Precoding of Orthogonal Space-Time Block Codes in Arbitrarily Correlated MIMO Channels: Iterative and Closed-Form Solutions

Are Hjørungnes, *Member, IEEE,* and David Gesbert, *Member, IEEE.*

Abstract

A memoryless precoder is designed for orthogonal space-time block codes (OSTBCs) for multiple-input multiple-output (MIMO) channels exhibiting joint transmit-receive correlation. Unlike most previous similar works which concentrate on transmit correlation only and pair-wise error probability (PEP) metrics, 1) the precoder is designed to minimize the *exact* symbol error rate (SER) as function of the channel correlation coefficients, which are fed back to the transmitter. 2) The correlation is arbitrary as it may or may not follow the so-called *Kronecker structure*. 3) The proposed method can handle general propagation settings including those arising from a cooperative macro-diversity (multi-base) scenario. We present two algorithms. The first is suboptimal, but provide a simple closed-form precoder that handles the case of uncorrelated transmitters, correlated receivers. The second is a fast-converging numerical optimization of the exact SER which covers the general case. Finally, a number of novel properties of the minimum SER precoder are derived.

Index Terms: MIMO, orthogonal space-time block code, precoder optimization, minimum exact symbol error rate, power constraint.

I. Introduction

In the area of efficient communications over non-reciprocal MIMO channels, recent research [2], [3], [4], [5] has demonstrated the value of feeding back to the transmitter information about channel state observed at the receiver. Among those, there has been a growing interest in transmitter schemes that can exploit low-rate long-term statistical channel state information in the form of antenna correlation coefficients. So far, emphasis has been on designing precoders for space-time block coded (STBC) [3]

Corresponding author: A. Hjørungnes is with UniK - University Graduate Center, University of Oslo, Instituttveien 25, P. O. Box 70, N-2027 Kjeller, Norway, email: `arehj@unik.no`, phone: +47-64844700, and fax: +47-63818146.

D. Gesbert is with the Mobile Communications Department, Eurécom Institute, 2229 Route des Crêtes, BP 193, F-06904 Sophia Antipolis Cédex, France, email: `david.gesbert@eurecom.fr`.

Part of this work was presented in [1]. This work is supported by the Research Council of Norway through project number 157716/432.

signals or spatially multiplexed streams that are adjusted based on the knowledge of the transmit correlation only while the receiving antennas are uncorrelated [4], [5], [6], [7]. These techniques are well suited to downlink situations where an elevated access point (situated above the surrounding clutter) transmits to a subscriber placed in a rich scattering environment. Although simple models exist for the joint transmit receiver correlation based on the well known Kronecker structure [3], the accuracy of these models has recently been questioned in the literature based on measurement campaigns [8]. Therefore, there is interest in investigating the precoding of OSTBC signals for MIMO channels that *do not* necessarily follow the Kronecker structure.

An upper bound of the PEP is minimized in [4], [5] for transmit-only correlation, and for full channel correlation in [2], [9]. In [10], the exact SER expressions were derived for when there is *no receiver correlation* and maximum ratio combining is used at the receiver. A bound of the exact error probability was used as the optimization criterion in [10]. The no receiver correlation assumption might be an unrealistic channel model for example in uplink communications, where the access point (receiver) is equipped with several receiver antennas and where the direction of arrival has a small spread at the receiver antennas. In [11], exact SER expressions were found for uncorrelated MIMO channels that are precoded with the identity matrix. Exact expressions where derived for correlated Rayleigh and Ricean Fading channels *without* precoding in [12].

In this paper, we address the problem of linear precoding of OSTBC signals launched over a jointly transmit-receive correlated MIMO channel when the transmitter knows the correlation matrix of the MIMO channel matrix and the receiver knows the channel realization exactly. Our main contributions are:

1) We derive easy to evaluate *exact* expressions for the average SER for a system where the transmitter has an OSTBC followed by a *full precoder matrix* and where the receiver also has multiple antennas and is using maximum likelihood decoding (MLD).
2) We propose an iterative numerical technique for minimizing the exact SER with respect to the precoder matrix. This is in contrast with previous precoders based on *bounds* of the SER or the PEP and also do not address the arbitrary non-Kronecker correlation case.
3) Several properties of the minimum SER precoder are presented. We identify cases in which the precoder is dependent *or not* of the *receive* correlation matrix. We show the dependency is strongly related to the Kronecker model being valid or not.

4) An analytical closed-form precoder is proposed as an approximation based on the hereby proposed *Equal diversity spread principle*, in the particular case of cooperative diversity. This solution is also easily interpretable.

The rest of this article is organized as follows: In Section II, the precoded OSTBC system is described. Exact SER expressions are derived in Section III. Section IV presents the optimization problem and several properties of the minimum SER precoder are derived. In Section V, a closed-form solution is proposed when no transmit correlation is present. A numerical optimization algorithm for the general case is proposed in Section VI. Section VII contains simulation results and comparisons to alternative solutions. Section VIII presents the conclusions and the proofs are given in the appendices.

II. System Description

A. OSTBC Signal Model

Figure 1 (a) shows the block MIMO system model with M_t and M_r transmitter and receiver antennas, respectively. The transmit symbol vector of size $K \times 1$ is denoted $\boldsymbol{x} = [x_0, x_1, \ldots, x_{K-1}]^T$, where $x_i \in \mathcal{A}$, where $\mathcal{A}$ is a signal constellation set such as uniform M-PAM, M-QAM, or M-PSK, satisfying $E\left[|x_i|^2\right] = \sigma_x^2$. This vector is transmitted by means of a given OSTBC matrix $\boldsymbol{C}(\boldsymbol{x})$ of size $B \times N$, where B and N are the space and time dimension of the OSTBC, respectively. If bits are used as inputs to the system, $K \log_2 M$ bits are used to produce the vector $\boldsymbol{x}$. Since the OSTBC is orthogonal, the following holds: $\boldsymbol{C}(\boldsymbol{x})\boldsymbol{C}^H(\boldsymbol{x}) = a\sum_{i=0}^{K-1}|x_i|^2\boldsymbol{I}_B$, where $a = 1$ if $\boldsymbol{C}(\boldsymbol{x}) = \mathcal{G}_2^T$, $\boldsymbol{C}(\boldsymbol{x}) = \mathcal{H}_3^T$, or $\boldsymbol{C}(\boldsymbol{x}) = \mathcal{H}_4^T$ in [13] and $a = 2$ if $\boldsymbol{C}(\boldsymbol{x}) = \mathcal{G}_3^T$ or $\boldsymbol{C}(\boldsymbol{x}) = \mathcal{G}_4^T$ in [13]. The rate of the code is K/N. The developed theory is valid for any OSTBC. A precoder $\boldsymbol{F}$ of size $M_t \times B$ is applied before the signal is sent over the channel MIMO channel $\boldsymbol{H}$ of size $M_r \times M_t$. The channel is corrupted by the additive block noise $\boldsymbol{V}$, of size $M_r \times N$, which is complex Gaussian circularly distributed with independent components having variance N_0 and zero mean. The $M_r \times N$ receive block signal $\boldsymbol{Y}$ becomes

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{F}\boldsymbol{C}(\boldsymbol{x}) + \boldsymbol{V}. \tag{1}$$

The receiver is assumed to know $\boldsymbol{H}$ and $\boldsymbol{F}$ exactly, and it performs MLD of blocks of size $M_r \times N$ to find an estimate of $\boldsymbol{x}$ denoted $\hat{\boldsymbol{x}}$.

B. Correlated Channel Models

A flat block-fading correlated Rayleigh fading channel model [3] is assumed. Let the channel $\boldsymbol{H}$ have zero mean, complex Gaussian circularly distribution with positive semi-definite autocorrelation given by $\boldsymbol{R} = E\left[\mathrm{vec}\left(\boldsymbol{H}\right)\mathrm{vec}^{H}\left(\boldsymbol{H}\right)\right]$ of size $M_t M_r \times M_t M_r$, where the operator $\mathrm{vec}(\cdot)$ stacks the columns of the matrix it is applied to into a long column vector [14]. A channel realization of the correlated channel can then be found by $\mathrm{vec}\left(\boldsymbol{H}\right) = \boldsymbol{R}^{1/2}\,\mathrm{vec}\left(\boldsymbol{H}_w\right)$, where $\boldsymbol{R}^{1/2}$ is the unique positive definite matrix square root [15] of $\boldsymbol{R}$ and $\boldsymbol{H}_w$ has size $M_r \times M_t$ and is complex Gaussian circularly distributed with independent components all having unit variance and zero mean.

Kronecker model: A special case of the model above is as follows [3]

$$\boldsymbol{R} = \boldsymbol{R}_t^T \otimes \boldsymbol{R}_r, \tag{2}$$

where the operator $(\cdot)^T$ denotes transposition, $\otimes$ is the Kronecker product, the matrices $\boldsymbol{R}_r$ and $\boldsymbol{R}_t$ are the correlations matrices of the receiver and transmitter, respectively, and their sizes are $M_r \times M_r$ and $M_t \times M_t$. Unlike (2), the general model considers that the receive (or transmit) correlation depends on at which transmit (or receive) antenna the measurements are performed.

C. Equivalent Single-Input Single-Output Model

Let $\boldsymbol{\Phi} \triangleq \boldsymbol{R}^{1/2}\left[\left(\boldsymbol{F}^{*}\boldsymbol{F}^{T}\right)\otimes \boldsymbol{I}_{M_r}\right]\boldsymbol{R}^{1/2}$ be a positive semidefinite matrix of size $M_t M_r \times M_t M_r$. Define the scalar $\alpha \triangleq \|\boldsymbol{H}\boldsymbol{F}\|_F^2 = \mathrm{vec}^{H}\left(\boldsymbol{H}_w\right)\boldsymbol{\Phi}\,\mathrm{vec}\left(\boldsymbol{H}_w\right)$, where $\|\cdot\|_F$ is the Frobenius norm. By generalizing the approach given in [11], [16] to include a *full* complex-valued precoder $\boldsymbol{F}$ of size $M_t \times B$ and having a *full* channel correlation matrix $\boldsymbol{R}$ the OSTBC system can be shown to be equivalent to a collapsed system having the following output input relationship

$$y'_k = \sqrt{\alpha}x_k + v'_k, \tag{3}$$

for $k \in \{0, 1, \ldots, K-1\}$, and where $v'_k \sim \mathcal{CN}(0, N_0/a)$ is complex circularly distributed. This signal is fed into a memoryless MLD that is designed from the signal constellation of the source symbols $\mathcal{A}$. The equivalent single-input single-output (SISO) model is shown in Figure 1 (b).

