Space–Time Spreading and Block Coding for Correlated Fading Channels in the Presence of Interference

Xiaodong Cai, *Member, IEEE*, Georgios B. Giannakis, *Fellow, IEEE*, and Michael D. Zoltowski, *Fellow, IEEE*

***Abstract*—We consider point-to-point wireless links with multiple antennas in the presence of interference, and exploit channel's spatial correlation and the temporal covariance of the interference to design multiantenna transmitters. We develop a space–time spreading scheme that maximizes average signal-to-interference-and-noise ratio, and an optimally power-loaded space–time beamforming (STBF) scheme which improves error-probability performance. In order to increase transmission rates, we combine orthogonal space–time block coding with STBF, optimize power loading across beams, and develop low-complexity receivers. Optimal training for least-squares error channel estimation, and STBF for minimum mean-square error channel estimation, are also studied. Our analytical and simulated results corroborate that STBF with optimal power loading can considerably reduce error probability and channel-estimation errors.**

***Index Terms*—Beamforming, covariance feedback, interference suppression, multiple-input multiple-output (MIMO), space–time coding (STC).**

I. Introduction

Multiantenna transmissions can boost spectral efficiency [1], [2], and enable the diversity provided by multiple-input multiple-output (MIMO) wireless channels. Without channel knowledge at the transmitter, space–time coding (STC) offers an effective fading countermeasure, and thereby reduces error probability [3], [4]. Whenever (even partial) channel state information (CSI) is available at the transmitter, it can be exploited to further improve performance. As the transmitter cannot acquire CSI perfectly, use of partial CSI at the transmitter has received considerable attention recently. Performance optimization in terms of average signal-to-noise ratio (SNR) and expected mutual information based on partial CSI was pursued in [5]. Based on a statistical model of partial CSI [6]–[8], a linear transformation (beamformer) was combined with orthogonal space–time block coding (STBC) to improve error probability in [6]. Since in frequency-division duplex (FDD) systems, the transmitter obtains CSI through a feedback channel, two cases of partial CSI, termed mean and covariance feedback, were used to maximize the ergodic capacity of multiple-input single-output (MISO) channels in [7], whereas outage or/and ergodic capacity based on covariance or/and mean feedback were investigated in [9] and [10]–[12]. Optimal beamforming and STBC minimizing the symbol-error rate (SER) were derived in [8] and [13], based on channel correlation and mean, respectively.

Besides fading, wireless links are also affected by intentional or unintentional interference. In multiple access, e.g., cellular systems, multiple users share the same radio spectrum, which typically causes multiuser interference as well as co-channel interference (CCI). In military communications operating in hostile environments, intentional jammers may severely affect the reliability of wireless links. Capacity of MIMO channels with interference were investigated in [14] and [15]. Receiver design for multiantenna systems with STBC in the presence of *spatially* correlated interference was studied in [16]. User code designs for CDMA systems exploiting interference covariance were recently investigated in [17] and [18] for single-input single-output (SISO) channels. While the channel's spatial correlation was exploited in [6] and [8] to improve error performance of STBC, use of the channel's spatial correlation and the interference's temporal correlation has not been considered.

In this paper, we consider multiantenna point-to-point links in the presence of *temporally* correlated interference. Exploiting the channel's spatial correlation along with the temporal correlation of the interference, we will first develop space–time spreading (STS) transmission schemes, and show that if we transmit along directions where the channels are "strong" and the interference is "weak," considerable performance gain can be achieved. While STS can exploit channel diversity and effectively suppress interference, it is not spectrally efficient. To improve spectral efficiency, we combine orthogonal STBC with STS to increase transmission rates. In the absence of interference, or, if the interference is temporally white, a maximum-likelihood (ML) receiver for detecting orthogonal STBC reduces to a simple symbol-by-symbol linear receiver [4], [16].

Paper approved by J. Wang, the Editor for Wireless Spread Spectrum of the IEEE Communications Society. Manuscript received February 13, 2004; revised June 16, 2004. The work of X. Cai and G. B. Giannakis was supported through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. The work of M. D. Zoltowski was supported by the Air Force Office of Scientific Research under Grant F49620-03-1-0149. This paper was presented in part at the Signal Processing Workshop on Advances in Wireless Communications, Lisbon, Portugal, July 2004.

X. Cai was with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA. He is now with the Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL 33124 USA (e-mail: x.cai@miami.edu).

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu).

M. D. Zoltowski is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: mikedz@ecn.purdue.edu).

Digital Object Identifier 10.1109/TCOMM.2005.843459

In the presence of temporally colored interference, however, the ML receiver needs to jointly detect all STBC symbols—a computationally complex task, which motivates low-complexity suboptimal receivers. To this end, and since CSI is required at the receiver for coherent detection, we also investigate optimal training for channel estimation when second-order statistics of the channel and the interference are available.

The rest of the paper is organized as follows. Section II designs an STS scheme that maximizes average signal-to-interference-and-noise ratio (SINR), as well as a space–time beamforming (STBF) scheme that improves error performance. Section III combines STBC with STBF to increase transmission rates, and investigates low-complexity receivers. Optimal training is considered in Section IV. Numerical and simulation results are presented in Section V, and conclusions are drawn in Section VI

Notation: Column vectors (matrices) are denoted by boldface lower- (upper–) case letters. Superscripts ^{T}, *, and $^{\mathcal{H}}$ stand for transpose, conjugate, and Hermitian transpose, respectively; $E[]$ denotes expectation over the random variables within the brackets; $\Re(x)$ and $\Im(x)$ represent the real and imaginary parts of x, respectively. We will use $\mathbf{I}_N$ to denote the $N \times N$ identity matrix, $\mathrm{Tr}(\mathbf{A})$ the trace of $\mathbf{A}$, and $[\mathbf{A}]_{m,n}$ the (m,n)th entry of $\mathbf{A}$.

II. STS Transmissions

A. Signal Model

We consider a wireless communication system with N_t transmit antennas signaling over frequency-flat fading channels. We will first focus on the single-receiver-antenna case, and then will extend our results to multiple receiver antennas. In this section, we consider STS-only transmissions, where the kth information-bearing symbol $s(k)$ is first spread by the code $\mathbf{c}_\mu = [c_\mu(1), \ldots, c_\mu(N)]^T$, and then is transmitted through the μth ($\mu \in [1, N_t]$) transmit antenna over N chips. Letting $\mathbf{C} := [\mathbf{c}_1, \ldots, \mathbf{c}_{N_t}]$ and the $N \times 1$ vector $\mathbf{y}(k)$ containing the baseband-equivalent received samples (at the chip rate) corresponding to the kth symbol, we have

$$\mathbf{y}(k) = \sqrt{\mathcal{E}_s}\mathbf{C}\mathbf{h}(k)s(k) + \mathbf{w}_i(k) + \mathbf{w}_n(k) \tag{1}$$

where $\mathbf{h}(k) := [h_1(k), \ldots, h_{N_t}(k)]^T$ with $h_\mu(k)$ denoting the fading-channel coefficient associated with the μth transmit antenna, the $N \times 1$ vector $\mathbf{w}_i(k)$ consists of interference, e.g., CCI and jamming signals, and $\mathbf{w}_n(k)$ denotes additive white Gaussian noise (AWGN) with zero mean and covariance $N_0\mathbf{I}_N$. With the constraints $\mathrm{Tr}(\mathbf{C}\mathbf{C}^{\mathcal{H}}) = 1$ and $E[|s|^2] = 1$, $\mathcal{E}_s$ in (1) controls the transmit energy per symbol. Although the fading channel $\mathbf{h}(k)$ is allowed to vary from block to block, we assume that it remains invariant during each block. We further assume a Rayleigh fading channel; i.e., $\mathbf{h}(k)$ is complex Gaussian distributed with zero mean and covariance $\mathbf{R}_h := E[\mathbf{h}(k)\mathbf{h}^{\mathcal{H}}(k)]$. Unlike the channel itself, which may change fast depending on the Doppler frequency, the channel correlation $\mathbf{R}_h$ varies slowly, since $\mathbf{R}_h$ is determined by the angle of arrival and the angle spread for a given antenna spacing and arrangement [19]. The interference $\mathbf{w}_i(k)$ is modeled as Gaussian distributed with zero mean and covariance $\mathcal{P}_i\mathbf{R}_i$, where $\mathcal{P}_i$ is the received interference power and $\mathbf{R}_i$ denotes the normalized interference covariance matrix with diagonal entries equal to one. Matrix $\mathbf{R}_i$ is determined by the interference's power spectrum density (PSD), or/and the interference's transmit antenna correlation, and thus it varies slowly with time. Since both $\mathbf{R}_h$ and $\mathbf{R}_i$ depend on long-term properties of the channel, they are easily estimated at the receiver and can be fed back to the transmitter. Our objective here is to find the optimal STS matrix $\mathbf{C}$ which improves error performance, capitalizing on knowledge of $\mathbf{R}_h$ and $\mathbf{R}_i$ at the transmitter.

B. Average SINR Maximizing STS

Since we will detect the transmitted symbols $\{s(k)\}$ on a per-block basis, we will omit the block index k and subsequently deal with the input–output relationship

$$\mathbf{y} = \sqrt{\mathcal{E}_s}\mathbf{C}\mathbf{h}s + \mathbf{w} \tag{2}$$

where $\mathbf{w} := \mathbf{w}_i + \mathbf{w}_n$. Supposing that the channel $\mathbf{h}$ is acquired perfectly, the optimal receiver is a whitening filter followed by a maximum ratio combiner (MRC), whose output can be written as

$$\begin{aligned} z &= \sqrt{\mathcal{E}_s}\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}(N_0\mathbf{R}_w)^{-1}\mathbf{y} \\ &= \left(\frac{\mathcal{E}_s}{N_0}\right)\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{h}s + \tilde{w} \end{aligned} \tag{3}$$

where $\mathbf{R}_w := E[\mathbf{w}\mathbf{w}^{\mathcal{H}}]/N_0 = (\mathcal{P}_i/N_0)\mathbf{R}_i + \mathbf{I}_N$, and $\tilde{w} := \sqrt{\mathcal{E}_s}\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}(N_0\mathbf{R}_w)^{-1}\mathbf{w}$ has zero mean and variance $(\mathcal{E}_s/N_0)\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{h}$. It readily follows from (3) that the SINR at the receiver output is given by

$$\gamma = \frac{\mathcal{E}_s}{N_0}\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{h}. \tag{4}$$

Depending on the chosen figure of merit, the STS matrix $\mathbf{C}$ may take different forms to optimize performance. Our goal in this subsection is to find the optimal $\mathbf{C}$ to maximize the average SINR

$$\begin{aligned} \bar{\gamma} := E[\gamma] &= \frac{\mathcal{E}_s}{N_0}E\left[\mathrm{Tr}\left(\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{h}\mathbf{h}^{\mathcal{H}}\right)\right] \\ &= \frac{\mathcal{E}_s}{N_0}\mathrm{Tr}\left(\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{R}_h\right) \end{aligned} \tag{5}$$

under the constraint $\mathrm{Tr}(\mathbf{C}\mathbf{C}^{\mathcal{H}}) = 1$.

