Optimal Transmitter Eigen-Beamforming and Space-Time Block Coding based on Channel Correlations*

Shengli Zhou and Georgios B. Giannakis
Dept. of ECE, Univ. of Minnesota, 200 Union Street SE, Minneapolis, MN 55455

***Abstract*—Optimal transmitter designs obeying the water-filling principle are well-documented, and widely applied when the propagation channel is deterministically known and regularly updated at the transmitter. Because channel state information may be costly or impossible to acquire in rapidly varying wireless environments, we develop in this paper statistical water-filling approaches for stationary random fading channels. The resulting optimal designs require only knowledge of the channel's second order statistics that do not require frequent updates, and can be easily acquired. Optimality refers to minimizing a tight bound on the symbol error rate. Applied to a multiple transmit-antenna paradigm, the optimal precoder turns out to be a generalized eigen-beamformer with multiple beams pointing to orthogonal directions along the eigenvectors of the channel's covariance matrix, and with proper power loading across the beams. Coupled with orthogonal space time block codes, two-directional eigen-beamforming emerges as a more attractive choice than conventional one-directional beamforming, with uniformly better performance, and without rate reduction or complexity increase.**

I. Introduction

Multi-antenna diversity is well motivated for wireless communications through fading channels. In certain applications, e.g., cellular downlink, multiple receive antennas may be expensive or impractical to deploy, which endeavors diversity systems relying on multiple transmit antennas.

When perfect or partial channel state information (CSI) is made available at the transmitter, multi-antenna systems can further enhance performance and capacity [7]. For slowly time-varying wireless channels, this amounts to feeding back to the transmitter the instantaneous channel estimates [7, 11]. But when the channel varies rapidly it is costly, yet not meaningful, to acquire CSI at the transmitter, because optimal transmissions tuned to previously acquired information become outdated quickly. Designing optimal transmitters based on statistical information about the underlying stationary random channel, is thus well motivated.

So long as the channel remains stationary, it has invariant statistics. Through field measurements, or theoretical models, the transmitter can acquire such statistical CSI a priori [8]. Alternatively, the receiver can estimate the channel correlations, and feed them back to the transmitter on line (this is referred to as covariance feedback in [5, 11]). Based on channel covariance information, optimal transmitter design has been pursued in [5, 11] based on a capacity criterion. Focusing on symbol by symbol detection, optimal precoding was designed in [2] to minimize the symbol error rate (SER) for differential BPSK transmissions, and in [4] for PSK based on channel estimation error, and conditional mutual information criteria.

In this paper, we design optimal transmit-diversity precoders for widely used constellations, and our performance-oriented approach relies on the Chernoff bound on SER. Optimal precoders turn out to be eigen-beamformers with multiple beams pointing to orthogonal directions along the eigenvectors of the channel's covariance matrix; hence, the name optimal transmitter eigen-beamforming. The optimal eigen-beams are power loaded according to a spatial water-filling principle. To increase the data rate without compromising the performance, we also propose parallel transmissions equipped with orthogonal space time block coding (STBC) [1, 3, 10] across optimally loaded eigen-beams. Interestingly, coupling optimal precoding with orthogonal STBC leads to a two-directional eigen-beamforming that enjoys uniformly better performance than the conventional one-directional beamforming without rate reduction, and without complexity increase.

Notation: Bold upper (lower) letters denote matrices (column vectors); $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^{\mathcal{H}}$ denote conjugate, transpose, and Hermitian transpose, respectively; $|\cdot|$ stands for the absolute value of a scalar and the determinant of a matrix; $\mathrm{E}\{\cdot\}$ stands for expectation, $\mathrm{tr}\{\cdot\}$ for the trace of a matrix; $\mathrm{Re}\{\cdot\}$ stands for the real part of a complex number; $\mathbf{I}_K$ denotes the identity matrix of size K; $\mathbf{0}_{K\times P}$ denotes an all-zero matrix with size $K \times P$; $\mathrm{diag}(\mathbf{x})$ stands for a diagonal matrix with $\mathbf{x}$ on its diagonal; $[\cdot]_p$ denotes the pth entry of a vector.