III. SER EXPRESSIONS

A. SER *Expressions for Given Received* SNR

By considering the SISO system in Figure 1 (b), it is seen that the instantaneous *received* SNR γ per source symbol is given by $\gamma \triangleq \frac{a\sigma_x^2\alpha}{N_0} = \delta\alpha$, where $\delta \triangleq \frac{a\sigma_x^2}{N_0}$. Define the following three signal constellation dependent constants $g_{\text{PSK}} \triangleq \sin^2 \frac{\pi}{M}$, $g_{\text{PAM}} \triangleq \frac{3}{M^2-1}$, and $g_{\text{QAM}} \triangleq \frac{3}{2(M-1)}$. The symbol error probability $\text{SER}_\gamma \triangleq \Pr\{\text{Error}|\gamma\}$ for a given γ for M-PSK, M-PAM, and M-QAM signaling are, respectively, given by [17]

$$\text{SER}_\gamma = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} e^{-\frac{g_{\text{PSK}}\gamma}{\sin^2(\theta)}} d\theta, \tag{4}$$

$$\text{SER}_\gamma = \frac{2}{\pi}\frac{M-1}{M}\int_0^{\frac{\pi}{2}} e^{-\frac{g_{\text{PAM}}\gamma}{\sin^2(\theta)}} d\theta, \tag{5}$$

$$\text{SER}_\gamma = \frac{4}{\pi}\left(1-\frac{1}{\sqrt{M}}\right)\left[\frac{1}{\sqrt{M}}\int_0^{\frac{\pi}{4}} e^{-\frac{g_{\text{QAM}}\gamma}{\sin^2(\theta)}} d\theta + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} e^{-\frac{g_{\text{QAM}}\gamma}{\sin^2(\theta)}} d\theta\right]. \tag{6}$$

B. Exact SER *Expressions*

The moment generating function of the probability density function $p_\gamma(\gamma)$ is defined as $\phi_\gamma(s) \triangleq \int_0^\infty p_\gamma(\gamma)e^{s\gamma}d\gamma$. Since all the K source symbols x_k go through the same SISO system in Figure 1 (b), the average SER of the MIMO system can be found as

$$\text{SER} \triangleq \Pr\{\text{Error}\} = \int_0^\infty \Pr\{\text{Error}|\gamma\}\, p_\gamma(\gamma)d\gamma = \int_0^\infty \text{SER}_\gamma\, p_\gamma(\gamma)d\gamma. \tag{7}$$

From the definition of α, it is seen that $\alpha = \text{vec}^H(\boldsymbol{H}'_w)\,\boldsymbol{\Lambda}\,\text{vec}(\boldsymbol{H}'_w)$ where $\boldsymbol{H}'_w$ and $\boldsymbol{H}_w$ has the same distribution, and from this expression of α, it follows [18] that $\phi_\alpha(s)$ is given by: $\phi_\alpha(s) = \frac{1}{\prod_{i=0}^{M_tM_r-1}(1-\lambda_i s)}$, where λ_i is eigenvalue number i of the positive semi-definite matrix $\boldsymbol{\Phi}$. Since $\gamma = \delta\alpha$, it follows that $\phi_\gamma(s)$ is given by:

$$\phi_\gamma(s) = \phi_\alpha(\delta s) = \frac{1}{\prod_{i=0}^{M_tM_r-1}(1-\delta\lambda_i s)}. \tag{8}$$

By using (7) and the definition of $\phi_\gamma(s)$ together with the result in (8), it is possible to express the exact SER for all the signal constellations in terms of the eigenvalues λ_i of $\boldsymbol{\Phi}$. When finding the necessary conditions for the optimal precoder, eigenvalues that are not simple[1] might cause difficulties in connection with calculations of derivatives. Therefore, it is useful to rewrite the SER expressions in terms of the matrix $\boldsymbol{\Phi}$. This can be done by utilizing its eigen-decomposition. The result of all these operations led to the following exact expressions for the SER for M-PSK, M-PAM, and M-QAM

[1]The matrix $\boldsymbol{\Phi} \in \mathbb{C}^{M_tM_r \times M_tM_r}$ has in general M_tM_r different non-negative eigenvalues. However, the roots of the characteristic equation, i.e., the eigenvalues need not to be distinct. The number of times an eigenvalue appears is equal to its multiplicity. If one eigenvalue appear only once, it is called a simple eigenvalue [19].

$$\mathrm{SER} = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_t M_r} + \delta\frac{g_{\mathrm{PSK}}}{\sin^2\theta}\boldsymbol{\Phi}\right)}, \tag{9}$$

$$\mathrm{SER} = \frac{2}{\pi}\frac{M-1}{M}\int_0^{\frac{\pi}{2}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_t M_r} + \delta\frac{g_{\mathrm{PAM}}}{\sin^2\theta}\boldsymbol{\Phi}\right)}, \tag{10}$$

$$\mathrm{SER} = \frac{4}{\pi}\frac{\sqrt{M}-1}{\sqrt{M}}\left[\frac{1}{\sqrt{M}}\int_0^{\frac{\pi}{4}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_t M_r} + \delta\frac{g_{\mathrm{QAM}}}{\sin^2\theta}\boldsymbol{\Phi}\right)} + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_t M_r} + \delta\frac{g_{\mathrm{QAM}}}{\sin^2\theta}\boldsymbol{\Phi}\right)}\right], \tag{11}$$

respectively. It is seen that (9) and (10) give the same result when $M = 2$, and this is not surprising, since, the 2-PSK and 2-PAM constellations are identical. When $M = 4$, it can be shown that (9) and (11) return the same result. If $\boldsymbol{R} = \boldsymbol{I}_{M_t M_r}$ and $\boldsymbol{F} = \boldsymbol{I}_{M_t}$, then the performance expressions are reduced to the results found in [11]. If $\delta \to 0^+$, it is seen from (9), (10), and (11), that $\mathrm{SER} \to \frac{M-1}{M}$ for any precoder $\boldsymbol{F}$, which clearly is the symbol error rate for a random symbol generator. The expressions in [12] are not as easy to evaluate as the proposed expressions since, in [12], the input signal constellation was arbitrary and then the SER expressions must be found by performing two-dimensional integrals over possibly complicated regions in the complex plane. The proposed expressions are very easy to evaluate.

IV. Precoding of OSTBC Signals for the General Case

A. *Optimal Precoder Problem Formulation*

By using the properties of OSTBCs the average power constraint on the transmitted block $\boldsymbol{Z} \triangleq \boldsymbol{F}\boldsymbol{C}(\boldsymbol{x})$ can be expressed as $aK\sigma_x^2 \operatorname{Tr}\left\{\boldsymbol{F}\boldsymbol{F}^H\right\} = P$, where P is the average power used by the transmitted block $\boldsymbol{Z}$. The goal is to find the matrix $\boldsymbol{F}$ such that the exact SER is minimized under the power constraint. We propose that the optimal precoder is given by the following optimization problem:

Problem 1:

$$\min_{\left\{\boldsymbol{F}\in\mathbb{C}^{M_t\times B} \;\middle|\; Ka\sigma_x^2 \operatorname{Tr}\left\{\boldsymbol{F}\boldsymbol{F}^H\right\}=P\right\}} \mathrm{SER}\,.$$

Remark 1: *In general, the optimal precoder is dependent on the value of* N_0 *and, therefore, also on the signal to noise ratio (*SNR*) defined as* $\mathrm{SNR} \triangleq 10\log_{10}\frac{P}{N_0}$.

B. *Upper Bound on* SER *and Connection to* PEP

If $\sin^2(\theta)$ is replaced with 1 in all the integrals in (9), (10), and (11), the following upper bound is found for SER for all the constellations considered:

$$\mathrm{SER} \leq \frac{M-1}{M}\frac{1}{\det\left(\boldsymbol{I}_{M_t M_r} + \delta g\boldsymbol{\Phi}\right)}, \tag{12}$$

where g is chosen according to the signal constellation. If this upper bound of SER is minimized under the power constraint, it is seen that this is equivalent to maximizing $\det\left(\boldsymbol{I}_{M_tM_r} + \delta g \boldsymbol{\Phi}\right)$ under the power constraint. In [10], this criterion was used when there is no correlation between the receiver antennas, and the criterion is equivalent to an upper bound on PEP used in [9] for a full correlation matrix $\boldsymbol{R}$.

Interestingly, the minimum SER and the PEP based precoders will perform similarly in the very low and very high SNR range. For medium values of SNR, some gain can be achieved by using the SER based method over the PEP.

C. Properties of the Optimal Precoder

Below, we give several properties to help characterize the optimal precoder in particular situations of interest.

Lemma 1: *If $\boldsymbol{F}$ is an optimal solution of Problem 1, then the precoder $\boldsymbol{F}\boldsymbol{W}$, where $\boldsymbol{W} \in \mathbb{C}^{B\times B}$ is unitary, is also optimal.*

The proof of this lemma can be found in Appendix I.

Proposition 1: *If $B = M_t$, then it is possible to chose the optimal precoder $\boldsymbol{F}$ Hermitian or symmetric.*

The proof of this proposition can be found in Appendix II.

Proposition 2: *If $\mathrm{SNR} \to \infty$, $B = M_t$, and $\boldsymbol{R}$ is non-singular, then the optimal precoder is given by the trivial identity-scaled precoder $\boldsymbol{F} = \sqrt{\frac{P}{Ka\sigma_x^2 M_t}}\boldsymbol{I}_{M_t}$ for the M-PSK, M-PAM, and M-QAM constellations.*

The proof of this proposition can be found in Appendix III. This comforts the information theoretic viewpoint by which channel-based transmitter optimization yields no benefit at high SNR in MIMO Rayleigh channels [20].

Remark 2: *If $\boldsymbol{R}$ is singular, examples can be constructed showing that, in general, the optimal precoder $\boldsymbol{F}$ is not proportional to the identity matrix when $\mathrm{SNR} \to +\infty$, see Scenario 2 in Section VII.*

Proposition 3: *If $M_t = B$ and $\boldsymbol{R} = \boldsymbol{I}_{M_tM_r}$, then the optimal precoder is given by the trivial precoder $\boldsymbol{F} = \sqrt{\frac{P}{Ka\sigma_x^2 M_t}}\boldsymbol{I}_{M_t}$ for the M-PSK, M-PAM, and M-QAM constellations.*

The proof of this proposition can be found in Appendix IV. Proposition 3 shows that there is no need for precoding in the absence of any correlation. The result in Proposition 3 is also given in [10].