Using the spectral representation of a Hermitian matrix, we can express $\mathbf{R}_h$ and $\mathbf{R}_w$ as $\mathbf{R}_h = \mathbf{U}_h\mathbf{\Lambda}_h\mathbf{U}_h^{\mathcal{H}}$ and $\mathbf{R}_w = \mathbf{U}_w\mathbf{\Lambda}_w\mathbf{U}_w^{\mathcal{H}}$, where the diagonal matrices $\mathbf{\Lambda}_h := \mathrm{diag}(\lambda_{h,1}, \ldots \lambda_{h,N_t})$ and $\mathbf{\Lambda}_w := \mathrm{diag}(\lambda_{w,1}, \ldots, \lambda_{w,N})$ contain the eigenvalues of $\mathbf{R}_h$ and $\mathbf{R}_w$, while the unitary matrices $\mathbf{U}_h := [\mathbf{u}_{h,1}, \ldots, \mathbf{u}_{h,N_t}]$ and $\mathbf{U}_w := [\mathbf{u}_{w,1}, \ldots, \mathbf{u}_{w,N}]$ comprise the eigenvectors of $\mathbf{R}_h$ and $\mathbf{R}_w$, respectively. For convenience and without loss of generality, we assume that the diagonal entries of $\mathbf{\Lambda}_h$ are arranged in nonincreasing order $\lambda_{h,1} \geq \cdots \geq \lambda_{h,N_t}$, while the diagonal entries of $\mathbf{\Lambda}_w$ are arranged in nondecreasing order $\lambda_{w,1} \leq \cdots \leq \lambda_{w,N}$. We then have the following proposition characterizing the optimal $\mathbf{C}$ that maximizes $\bar{\gamma}$ (see Appendix I for the proof).

Proposition 1: The optimal STS matrix $\mathbf{C}$ maximizing the average SINR under the constraint $\mathrm{Tr}(\mathbf{C}\mathbf{C}^{\mathcal{H}}) = 1$ is given by the rank-one matrix $\mathbf{C}_{\text{opt}} = \alpha \mathbf{u}_{w,1}\mathbf{u}_{h,1}^{\mathcal{H}}$, where α is a constant satisfying $|\alpha| = 1$, $\mathbf{u}_{h,1}$ is the eigenvector of $\mathbf{R}_h$ corresponding to the largest eigenvalue $\lambda_{h,1}$, and $\mathbf{u}_{w,1}$ is the eigenvector of $\mathbf{R}_w$ corresponding to the smallest eigenvalue $\lambda_{w,1}$. The maximum average SINR is given by $\bar{\gamma}_{z,\max} = (\mathcal{E}_s/N_0)(\lambda_{h,1}/\lambda_{w,1})$.

Proposition 1 reveals that we should concentrate transmit power on the strongest eigen-direction of the channel and the weakest eigen-direction of the interference to maximize the average received SINR, which is intuitively reasonable. Note that *Proposition 1* generalizes the SNR maximizing beamforming derived in [5] in the *absence* of interference.

Another more important performance metric for a communication system is error probability. Recall that maximizing the average received SINR does not necessarily minimize error probability in fading channels, because error probability is not only determined by the average received SINR but is also affected by the probability density function (pdf) of the received SINR. We are thus motivated to look for the STS matrix $\mathbf{C}$ that improves error probability performance, which leads us to Section II-C.

C. STBF

For a given $\mathbf{C}$, the SER can be found in closed form from the SINR in (4) [20, Sec. 9.2.3]. However, it is difficult to use this SER expression to optimize $\mathbf{C}$. For this reason, we resort to a tractable error-probability bound. Let us define the conditional pairwise error probability (PEP), $P(s \to \tilde{s}|\mathbf{h})$, as the probability that the receiver decides in favor of $\tilde{s}$ ($\tilde{s} \neq s$), when s is actually transmitted, given a channel realization $\mathbf{h}$. Then from (3), we obtain [21, p. 256]

$$P(s \to \tilde{s}|\mathbf{h}) = Q\left(\sqrt{\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{h}\mathcal{E}_s/(2N_0)}|s-\tilde{s}|\right)$$

where $Q(x) := (1/\sqrt{2\pi})\int_x^\infty e^{-t^2/2}dt$. Applying the Chernoff bound, we have $P(s \to \tilde{s}|\mathbf{h}) < \exp(-\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{h}|s - \tilde{s}|^2\mathcal{E}_s/(4N_0))$. Since $\mathbf{h}$ is normally distributed, the Chernoff bound on the unconditional PEP $P(s \to \tilde{s}) := E[P(s \to \tilde{s}|h)]$ can be found as [22, p. 595]

$$\begin{aligned} P(s \to \tilde{s}) &< \left|\mathbf{I}_{N_t} + \frac{\mathcal{E}_s|s-\tilde{s}|^2}{4N_0}\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{R}_h\right|^{-1} \\ &\le \underbrace{\left|\mathbf{I}_{N_t} + \frac{\mathcal{E}_s d_{\min}^2}{4N_0}\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{R}_h\right|^{-1}}_{P_{\text{bound}}} \end{aligned} \tag{6}$$

where $d_{\min}^2$ is the minimum of $|s - \tilde{s}|^2$ over the set $\{(s, \tilde{s})|s \neq \tilde{s}\}$. With proper scaling, the Chernoff bound P_{bound} can also serve as a lower bound on the average SER, and the gap between the upper and lower bounds is relatively small [8]. Hence, in order to reduce the SER, it is reasonable to seek the $\mathbf{C}$ that minimizes P_{bound}.

Using singular value decomposition (SVD), we can write $\mathbf{C}$ as $\mathbf{C} = \mathbf{U}_c\mathbf{D}\mathbf{V}_c^{\mathcal{H}}$, where matrices $\mathbf{U}_c$ and $\mathbf{V}_c$ are unitary, and $\mathbf{D}$ contains the singular values of $\mathbf{C}$. When there is no interference, we have $\mathbf{R}_w = \mathbf{I}_N$; in this case, it is shown in [8] that we must have $\mathbf{V}_c = \mathbf{U}_h$ for the optimal $\mathbf{C}$ that minimizes P_{bound}, while $\mathbf{U}_c$ can be an arbitrary unitary matrix and the optimal $\mathbf{D}$ can be found analytically. When the entries of $\mathbf{h}$ are independent, identically distributed (i.i.d.), we have $\mathbf{R}_h = \mathbf{I}_{N_t}$. And following the arguments in [8], it is not difficult to prove that we must have $\mathbf{U}_c = \mathbf{U}_w$ for the optimal $\mathbf{C}$, while $\mathbf{V}_c$ can be an arbitrary unitary matrix and the optimal $\mathbf{D}$ can be found in closed form, as in [8]. When both $\mathbf{R}_h$ and $\mathbf{R}_w$ are not proportional to the identity matrix, we will choose the STS matrix as

$$\mathbf{C} = \mathbf{U}_w\mathbf{P}\mathbf{D}\mathbf{U}_h^{\mathcal{H}} \tag{7}$$

where $\mathbf{P}$ is an $N \times N$ permutation matrix. Letting $N_{\min} = \min(N, N_t)$, we can express $\mathbf{D}$ as $\mathbf{D} := [\tilde{\mathbf{D}}, \mathbf{0}]^T$ if $N > N_t$, and $\mathbf{D} := [\tilde{\mathbf{D}}, \mathbf{0}]$ if $N < N_t$ with $\tilde{\mathbf{D}}$ being an $N_{\min} \times N_{\min}$ diagonal matrix. With $\mathbf{C}$ chosen as in (7), the error probability bound P_{bound} in (6) reduces to

$$P_{\text{bound}} = \left|\mathbf{I}_{N_t} + \frac{\mathcal{E}_s d_{\min}^2}{4N_0}\mathbf{D}^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\mathbf{D}\mathbf{\Lambda}_h\right|^{-1}. \tag{8}$$

The following proposition proved in Appendix II specifies the optimal $\mathbf{P}$.

Proposition 2: With the diagonal entries of $\mathbf{\Lambda}_h$ arranged in a nonincreasing order and those of $\mathbf{\Lambda}_w$ in a nondecreasing order, the bound P_{bound} in (8) is minimized when $\mathbf{P} = \mathbf{I}_N$.

While we can view a signaling scheme with a general matrix $\mathbf{C}$ as STS, we can also interpret $\mathbf{C}$ in (7) as STBF. We transmit along the eigen-directions of the channel in the space domain and along the eigen-directions of the interference in the time domain. By choosing $\mathbf{P} = \mathbf{I}_N$, strong channels coincide with weak interference, thereby improving performance. Diagonal entries of $\mathbf{D}$ determine the power loaded on different eigen-beams, which we will later optimize to minimize P_{bound}. Note that the average SINR maximizing $\mathbf{C}$ in *Proposition 1* is a special case of (7) with $|[\mathbf{D}]_{11}| = 1$ and $[\mathbf{D}]_{ii} = 0\ \forall\ i \neq 1$; thus, it cannot have better SER performance than an optimally power-loaded STBF. It follows from the proof of *Proposition 2* that when $\{\lambda_{w,i}\}_{i=1}^{N_{\min}}$ are distinct, the optimal $\mathbf{P}$ is unique, i.e., P_{bound} is minimized if and only if $\mathbf{P} = \mathbf{I}_N$. Hence, if we combine STBC with STBF to increase transmission rates, as we will do in the ensuing section, SER performance will degrade, because as we will show later, each symbol in a space–time codeword corresponds to a different permutation matrix.

With $\mathbf{P} = \mathbf{I}_N$, we can write P_{bound} in (8) as

$$P_{\text{bound}} = \prod_{i=1}^{N_{\min}}\left(1 + \frac{\mathcal{E}_s d_{\min}^2}{4N_0}\frac{\lambda_{h,i}[\mathbf{D}]_{ii}^2}{\lambda_{w,i}}\right)^{-1}. \tag{9}$$

Since $\log(x)$ is a monotonically increasing function, minimizing P_{bound} is equivalent to minimizing $\log(P_{\text{bound}})$, or, maximizing $-\log(P_{\text{bound}})$; hence, we can formulate the following optimization problem to find the optimal $\mathbf{D}$:

$$\begin{aligned} \text{maximize} \quad & -\log(P_{\text{bound}}) \\ & = \sum_{i=1}^{N_{\min}}\log\left(1 + \frac{\mathcal{E}_s d_{\min}^2}{4N_0}\frac{\lambda_{h,i}[\mathbf{D}]_{ii}^2}{\lambda_{w,i}}\right) \\ \text{subject to} \quad & \sum_{i=1}^{N_{\min}}[\mathbf{D}]_{ii}^2 = 1. \end{aligned} \tag{10}$$

Using the Lagrange multiplier method, we can find the optimal $\mathbf{D}$ as follows:

$$[\mathbf{D}]_{ii}^2 = \left[\frac{1}{\bar{N}} + \frac{\mathcal{E}_s d_{\min}^2}{4N_0}\left(\frac{1}{\bar{N}}\sum_{j=1}^{\bar{N}}\frac{\lambda_{w,j}}{\lambda_{h,j}} - \frac{\lambda_{w,i}}{\lambda_{h,i}}\right)\right]_+ \quad (11)$$

where $\bar{N} \in [1\ N_{\min}]$ is the maximum number of $\{[\mathbf{D}]_{ii}\}$ that take nonzero values and $[x]_+ := \max(x, 0)$.