II. System Model

Fig. 1 depicts the block diagram of a transmit diversity system with a single receive- and N_t transmit- antennas. In the μth transmit-antenna, the information-bearing signal $s(i)$ is first spread (or, precoded) by the code $\mathbf{c}_\mu := [c_\mu(0), \ldots, c_\mu(P-1)]^T$ of length P to obtain the chip sequence: $u_\mu(n) = \sum_{i=-\infty}^{\infty} s(i)c_\mu(n-iP)$. The transmission channels are flat faded (frequency non-selective) with complex fading coefficients h_μ, $\mu = 1, \ldots, N_t$. The received samples in the presence of additive Gaussian noise $w(n)$ are thus given by:

$$x(n) = \sum_{i=-\infty}^{\infty} \sum_{\mu=1}^{N_t} h_\mu s(i) c_\mu(n-iP) + w(n). \quad (1)$$

To cast (1) into a convenient matrix-vector form, we define the $P\times 1$ vectors $\mathbf{x}(i) := [x(iP+0), \ldots, x(iP+P-1)]^T$ (likewise for $\mathbf{w}(i)$), the $N_t \times 1$ channel vector $\mathbf{h} := [h_1, \ldots, h_{N_t}]^T$, and the $P \times N_t$ code matrix $\mathbf{C} := [\mathbf{c}_1, \ldots, \mathbf{c}_{N_t}]$. Eq. (1) can then be re-written as: $\mathbf{x}(i) = \mathbf{C}\mathbf{h}s(i) + \mathbf{w}(i)$. Because we will focus on symbol by symbol detection, we omit the symbol index i, and subsequently deal with the input-output model

$$\mathbf{x} = \mathbf{C}\mathbf{h}s + \mathbf{w}. \quad (2)$$

At the receiver, the channel $\mathbf{h}$ is acquired first to enable maximum ratio combining (MRC) using

$$\mathbf{g}_{opt}^{\mathcal{H}} := [g(0), \ldots, g(P-1)] = (\mathbf{C}\mathbf{h})^{\mathcal{H}}. \quad (3)$$

*This work was supported by the NSF Wireless Initiative grant no. 99-79443, the NSF grant no. 0105612, and by the ARL/CTA grant no. DAAD19-01-2-011.

0-7803-7400-2/02/$17.00 © 2002 IEEE

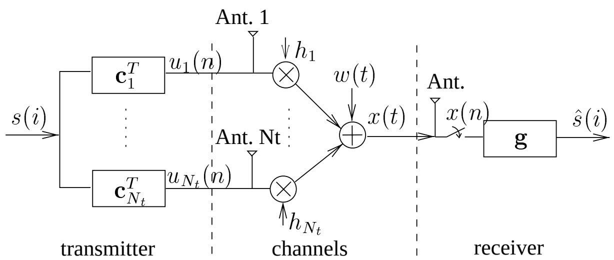

Fig. 1. Discrete-time baseband equivalent model

The MRC receiver maximizes the signal to noise ratio (SNR) at its output, and yields $\hat{s} = \mathbf{g}_{opt}^{\mathcal{H}}\mathbf{x} = \mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{x}$.

For a given precoder $\mathbf{C}$, eq. (3) specifies the optimal receiver $\mathbf{g}$ in the sense of maximizing the output SNR. The question that arises is how to select an optimal precoder $\mathbf{C}$. In the following, we design optimal $\mathbf{C}$ for random fading channels, based on knowledge of the channel's second-order statistics: namely $\mathbf{R}_{hh} := \mathrm{E}\left\{\mathbf{h}\mathbf{h}^{\mathcal{H}}\right\}$, and $\mathbf{R}_{ww} := \mathrm{E}\left\{\mathbf{w}\mathbf{w}^{\mathcal{H}}\right\}$.

III. Optimal Eigen-Beamforming

Throughout this paper, we adopt the following assumptions:
a0) *the channel* $\mathbf{h}$ *is complex Gaussian distributed, with zero-mean, and covariance matrix* $\mathbf{R}_{hh}$*;*
a1) *the noise* $\mathbf{w}$ *is zero-mean, white, complex Gaussian with each entry having variance* $N_0/2$ *per real and imaginary dimension, i.e.,* $\mathbf{R}_{ww} = N_0\mathbf{I}_P$*;*
a2) *the channel* $\mathbf{h}$ *and the noise* $\mathbf{w}$ *are uncorrelated.*
a3) *channel correlation information* $(\mathbf{R}_{hh}, N_0\mathbf{I}_P)$ *is available at the transmitter.*

Our performance metric for optimal precoder design will be symbol error rate (SER). We next derive a closed-form SER expression. The SNR at the MRC output for a fixed channel realization is $\gamma = \mathrm{E}\{|\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\mathbf{h}s|^2\}/\mathrm{E}\{\mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{w}\mathbf{w}^{\mathcal{H}}\mathbf{C}\mathbf{h}\}$. Denoting $E_s := \mathrm{E}\{|s|^2\}$ as the average energy of the underlying signal constellation, the SNR γ becomes:

$$\gamma = \mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\mathbf{h}\, E_s/N_0. \tag{4}$$