Proposition 4: *The diversity of a system using a precoder satisfying $\mathrm{rank}\left(\boldsymbol{F}\right) = M_t$ is $\mathrm{rank}\left(\boldsymbol{R}\right)$.*

For the proof see Appendix V.

Proposition 5: *If $\boldsymbol{R}$ has full rank, then some diversity is lost by using $B < M_t$.*

For the proof see Appendix VI. In this case, this makes sense since some spatial degrees of freedom are not being excited at the transmitter.

We now give an important result, which extends one of the results given in [4].

Theorem 1: *Let $\boldsymbol{R}$ satisfy (2), and let the transmitter correlation matrix have the following eigen-decomposition $\boldsymbol{R}_t = \boldsymbol{U}_t \boldsymbol{\Lambda}_t \boldsymbol{U}_t^H$, where $\boldsymbol{U}_t \in \mathbb{C}^{M_t \times M_t}$ is unitary and $\boldsymbol{\Lambda}_t$ is diagonal of size $M_t \times M_t$. The optimal SER precoder can be expressed as $\boldsymbol{F} = \boldsymbol{U}_t \boldsymbol{\Delta}$, where $\boldsymbol{\Delta}$ is a diagonal matrix of size $M_t \times B$.*

The proof of this theorem can be found in Appendix VII. According to this result, the optimal precoder is built from a singular vector transmit matrix obtained in closed form, along with a power allocation scheme. In [4], a PEP criterion is used to obtain the optimal power allocation based on a water-filling procedure when the receiver correlation matrix was equal to the identity matrix. Theorem 1 is valid for any receiver correlation matrix in the Kronecker model and, therefore, it extends one of the main result in [4] to the SER criterion with arbitrary receiver correlation in the Kronecker model.

Note that, signaling on the eigenvectors of the transmitter correlation matrix was also used in [4], [5] for the precoder that minimizes an upper bound of the PEP when $\boldsymbol{R}_r = \boldsymbol{I}_{M_r}$ was used in the Kronecker model in (2). A similar factorization result was found for maximizing the *capacity* with a full Kronecker model in [23]. If the full correlation matrix does *not* follow the Kronecker product assumption, then the transmitter correlation matrix is simply not definable. In this case, the general iterative optimization technique presented in Section VI can be applied to find the minimum SER precoder.

V. Precoding of OSTBC for Zero Transmit Correlation

We now focus on the case where the transmit antennas are uncorrelated, yet the receive antennas are. What is intriguing in this case is that it is not intuitively clear what role a transmit precoder can play to improve performance when there is no transmit correlation structure to be exploited. However, we give a key result here showing that precoding will help simply in all cases where the Kronecker assumption (2) does not hold, a key example of which is addressed below.

A. Distributed Space Time Coding

In distributed (or cooperative) space time coding [21], [22], the code word is transmitted from antennas belonging to distinct access points, toward the users. Thus, the transmitter antennas are widely separated and typically experience different channel correlation at the same receiving array, see Figure 2.

If only receiver correlation is present, the total correlation matrix can be expressed as

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_{r_0} & \mathbf{0}_{M_r \times M_r} & \cdots & \mathbf{0}_{M_r \times M_r} \\ \mathbf{0}_{M_r \times M_r} & \boldsymbol{R}_{r_1} & \cdots & \mathbf{0}_{M_r \times M_r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{M_r \times M_r} & \mathbf{0}_{M_r \times M_r} & \cdots & \boldsymbol{R}_{r_{M_t-1}} \end{bmatrix}, \tag{13}$$

where $\boldsymbol{R}_{r_i}$ is the receive correlation matrix "seen" from transmitter number i and the matrix $\mathbf{0}_{k \times l}$ of size $k \times l$, contains only zeroes.

We now proceed to prove two key results. First, claims are formulated, then interpretations are given.

Theorem 2: *Let $B = M_t$. If (13) holds, the optimal $\boldsymbol{F}$ can be chosen diagonal with real and non-negative diagonal elements.*

The proof of this theorem can be found in Appendix VIII.

Theorem 3: *Let $B = M_t$ and let $\boldsymbol{R}$ satisfy (13) with $\boldsymbol{R}_{r_i} = \boldsymbol{R}_r$ for all $i \in \{0, 1, \ldots, M_t - 1\}$. This is the same as using $\boldsymbol{R}_t = \boldsymbol{I}_{M_t}$ in (2). Then the optimal precoder is independent of the receiver correlation matrix $\boldsymbol{R}_r$ and the precoder is given by $\boldsymbol{F} = \sqrt{\frac{P}{K a \sigma_x^2 M_t}} \boldsymbol{I}_{M_t}$.*

The proof of this theorem can be found in Appendix IX.

Interpretations: Theorem 2 tells us that despite the lack of transmitter correlation, a transmit precoder makes sense whenever the Kronecker structure for the overall correlation matrix does not hold, a practical case of which is seen in cooperative/distributed OSTBC. It also tells us that precoding takes the form of power allocation across the transmit antennas. In the next subsection, we propose a closed form approach to derive the power weights. Theorem 3 indicates that if the Kronecker structure holds (in addition to having uncorrelated transmitters, correlated receivers), then the precoder has no useful impact.

B. Solution for a Closed-Form Precoder

In this subsection, we derive a method to obtain a closed-form expression for the precoder in the particular case when the transmit antennas are uncorrelated but the receive antennas are not, i.e., $\boldsymbol{R}$ satisfies (13). From Theorem 2, the optimal precoder boils down to a diagonal precoder, i.e., the precoder amounts to a power allocation strategy. For the sake of space, we limit ourselves to the case of two transmitters. We also assume the two transmitters experience the same average path loss to the receiver. Generalizations to $M_t > 2$ and unequal path loss cases are addressed in a separate paper [24]. The number of receive antenna remains arbitrary. We also take the following normalization: $\frac{P}{aK\sigma_x^2} = 1$. For the diagonal 2×2 precoder $\boldsymbol{F}$ to satisfy the power constraint, it follows that $f_0^2 + f_1^2 = 1$, where $f_i \triangleq (\boldsymbol{F})_{i,i}$.

1) Equivalent SISO Channel Formulation: Let $\boldsymbol{H} = [\boldsymbol{h}_0, \;\; \boldsymbol{h}_1]$ and $\boldsymbol{R}_{r_i} = E\left[\boldsymbol{h}_i \boldsymbol{h}_i^H\right]$ have the following eigen-decomposition: $\boldsymbol{R}_{r_i} = \boldsymbol{V}_{r_i} \boldsymbol{\Lambda}_{r_i} \boldsymbol{V}_{r_i}^H$. From $\mathrm{vec}\,(\boldsymbol{H}) = \boldsymbol{R}^{1/2}\,\mathrm{vec}\,(\boldsymbol{H}_w)$, it follows that $\boldsymbol{h}_i = \boldsymbol{R}_{r_i}^{1/2} \boldsymbol{h}_{w_i}$, where $\boldsymbol{h}_{w_i}$ is an $M_r \times 1$ vector containing zero-mean complex Gaussian i.i.d. components. From the equivalent SISO model in (3), it is seen that all the K original symbols are going through the same SISO system. From the definitions of $\boldsymbol{\Phi}$ and α, it is seen that α, in the equivalent SISO model, can be expressed as:

$$\alpha = f_0^2 \|\boldsymbol{h}_0\|^2 + f_1^2 \|\boldsymbol{h}_1\|^2 = f_0^2 \sum_{j=0}^{M_r-1} \lambda_{r_{0_j}} |h'_{w_{0_j}}|^2 + f_1^2 \sum_{j=0}^{M_r-1} \lambda_{r_{1_j}} |h'_{w_{1_j}}|^2, \tag{14}$$

where the variable $h'_{w_{i_j}}$ is the jth component of the vector $\boldsymbol{V}_{r_i}^H \boldsymbol{h}_{w_i}$. Since $\boldsymbol{V}_{r_i}$ is unitary, each of the variables $h'_{w_{i_j}}$ has the same distribution as the variables $h_{w_{i_j}} \triangleq (\boldsymbol{h}_{w_i})_j$.

2) Equal Diversity Spread Principle: In this subsection, we examine the expression for α and propose a simple framework coined *equal diversity spreading* that allows us to determine the power weights f_0 and f_1 in closed form, thus serving as a practical alternative to the numerical-based optimization of the symbol error rate. Note that, we do not claim optimality of the approach below in terms of error rate, although we do conjecture the obtained coefficients are close to optimal, which is confirmed by our simulations later.

From (14), α is a sum of $2M_r$ uncorrelated diversity branches $h'_{w_{i_j}}$ weighted by $f_i^2 \lambda_{r_{i_j}}$. According to our proposed principle, we make these weights as similar to each other as possible in order to spread the symbol energy evenly across all diversity branches. This translates simply into a minimum variance problem.

Interestingly, the mean m of the weighing factors $f_i^2 \lambda_{r_{i_j}}$ is constant, given by

$$m = \frac{1}{2M_r} \sum_{i=0}^{1} \sum_{j=0}^{M_r-1} f_i^2 \lambda_{r_{i_j}} = \frac{f_0^2}{2M_r} \sum_{j=0}^{M_r-1} \lambda_{r_{0_j}} + \frac{f_1^2}{2M_r} \sum_{j=0}^{M_r-1} \lambda_{r_{1_j}} = \frac{1}{2M_r}(f_0^2 M_r + f_1^2 M_r) = \frac{1}{2},$$

where it is assumed that $\mathrm{Tr}\{\boldsymbol{R}_{r_i}\} = M_r \quad \forall \quad i$. The weights are now obtained from minimizing the variance (under mean constraint):

Problem 2:

$$\min_{\{f_0, f_1 \geq 0 \;|\; f_0^2 + f_1^2 = 1\}} \sum_{i=0}^{1} \sum_{j=0}^{M_r-1} \left(f_i^2 \lambda_{r_{i_j}} - \frac{1}{2} \right)^2.$$

Fortunately, this problem admits a simple closed-form solution which is detailed in the theorem below.