While our main focus in this paper is on MISO systems, we can extend our transmission schemes to MIMO systems. Suppose that there are N_r receive antennas, and denote the $N_t \times 1$ vector $\mathbf{h}_\nu$ as the channel associated with the νth receiver antenna. Similar to [8] and [9], we assume that receive antennas are uncorrelated, while transmit antennas are correlated, i.e., $E\left[\mathbf{h}_\nu \mathbf{h}_{\tilde{\nu}}^{\mathcal{H}}\right] = \mathbf{R}_h \delta(\nu - \tilde{\nu})$. This assumption is valid when the transmit antennas are unobstructed and the receive antennas are surrounded by rich local scatterers [19], as in the downlink of a typical cellular system deployed in urban areas. In this case, the receive SINR in (4) becomes $\gamma = (\mathcal{E}_s/N_0)\sum_{\nu=1}^{N_r}\mathbf{h}_\nu^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{h}_\nu$, and the average SINR is then given by $\bar{\gamma} = N_r(\mathcal{E}_s/N_0)\mathrm{Tr}\left(\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{R}_h\right)$. From this average SINR, it is clear that the optimal $\mathbf{C}$ given in *Proposition 1* for $N_r = 1$ still holds for $N_r > 1$. For $N_r > 1$, the error probability bound in (6) becomes $P_{\mathrm{bound}} = \left|\mathbf{I}_{N_t} + (\mathcal{E}_s d_{\min}^2)/(4N_0)\mathbf{C}^{\mathcal{H}}\mathbf{R}_w^{-1}\mathbf{C}\mathbf{R}_h\right|^{-N_r}$. Hence, we still can choose $\mathbf{C}$ as in (7), and it is not difficult to show that the results in *Proposition 2* for the optimal $\mathbf{P}$ and for the optimal $\mathbf{D}$ in (11) are valid for $N_r > 1$. If the channels associated with different receive antennas are correlated, their correlation may also be exploited in designing a MIMO transceiver—a subject going beyond the scope of this paper.

While STS transmissions can effectively suppress interference and exploit channel-induced diversity, they are not spectrally efficient because the transmission rate is $1/N$ symbol per channel use; and thus, they are more suitable for, e.g., military communications, where reliability is the major concern. In the ensuing section, we will invoke STBC to increase transmission rates. In the absence of interference, it is shown in [8] that combining STBC with eigen-beamforming can increase the transmission rate without sacrificing error-probability performance. In the presence of interference, although the error performance will degrade when increasing the rate, combining STBF with STBC will reduce error probability considerably, compared with the transmission scheme where only STBC is used.

III. Space–Time Beamforming and Block Coding

In this section, we will first combine STBC with STBF to increase the transmission rate, and then investigate low-complexity receivers.

A. STBC and STBF

Every M incoming bits are first mapped to a symbol s drawn from an M-ary constellation, e.g., M-quadrature amplitude modulation (QAM), and then every K symbols $\{s_k\}_{k=1}^K$ are encoded into an $N \times N_t$ ($N \geq N_t$) STBC matrix $\mathbf{S}$ using orthogonal STBC [4], [23], where we have $\mathbf{S}^{\mathcal{H}}\mathbf{S} = \sum_{k=1}^K |s_k|^2 \mathbf{I}_{N_t}$, and we assume that $E[|s_k|^2] = 1$. An orthogonal space–time block codeword constructed from complex symbols $\left\{s_k = s_k^R + js_k^I\right\}_{k=1}^K$ can be represented as [23], [24]

$$\mathbf{S} = \sum_{k=1}^K \left(\mathbf{\Phi}_k s_k^R + j\mathbf{\Psi}_k s_k^I\right) \quad (12)$$

or equivalently, as [23]

$$\mathbf{S} = [\mathbf{A}_1\mathbf{s} + \mathbf{B}_1\mathbf{s}^*, \ldots, \mathbf{A}_{N_t}\mathbf{s} + \mathbf{B}_{N_t}\mathbf{s}^*] \quad (13)$$

where matrices $\{\mathbf{\Phi}_k\}_{k=1}^K$, $\{\mathbf{\Psi}_k\}_{k=1}^K$, $\{\mathbf{A}_k\}_{k=1}^K$, and $\{\mathbf{B}_k\}_{k=1}^K$ have entries taken from $\{\pm 1, 0\}$ and satisfy certain conditions [23, eqs. (18) and (19)], [24, eq. (33)], and $\mathbf{s} := [s_1, \ldots, s_K]^T$. These two equivalent representations of a space–time block codeword will be useful later in analyzing error probability and in deriving low-complexity receivers.

Without any channel knowledge, we transmit each column of $\mathbf{S}$ over each antenna in N symbol periods. With $\mathbf{R}_h$ and $\mathbf{R}_w$ available at the transmitter, we combine STBC with STBF and form the transmitted signal matrix as

$$\mathbf{X} = \sqrt{\mathcal{E}_s}\mathbf{U}_w\mathbf{P}\mathbf{S}\mathbf{D}\mathbf{U}_h^{\mathcal{H}} \quad (14)$$

where the $N_t \times N_t$ diagonal matrix $\mathbf{D}$ and the $N \times N$ permutation matrix $\mathbf{P}$ will be chosen to reduce error probability later. The received block can be written as

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{w} \quad (15)$$

where $\mathbf{w}$ contains both interference and AWGN, as described in Section II. Based on (15), the decision of the ML receiver can be expressed as

$$\hat{\mathbf{S}} = \arg_{\mathbf{S}}\min\left[(\mathbf{y} - \mathbf{X}\mathbf{h})^{\mathcal{H}}\mathbf{R}_w^{-1}(\mathbf{y} - \mathbf{X}\mathbf{h})\right] \quad (16)$$

which can also be written as

$$\hat{\mathbf{S}} = \arg_{\mathbf{S}}\min\left|\mathbf{\Lambda}_w^{-1/2}\left(\tilde{\mathbf{y}} - \sqrt{\mathcal{E}_s}\mathbf{P}\mathbf{S}\mathbf{D}\mathbf{U}_h^{\mathcal{H}}\mathbf{h}\right)\right|^2 \quad (17)$$

where $\tilde{\mathbf{y}} = \mathbf{U}_w^{\mathcal{H}}\mathbf{y}$. Without interference, we have $\mathbf{R}_w = N_0\mathbf{I}_N$ and $\mathbf{\Lambda}_w = \mathbf{I}_N$, and thus, the ML receiver in (17) reduces to a simple symbol-by-symbol linear receiver [4]. In the presence of interference, however, either exhaustive search or a low-complexity search we will develop in Section III-B is required to find $\hat{\mathbf{S}}$.

Based on (15) and (17), the conditional PEP, $P(\mathbf{S} \to \tilde{\mathbf{S}}|\mathbf{h})$ with $\mathbf{S} \neq \tilde{\mathbf{S}}$, can be found as

$$P(\mathbf{S} \to \tilde{\mathbf{S}}|\mathbf{h}) = Q\left(\sqrt{|\mathbf{\Lambda}_w^{-1/2}\mathbf{P}(\mathbf{S} - \tilde{\mathbf{S}})\mathbf{D}\mathbf{U}_h^{\mathcal{H}}\mathbf{h}|^2\mathcal{E}_s/(2N_0)}\right)$$

which can be bounded by its Chernoff bound as

$$P(\mathbf{S} \to \tilde{\mathbf{S}}|\mathbf{h}) < \exp\left(-|\mathbf{\Lambda}_w^{-1/2}\mathbf{P}(\mathbf{S} - \tilde{\mathbf{S}})\mathbf{D}\mathbf{U}_h^{\mathcal{H}}\mathbf{h}|^2\mathcal{E}_s/(4N_0)\right).$$

Hence, the PEP $P(\mathbf{S} \to \tilde{\mathbf{S}})$ can be bounded as [22, p. 595]

$$P(\mathbf{S} \to \tilde{\mathbf{S}}) < \left|\mathbf{I}_{N_t} + \frac{\mathcal{E}_s}{4N_0} \times\mathbf{D}(\mathbf{S} - \tilde{\mathbf{S}})^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}(\mathbf{S} - \tilde{\mathbf{S}})\mathbf{D}\mathbf{\Lambda}_h\right|^{-1}. \quad (18)$$

Without interference, we have $\mathbf{\Lambda}_w = \mathbf{I}_N$, and thus $(\mathbf{S} - \tilde{\mathbf{S}})^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}(\mathbf{S} - \tilde{\mathbf{S}}) = \sum_{k=1}^K |e_k|^2\mathbf{I}_{N_t}$, where $e_k := s_k - \tilde{s}_k$. Hence, the PEP $P(\mathbf{S} \to \tilde{\mathbf{S}})$ can be bounded by a simple upper bound: $P(\mathbf{S} \to \tilde{\mathbf{S}}) < \left|\mathbf{I}_{N_t} + (d_{\min}^2 \mathcal{E}_s)/(4N_0)\mathbf{D}^2\mathbf{\Lambda}_h\right|^{-1}$; and the optimal $\mathbf{D}$ can be found by minimizing this PEP bound as in [8]. In the presence of interference, however, there is no such simple upper bound on $P(\mathbf{S} \to \tilde{\mathbf{S}})$. For this reason, we consider the single-error PEP $P_k := P(\mathbf{S} \to \tilde{\mathbf{S}})$ with $\Delta\mathbf{S}_k = \mathbf{S} - \tilde{\mathbf{S}} = \mathbf{\Phi}_k e_k^R + j\mathbf{\Psi}_k e_k^I$ and $\left|e_k^R + je_k^I\right|^2 = d_{\min}^2$, since the overall error probability is strongly affected by these single-error PEPs. The Chernoff bound on P_k can be found from (18) as

$$P_{k,\mathrm{bound}} = \left|\mathbf{I}_{N_t} + \frac{\mathcal{E}_s}{4N_0}\mathbf{D}\Delta\mathbf{S}_k^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\Delta\mathbf{S}_k\mathbf{D}\mathbf{\Lambda}_h\right|^{-1}. \tag{19}$$

Since $\mathbf{P}$ is a permutation matrix, $\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}$ is a diagonal matrix whose diagonal entries are obtained by arranging those of $\mathbf{\Lambda}_w^{-1}$ possibly in a different order. Due to the special structure of orthogonal STBC [23], [24], $\mathbf{\Phi}_k^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\mathbf{\Phi}_k = \mathbf{\Psi}_k^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\mathbf{\Psi}_k$ and $\mathbf{\Phi}_k^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\mathbf{\Psi}_k = \mathbf{\Psi}_k^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\mathbf{\Phi}_k$ are diagonal. Hence, $P_{k,\mathrm{bound}}$ can be easily calculated. In the STS scheme of Section II, we showed that $\mathbf{P} = \mathbf{I}_N$ is optimal. With STBC, we see from (19) that it is impossible to choose a matrix $\mathbf{P}$ that minimizes $P_{k,\mathrm{bound}}, \forall\, k$, because it is impossible for diagonal entries of $\mathbf{Q}_k := \left(\mathbf{\Phi}_k e_k^R + j\mathbf{\Psi}_k e_k^I\right)^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\left(\mathbf{\Phi}_k e_k^R + j\mathbf{\Psi}_k e_k^I\right)$ to appear in a nonincreasing order $\forall\, k$. To improve error-probability performance, we thus choose $\mathbf{P}$ according to the following heuristic rule.

