mathbf{I}_{N_\mathrm{T}}$ and $\mathbf{R} = \mathbf{I}_{N_\mathrm{R}}$ and we need to maximize the minimum rank of $\mathbf{Z}_{ij} = ((\mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger) \circ \mathbf{\Sigma}) \otimes \mathbf{I}_{N_\mathrm{R}}$. This reduces to maximizing 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 Oppenheim's inequality

$$\det(\mathbf{A}) \cdot \prod_{i=1}^n \mathbf{B}_{ii} \leq \det(\mathbf{A} \circ \mathbf{B}) \tag{10}$$

for any two positive semidefinite matrices $\mathbf{A}$ and $\mathbf{B}$. Applying (10) with $\mathbf{B} = \mathbf{\Sigma}$ and $\mathbf{A} = \mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger$ and noting that the diagonal entries of $\mathbf{\Sigma}$ are all 1, we see that ensuring $T \leq N_\mathrm{T}$ 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 $\mathbf{Z}_{ij}$ has rank TN_R. On the other hand, if $\mathbf{\Sigma}$ is known to have full rank, then application of (10) with $\mathbf{A} = \mathbf{\Sigma}$ and $\mathbf{B} = \mathbf{X}_{ij}\mathbf{X}_{ij}^\dagger$ shows that a diversity order of TN_R can be achieved if all rows of $\mathbf{X}_{ij}$ are non-zero and that maximizing the minimum product of the norms of the rows of $\mathbf{X}_{ij}$,among all $i \neq j$, maximizes a lower bound on the coding gain. The diversity order results under these special cases have also been noted in [10]. However, no insights are provided therein for the most general case.

When $T > N_\mathrm{T}$ and $\mathbf{\Sigma}$ is not full rank, one must look at the structure of the correlation matrix $\mathbf{\Sigma}$ to arrive at a design criterion. To obtain insights into the design method for an arbitrary correlation matrix, we first state the following result.

Lemma 1: 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

$$\mathbf{A}\mathbf{A}^\dagger \circ \mathbf{B}\mathbf{B}^\dagger = \mathbf{C}\mathbf{C}^\dagger, \tag{11}$$

where $\mathbf{C} = [\text{diag}(\mathbf{B}_1)\mathbf{A}, \text{diag}(\mathbf{B}_2)\mathbf{A}, \ldots, \text{diag}(\mathbf{B}_n)\mathbf{A}]$.

Using the Cholesky decomposition [9], we write $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^\dagger$, where $\mathbf{L}$ is lower triangular. 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 is the row index of the first non-zero entry of the k-th column $\mathbf{L}_k$ of $\mathbf{L}$, then we require that $1 = p_1 < p_2 \ldots < p_\sigma \leq T$. Let d_k denote the number of non-zero entries in $\mathbf{L}_k$ from row index p_k to $p_{k+1} - 1$ for $1 \leq k \leq \sigma - 1$. Let d_σ be the number of non-zero rows of $\mathbf{L}_\sigma$. The proposed code design strategy and a guarantee on the diversity order achieved is given in the following proposition.

Proposition 1: Design the codebook such that for each signal pair $\{\mathbf{X}_i(t)\}$ and $\{\mathbf{X}_j(t)\}, i \neq j$, the submatrix of $\mathbf{X}_{ij}$ obtained by taking the rows corresponding to the non-zero row indices of $\mathbf{L}_k$ between p_k and $p_{k+1} - 1$ has full rank $\min(d_k, N_\mathrm{T})$ for each $1 \leq k \leq \sigma - 1$. Also, ensure that the submatrix of $\mathbf{X}_{ij}, \forall i \neq j$, obtained from the rows corresponding to the non-zero row indices of $\mathbf{L}_\sigma$ has full rank $\min(d_\sigma, N_\mathrm{T})$. Then, the diversity order achieved is at least

$$N_{\mathrm{R}} \times \sum_{k=1}^{\sigma} \min(d_k, N_{\mathrm{T}}). \tag{12}$$

The proof of Proposition 1 makes use of Lemma 1 which implies that the rank of $\mathbf{X}_{ij}\mathbf{X}_{ij}^{\dagger} \circ \mathbf{\Sigma}$ is equal to the rank of

$$\widehat{\mathbf{X}}_{ij} = [\mathrm{diag}(\mathbf{L}_1)\mathbf{X}_{ij}, \ldots, \mathrm{diag}(\mathbf{L}_{\sigma})\mathbf{X}_{ij}]. \tag{13}$$

The lower triangular nature of $\mathbf{L}$ and the imposed structure on the space-time codebook can then be exploited to conclude that the rank of $\widehat{\mathbf{X}}_{ij}$ is at least the sum of the ranks of the diagonal blocks of $\widehat{\mathbf{X}}_{ij}$, which is given by the summation in (12).