Theorem 4: *We parametrize the precoder according to* $f_0 = \cos(\theta)$ *and* $f_1 = \sin(\theta)$, *where* θ *is arbitrary in* $\left[0, \frac{\pi}{2}\right]$. *The solution to Problem 2 is given in terms of* θ *by:*

$$\tan\theta = \sqrt{\sum_{j=0}^{M_r-1} \lambda_{r_{0_j}}^2 \Bigg/ \sum_{j=0}^{M_r-1} \lambda_{r_{1_j}}^2}\,. \tag{15}$$

The proof of this theorem can be found in Appendix X.

Interpretations: Theorem 4 can be interpreted as follows: The power allocation scheme above assigns more power on the transmit branch experiencing less receiver correlation and less power on the other one. However, note that, this is *not* a water-filling strategy (unlike [4]) the power levels are always strictly bounded away from zero.

Interestingly, it can be shown that the principle above, beyond simple intuition, bears close connection to symbol error rate optimization and thus can be formally justified [24].

We now examine two scenario examples of application of this result.

Example 1 (Precoding for Kronecker Correlation): We can make the model used in (13) a Kronecker one by setting $\boldsymbol{R}_{r_0} = \boldsymbol{R}_{r_1}$, in which case the eigenvalues are characterized by $\lambda_{r_{0_j}} = \lambda_{r_{1_j}}$ which according to (15) yields $f_0^2 = f_1^2 = \frac{1}{2}$. In other words, if the transmit antennas are uncorrelated and the receive antenna are correlated but in a way that is independent of which transmit antenna is taken, then the best strategy is to pour power equally across the transmit antennas, which makes good intuitive sense. It means that the fact that the receive antennas are correlated when the transmitter antennas are uncorrelated, cannot be compensated for at the transmitter through precoding of the OSTBC signals in the Kronecker case.

Example 2 (Precoding for Non-Kronecker Correlation): Consider the distributed space time coding of signals with $M_t = 2$ with the case where the two transmit antennas see two widely different receive correlation matrices. Transmit antenna number 0 sees an uncorrelated receiver $\boldsymbol{R}_{r_0} = \boldsymbol{I}_{M_r}$. This corresponds to a link with M_r orders of diversity with a wide angle spread in the direction of arrival. While antenna number 1 sees a fully correlated receiver $\boldsymbol{R}_{r_1} = \mathbf{1}_{M_r \times M_r}$, where the matrix $\mathbf{1}_{M_r \times M_r}$ contains only ones and has size $M_r \times M_r$. Hence, transmit link from transmitter antenna number 1 corresponds to a link with no receive diversity, due to, e.g., a small angle spread in the direction of arrival, see

Figure 2. However, the overall MIMO channel still exhibits transmit diversity of order two. In this example, $\lambda_{r_{0_i}} = 1 \;\; \forall \;\; i \in \{0, 1, \ldots, M_r - 1\}$ and $\lambda_{r_{1_0}} = M_r$ and $\lambda_{r_{1_i}} = 0 \;\; \forall \;\; i \in \{1, \ldots, M_r - 1\}$. Theorem 4 yields directly $\tan\theta = \sqrt{\frac{1}{M_r}}$, thus, $f_0^2 = \frac{M_r}{M_r+1}$ and $f_1^2 = \frac{1}{M_r+1}$. Interestingly, this result is reminiscent of the classical water-filling result in information theory. Here, the channel quality is measured in terms of receive diversity order instead of average SNR. Hence, more power is poured into the transmit channel that exhibits more diversity and less into the other transmit channel.

VI. Optimization Algorithm for General Correlation Case

We now return to general case of arbitrary joint transmit-receive correlation, where a closed-form precoder is difficult to derive. Instead, we focus on a fast-converging numerical algorithm for the SER minimization problem.

Let $\boldsymbol{K}_{k,l}$ be the commutation[2] matrix [14] of size $kl \times kl$. The constrained maximization Problem 1 can be converted into an unconstrained optimization problem by introducing a Lagrange multiplier μ':

$$\mathcal{L}(\boldsymbol{F}) = \mathrm{SER} + \mu' \operatorname{Tr}\left\{\boldsymbol{F}\boldsymbol{F}^H\right\}. \tag{16}$$

Since the objective function should be minimized, $\mu' > 0$. Define the $M_t^2 \times M_t^2 M_r^2$ matrix $\boldsymbol{\Pi}$ as

$$\boldsymbol{\Pi} = \left[\boldsymbol{I}_{M_t^2} \otimes \operatorname{vec}^T\left(\boldsymbol{I}_{M_r}\right)\right]\left[\boldsymbol{I}_{M_t} \otimes \boldsymbol{K}_{M_t,M_r} \otimes \boldsymbol{I}_{M_r}\right]. \tag{17}$$

In order to present the results compactly, define the following $BM_t \times 1$ vector $\boldsymbol{s}(\boldsymbol{F}, \theta, g, \mu)$:

$$\boldsymbol{s}(\boldsymbol{F}, \theta, g, \mu) = \mu\left[\boldsymbol{F}^T \otimes \boldsymbol{I}_{M_t}\right]\boldsymbol{\Pi}\left[\boldsymbol{R}^{1/2} \otimes \left(\boldsymbol{R}^{1/2}\right)^*\right]\frac{\operatorname{vec}\left(\left[\boldsymbol{I}_{M_tM_r} + \delta\frac{g}{\sin^2(\theta)}\boldsymbol{\Phi}^*\right]^{-1}\right)}{\sin^2(\theta)\det\left(\boldsymbol{I}_{M_tM_r} + \delta\frac{g}{\sin^2(\theta)}\boldsymbol{\Phi}\right)}. \tag{18}$$

Theorem 5: *The precoder that is optimal for Problem 1 must satisfy:*

$$\operatorname{vec}(\boldsymbol{F}) = \int_0^{\frac{M-1}{M}\pi} \boldsymbol{s}(\boldsymbol{F}, \theta, g_{PSK}, \mu)d\theta, \tag{19}$$

$$\operatorname{vec}(\boldsymbol{F}) = \int_0^{\frac{\pi}{2}} \boldsymbol{s}(\boldsymbol{F}, \theta, g_{PAM}, \mu)d\theta, \tag{20}$$

$$\operatorname{vec}(\boldsymbol{F}) = \frac{1}{\sqrt{M}}\int_0^{\frac{\pi}{4}} \boldsymbol{s}(\boldsymbol{F}, \theta, g_{QAM}, \mu)d\theta + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \boldsymbol{s}(\boldsymbol{F}, \theta, g_{QAM}, \mu)d\theta. \tag{21}$$

for the M-PSK, M-PAM, and M-QAM constellations, respectively. μ is a positive scalar chosen such that the power constraint in Problem 1 is satisfied.

The proof of this theorem can be found in Appendix XI.

[2] The *commutation matrix* $\boldsymbol{K}_{k,l}$ is the *unique* $kl \times kl$ permutation matrix satisfying $\boldsymbol{K}_{k,l}\operatorname{vec}(\boldsymbol{A}) = \operatorname{vec}\left(\boldsymbol{A}^T\right)$ for all matrices $\boldsymbol{A} \in \mathbb{C}^{k\times l}$.

(19), (20), and (21) can be used in a fixed point iteration for finding the precoder that solves Problem 1. Notice that, the positive constants μ' and μ are in general different. When the fixed point iterations were used to find solutions, convergence was always observed.

VII. Results and Comparisons

Comparisons are made against a system using trivial precoding, i.e., $\boldsymbol{F} = \sqrt{\frac{P}{Ka\sigma_x^2 M_t}}\boldsymbol{I}_{M_t}$ and the system minimizing an upper bound of the PEP [9]. In the simulations, $\sigma_x^2 = 1/2$, $P = 1$, and $M_r = 6$.

Scenario 1: The following parameters are used in Scenario 1: The signal constellation was 8-PAM. The OSTBC $\boldsymbol{C}(\boldsymbol{x}) = \mathcal{G}_4^T$ in [13] was used, such that $a = 2$, $K = M_t = B = 4$, and $N = 8$. Let the correlation matrix $\boldsymbol{R}$ be given by $(\boldsymbol{R})_{k,l} = 0.9^{|k-l|}$, where the notation $(\cdot)_{k,l}$ picks out element with row number k and column number l.

Scenario 2: The OSTBC $\boldsymbol{C}(\boldsymbol{x}) = \mathcal{G}_2^T$ in [13] was used, such that $a = 1$ and $K = M_t = B = N = 2$. Let the correlation matrix $\boldsymbol{R}$ be given by (13) with $M_t = 2$, $\boldsymbol{R}_{r_0} = \boldsymbol{I}_{M_r}$, and $\boldsymbol{R}_{r_1} = \boldsymbol{1}_{M_r \times M_r}$. 9-QAM was used.

Scenario 3: The OSTBC $\boldsymbol{C}(\boldsymbol{x}) = \mathcal{G}_4^T$ in [13] was used, such that $a = 2$, $K = M_t = B = 4$, and $N = 8$. Let the correlation matrix $\boldsymbol{R}$ be given by (2) with $(\boldsymbol{R}_t)_{k,l} = 0.5^{|k-l|}$ and $(\boldsymbol{R}_r)_{k,l} = \rho^{|k-l|}$, where ρ is a scalar. In [4], [5], the optimization criterion used was an upper bound of the pairwise error probability when the Kronecker model is valid with $\boldsymbol{R}_r = \boldsymbol{I}_{M_r}$. If the notation used in this article is used the criterion used in [4], [5] can be written as $\det\left(\boldsymbol{I}_{M_t} + \delta g \boldsymbol{R}_t^{1/2}\boldsymbol{F}\boldsymbol{F}^H\boldsymbol{R}_t^{1/2}\right)$. In [9], this criterion was extended to any $\boldsymbol{R}$, and it is equivalent to maximizing $\det\left(\boldsymbol{I}_{M_t M_r} + \delta g \boldsymbol{\Phi}\right)$. Let $\mathrm{SNR} = 10$ dB and 9-QAM constellation was used.