Rule 1: Choose $\mathbf{P}$ so that the large diagonal entries of $\mathbf{\Lambda}_h$ coincide with the large diagonal entries of $\mathbf{Q}_k, \forall\, k$.

Note that we only require knowledge of the STBC's structure determined by $\{\mathbf{\Phi}_k, \mathbf{\Psi}_k\}_{k=1}^K$ to choose $\mathbf{P}$; thus, $\mathbf{P}$ is predetermined offline. Now, we formulate the following optimization problem to be solved for $\mathbf{D}$:

$$\mathbf{D}_{\mathrm{opt}} = \arg_{\mathbf{D}} \min \sum_{k=1}^{K} P_{k,\mathrm{bound}} \quad \text{subject to} \quad \mathrm{Tr}(\mathbf{D}^2) = 1. \tag{20}$$

Letting $x_i := [\mathbf{D}]_{ii}^2$, the error-probability bound $P_{k,\mathrm{bound}}$ is a function of $\{x_i\}$ in the form $P_{k,\mathrm{bound}} = \prod_{i=1}^{N_t}(1+a_i x_i)$, where positive constant a_i can be obtained from (19). In Appendix III, we prove the following.

Fact 1: The error-probability bound $P_{k,\mathrm{bound}}$ is a convex function of $\{x_i\}$.

Assured that the objective function in (20) is convex, the global minimum is unique. We can thus use numerical search based on, e.g., sequential quadratic programming (SQP) [25], to solve (20) for the optimal $\mathbf{D}$.

To illustrate how $\mathbf{P}$ and $\mathbf{D}$ are obtained, let us consider an example, where the STBC is given by [23, eq. (62)]

$$\mathbf{S} = \begin{pmatrix} s_3 & 0 & s_2 & s_1 \\ 0 & s_3 & s_1^* & -s_2^* \\ s_2^* & s_1 & -s_3^* & 0 \\ s_1^* & -s_2 & 0 & -s_3^* \end{pmatrix}. \tag{21}$$

In this case, we have $\mathbf{\Phi}_3^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\mathbf{\Phi}_3 = \mathbf{\Psi}_3^{\mathcal{H}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\mathbf{\Psi}_3 = \mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}$, and thus $\mathbf{P} = \mathbf{I}_N$ is optimal for s_3. On the other hand, $\mathbf{P} = \mathbf{I}_N$ will cause the worst error performance to s_1, because with $\mathbf{P} = \mathbf{I}_N$, we have $\mathbf{\Phi}_1^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{\Phi}_1 = \mathbf{\Psi}_1^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{\Psi}_1 = \mathrm{diag}(1/\lambda_{w,N}, \ldots, 1/\lambda_{w,1})$, which implies that the strongest channel coincides with the strongest interference. Based on **Rule 1**, we choose $\mathbf{P}$ as

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{22}$$

Letting $d_i = [\mathbf{D}]_{ii}$ and $\rho = \mathcal{E}_s d_{\min}^2/(4N_0)$, the upper bound of the single-error PEP can be found as $P_{1,\mathrm{bound}} = [(1+\rho d_1^2\lambda_{h,1}/\lambda_{w,2})(1+\rho d_2^2\lambda_{h,2}/\lambda_{w,1})\,(1+\rho d_3^2\lambda_{h,3}/\lambda_{w,4})(1+\rho d_4^2\lambda_{h,4}/\lambda_{w,3})]^{-1}$, $P_{2,\mathrm{bound}} = [(1+\rho d_1^2\lambda_{h,1}/\lambda_{w,1})(1+\rho d_2^2\lambda_{h,2}/\lambda_{w,2})\,(1+\rho d_3^2\lambda_{h,3}/\lambda_{w,3})(1+\rho d_4^2\lambda_{h,4}/\lambda_{w,4})]^{-1}$, and $P_{3,\mathrm{bound}} = [(1+\rho d_1^2\lambda_{h,1}/\lambda_{w,3})(1+\rho d_2^2\lambda_{h,2}/\lambda_{w,4})\,(1+\rho d_3^2\lambda_{h,3}/\lambda_{w,1})(1+\rho d_4^2\lambda_{h,4}/\lambda_{w,2})]^{-1}$; and the optimal $\mathbf{D}$ can be found by minimizing $\sum_{k=1}^{3} P_{k,\mathrm{bound}}$.

B. Low-Complexity Receivers

ML detection of $\mathbf{S}$ based on exhaustive search requires very high computational complexity. In this subsection, we will consider low-complexity alternatives, after we recast the input–output relationship into a linear model.

Recently, the sphere decoding (SD) algorithm, which was originally introduced to determine short vectors belonging to an arbitrary lattice, has been applied to lattice-code decoding [26], and space–time detection [27]. If the input–output relationship is given by the linear model $\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{w}$, where the $N_t \times 1$ vector $\mathbf{s}$ contains transmitted symbols, $\mathbf{H}$ is a known $N \times N_t$ $(N \geq N_t)$ channel matrix, and $\mathbf{w}$ is AWGN, then SD can be used to obtain a near-ML solution of $\min_{\mathbf{s}} |\mathbf{y} - \mathbf{H}\mathbf{s}|^2$ for $\mathbf{s}$ belonging to a finite alphabet with an average complexity on the order of $\mathcal{O}\left(N_t^3\right)$.

Since $\mathbf{P}^T\mathbf{P} = \mathbf{I}_N$, we can think of the ML decision rule in (17) as based on the following input–output relationship:

$$\mathbf{P}^T\tilde{\mathbf{y}} = \sqrt{\mathcal{E}_s}\mathbf{S}\mathbf{D}\mathbf{U}_h^{\mathcal{H}}\mathbf{h} + \tilde{\mathbf{w}} \tag{23}$$

where $\tilde{\mathbf{w}} := \mathbf{P}^T\mathbf{U}_w\mathbf{w}$. Letting $\tilde{\mathbf{h}} := \sqrt{\mathcal{E}_s}\mathbf{D}\mathbf{U}_h^{\mathcal{H}}\mathbf{h} = [\tilde{h}_1, \ldots, \tilde{h}_{N_t}]^T$ and using (13), we have $\mathbf{x} := \mathbf{S}\tilde{\mathbf{h}} = \sum_{i=1}^{N_t}(\mathbf{A}_i\mathbf{s} + \mathbf{B}_i\mathbf{s}^*)\tilde{h}_i$. Hence, we obtain $\Re(\mathbf{x}) = \sum_{i=1}^{N_t}(\mathbf{A}_i + \mathbf{B}_i)\Re(\tilde{h}_i)\Re(\mathbf{s}) + \sum_{i=1}^{N_t}(-\mathbf{A}_i + \mathbf{B}_i)\Im(\tilde{h}_i)\Im(\mathbf{s})$, and $\Im(\mathbf{x}) = \sum_{i=1}^{N_t}(\mathbf{A}_i + \mathbf{B}_i)\Im(\tilde{h}_i)\Re(\mathbf{s}) + \sum_{i=1}^{N_t}(\mathbf{A}_i - \mathbf{B}_i)\Re(\tilde{h}_i)\Im(\mathbf{s})$. Defining $\bar{\mathbf{y}} := [\Re(\mathbf{P}^T\tilde{\mathbf{y}})^T, \Im(\mathbf{P}^T\tilde{\mathbf{y}})^T]^T$, $\bar{\mathbf{s}} := [\Re(\mathbf{s})^T, \Im(\mathbf{s})^T]^T$, $\bar{\mathbf{w}} := [\Re(\tilde{\mathbf{w}})^T, \Im(\tilde{\mathbf{w}})^T]^T$, $\mathbf{H}_1 := \left[\sum_{i=1}^{N_t}(\mathbf{A}_i + \mathbf{B}_i)\Re(\tilde{h}_i),\ \sum_{i=1}^{N_t}(-\mathbf{A}_i + \mathbf{B}_i)\Im(\tilde{h}_i)\right]$, $\mathbf{H}_2 := \left[\sum_{i=1}^{N_t}(\mathbf{A}_i + \mathbf{B}_i)\Im(\tilde{h}_i),\ \sum_{i=1}^{N_t}(\mathbf{A}_i - \mathbf{B}_i)\Re(\tilde{h}_i)\right]$, and $\mathbf{H} := \left[\mathbf{H}_1^T\ \mathbf{H}_2^T\right]^T$, we have the following linear input–output relationship:

$$\bar{\mathbf{y}} = \mathbf{H}\bar{\mathbf{s}} + \bar{\mathbf{w}}. \tag{24}$$

Note that the covariance of $\bar{\mathbf{w}}$ is given by the block diagonal matrix $\mathbf{\Lambda}_{\bar{w}} = \mathrm{diag}(\mathbf{P}^T\mathbf{\Lambda}_w\mathbf{P}/2, \mathbf{P}^T\mathbf{\Lambda}_w\mathbf{P}/2)$. Based on the signal model (24), we can apply SD to detect the transmitted symbol $\mathbf{s}$ after whitening $\bar{\mathbf{w}}$.

We can further reduce complexity by using a linear receiver, e.g., a matched filter (MF), minimum mean-square error

(MMSE), or zero-forcing (ZF) receiver, at the price of sacrificing performance. Using the properties of matrices $\{\mathbf{A}_i, \mathbf{B}_i\}$ [23], it is easy to show that $\mathbf{H}^T\mathbf{H} = \sum_{i=1}^{N_t} |h_i|^2 \mathbf{I}_{2N_t}$. Hence, ZF and MF receivers have identical performance, and they both achieve full diversity; however, they cannot effectively suppress interference, and may degrade error performance considerably.

IV. Optimal Training

In Sections II and III, we assumed that the channel is perfectly known at the receiver. In practice, however, the channel needs to be estimated through, e.g., training symbols. In this section, we will consider channel estimation based on the following signal model:

$$\mathbf{y} = \sqrt{N\mathcal{E}_p}\mathbf{C}\mathbf{h} + \mathbf{w} \tag{25}$$

where the $N \times N_t$ $(N \geq N_t)$ training matrix (symbol) $\mathbf{C}$ is known at the receiver, and satisfies $\mathrm{Tr}(\mathbf{C}\mathbf{C}^{\mathcal{H}}) = 1$, and thus SNR $= \mathcal{E}_p/N_0$. When training symbols are periodically transmitted, the receiver can exploit the channel's temporal correlation, and estimate the channel based on multiple received training blocks, as done in the pilot-symbol-assisted modulation (PSAM) [28]. In this paper, we will consider channel estimation based on a *single* received training block $\mathbf{y}$ as in (25). Channel estimation on a per-block basis is known to maximize a lower bound on ergodic capacity [29]. Furthermore, if temporal channel correlations are available, temporal interpolation can be used to reduce estimation errors.