The rank of $\widehat{\mathbf{X}}_{ij}$ given by (13) is exactly equal to the summation in (12) if $\widehat{\mathbf{X}}_{ij}$ is block diagonal but in general it could be more. However, the rank of $\widehat{\mathbf{X}}_{ij}$ is at most

$$\min(\sigma N_{\mathrm{T}}, T) \tag{14}$$

so that the maximum diversity order under spatially white correlation is given by $N_{\mathrm{R}} \min(\sigma N_{\mathrm{T}}, T)$, a fact that was arrived at in a different fashion in [10]. In Section V, we generalize the expression for maximum achievable diversity under arbitrary spatial and temporal correlation structure.

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 two cases corresponding to the quasi-static scenario ($\mathbf{\Sigma} = \mathbf{1}\mathbf{1}^{\mathsf{T}}$) and the fast fading scenario ($\mathbf{\Sigma} = \mathbf{I}_T$) [5] . 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 the Oppenheim's inequality.

IV. Spatial Correlation with Quasi-static fading

In the quasi-static scenario with spatial correlation, we have $\mathbf{\Sigma} = \mathbf{1}\mathbf{1}^{\mathsf{T}}$ and the diversity order is given by the minimum rank of $\mathbf{Z}_{ij} = (\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^{\dagger}\mathbf{X}_{ij}^{\dagger}) \otimes (\mathbf{R}^{\dagger}\mathbf{R})$. By the properties of Kronecker products, it is easy to see that 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 $\min(T, \mathrm{rank}(\mathbf{S}))\mathrm{rank}(\mathbf{R})$. It was noted in [8] that choosing a space-time code that leads to full diversity N_{T} under spatially white fading will lead to a diversity of $\mathrm{rank}(\mathbf{S})\mathrm{rank}(\mathbf{R})$ under arbitrary spatial correlation. Full rank space-time codes were also used for simulations on spatially correlated channels in [11]. It was also stated in [8] that if the space-time code is rank deficient under spatially white fading, then 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 a diversity order of $\min(T, \mathrm{rank}(\mathbf{S}))\mathrm{rank}(\mathbf{R})$. We prove that 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 [12,13]. 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.

Averaging the conditional PEP in (5) over the density of $\mathbf{h}$ and using (9), one obtains the exact expression

$$\begin{aligned} P(\{\mathbf{X}_i(t)\} \to \{\mathbf{X}_j(t)\}) &= \frac{1}{\pi}\int_0^{\pi/2} \frac{d\theta}{\left|\mathbf{I} + \frac{\rho}{4\sin^2(\theta)}\mathbf{G}_{ij}^{\dagger}\mathbf{G}_{ij}\right|} \\ &= \frac{1}{\pi}\int_0^{\pi/2} \prod_{n=1}^{r}\prod_{m=1}^{v} \left(1 + \frac{\rho}{4\sin^2(\theta)}\lambda_m(\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^{\dagger}\mathbf{X}_{ij}^{\dagger})\lambda_n(\mathbf{R}^{\dagger}\mathbf{R})\right)^{-1} d\theta, \end{aligned} \tag{15}$$

where 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 $\mathrm{rank}(\mathbf{S})\mathrm{rank}(\mathbf{R})$ 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 2: 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

$$\alpha = \left(\frac{\mathrm{E}_{\mathbf{C}\in\mathcal{C}}[\|\mathbf{C}\|^2]}{\mathrm{E}_{\mathbf{C}'\in\mathcal{C}'}[\|\mathbf{C}'\|^2]}\right)^{1/2}, \tag{16}$$

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 a pair of codewords in $\mathcal{X}$ and $\mathcal{X}'$, the pairwise error probability

$$P(\mathbf{X}'_i \to \mathbf{X}'_j) \leq P(\mathbf{X}_i \to \mathbf{X}_j) \tag{17}$$