Figures 3 and 4 show the SER versus SNR performance for Scenario 1 and 2, respectively, for the trivial precoder, the minimum upper bound PEP precoder [9], and the proposed minimum SER precoder in Theorem 5. For Scenario 2, the analytical precoder in Theorem 4 is also shown. From Figure 3, it is seen that the proposed minimum SER precoder outperforms the trivial precoder system systems for all values of SNR, however, the performance of the proposed system is similar to the minimum PEP precoder for low and high values of SNR. For moderate values of SNR, a gain up to 0.13 dB can be achieved, as seen from magnified version of the results within Figure 3. In Figure 4, the performance of the minimum SER, PEP and the precoder in Theorem 4 are indistinguishable in this example, and the performance of these three precoders are up to 1 dB better than the trivial precoder.

In Figure 5, the SER versus ρ performance of the systems are shown for Scenario 3. The proposed minimum SER precoder performs best, however, the maximum determinant precoder in [9] performs very close to the minimum SER coder. It is observed from Figure 5, that when the $\boldsymbol{R} = \mathbf{1}_{M_r \times M_r}$, i.e., $\rho = 1$, then the trivial precoder performs better than the precoder in [4], [5] which was designed based on the transmitter correlation only.

Monte Carlo simulations verify the exact theoretical SER expressions.

VIII. Conclusions

For an arbitrary given OSTBC, exact SER expressions, which are easy to evaluate, have been derived for a precoded MIMO system with *arbitrary* joint correlations in the transmitter and the receiver, for a ML receiver. Several key properties of the optimal precoder were derived. In particular, we show that receive correlation has an impact on precoding in general, to the sole exception of the case with zero transmit correlation *and* Kronecker-based receiver correlation. In the special case of cooperative diversity with two transmitters, we present a closed-form precoder which approximates well the optimal precoder. In the general case, an iterative method was proposed for finding the minimum SER precoder for M-PSK, M-PAM, and M-QAM signaling.

Appendix I
Proof of Lemma 1

Proof: Let $\boldsymbol{F}$ be an optimal solution of Problem 1 and $\boldsymbol{W} \in \mathbb{C}^{B \times B}$ be an arbitrary unitary matrix. It is then seen by insertion of $\boldsymbol{FW}$ as the precoder that the objective function and the power constraint are unaltered by the unitary matrix. ■

Appendix II
Proof of Proposition 1

Proof: Let $B = M_t$ and assume that the singular value decomposition of the optimal precoder can be expressed as: $\boldsymbol{F} = \boldsymbol{V}_0 \boldsymbol{\Sigma} \boldsymbol{V}_1^H$, where $\boldsymbol{V}_i \in \mathbb{C}^{M_t \times M_t}$ is unitary and the $M_t \times M_t$ matrix $\boldsymbol{\Sigma}$ contains the (non-negative) singular values of the precoder on its main diagonal and zeros elsewhere. It follows from Lemma 1, that the precoder $\boldsymbol{F}\boldsymbol{V}_1\boldsymbol{V}_0^H = \boldsymbol{V}_0\boldsymbol{\Sigma}\boldsymbol{V}_0^H$ is also optimal. But this precoder is Hermitian. Symmetry follows in a similar way since $\boldsymbol{F}\boldsymbol{V}_1\boldsymbol{V}_0^T = \boldsymbol{V}_0\boldsymbol{\Sigma}\boldsymbol{V}_0^T$ is also optimal and symmetric. ■

APPENDIX III

PROOF OF PROPOSITION 2

Proof: Let $\mathrm{SNR} = P/N_0 \to \infty$. By studying the expressions for SER in (9), (10), and (11), it follows that the identity matrices inside the determinants in these equations can be eliminated if $\boldsymbol{R}$ is non-singular. Then, the problem can be rewritten as finding the maximum of $\det(\boldsymbol{\Phi})$ under the power constraint. This problem is again equivalent to maximize $\det\left(\boldsymbol{F}\boldsymbol{F}^H\right)$ subject to $\mathrm{Tr}\left\{\boldsymbol{F}\boldsymbol{F}^H\right\} = \frac{P}{aK\sigma_x^2}$. It can be shown that the solution of this symmetrical equivalent problem is the trivial precoder. ■

APPENDIX IV

PROOF OF PROPOSITION 3

Proof: Let $\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{U}_{\boldsymbol{F}\boldsymbol{F}^H}\boldsymbol{\Lambda}_{\boldsymbol{F}\boldsymbol{F}^H}\boldsymbol{U}_{\boldsymbol{F}\boldsymbol{F}^H}^H$ be the eigen-decomposition of $\boldsymbol{F}\boldsymbol{F}^H$. Observe that $\boldsymbol{R} = \boldsymbol{I}_{M_tM_r} = \boldsymbol{I}_{M_t} \otimes \boldsymbol{I}_{M_r}$. For M-PAM and M-PSK signal constellations, the optimization problem can be formulated as to find the minimum of

$$\int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_t} \otimes \boldsymbol{I}_{M_r} + \frac{\delta g}{\sin^2\theta}\left(\boldsymbol{F}^*\boldsymbol{F}^T\right) \otimes \boldsymbol{I}_{M_r}\right)} = \int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det^{M_r}\left(\boldsymbol{I}_{M_t} + \frac{\delta g}{\sin^2\theta}\boldsymbol{F}^*\boldsymbol{F}^T\right)}$$
$$= \int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det^{M_r}\left(\boldsymbol{I}_{M_t} + \frac{\delta g}{\sin^2\theta}\boldsymbol{\Lambda}_{\boldsymbol{F}\boldsymbol{F}^H}\right)}, \tag{22}$$

under the constraint that $\|\boldsymbol{F}\|_F^2 = \frac{P}{Ka\sigma_x^2} = \mathrm{Tr}\left\{\boldsymbol{\Lambda}_{\boldsymbol{F}\boldsymbol{F}^H}\right\}$. This is a symmetrical problems where all the independent variables can be interchanged without altering the objective function and the constraint. For this reason, in the optimum, all independent variables will be equal in the optimum point. Because of the power constraint, the optimal solution can be written as: $\boldsymbol{F}\boldsymbol{F}^H = \frac{P}{Ka\sigma_x^2 M_t}\boldsymbol{I}_{M_t}$, and then the results follows from Lemma 1 for M-PSK and M-PAM signaling. For M-QAM signaling a similar proof can be given with the only difference that the objective function has two integrals with similar form. ■

APPENDIX V

PROOF OF PROPOSITION 4

Proof: Let $\mathrm{rank}(\boldsymbol{R}) = d$, and let the eigen-decomposition of $\boldsymbol{R}$ be given by:

$$\boldsymbol{R} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H = [\boldsymbol{U}_+ \;\; \boldsymbol{U}_0] \begin{bmatrix} \boldsymbol{\Lambda}_+ & \boldsymbol{0}_{d\times(M_tM_r-d)} \\ \boldsymbol{0}_{(M_tM_r-d)\times d} & \boldsymbol{0}_{(M_tM_r-d)\times(M_tM_r-d)} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_+^H \\ \boldsymbol{U}_0 \end{bmatrix}, \tag{23}$$

where $\boldsymbol{\Lambda}_+$ of size $d \times d$, contains the positive eigenvalues of $\boldsymbol{R}$, the $M_tM_r \times d$ matrix $\boldsymbol{U}_+$ contains the normalized eigenvectors corresponding to the positive eigenvalues, and $\boldsymbol{U}_0$ of size $M_tM_r \times (M_tM_r - d)$, contains the normalized eigenvectors corresponding the the eigenvalue 0. The eigen-decomposition of $\boldsymbol{R}^{1/2}$ follows from (23), and with this, the integral inside the SER expression can be written as

$$\int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_tM_r} + \frac{\delta g}{\sin^2\theta}\boldsymbol{R}^{1/2}\left[\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)\otimes \boldsymbol{I}_{M_r}\right]\boldsymbol{R}^{1/2}\right)}$$
$$= \int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\boldsymbol{I}_d + \frac{\delta g}{\sin^2\theta}\left[\boldsymbol{\Lambda}_+^{1/2} \;\; \mathbf{0}_{d\times(M_tM_r-d)}\right]\boldsymbol{U}^H\left[\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)\otimes \boldsymbol{I}_{M_r}\right]\boldsymbol{U}\begin{bmatrix}\boldsymbol{\Lambda}_+^{1/2}\\ \mathbf{0}_{(M_tM_r-d)\times d}\end{bmatrix}\right)}. \tag{24}$$

The diversity is found by letting $\mathrm{SNR}\rightarrow+\infty$. This implies that in the matrix sum within the determinant of (24), the identity matrix can be eliminated and (24) approaches:

$$\int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\frac{\delta g}{\sin^2\theta}\left[\boldsymbol{\Lambda}_+^{1/2} \;\; \mathbf{0}_{d\times(M_tM_r-d)}\right]\boldsymbol{U}^H\left[\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)\otimes \boldsymbol{I}_{M_r}\right]\boldsymbol{U}\begin{bmatrix}\boldsymbol{\Lambda}_+^{1/2}\\ \mathbf{0}_{(M_tM_r-d)\times d}\end{bmatrix}\right)}$$
$$= \delta^{-d}\frac{\int_{\theta_{\min}}^{\theta_{\max}}\sin^{2d}\theta d\theta}{\det\left(g\left[\boldsymbol{\Lambda}_+^{1/2} \;\; \mathbf{0}_{d\times(M_tM_r-d)}\right]\boldsymbol{U}^H\left[\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)\otimes \boldsymbol{I}_{M_r}\right]\boldsymbol{U}\begin{bmatrix}\boldsymbol{\Lambda}_+^{1/2}\\ \mathbf{0}_{(M_tM_r-d)\times d}\end{bmatrix}\right)}, \tag{25}$$

which shows that SNR is proportional with δ^{-d}. This means that the diversity of the system with a precoder satisfying $\mathrm{rank}\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)=\mathrm{rank}\left(\boldsymbol{F}\right)=M_t$ is $d=\mathrm{rank}\left(\boldsymbol{R}\right)$. ∎

Remark 3: *The expression in (25) gives an asymptotic expression for the* SER *for large values of* SNR *and it shows how* SER *depends on the correlation matrix* $\boldsymbol{R}$ *through* $\boldsymbol{U}$ *and* $\boldsymbol{\Lambda}_+^{1/2}$ *for high values of* SNR*. (25) can be used to find the explicit dependency of* SER *on the transmitter and receiver correlation matrix in the Kronecker model (2) for high* SNR *values. For full ranked matrices* $\boldsymbol{R}$*, the asymptotic expression in (25) can be simplified.*

APPENDIX VI

PROOF OF PROPOSITION 5

Proof: If $B<M_t$, then $\mathrm{rank}\left(\boldsymbol{F}\right)=\mathrm{rank}\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)\leq B<M_t$. This means that $\mathrm{rank}(\boldsymbol{\Phi})=M_r\,\mathrm{rank}\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)<M_rM_t$, and, therefore, diversity is lost in comparison the case where $B=M_t$. ∎

APPENDIX VII

PROOF OF THEOREM 1

First, a lemma is derived to help deriving the main result.