A. LSE Channel Estimation

Letting $\tilde{\mathbf{C}} := \sqrt{N\mathcal{E}_p}\mathbf{C}$, the least-squares error (LSE) channel estimate is given by

$$\hat{\mathbf{h}} = (\tilde{\mathbf{C}}^{\mathcal{H}}\tilde{\mathbf{C}})^{-1}\tilde{\mathbf{C}}^{\mathcal{H}}\mathbf{y} = \mathbf{h} + (\tilde{\mathbf{C}}^{\mathcal{H}}\tilde{\mathbf{C}})^{-1}\tilde{\mathbf{C}}^{\mathcal{H}}\mathbf{w}. \tag{26}$$

From (26), the covariance matrix of the estimated channel can be written as $\mathbf{C}_{\hat{h}} = (\mathbf{C}^{\mathcal{H}}\mathbf{C})^{-1} \mathbf{C}^{\mathcal{H}}(N_0\mathbf{R}_w)\mathbf{C}(\mathbf{C}^{\mathcal{H}}\mathbf{C})^{-1}/(N\mathcal{E}_p)$. The mean-square error (MSE) of the estimated channel is then given by $\zeta = \mathrm{Tr}(\mathbf{C}_{\hat{h}})$. Using the SVD of $\mathbf{C}$, we can express $\mathbf{C}_{\hat{h}}$ as $\mathbf{C}_{\hat{h}} = N_0\mathbf{V}_c\tilde{\mathbf{D}}^{-1}\mathbf{U}_c^{\mathcal{H}}\mathbf{R}_w\mathbf{U}_c\tilde{\mathbf{D}}^{-1}\mathbf{V}_c^{\mathcal{H}}/(N\mathcal{E}_p)$, and thus, we have

$$\zeta = \frac{N_0}{N\mathcal{E}_p}\mathrm{Tr}\left(\tilde{\mathbf{U}}\mathbf{\Lambda}_w\tilde{\mathbf{U}}^{\mathcal{H}}\tilde{\mathbf{D}}^{-2}\right) \tag{27}$$

where $\tilde{\mathbf{U}} := \mathbf{U}_c^{\mathcal{H}}\mathbf{U}_w$. If the diagonal entries of $\tilde{\mathbf{D}}^2$ appear in nondecreasing order, which we will verify later, then based on (27) and using the results of [13, Lemma 1] and [9, eq. (19)], we deduce that ζ is minimized when $\tilde{\mathbf{U}} = \mathbf{I}_N$, which implies that $\mathbf{U}_c = \mathbf{U}_w$. Thus, we have established the following proposition.

Proposition 3: To minimize the MSE of the LSE channel estimator, the optimal training matrix $\mathbf{C}$ is given by $\mathbf{C} = \mathbf{U}_w\mathbf{D}\mathbf{V}_c^{\mathcal{H}}$, where $\mathbf{V}_c$ is an arbitrary unitary matrix.

With $\mathbf{U}_c = \mathbf{U}_w$, the channel MSE ζ in (27) reduces to

$$\zeta = \frac{N_0}{N\mathcal{E}_p}\sum_{i=1}^{N_t}\frac{\lambda_{w,i}}{[\mathbf{D}]_{ii}^2} \tag{28}$$

and we can find the optimal $\mathbf{D}$ by minimizing ζ subject to the constraint $\mathrm{Tr}(\mathbf{D}\mathbf{D}^{\mathcal{H}}) = 1$, which leads to

$$[\mathbf{D}]_{ii}^2 = \frac{\sqrt{\lambda_{w,i}}}{\sum_{j=1}^{N_t}\sqrt{\lambda_{w,j}}} \tag{29}$$

and

$$\zeta = \frac{N_0}{N\mathcal{E}_p}\left(\sum_{i=1}^{N_t}\sqrt{\lambda_{w,i}}\right)^2. \tag{30}$$

It is seen from (30) that $\{[\mathbf{D}]_{ii}^2\}$ are in a nondecreasing order, because $\{\lambda_{w,i}\}$ are in a nondecreasing order. Without interference, we have $\lambda_{w,1} = \cdots = \lambda_{w,N}$, and thus, we obtain from (29) that $[\mathbf{D}]_{ii}^2 = 1/N_t, \forall\, i$. Hence, the optimal training matrix satisfies $\mathbf{C}^{\mathcal{H}}\mathbf{C} = \mathbf{I}_{N_t}/N$, i.e., columns of $\mathbf{C}$ are orthogonal.

B. Linear MMSE Channel Estimation

The LSE channel estimator does not exploit the statistics of the channel and the interference. Although LSE leads to robust channel estimation, it does not minimize the channel MSE. We will next consider linear MMSE estimation, where statistics of both channel and interference are exploited to minimize channel-estimation errors.

Based on the signal model (25), the linear MMSE channel estimator is given by [30, p. 391]

$$\hat{\mathbf{h}} = \sqrt{N\mathcal{E}_p}\mathbf{R}_h\mathbf{C}^{\mathcal{H}}(N\mathcal{E}_p\mathbf{C}\mathbf{R}_h\mathbf{C} + N_0\mathbf{R}_w)^{-1}\mathbf{y} \tag{31}$$

and the covariance matrix of the channel estimate can be written as

$$\mathbf{C}_{\hat{h}} = \mathbf{R}_h - N\mathcal{E}_p\mathbf{R}_h\mathbf{C}^{\mathcal{H}}\left(N_0\mathbf{R}_w + N\mathcal{E}_p\mathbf{C}\mathbf{R}_h\mathbf{C}^{\mathcal{H}}\right)^{-1}\mathbf{C}\mathbf{R}_h. \tag{32}$$

Similar to (7), we will choose the training matrix based on the STBF matrix, $\mathbf{C} = \mathbf{U}_w\mathbf{D}\mathbf{U}_h^{\mathcal{H}}$. It is not difficult to find the resulting channel MSE $\zeta = \mathrm{Tr}(\mathbf{C}_{\hat{h}})$ as

$$\zeta = \mathrm{Tr}\Bigg(\mathbf{\Lambda}_h - N\mathcal{E}_p\mathbf{\Lambda}_h\mathbf{D}^{\mathcal{H}} \times\left(N_0\mathbf{\Lambda}_w + N\mathcal{E}_p\mathbf{D}\mathbf{\Lambda}_h\mathbf{D}^{\mathcal{H}}\right)^{-1}\mathbf{D}\mathbf{\Lambda}_h\Bigg). \tag{33}$$

Letting $N_h = \mathrm{rank}(\mathbf{\Lambda}_h)$, we can also simplify (33) as

$$\zeta = \sum_{i=1}^{N_h}\left(\frac{1}{\lambda_{h,i}} + \frac{N\mathcal{E}_p}{N_0}\frac{[\mathbf{D}]_{ii}^2}{\lambda_{w,i}}\right)^{-1}. \tag{34}$$

Minimizing ζ with respect to $\mathbf{D}$ under the constraint $\mathrm{Tr}(\mathbf{D}\mathbf{D}^{\mathcal{H}}) = 1$, we obtain

$$[\mathbf{D}]_{ii}^2 = \left[\frac{1 + \frac{N_0}{N\mathcal{E}_p}\sum_{j\in\mathcal{I}}\frac{\lambda_{w,j}}{\lambda_{h,j}}}{\sum_{j\in\mathcal{I}}\sqrt{\lambda_{w,j}}}\sqrt{\lambda_{w,i}} - \frac{N_0}{N\mathcal{E}_p}\frac{\lambda_{w,i}}{\lambda_{h,i}}\right]_+ \tag{35}$$

where $\mathcal{I}$ denotes the set $\mathcal{I} := \{[\mathbf{D}]_{ii} | [\mathbf{D}]_{ii} \neq 0\}$. Note that the power loading in (35) which minimizes channel estimation errors is different from that in (11), which minimizes the error-probability bound.

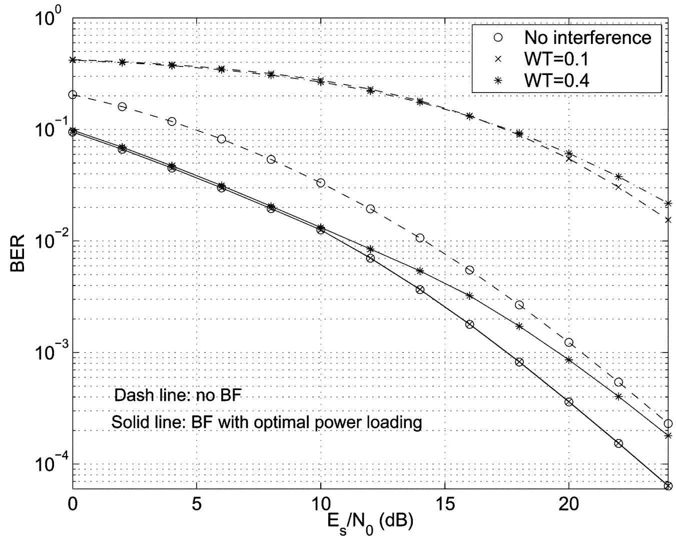

Fig. 1. BER of STBF, channel 1, INR = 10 dB.

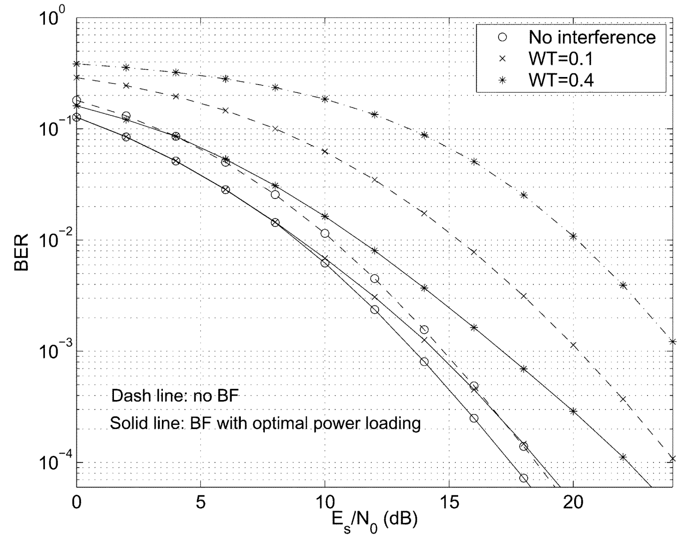

Fig. 2. BER of STBF, channel 2, INR = 10 dB.