To simplify (4), we diagonalize $\mathbf{R}_{hh}$ as:

$$\mathbf{R}_{hh} = \mathbf{U}_h\mathbf{D}_h\mathbf{U}_h^{\mathcal{H}}, \quad \mathbf{D}_h := \mathrm{diag}(\lambda_1, \ldots, \lambda_{N_t}), \tag{5}$$

where $\mathbf{U}_h$ is unitary, and λ_μ denotes the μth eigenvalue of $\mathbf{R}_{hh}$ that is non-negative. Without loss of generality, we assume that λ_μ's are arranged in a non-increasing order: $\lambda_1 \geq \cdots \geq \lambda_{N_t} \geq 0$. Using (5), we can pre-whiten $\mathbf{h}$ to $\tilde{\mathbf{h}}$, so that $\mathbf{h} = \mathbf{U}_h\mathbf{D}_h^{\frac{1}{2}}\tilde{\mathbf{h}}$, and the entries of $\tilde{\mathbf{h}}$ are i.i.d with unit variance: $\mathrm{E}\{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathcal{H}}\} = \mathbf{I}_{N_t}$. Therefore, the SNR of (4) reduces to $\gamma = \tilde{\mathbf{h}}^{\mathcal{H}}\mathbf{D}_h^{\frac{1}{2}}\mathbf{U}_h^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\mathbf{U}_h\mathbf{D}_h^{\frac{1}{2}}\tilde{\mathbf{h}}\, E_s/N_0$. Let us now define $\mathbf{A} := \mathbf{D}_h^{\frac{1}{2}}\mathbf{U}_h^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\mathbf{U}_h\mathbf{D}_h^{\frac{1}{2}}$, which is non-negative definite, and thus it can be decomposed as: $\mathbf{A} = \mathbf{U}_A\mathbf{D}_A\mathbf{U}_A^{\mathcal{H}}$, where $\mathbf{D}_A := \mathrm{diag}(\delta_1, \ldots, \delta_{N_t})$ contains the N_t non-negative eigenvalues of $\mathbf{A}$. Because $\mathbf{U}_A$ is unitary, the vector $\tilde{\mathbf{h}}' := \mathbf{U}_A^{\mathcal{H}}\tilde{\mathbf{h}}$ has i.i.d entries (denoted by $\tilde{h}'_\mu := [\tilde{\mathbf{h}}']_\mu$), with covariance matrix $\mathbf{I}_{N_t}$. The SNR can then be further simplified to

$$\gamma = (\tilde{\mathbf{h}}')^{\mathcal{H}}\mathbf{D}_A\tilde{\mathbf{h}}' E_s/N_0 = \textstyle\sum_{\mu=1}^{N_t} \delta_\mu|\tilde{h}'_\mu|^2 E_s/N_0. \tag{6}$$

Notice that the SNR expression (6) coincides with that of the MRC output for N_t independent channels [9], with $\delta_\mu|\tilde{h}'_\mu|^2E_s/N_0$ denoting the μth subchannel's SNR. Averaging over the Rayleigh distributed $|\tilde{h}'_\mu|$, closed form SER expressions are found in [9] for M-ary phase shift keying (M-PSK), and square M-ary quadrature amplitude modulation (M-QAM) constellations, as:

$$P_{s,PSK} = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} \prod_{\mu=1}^{N_t} I_\mu(\delta_\mu E_s/N_0, g_{PSK}, \theta)d\theta, \tag{7}$$

$$P_{s,QAM} = \frac{b_{QAM}}{\sqrt{M}}\int_0^{\frac{\pi}{4}} \prod_{\mu=1}^{N_t} I_\mu(\delta_\mu E_s/N_0, g_{QAM}, \theta)d\theta + b_{QAM}\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \prod_{\mu=1}^{N_t} I_\mu(\delta_\mu E_s/N_0, g_{QAM}, \theta)d\theta, \tag{8}$$

where $b_{QAM} := 4(1-1/\sqrt{M})/\pi$, $I_\mu(x, g, \theta)$ is the moment generating function of the probability density function (p.d.f) of $|\tilde{h}'(\mu)|$ evaluated at $-gx/\sin^2\theta$ [9, eq. (24)], and the constellation-specific g is given respectively, by:

$$g_{PSK} = \sin^2(\pi/M), \quad \text{and} \quad g_{QAM} = 3/[2(M-1)]. \tag{9}$$

Because $|\tilde{h}'_\mu|$ is Rayleigh, $I_\mu(x, g, \theta)$ has the form: [9]:

$$I_\mu(x, g, \theta) = (1 + gx/\sin^2\theta)^{-1}. \tag{10}$$

A. Chernoff Bound Criterion

Our ultimate goal is to minimize the SER in (7), or (8), with respect to $\mathbf{C}$. However, direct optimization based on the exact SER turns out to be difficult because of the integration involved. Instead, we will design the optimal precoder $\mathbf{C}$ based on a tight Chernoff bound on SER.