Lemma 2: *Let $\boldsymbol{A}(\theta) \in \mathbb{C}^{N\times N}$ be a positive definite matrix for all $\theta \in [\theta_{\min}, \theta_{\max}]$, and let the operator $\mathrm{dg} : \mathbb{C}^{N\times N} \rightarrow \mathbb{C}^{N\times N}$ return the matrix with zero off-diagonal elements and with the same diagonal elements as the matrix it is applied to. Then the following inequality holds:*

$$\int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\boldsymbol{A}(\theta)\right)} \geq \int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\mathrm{dg}\left(\boldsymbol{A}(\theta)\right)\right)}, \tag{26}$$

with equality if and only if $\boldsymbol{A}(\theta)$ is diagonal for all $\theta \in [\theta_{\min}, \theta_{\max}]$.

Proof: Since $\boldsymbol{A}(\theta)$ is positive definite, it follows from Theorem 1.28 in [14] that

$$\det\left(\boldsymbol{A}(\theta)\right) \leq \det\left(\mathrm{dg}\left(\boldsymbol{A}(\theta)\right)\right), \tag{27}$$

for all $\theta \in [\theta_{\min}, \theta_{\max}]$ and with equality if and only if $\boldsymbol{A}(\theta)$ is diagonal. Since $\boldsymbol{A}(\theta)$ is positive definite, (27) is equivalent to

$$\frac{1}{\det\left(\boldsymbol{A}(\theta)\right)} \geq \frac{1}{\det\left(\mathrm{dg}\left(\boldsymbol{A}(\theta)\right)\right)}, \tag{28}$$

for all $\theta \in [\theta_{\min}, \theta_{\max}]$ and equality is achieved if and only if $\boldsymbol{A}(\theta)$ is a diagonal matrix. Since (28) is valid for all $\theta \in [\theta_{\min}, \theta_{\max}]$, it follows by considering integrals as a sum, that (26) holds. If $\boldsymbol{A}(\theta)$ is diagonal for all $\theta \in [\theta_{\min}, \theta_{\max}]$, it follows from Theorem 1.28 in [14] that (26) holds with equality. Assume now that (26) holds with equality. From (28), it follows that one of the integrands are always less than or equal to the other integrand for all values of θ. Then, the only way the two integrals can be equal is that the integrands are equal for all values of $\theta \in [\theta_{\min}, \theta_{\max}]$. ■

Proof: Let the receiver correlation matrix have the following eigen-decomposition: $\boldsymbol{R}_r = \boldsymbol{U}_r \boldsymbol{\Lambda}_r \boldsymbol{U}_r^H$, where $\boldsymbol{U}_r \in \mathbb{C}^{M_r\times M_r}$ is unitary and and $\boldsymbol{\Lambda}_r$ is diagonal of size $M_r \times M_r$. The integral in the SER, using $\boldsymbol{R}$ from (2), can be rewritten as:

$$\int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_t M_r} + \frac{\delta g}{\sin^2\theta}\boldsymbol{\Phi}\right)} = \int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_t M_r} + \frac{\delta g}{\sin^2\theta}\left[\boldsymbol{\Lambda}_t^{1/2}\boldsymbol{U}_t^T\boldsymbol{F}^*\boldsymbol{F}^T\boldsymbol{U}_t^*\boldsymbol{\Lambda}_t^{1/2}\right]\otimes\boldsymbol{\Lambda}_r\right)}. \tag{29}$$

Using Lemma 2, it is seen that the SER is minimized if and only if $\boldsymbol{\Lambda}_t^{1/2}\boldsymbol{U}_t^T\boldsymbol{F}^*\boldsymbol{F}^T\boldsymbol{U}_t^*\boldsymbol{\Lambda}_t^{1/2}$ is diagonal. Therefore, it follows that there is no loss of optimality to restrict $\boldsymbol{F}^*\boldsymbol{F}^T$ to the following form

$$\boldsymbol{F}^*\boldsymbol{F}^T = \boldsymbol{U}_t^*\boldsymbol{D}\boldsymbol{U}_t^T, \tag{30}$$

where $\boldsymbol{D}$ is a diagonal $M_t \times M_t$ matrix. Since $\boldsymbol{F}$ is of size $M_t \times B$ is satisfying (30), it is seen that the theorem follows with $(\boldsymbol{\Delta})_{i,i} = \sqrt{(\boldsymbol{D})_{i,i}}$. ■

Appendix VIII

Proof of Theorem 2

We need a lemma to establish the proof below. First, let the eigenvalue decomposition of $\boldsymbol{R}_{r_i}$ be given by $\boldsymbol{R}_{r_i} = \boldsymbol{V}_{r_i}\boldsymbol{\Lambda}_{r_i}\boldsymbol{V}_{r_i}^H$, where $\boldsymbol{V}_{r_i} \in \mathbb{C}^{M_r \times M_r}$ is unitary and $\boldsymbol{\Lambda}_{r_i} \in \mathbb{R}^{M_r \times M_r}$ is diagonal with non-negative diagonal elements. It follows that the eigenvalue decomposition of $\boldsymbol{R} = \boldsymbol{V}_{\boldsymbol{R}}\boldsymbol{\Lambda}_{\boldsymbol{R}}\boldsymbol{V}_{\boldsymbol{R}}^H$ is given by the matrices

$$\boldsymbol{V}_{\boldsymbol{R}} = \begin{bmatrix} \boldsymbol{V}_{r_0} & \boldsymbol{0}_{M_r\times M_r} & \cdots & \boldsymbol{0}_{M_r\times M_r} \\ \boldsymbol{0}_{M_r\times M_r} & \boldsymbol{V}_{r_1} & \cdots & \boldsymbol{0}_{M_r\times M_r} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0}_{M_r\times M_r} & \boldsymbol{0}_{M_r\times M_r} & \cdots & \boldsymbol{V}_{r_{M_t-1}} \end{bmatrix}, \quad \boldsymbol{\Lambda}_{\boldsymbol{R}} = \begin{bmatrix} \boldsymbol{\Lambda}_{r_0} & \boldsymbol{0}_{M_r\times M_r} & \cdots & \boldsymbol{0}_{M_r\times M_r} \\ \boldsymbol{0}_{M_r\times M_r} & \boldsymbol{\Lambda}_{r_1} & \cdots & \boldsymbol{0}_{M_r\times M_r} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0}_{M_r\times M_r} & \boldsymbol{0}_{M_r\times M_r} & \cdots & \boldsymbol{\Lambda}_{r_{M_t-1}} \end{bmatrix}. \tag{31}$$

Lemma 3: *Let the correlation matrix $\boldsymbol{R}$ satisfy (13) and let $\boldsymbol{W} \in \mathbb{C}^{M_t \times M_t}$. The matrix given by $\boldsymbol{B} \triangleq \frac{\delta g}{\sin^2\theta}\boldsymbol{\Lambda}_{\boldsymbol{R}}^{1/2}\boldsymbol{V}_{\boldsymbol{R}}^H\left[\boldsymbol{W}\otimes\boldsymbol{I}_{M_r}\right]\boldsymbol{V}_{\boldsymbol{R}}\boldsymbol{\Lambda}_{\boldsymbol{R}}^{1/2}$ is diagonal if and only if $\boldsymbol{W}$ is diagonal.*

Proof: Block element number (k,l) of size $M_r \times M_r$ of $\boldsymbol{B}$ is given by $\frac{\delta g}{\sin^2\theta}\boldsymbol{\Lambda}_{\boldsymbol{r}_k}^{1/2}\boldsymbol{V}_{\boldsymbol{r}_k}^H\boldsymbol{V}_{\boldsymbol{r}_l}\boldsymbol{\Lambda}_{\boldsymbol{r}_l}^{1/2}\left(\boldsymbol{W}\right)_{k,l}$. From this expression it is seen that $\boldsymbol{B}$ is diagonal if and only if $\boldsymbol{W}$ is diagonal. ■

Proof: The SER objective function that should be minimized is proportional to:

$$\int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_tM_r} + \frac{\delta g}{\sin^2(\theta)}\boldsymbol{\Lambda}_{\boldsymbol{R}}^{1/2}\boldsymbol{V}_{\boldsymbol{R}}^H\left[\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)\otimes\boldsymbol{I}_{M_r}\right]\boldsymbol{V}_{\boldsymbol{R}}\boldsymbol{\Lambda}_{\boldsymbol{R}}^{1/2}\right)}. \tag{32}$$

By using (32) together with Lemmas 3 and 2, it follows that that SER is minimized under the power constraint if and only if $\boldsymbol{F}^*\boldsymbol{F}^T$ is diagonal. Hence, from Lemma 1, it follows that $\boldsymbol{F}$ is diagonal. From the expressions for SER and the power constraint, it is seen that without loss of optimality, the diagonal elements of $\boldsymbol{F}$ can be chosen real and non-negative. ■

Appendix IX

Proof of Theorem 3

Proof: The correlation matrix is given by (13) with equal block diagonal element matrices. Let $\boldsymbol{R}_r = \boldsymbol{V}_r\boldsymbol{\Lambda}_r\boldsymbol{V}_r^H$. From Theorem 2, it follows that $\boldsymbol{F}$ can be chosen diagonal without loss of optimality. By inserting a diagonal $\boldsymbol{F}^*\boldsymbol{F}^T$ into (32), the goal is to minimize:

$$\int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\prod_{i=0}^{M_t-1}\det\left(\boldsymbol{I}_{M_r} + \frac{\delta g}{\sin^2(\theta)}\left(\boldsymbol{F}^*\boldsymbol{F}^T\right)_{i,i}\boldsymbol{\Lambda}_r\right)}, \tag{33}$$

subject to the power constraint $\mathrm{Tr}\left\{\boldsymbol{F}^*\boldsymbol{F}^T\right\} = P/\left(Ka\sigma_x^2\right)$. This is a symmetrical optimization problem in the unknown diagonal elements of the precoder $\boldsymbol{F}$, it follows that all the diagonal elements of $\boldsymbol{F}$ should be made equal and then the theorem follows from the power constraint. ■