V. Simulation and Numerical Results

We consider a linear array with $N_t = 4$ transmit antennas, $N_r = 1$ receive antenna, and an STS matrix $\mathbf{C}$ of size $N_t \times N_t$. The N_t transmit antennas are equispaced by d. We assume that the direction of arrival is perpendicular to the transmit antenna array. Let λ be the wavelength of the transmitted signal, and Δ denote the angle spread. When Δ is small, the channel correlation can be calculated from the "one-ring" channel model as $[\mathbf{R}_h]_{m,n} \approx 1/2\pi \int_0^{2\pi} \exp[-j2\pi(m-n)d\Delta \sin\theta/\lambda]d\theta$ [19]. In our analysis and simulations, we will consider two channels with different correlations. Channel 1 has $d = 0.5\lambda$ and $\Delta = 5^\circ$, while channel 2 has $d = 0.5\lambda$ and $\Delta = 25^\circ$. Channels are normalized so that $\mathrm{Tr}(\mathbf{R}_h) = N_t$. For channel 1, the eigenvalues of $\mathbf{R}_h$ are $\mathbf{\Lambda}_{h_1} = \mathrm{diag}(3.81849, 0.18079, 0.00071, 0.00001)$; and for channel 2, we have $\mathbf{\Lambda}_{h_2} = \mathrm{diag}(1.790, 1.741, 0.454, 0.015)$. While channel 1 is highly correlated, channel 2 is less correlated and provides more diversity. We will consider partial band interference with a normalized power spectral density $S(f) = \sin(\pi W f)/(\pi W f)$, where W denotes the interference bandwidth. We assume that the interference center frequency coincides with the carrier frequency. We will use two different values of W: $W = 0.1/T$ and $W = 0.4/T$, where T is the sampling period that is used to obtain the discrete signal model (1). When $W = 0.1/T$, the eigenvalues of the 4×4 interference covariance matrix $\mathbf{R}_i$ are in $\mathbf{\Lambda}_{i_1} = \mathrm{diag}(3.8412, 0.1579, 0.0009, 0.0000)$; and when $W = 0.4/T$, the eigenvalues of $\mathbf{R}_i$ are in $\mathbf{\Lambda}_{i_2} = \mathrm{diag}(2.391, 1.394, 0.210, 0.005)$. Note that when $W = 0.1/T$, interference concentrates on two eigen-vectors, and is relatively easy to avoid. In all plots, we will adopt quaternary phase-shift keying (QPSK) modulation, and we define the interference-to-noise ratio (INR) as $\mathcal{P}_I/N_0$.

Figs. 1 and 2 depict the bit-error rate (BER) of the STS transmissions in Section II for channels 1 and 2, calculated from the SINR in (4) using the formula of [20, eq. (9.11)]. We plot BER curves without beamforming ($\mathbf{C} = \mathbf{I}_{N_t}$), and with optimally power-loaded STBF [$\mathbf{C}$ given in (7)]. When $W = 0.1/T$, performance of STBF in the presence of interference comes very close to its lower bound in the absence of interference. This is because most interference power lies in the weakest channel that conveys no or very little signal power. When $W = 0.4/T$, the gap between STBF with interference and its lower bound increases, because the interference power is spread over more eigen-directions. In both cases, it is observed that optimally power-loaded STBF has considerably larger performance gain, relative to the transmission without beamforming. Fig. 3 depicts BER versus INR for STS transmissions. STBF can tolerate interference power more than 25 dB higher than transmissions without beamforming when $W = 0.1/T$, or more than 10 dB when $W = 0.4/T$.

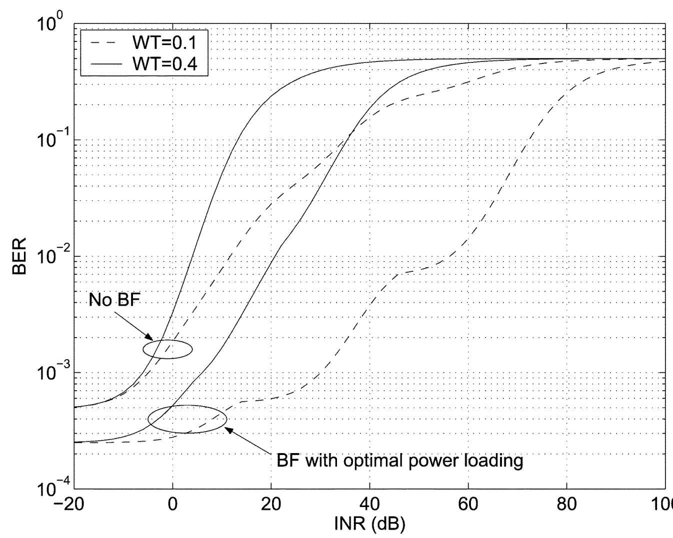

Fig. 3. BER of STBF versus INR, $\mathcal{E}_s/N_0 = 16$ dB, channel 2.

Figs. 4–7 describe BER of STBC with ML and MMSE receivers. The transmitted block $\mathbf{X}$ is given in (14) with STBF, and $\mathbf{X} = \mathbf{S}$ without beamforming, where the STBC matrix is given by (21). The ML receiver is implemented with the

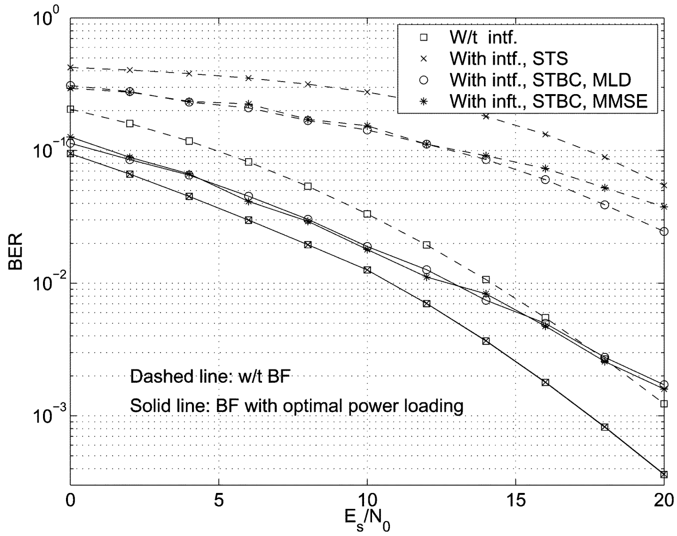

Fig. 4. Simulated BER of STBC, INR $= 10$ dB, channel 1, $WT = 0.1$.

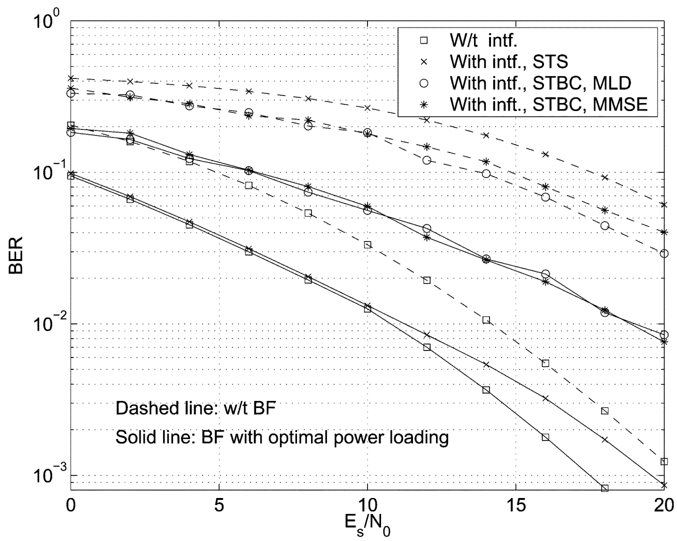

Fig. 5. Simulated BER of STBC, INR $= 10$ dB, channel 1, $WT = 0.4$.

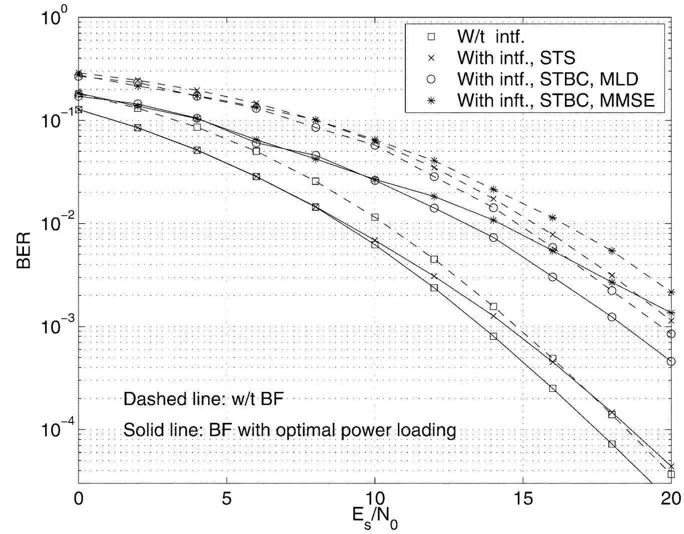

Fig. 6. Simulated BER of STBC, INR $= 10$ dB, channel 2, $WT = 0.1$.

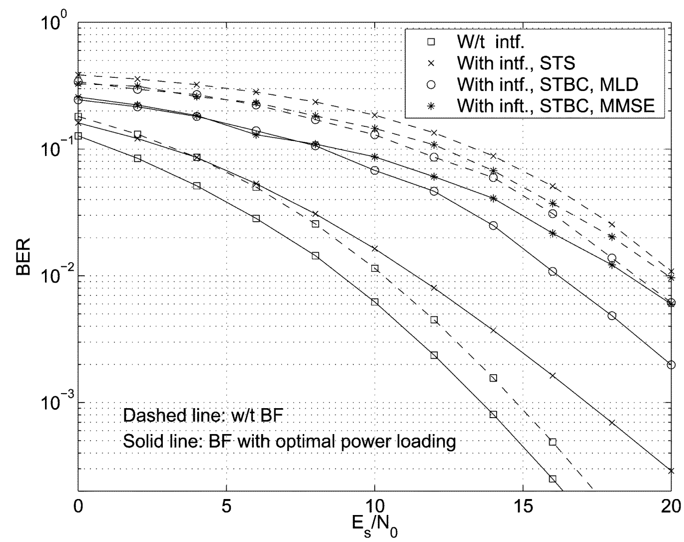

Fig. 7. Simulated BER of STBC, INR $= 10$ dB, channel 2, $WT = 0.4$.

low-complexity SD algorithm, whose performance has no noticeable difference relative to that of exhaustive search, and thus, we do not plot the BER with exhaustive search. As observed in Figs. 4 and 5, STBF exhibits a large performance gain relative to the transmission without beamforming; although this performance gain reduces in channel 2, it is still considerably large, as shown in Figs. 6 and 7. Comparing Fig. 4 with Fig. 5, and Fig. 6 with Fig. 7, we see that when the bandwidth of interference increases, BER degrades. One surprising observation is that without beamforming, the BER of STS is worse than that of STBC. This is actually reasonable, because each symbol in a space–time block codeword has a different permutation matrix. Some of these permutations can reduce the single-error PEP, as analyzed in Section III-A, and thereby decrease the overall error probability.