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 $\mathbf{X}_{ij} = \mathbf{X}_i - \mathbf{X}_j, \mathbf{X}'_{ij} = \mathbf{X}'_i - \mathbf{X}'_j, \mathbf{C}_{ij} = \mathbf{C}_i - \mathbf{C}_j$ and $\mathbf{C}'_{ij} = \mathbf{C}'_i - \mathbf{C}'_j$ be corresponding codeword differences in $\mathcal{X}, \mathcal{X}', \mathcal{C}$ and $\mathcal{C}'$, respectively. Looking at the PEP expression in (15), 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

$$\lambda_m(\mathbf{X}'_{ij}\mathbf{S}\mathbf{S}^{\dagger}\mathbf{X}'_{ij}) = \lambda_m(\alpha[\mathbf{C}'_{ij}\mathbf{0}]\Lambda\alpha[\mathbf{C}'_{ij}\mathbf{0}]^{\dagger})$$

$$= \alpha^2 \lambda_m(\mathbf{C}'_{ij}\Lambda_s\mathbf{C}'^{\dagger}_{ij}) = \alpha^2\lambda_m(\mathbf{C}_{ij}\Lambda\mathbf{C}^{\dagger}_{ij})$$
$$= \alpha^2\lambda_m(\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^{\dagger}\mathbf{X}^{\dagger}_{ij}).$$

Since $\alpha > 1$, $\lambda_m(\mathbf{X}'_{ij}\mathbf{S}\mathbf{S}^{\dagger}\mathbf{X}'_{ij}) > \lambda_m(\mathbf{X}_{ij}\mathbf{S}\mathbf{S}^{\dagger}\mathbf{X}^{\dagger}_{ij})$ and so from (15), we get that $P(\mathbf{X}'_i \to \mathbf{X}'_j) \leq P(\mathbf{X}_i \to \mathbf{X}_j)$. ■

Proposition 3: Let $T \geq N_{\mathrm{T}}$ and let $\mathcal{X}$ be a space-time code with $\mathrm{rank}(\mathbf{X}_i - \mathbf{X}_j) = N_{\mathrm{T}}, \forall \mathbf{X}_i, \mathbf{X}_j \in \mathcal{X}, i \neq j$. 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 $\mathcal{X} \equiv \mathcal{C}\mathbf{U}^{\dagger}$, where $\mathcal{C} \equiv \mathcal{X}\mathbf{U}$ is a code that meets the conditions of Proposition 2. 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 3, we can conclude that a space-time code that leads to full diversity under spatially white fading will provide full diversity order of $\mathrm{rank}(\mathbf{S})\mathrm{rank}(\mathbf{R})$ but it is not optimum in terms of coding gain and PEP minimization when $\mathrm{rank}(\mathbf{S}) < N_{\mathrm{T}}$. The strategy of precoding a reduced dimensional space-time code as $[\mathcal{C}'\mathbf{0}]\mathbf{U}^{\dagger}$ also leads to the maximum diversity of $\mathrm{rank}(\mathbf{S})\mathrm{rank}(\mathbf{R})$ and the optimum selection of such a space-time code will necessarily lead to smaller PEPs.

We now turn to the question of choosing the space-time codebook $\mathcal{C}'$ of size $T \times s$ for use in the precoding strategy for the spatially correlated fading scenario. One may consider the space-time codebook $\mathcal{T}\mathbf{P}$, where $\mathcal{T}$ is a $T \times s$ code that could either be a rectangular LPST code [14], a stacked extension of a code [15] or a rectangular TAST code [16]. The matrix $\mathbf{P} = \mathrm{diag}(\sqrt{P_1}, \ldots, \sqrt{P_s})$ is a diagonal matrix describing the power allocation on each of the s columns of the code $\mathcal{T}$. 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 [12,13]. 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 [17].

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. 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.

Proposition 4: 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 $\mathbf{C}_{ij} = \mathbf{C}_i - \mathbf{C}_j, \mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, i \neq j$, be such that the submatrix of $\mathbf{C}_{ij}$ obtained by taking the rows corresponding to the non-zero row indices of $\mathbf{L}_k$ between p_k and $p_{k+1} - 1$ has full rank $\min(d_k, s)$ for each $1 \leq k \leq \sigma - 1$. Also, let the submatrix of $\mathbf{C}_{ij}, \forall i \neq j$, obtained from the rows corresponding to the non-zero row indices of $\mathbf{L}_{\sigma}$ have full rank $\min(d_{\sigma}, s)$. Then, the diversity order of the error probability with $\mathcal{X}$ used on a channel with general spatio-temporal correlation is lower bounded by

$$r \times \sum_{k=1}^{\sigma} \min(d_k, s). \tag{18}$$

For a general space-time code, the maximum rank of $\widehat{\mathbf{X}}_{ij}\mathbf{S}$, for $\widehat{\mathbf{X}}_{ij}$ given in Section III, cannot be more than $\min(T, \sigma \cdot s)$. Hence, the maximum achievable diversity order under arbitrary spatio-temporal correlation is

$$\mathrm{rank}(\mathbf{R}) \times \min(T, \mathrm{rank}(\mathbf{S}) \times \mathrm{rank}(\mathbf{\Sigma})), \tag{19}$$

which is a generalization of the results in both [10] and [8].