APPENDIX X

PROOF OF THEOREM 4

Proof: We rewrite the cost function in Problem 2 in terms of θ (unconstrained):

$$J(\theta) = \sum_{j=0}^{M_r-1} \left[\left(\cos^2(\theta)\lambda_{r_{0_j}} - \frac{1}{2} \right)^2 + \left(\sin^2(\theta)\lambda_{r_{1_j}} - \frac{1}{2} \right)^2 \right]. \tag{34}$$

The minimum is reached when $\frac{\partial J}{\partial \theta} = 0$. By eliminating false uninteresting extremal points, we find: $\sum_{j=0}^{M_r-1} \left(\cos^2(\theta)\lambda_{r_{0_j}} - \frac{1}{2} \right) \lambda_{r_{0_j}} = \sum_{j=0}^{M_r-1} \left(\sin^2(\theta)\lambda_{r_{1_j}} - \frac{1}{2} \right) \lambda_{r_{1_j}}$, which, knowing $\sum_{j=0}^{M_r-1} \lambda_{r_{0_j}} = \sum_{j=0}^{M_r-1} \lambda_{r_{1_j}} = M_r$, gives the result in (15). Since the right hand side of (15) is positive and the function $\tan(\cdot)$ is periodic with period π, it is enough to consider $\theta \in \left[0, \frac{\pi}{2}\right]$. ■

APPENDIX XI

PROOF OF THEOREM 5

Proof: The necessary condition for the optimality of Problem 1 is found by setting the derivative of the Lagrangian in (16) with respect to $\mathrm{vec}\left(\boldsymbol{F}^*\right)$ equal to zero. Finding the derivative with respect to the complex valued vector $\mathrm{vec}\left(\boldsymbol{F}^*\right)$ can be done by generalizing the works in [14], [25]. The following two expressions, which are found after several matrix manipulations, are useful:

$$\frac{\partial}{\partial\, \mathrm{vec}\left(\boldsymbol{F}^*\right)} \mathrm{Tr}\left\{ \boldsymbol{F}\boldsymbol{F}^H \right\} = \mathrm{vec}\left(\boldsymbol{F}\right), \tag{35}$$

$$\begin{aligned} \frac{\partial}{\partial\, \mathrm{vec}\left(\boldsymbol{F}^*\right)} \int_{\theta_{\min}}^{\theta_{\max}} \frac{d\theta}{\det\left(\boldsymbol{I}_{M_t M_r} + \frac{\delta g}{\sin^2\theta}\boldsymbol{\Phi} \right)} &= -\,\delta g \left[\boldsymbol{F}^T \otimes \boldsymbol{I}_{M_t} \right] \boldsymbol{\Pi} \left[\boldsymbol{R}^{1/2} \otimes \left(\boldsymbol{R}^{1/2} \right)^* \right] \\ &\times \int_{\theta_{\min}}^{\theta_{\max}} \frac{\mathrm{vec}\left(\left[\boldsymbol{I}_{M_t M_r} + \frac{\delta g}{\sin^2\theta}\boldsymbol{\Phi}^* \right]^{-1} \right)}{\sin^2\theta \det\left(\boldsymbol{I}_{M_t M_r} + \frac{\delta g}{\sin^2\theta}\boldsymbol{\Phi} \right)} d\theta. \end{aligned} \tag{36}$$

The necessary condition for optimality is found by utilizing the results from (35) and (36) and setting the derivative of the Lagrangian in (16) equal to zero. If this is done, and scalar factors are collected into the scalar named μ, the results in (19), (20), and (21) are found. Since the precoder matrix $\boldsymbol{F}$ should be scaled according to the power constraint in Problem 1, it is not necessary to decide the exact value of the scalar μ. This scalar can be found by adjusting the norm of the precoder according to the power constraint after each time the fixed point iteration is used. ■

REFERENCES

[1] A. Hjørungnes and D. Gesbert, "Minimum exact SER precoding of orthogonal space-time block codes for correlated MIMO channels," in *Proc. IEEE GLOBECOM*, vol. 1, Dallas, USA, Nov. - Dec. 2004, pp. 111–115.

[2] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.

[3] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, United Kingdom: Cambridge University Press, May 2003.

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

[5] ——, "Corrections to: 'Linear precoding for space-time coded systems with known fading correlations'," *IEEE Communications Letters*, submitted 2004.

[6] R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Cut-off rate based transmit optimization for spatial multiplexing on general MIMO channels," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc.*, vol. 5, 2003, pp. 61–64.

[7] R. U. Nabar, H. Bolcskei, and A. J. Paulraj, "Transmit optimization for spatial multiplexing in the presence of spatial fading correlation," in *Proc. IEEE GLOBECOM*, vol. 1, Nov. 2001, pp. 131–135.

[8] E. Bonek, H. Özcelik, M. Herdin, W. Weichselberger, and J. Wallace, "Deficiencies of a popular stochastic MIMO radio channel model," in *Proc. Int. Symp. on Wireless Personal Multimedia Communications*, Yokosuka, Japan, Oct. 2003.

[9] A. Hjørungnes, D. Gesbert, and J. Akhtar, "Precoding of space-time block coded signals for joint transmit-receive correlated MIMO channels," *IEEE Trans. Wireless Communications*, accepted for publication.

[10] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1673–1690, July 2003.

[11] H. Shin and J. H. Lee, "Exact symbol error probability of orthogonal space-time block codes," in *Proc. IEEE GLOBECOM*, vol. 2, Nov. 2002, pp. 1197–1201.

[12] M. Gharavi-Alkhansari, A. B. Gershman, and M. Haardt, "Exact error probability analysis of orthogonal space-time block codes over correlated rician fading channels," in *Proceedings of the ITG Workshop on Smart Antennas (ITG 2004)*, Munich, Germany, Mar. 2004.

[13] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 451–460, Mar. 1999.

[14] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Application in Statistics and Econometrics*. Essex, England: John Wiley & Sons, Inc., 1988.

[15] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press Cambridge, UK, 1991, reprinted 1999.

[16] X. Li, T. Luo, G. Yue, and C. Yin, "A squaring method to simplify the decoding of orthogonal space-time block codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1700–1703, Oct. 2001.

[17] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*. Wiley Series in Telecommunications and Signal Processing, 2000.

[18] D. Raphaeli, "Distribution of noncentral indefinite quadratic forms in complex normal variables," *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 1002–1007, May 1996.

[19] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press Cambridge, UK, 1985, reprinted 1999.

[20] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[21] P. A. Anghel and M. Kaveh, "On the diversity of cooperative systems," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc.*, vol. IV, Montreal, Canada, May 2004, pp. 577–580.

[22] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[23] S. A. Jafar and A. Goldsmith, "Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information," *IEEE Trans. Wireless Communications*, accepted for publications. [Online]. Available: http://systems.stanford.edu/~syed/journal1.pdf

[24] H. Skjevling, D. Gesbert, and A. Hjørungnes, "Precoding for distributed space-time coded downlink," 2005, in preperation.

[25] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proc., Parts F and H*, vol. 130, no. 1, pp. 11–16, Feb. 1983.

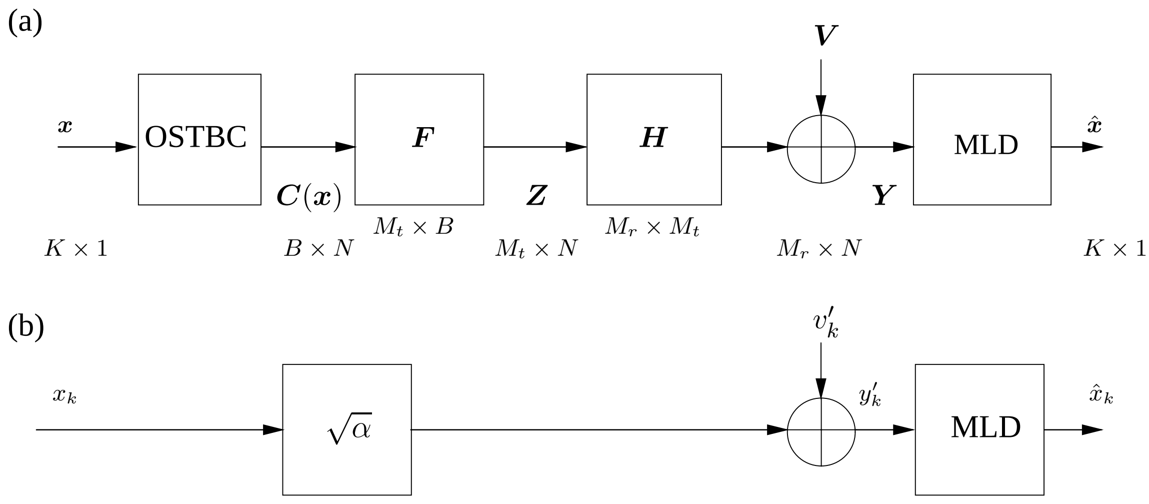

Fig. 1. (a) Block model of the linear precoded OSTBC MIMO system. (b) Equivalent SISO system.

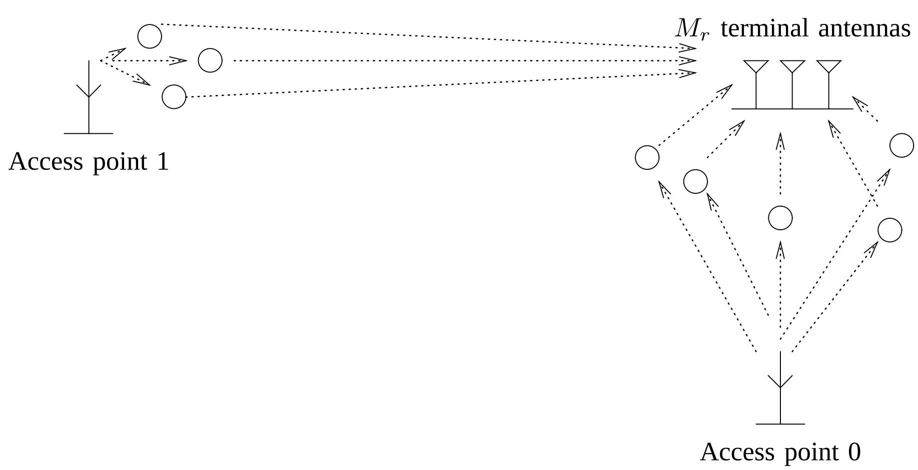

Fig. 2. Illustration of non-Kronecker correlation in distributed space time coding. Circles indicate scatterers. Unlike Access point 0, Access point 1 experiences small angle spread at the receiver, yielding high receive correlation. The two access points, remotely located, are uncorrelated.