Figs. 8 and 9 depict the MSE of LSE and MMSE channel estimators. We plot the channel MSE for cases without beamforming $(\mathbf{C} = \mathbf{I}_{N_t})$ and with optimally power-loaded beamforming. Without beamforming, LSE and MMSE estimators have almost the same MSE when $\mathcal{E}_p/N_0$ is moderately large. With optimal training, the LSE channel estimator has slightly better MSE than that without beamforming, because channel correlations are not exploited by the LSE channel estimator. On the other hand, STBF considerably improves the performance of the MMSE channel estimator.

VI. Conclusions

We have exploited second-order spatial statistics of the channel, and temporal statistics of the interference, to design transceivers for multiantenna wireless communication systems. Based on STS, we showed that if signals are transmitted along the strongest eigen-direction of the channel and the weakest eigen-direction of the interference, the average SINR is maximized. We also derived optimally power-loaded STBF schemes, and showed that if strong channels coincide with weak interference, then error probability reduces considerably.

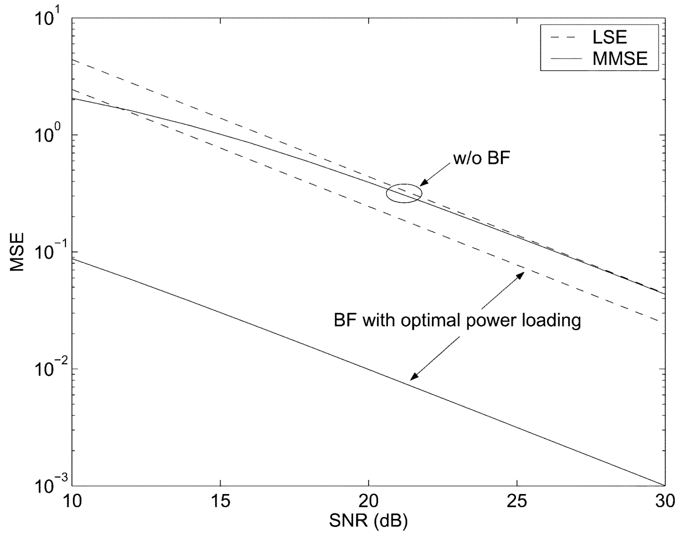

Fig. 8. MSE of channel estimators, INR = 10 dB, channel 1, $WT = 0.1$.

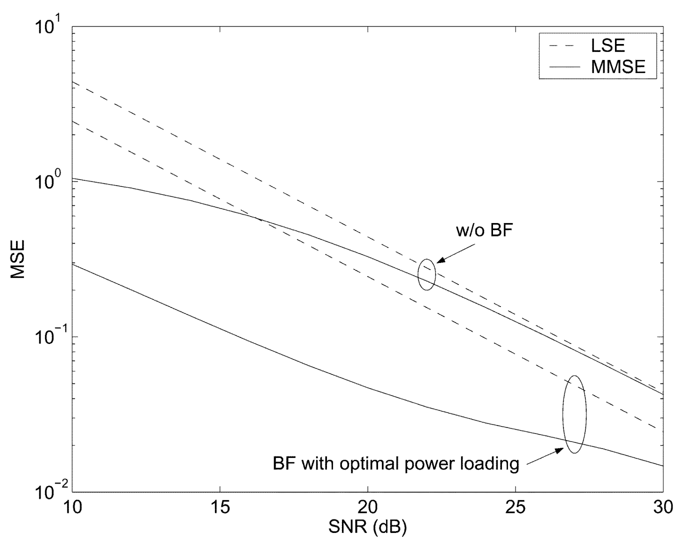

Fig. 9. MSE of channel estimators, INR = 10 dB, channel 2, $WT = 0.1$.

In order to increase transmission rates, we combined STBC with STBF and derived power-loading schemes and low-complexity receivers. We also investigated optimal training for LSE channel estimation and STBF for MMSE channel estimation. Our analytical and simulated results corroborate that STBF with optimal power loading can considerably reduce error-probability and channel-estimator errors.

APPENDIX I
PROOF OF *PROPOSITION 1*

Letting $\tilde{\mathbf{C}} := \mathbf{U}_w^{\mathcal{H}}\mathbf{C}\mathbf{U}_h$, the average SINR can also be written as $\bar{\gamma} = (\mathcal{E}_s/N_0)\mathrm{Tr}(\mathbf{\Lambda}_w^{-1}\tilde{\mathbf{C}}\,\mathbf{\Lambda}_h\tilde{\mathbf{C}}^{\mathcal{H}})$, and we have $\mathrm{Tr}(\tilde{\mathbf{C}}\tilde{\mathbf{C}}^{\mathcal{H}}) = \mathrm{Tr}(\mathbf{C}\mathbf{C}^{\mathcal{H}})$. Since both $\mathbf{\Lambda}_w^{-1}$ and $\mathbf{\Lambda}_h$ are diagonal, we can express $\bar{\gamma}$ as $\bar{\gamma} = (\mathcal{E}_s/N_0)\sum_{i=1}^{N}\sum_{j=1}^{N_t}(\lambda_{h,j})/(\lambda_{w,i})|[\tilde{\mathbf{C}}]_{ij}|^2$. To satisfy the constraint $\mathrm{Tr}(\mathbf{C}\mathbf{C}^{\mathcal{H}}) = 1$, we must have $\mathrm{Tr}(\tilde{\mathbf{C}}\tilde{\mathbf{C}}^{\mathcal{H}}) = \sum_{i=1}^{N}\sum_{j=1}^{N_t}|[\tilde{\mathbf{C}}]_{ij}|^2 = 1$. Using the convex combination inequality [31, p. 535], we infer that the maximum value of $\bar{\gamma}$ is given by

$$\bar{\gamma}_{\max} = \frac{\mathcal{E}_s}{N_0}\max_{\{i,j\}}\left\{\frac{\lambda_{h,j}}{\lambda_{w,i}}\right\} = \frac{\mathcal{E}_s}{N_0}\frac{\lambda_{h,1}}{\lambda_{w,1}} \tag{36}$$

and that the maximum is achieved when $|[\tilde{\mathbf{C}}]_{1,1}|^2 = 1$, and $[\tilde{\mathbf{C}}]_{i,j} = 0, \forall\, i \neq 1, \forall\, j \neq 1$, which implies that $\mathbf{C}_{\mathrm{opt}} = \alpha\mathbf{u}_{w,1}\mathbf{u}_{h,1}^{\mathcal{H}}$, where α is a constant satisfying $|\alpha| = 1$.

APPENDIX II
PROOF OF *PROPOSITION 2*

For clarity, let us first consider the case where $N = N_t$. Since $\mathbf{P}$ is a permutation matrix, $\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}$ is a diagonal matrix with its diagonal entries obtained by permuting those of $\mathbf{\Lambda}_w^{-1}$. Letting $[\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}]_{ii} = 1/\lambda_{w,n_i}$ where $\{n_i\}_{i=1}^N$ is a permutation of the set $\{i\}_{i=1}^N$, we can express P_{bound} in (8) as

$$P_{\mathrm{bound}} = \prod_{i=1}^{N}\left(1 + \frac{\mathcal{E}_s d_{\min}^2}{4N_0}\frac{\lambda_{h,i}[\mathbf{D}]_{ii}^2}{\lambda_{w,n_i}}\right)^{-1}. \tag{37}$$

Since minimizing P_{bound} is equivalent to maximizing $1/P_{\mathrm{bound}}$, we will show that $1/P_{\mathrm{bound}}$ is maximizing if $n_i = i$, or equivalently, $\mathbf{P} = \mathbf{I}_N$. But first, let us state the following lemma, which can be easily verified.

Lemma 1: Supposing that four variables x_1, x_2, y_1, and y_2 satisfy $x_1 \geq x_2 \geq 0$ and $y_1 \geq y_2 \geq 0$, we have $(1+x_1y_1)(1+x_2y_2) \geq (1+x_1y_2)(1+x_2y_1)$.

We also need the following lemma.

Lemma 2: Suppose that we have $2N$ variables, $\{x_i\}_{i=1}^N$ and $\{y_i\}_{i=1}^N$, satisfying $x_1 \geq \cdots \geq x_N \geq 0$, and $y_1 \geq \cdots \geq y_N \geq 0$, and another set of N positive variables $\{\alpha_i\}_{i=1}^N$. If $\{y_{n_i}\}_{i=1}^N$ is a set formed by permuting $\{y_i\}_{i=1}^N$, then there exists a set $\{\alpha_{m_i}\}_{i=1}^N$ obtained by permuting $\{\alpha_i\}_{i=1}^N$, so that $\prod_{i=1}^N(1 + x_iy_{n_i}\alpha_i) \leq \prod_{i=1}^N(1 + x_iy_i\alpha_{m_i})$.

Proof: Consider the case where $N = 2$. If $\alpha_1 \geq \alpha_2$, then from *Lemma 1*, we have $(1 + x_1y_2\alpha_1)(1 + x_2y_1\alpha_2) \leq (1 + x_1y_1\alpha_1)(1 + x_2y_2\alpha_2)$, since $x_1\alpha_1 \geq x_2\alpha_2$. If $\alpha_1 < \alpha_2$, then again using *Lemma 1*, we have $(1+x_1y_2\alpha_1)(1+x_2y_1\alpha_2) \leq (1 + x_1y_1\alpha_2)(1 + x_2y_2\alpha_1)$, since $y_1\alpha_2 \geq y_2\alpha_1$. For a given set $\{y_{n_i}\}_{i=1}^N$ with $N > 2$, we can apply the procedure in the $N = 2$ case to every term of $(1 + x_iy_{n_i}\alpha_i)(1 + x_jy_{n_j}\alpha_j)$ in $\prod_{i=1}^N(1 + x_iy_{n_i}\alpha_i)$, and find $\{\alpha_{m_i}\}_{i=1}^N$, so that $\prod_{i=1}^N(1 + x_iy_{n_i}\alpha_i) \leq \prod_{i=1}^N(1 + x_iy_i\alpha_{m_i})$. ■

Recall that the diagonal entries of $\mathbf{\Lambda}_h$ are arranged in nonincreasing order, and that the diagonal entries of $\mathbf{\Lambda}_w$ are in nondecreasing order. Hence, $1/P_{\mathrm{bound}}$ is maximized if $\mathbf{P} = \mathbf{I}_N$, because if $\mathbf{P} \neq \mathbf{I}_N$, according to *Lemma 2*, we can always find another diagonal matrix $\bar{\mathbf{D}}$, whose diagonal entries are obtained by permuting the diagonal entries of $\mathbf{D}$, so that

$$\left|\mathbf{I}_N + (\mathcal{E}_s d_{\min}^2)/(4N_0)\mathbf{D}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\mathbf{D}\mathbf{\Lambda}_h\right| \leq \left|\mathbf{I}_N + (\mathcal{E}_s d_{\min}^2)/(4N_0)\bar{\mathbf{D}}\mathbf{\Lambda}_w^{-1}\bar{\mathbf{D}}\mathbf{\Lambda}_h\right|$$

without changing the constraint on $\mathbf{C}$.