VI. Code Constructions

The first two examples in this section show the advantage of the precoding strategy over using a space-time code that is designed for spatially white channels. The third example describes a code designed for a channel with non-trivial spatial and temporal fading correlation.

A. $N_{\mathrm{T}} = N_{\mathrm{R}} = 2, T = 2, \mathbf{\Sigma} = \mathbf{1}\mathbf{1}^{\mathsf{T}}, \mathbf{R} = \mathbf{I}$

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 given in [11]. The coding scheme employed in [11] is a 2×2 full layer TAST code [18] with 4-QAM information symbols providing a rate of 4 bpcu. Each of the correlation matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ have a rank of 1, equal non-zero eigenvalue but different eigenvectors. 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 $\mathbf{S}\mathbf{S}^{\dagger}$. The performance of both coding schemes under each transmit correlation structure is shown in Figure 1. It is seen that the precoding scheme is much better than the TAST scheme when the transmit correlation matrix is of rank 1. 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. $N_{\mathrm{T}} = 4, N_{\mathrm{R}} = 3, T = 4, \mathbf{\Sigma} = \mathbf{1}\mathbf{1}^{\mathsf{T}}, \mathbf{R} = \mathbf{I}$

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. 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 [15]. The rate of each code is fixed at 12 bpcu and the performance of these codes is shown in Figure 2. Once again, the precoding strategy leads to a higher coding gain compared to the TAST codes.

C. $N_{\rm T} = 4, N_{\rm R} = 2, T = 4, \mathbf{R} = \mathbf{I}$

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 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 4×4 TAST code. This code provides the full diversity of $2\min(4, 3 * 2) = 8$ inherent in the system. The second code is the precoded stacked space-time code used in Example VI-B. This code provides a diversity of only 4. Finally, we consider a third code that is obtained by precoding a 4×2 space-time code which is a TAST code designed for coding over two independent fading blocks [18]. This code satisfies the conditions proposed in Section V and provides the full diversity of 8. The performance of these codes is shown in Figure 3. The third code meets all our criterion and provides the best performance. It exploits the entire diversity available in the system and also exhibits a coding gain over the 4×4 TAST code due to the precoding strategy.

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.

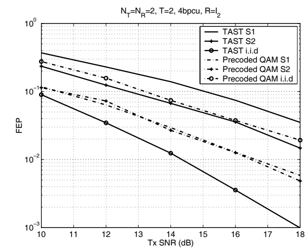

Fig. 1. Precoded QAM scheme vs. TAST scheme (Example A)

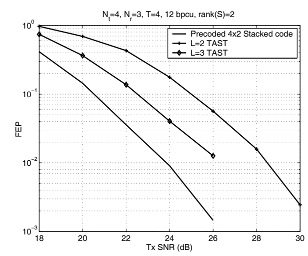

Fig. 2. Precoded Stacked code vs. TAST scheme (Example B)

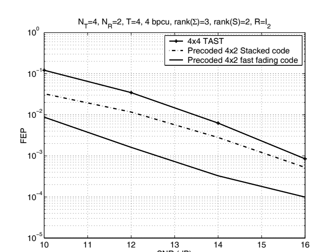

Fig. 3. Non-trivial spatial and temporal correlations (Example C)

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, Sept. 1988.

[2] J. Boutros and E. Viterbo, "Signal Space Diversity: A Power- and Bandwidth–Effi cient 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. Belfi ore, "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. Belfi ore, 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 unifi ed analysis of coherent and noncoherent space–time receivers," *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 2383–2399, Sept. 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*, Pacifi c Grove, CA, Nov. 2000, pp. 687–693.

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

[10] 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.

[11] 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.

[12] 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*, 2001, pp. 2266–2270.

[13] 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, Sept. 2004.

[14] P. Dayal and M. K. Varanasi, "Algebraic space-time codes that achieve maximal diversity and/or capacity optimality with low peak-to-mean power ratio," submitted to *IEEE Trans. Inform. Theory*, May 2003.

[15] 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*, Monterey, CA, Nov. 2003.

[16] 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.

[17] 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.

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