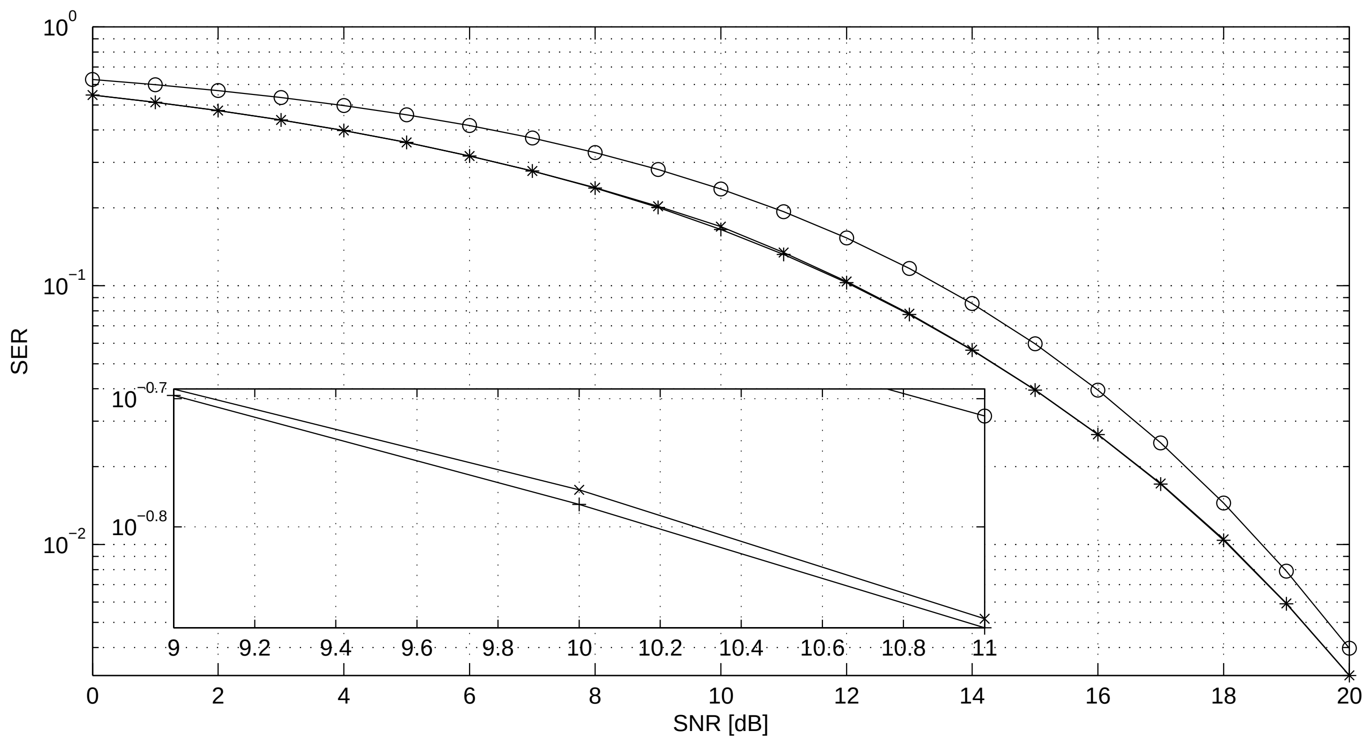

Fig. 3. Scenario 1: SER versus SNR performance of the proposed minimum SER precoder $-+-$ in Theorem 5, the trivial precoder $-\circ-$, and the minimum PEP precoder $-\times-$ proposed in [9]. A magnified portion of the curves is shown to illustrate the differences in performance.

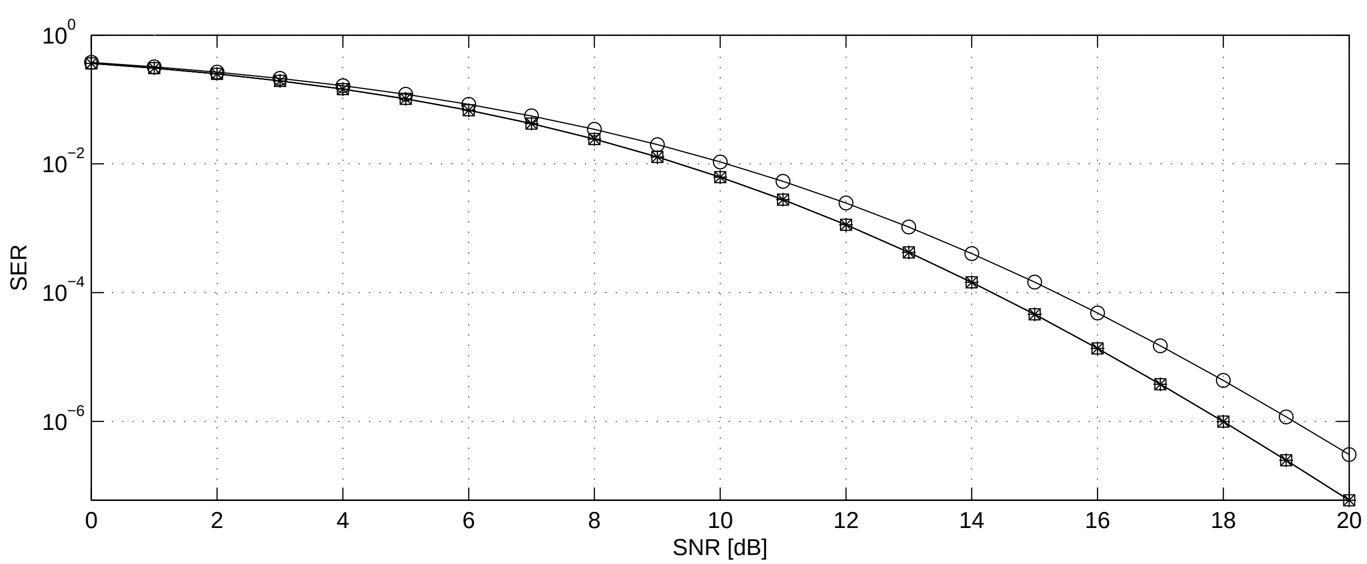

Fig. 4. Scenario 2: SER versus SNR performance of the proposed precoder $-+-$ in Theorem 5, the PEP precoder $-\times-$, the precoder in Theorem 4 $-\Box-$, and the trivial precoder $-\circ-$.

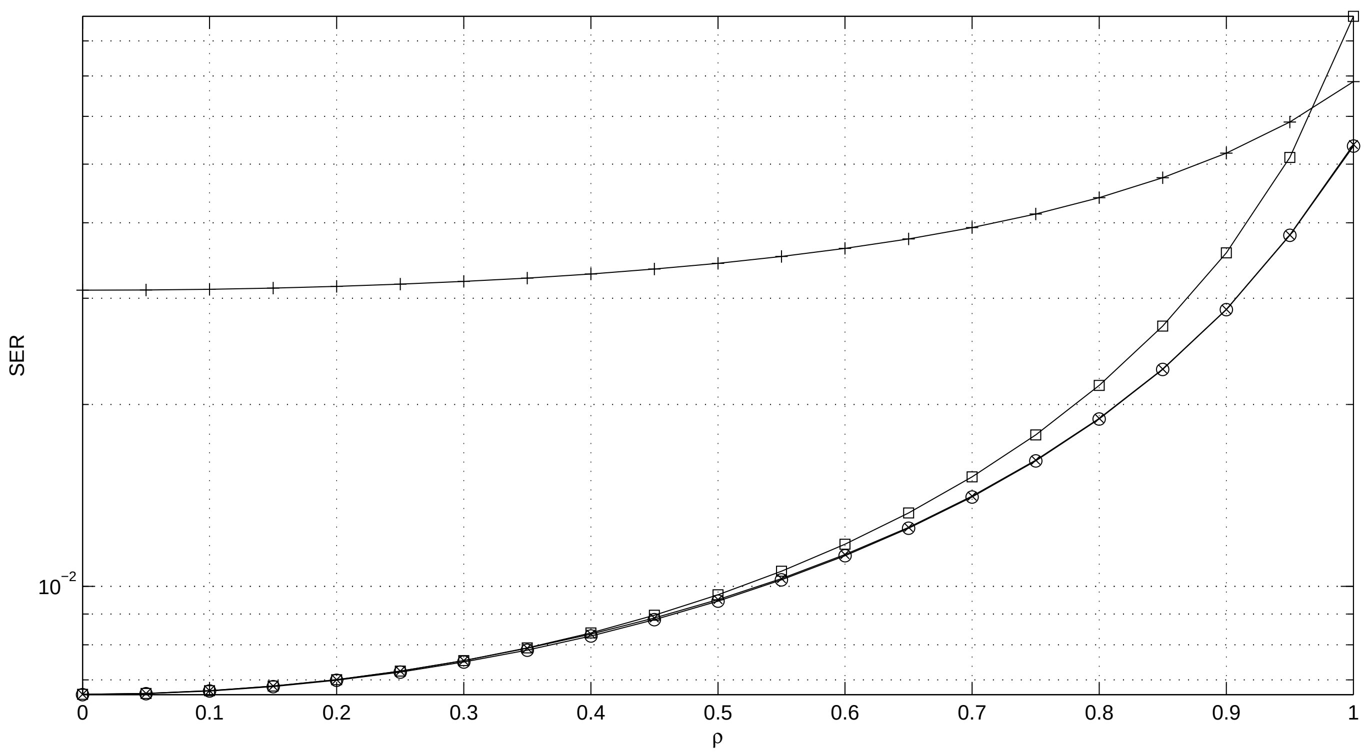

Fig. 5. Scenario 3: The SER versus ρ performances are shown for the following four systems: The proposed minimum SER precoder: $-\circ-$ in Theorem 5, the PEP precoder $-\times-$ in [9], the analytical precoder $-\square-$ in [4], [5], and the trivial precoder $-+-$.