Now consider the case where $N>N_t$. Let the $N_t\times N_t$ matrix $\tilde{\mathbf{\Lambda}}_w^{-1}$ contain the first N_t rows and columns of $\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}$, then P_{bound} in (8) becomes

$$P_{\text{bound}} = \left|\mathbf{I}_{N_t} + (\mathcal{E}_s d_{\min}^2)/(4N_0)\tilde{\mathbf{D}}\tilde{\mathbf{\Lambda}}_w^{-1}\tilde{\mathbf{D}}\mathbf{\Lambda}_h\right|^{-1}.$$

Hence, it is not difficult to show that P_{bound} is minimized if

$$\mathbf{P} = \begin{pmatrix} \mathbf{I}_{N_t} & \mathbf{0} \\ \mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix} \tag{38}$$

where $\mathbf{P}_1$ and $\mathbf{P}_2$ are arbitrary matrices, which implies that $\mathbf{P} = \mathbf{I}_N$ also minimizes P_{bound}. If $N < N_t$, then

$$P_{\text{bound}} = \left|\mathbf{I}_N + (\mathcal{E}_s d_{\min}^2)/(4N_0)\tilde{\mathbf{D}}\mathbf{P}^{\mathcal{H}}\mathbf{\Lambda}_w^{-1}\mathbf{P}\tilde{\mathbf{D}}\tilde{\mathbf{\Lambda}}_h\right|^{-1}$$

where the $N \times N$ matrix $\tilde{\mathbf{\Lambda}}_h$ contains the first N rows and columns of $\mathbf{\Lambda}_h$; and then it is easy to show that $\mathbf{P} = \mathbf{I}_N$ minimizes P_{bound}.

Appendix III
Proof of Fact 1

As stated in [32, p. 67], a function is convex if and only if it is convex when restricted to any line that intersects its domain. Suppose that $\mathbf{x}_1 = [x_{11},\ldots,x_{1N_t}]^T$ and $\mathbf{x}_2 = [x_{21},\ldots,x_{2N_t}]^T$ satisfy constraints $\sum_{i=1}^{N_t} x_{1i} = 1$ and $\sum_{i=1}^{N_t} x_{2i} = 1$. Letting $\mathbf{x} = t\mathbf{x}_1 + (1-t)\mathbf{x}_2$ with $0 \le t \le 1$, we need to prove that

$$f(t) = \frac{1}{\prod_{i=1}^{N_t}(1+a_i x_i)} = \frac{1}{\prod_{i=1}^{N_t}(1+a_i x_{2i} + a_i \Delta x_i t)} \tag{39}$$

where $\Delta x_i = x_{1i} - x_{2i}$ is a convex function of t. The first derivative of $f(t)$ can be written as

$$\frac{df}{dt} = \frac{1}{\prod_{i=1}^{N_t}(1+a_i x_{2i} + a_i \Delta x_i t)} \sum_{i=1}^{N_t} \frac{-a_i \Delta x_i}{1+a_i x_{2i} + a_i \Delta x_i t}. \tag{40}$$

Then the second derivative of $f(t)$ can be found as

$$\frac{d^2 f}{dt^2} = \frac{1}{\prod_{i=1}^{N_t}(1+a_i x_{2i} + a_i \Delta x_i t)} \times \left[\sum_{i=1}^{N_t} \frac{a_i^2 \Delta x_i^2}{(1+a_i x_{2i} + a_i \Delta x_i t)^2} + \left(\sum_{i=1}^{N_t} \frac{a_i \Delta x_i}{1+a_i x_{2i} + a_i \Delta x_i t}\right)^2\right] > 0 \tag{41}$$

which implies that $f(t)$ is a convex function of t.[1]

[1]The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

References

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

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

[4] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1456–1467, Jul. 1999.

[5] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[6] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 48, pp. 611–627, Mar. 2002.

[7] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2632–2639, Sep. 2001.

[8] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1673–1690, Jul. 2003.

[9] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. Int. Conf. Commun.*, Helsinki, Finland, May 2001, pp. 2266–2270.

[10] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback gain in multiple-antenna systems," *IEEE Trans. Commun.*, vol. 50, pp. 785–798, May 2002.

[11] A. L. Moustakas and S. H. Simon, "Optimizing multiple-input single-output (MISO) communication systems with general Gaussian channels: Nontrivial covariance and nonzero mean," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2770–2780, Oct. 2003.

[12] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 406–417, Apr. 2003.

[13] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Process.*, vol. 50, pp. 2599–2613, Oct. 2002.

[14] R. S. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 793–801, Jun. 2003.

[15] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large *n* analysis," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2545–2561, Oct. 2003.

[16] E. G. Larsson, P. Stoica, and J. Li, "Orthogonal space-time block codes: Maximum-likelihood detection for unknown channels and unstructured interferences," *IEEE Trans. Signal Process.*, vol. 51, pp. 362–372, Feb. 2003.

[17] A. Tarr and M. D. Zoltowski, "Design of codes for multicode CDMA based on correlation feedback from receiver," in *Proc. SPIE Int. Symp. AeroSense: Digital Wireless Commun.*, vol. 5100, Orlando, FL, Apr. 2003, pp. 261–266.

[18] ——, "MMSE receiver with reduced cross-talk in CDMA via eigenvector codes," in *Proc. 41st Allerton Conf.*, Monticello, IL, Oct. 2003, pp. 1807–1808.

[19] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[20] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels: A Unified Approach to Performance Analysis*. New York: Wiley, 2000.

[21] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.

[22] M. Schwartz, W. R. Benett, and S. Stein, *Communication Systems and Techniques*. Piscataway, NJ: IEEE Press, 1996.

[23] X.-B. Liang, "Orthogonal designs with maximal rates," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2468–2503, Oct. 2003.

[24] G. Ganesan and P. Stoica, "Space-time block codes: A maximum SNR approach," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1650–1656, May 2001.

[25] R. Fletcher, *Practical Methods of Optimization: Constrained Optimization*. New York: Wiley, 1980, vol. 2.

[26] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channel," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1639–1642, Jul. 1999.

[27] B. M. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, pp. 389–399, Mar. 2003.

[28] J. K. Cavers, "An analysis of pilot-symbol-assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.

[29] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, pp. 951–963, Apr. 2003.

[30] S. T. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993, vol. 1.

[31] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Xiaodong Cai (S'00–M'01) received the B.S. degree from Zhejiang University, Hangzhou, China, the M.Eng. degree from the National University of Singapore, Singapore, and the Ph.D. degree from the New Jersey Institute of Technology, Newark, in 2001, all in electrical engineering.

From February 2001 to June 2001, he was a Member of Technical Staff with Lucent Technologies, NJ. From July 2001 to October 2001, he was a Senior System Engineer with Sony Technology Center, San Diego, CA. From November 2001 to August 2004, he was a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis. Since September 2004, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL. His research interests lie in the areas of communication theory, signal processing, and networking. Current research focuses on cooperative communications and collaborative signal processing for wireless ad hoc and sensor networks.

Georgios B. Giannakis (S'84–M'86–SM'91–F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Greece, in 1981. He received the MSc. degree in electrical engineering in 1983, M.Sc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986, from the University of Southern California (USC), Los Angeles.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in 1987, where he became a Professor of Electrical Engineering in 1997. Since 1999, he has been with the University of Minnesota, Minneapolis, as a Professor in the Department of Electrical and Computer Engineering, and holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification, subjects on which he has published more than 200 journal papers, 350 conference papers, and two edited books. Current research focuses on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, complex-field and space–time coding, multicarrier, ultra-wideband wireless communication systems, cross-layer designs, and distributed sensor networks.

Dr. Giannakis is the corecipient of six Best Paper Awards from the IEEE Signal Processing (SP) Society (1992, 1998, 2000, 2001, 2003, and 2004). He also received the Society's Technical Achievement Award in 2000. He co-organized three IEEE-SP Workshops, and guest co-edited four special issues. He has served as Editor in Chief for the IEEE SIGNAL PROCESSING L ETTERS, as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, as secretary of the SP Conference Board, as member of the SP Publications Board, as member and vice-chair of the Statistical Signal and Array Processing Technical Committee, and as chair of the SP for Communications Technical Committee. He has served as a member of the Editorial Board for the PROCEEDINGS OF THE IEEE, and the steering committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He is a member of the IEEE Fellows Election Committee, and has served on the IEEE-SP Society's Board of Governors.

Michael D. Zoltowski (F'98) was born in Philadelphia, PA, on August 12, 1960. He received both the B.S. and M.S. degrees in electrical engineering with highest honors from Drexel University, Philadelphia, PA, in 1983, and the Ph.D. degree in systems engineering from the University of Pennsylvania, Philadelphia, in 1986.

From 1982 to 1986, he was an Office of Naval Research Graduate Fellow. During 1987, he was a Summer Faculty Research Fellow at the Naval Ocean Systems Center, San Diego, CA. In Fall 1986, he joined the faculty of Purdue University, West Lafayette, IN, where he currently holds the position of Professor of Electrical and Computer Engineering. He is a contributing author to *Adaptive Radar Detection and Estimation* (New York: Wiley, 1991), *Advances in Spectrum Analysis and Array Processing* (Englewood Cliffs, NJ: Prentice-Hall, 1994, vol. III), and the *CRC Handbook on Digital Signal Processing* (Boca Raton, FL: CRC Press, 1996). He has served as a consultant to several companies in the telecommunications industry. His present research interests include space–time adaptive processing for all areas of mobile and wireless communications, GPS, and radar.

Dr. Zoltowski was the recipient of the IEEE Outstanding Branch Counselor Award for 1989–1990, the Ruth and Joel Spira Outstanding Teacher Award for 1990–1991, and the 2001–2002 Wilfred Hesselberth Award for Teaching Excellence. In 2001, he was named a University Faculty Scholar by Purdue University. He is the recipient of a 2002 Technical Achievement Award from the IEEE Signal Processing Society. He was also selected as a 2003 Distinguished Lecturer for the IEEE Signal Processing Society. He was a co-recipient of the IEEE Communications Society 2001 Leonard G. Abraham Prize Paper Award in the Field of Communications Systems. He was also the recipient of the IEEE Signal Processing Society's 1991 Paper Award (Statistical Signal and Array Processing Area), "The Fred Ellersick MILCOM Award for Best Paper in the Unclassified Technical Program" at the 1998 IEEE Military Communications Conference, and a Best Paper Award at the 2000 IEEE International Symposium on Spread Spectrum Techniques and Applications. He has served as an Associate Editor for both the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE COMMUNICATIONS LETTERS. He currently serves as an Area Editor in charge of feature articles for the *IEEE Signal Processing Magazine*. Within the IEEE Signal Processing Society, he has been a member of both the Technical Committee for the Statistical Signal and Array Processing Area and the Technical Committee for DSP Education. He is currently a member of both the Technical Committee on Signal Processing for Communications (SPCOM) and the Technical Committee on Sensor and Multichannel (SAM) Processing. In addition, from 1998 to 2001, he was a Member-at-Large of the Board of Governors and Secretary of the IEEE Signal Processing Society.