Using the definite integral form for the Gaussian Q-function, the well-known Chernoff bound can be easily expressed as:

$$Q(x) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \exp\left(-\frac{x^2}{2\sin^2\theta}\right)d\theta \leq \frac{1}{2}\exp(-x^2/2), \tag{11}$$

for any $x \geq 0$, by observing that the maximum of the integrand occurs at $\theta = \pi/2$ [9]. Likewise, $I_\mu(x, g, \theta)$ in (10) peaks at $\theta = \pi/2$, and thus the Chernoff bound for the SER in (7) and (8) can be obtained in a unifying form:

$$P_{s,bound} = \alpha\prod_{\mu=1}^{N_t} I_\mu\left(\frac{\delta_\mu \mathrm{E}_s}{N_0}, g, \frac{\pi}{2}\right) = \alpha\left|\mathbf{I}_{N_t} + \mathbf{A}g\frac{E_s}{N_0}\right|^{-1}, \tag{12}$$

where $\alpha := (M-1)/M$, and g takes on constellation-specific values as in (9). The upper bound in (12) can also serve (within a scale) as a lower bound of the SER, e.g., $0.48P_{s,bound} \leq P_{s,PSK} \leq P_{s,bound}$ for QPSK and $N_t = 2$ [12].

The optimal precoder $\mathbf{C}$ will be chosen to maximize:

$$\mathcal{E}(\mathbf{C}) = \left|\mathbf{I}_{N_t} + \mathbf{A}\frac{gE_s}{N_0}\right| = \left|\mathbf{I}_{N_t} + \mathbf{D}_h^{\frac{1}{2}}\mathbf{U}_h\mathbf{C}^{\mathcal{H}}\mathbf{C}\mathbf{U}_h\mathbf{D}_h^{\frac{1}{2}}\frac{gE_s}{N_0}\right|,$$

under the constraint tr$\{\mathbf{C}^{\mathcal{H}}\mathbf{C}\} = 1$, i.e., the average transmitted power per symbol is E_s.

The cost function $\mathcal{E}(\mathbf{C})$ is maximized when the matrix $\mathbf{U}_h^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\mathbf{U}_h$ is diagonal [12]. We subsequently express:

$$\mathbf{U}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\mathbf{U}_h = \mathbf{D}_f^2, \text{ where } \mathbf{D}_f := \text{diag}(f_1, \ldots, f_{N_t}), \quad (13)$$

with $f_\mu \geq 0, \quad \forall \mu \in [1, N_t]$. Since $\log_2(\cdot)$ is a monotonically increasing function, we can equivalently optimize the cost function $\mathcal{E}'(\mathbf{C}) = \log_2 \mathcal{E}(\mathbf{C}) = \log_2 |\mathbf{I}_{N_t} + \mathbf{D}_f^2 \mathbf{D}_h g E_s / N_0|$, that will turn out to be more convenient. The equivalent constrained optimization problem is simplified to

$$\max_{\mathbf{D}_f} \mathcal{E}'(\mathbf{C}) \quad \text{subject to} \quad \mathcal{C} := \textstyle\sum_{\mu=1}^{N_t} f_\mu^2 - 1 = 0. \quad (14)$$

Differentiating the Lagrangian $\mathcal{E}'(\mathbf{C}) + \nu\mathcal{C}$ with respect to f_μ^2, where ν denotes the Lagrange multiplier, and equating it to zero, we obtain:

$$f_\mu^2 = [-1/(\nu \ln 2) - N_0/(g\lambda_\mu E_s)]_+ , \quad (15)$$

with the special notation $[x]_+ := \max(x, 0)$. Suppose that the given power budget E_s supports $\bar{N}_t$ non-zero f_μ^2's. Solving for ν using the power constraint, we arrive at the optimal loading:

$$f_\mu^2 = \left[\frac{1}{\bar{N}_t} + \frac{N_0}{gE_s} \left(\frac{1}{\bar{N}_t} \sum_{l=1}^{\bar{N}_t} \frac{1}{\lambda_l} - \frac{1}{\lambda_\mu} \right) \right]_+ . \quad (16)$$

The non-increasing order of the $\mathbf{R}_{hh}$ eigenvalues implies that: $f_1^2 \geq f_2^2 \cdots \geq f_{N_t}^2$, as confirmed by (16). We first set $\bar{N}_t = N_t$, and test if $f_{N_t}^2 \geq 0$. The entry $f_{N_t}^2 \geq 0$ in (16) with $\bar{N}_t = N_t$ imposes the following lower bound on the required SNR: $E_s/N_0 > (1/g)(N_t/\lambda_{N_t} - \sum_{\mu=1}^{N_t} 1/\lambda_\mu) := \bar{\gamma}_{th,N_t}$. If E_s is not large enough to afford optimal power allocation across all N_t beams, causing $f_{N_t}^2 < 0$, eq. (16) suggests that we should turn off the N_tth beam by setting $f_{N_t}^2 = 0$, and set $\bar{N}_t = \bar{N}_t - 1$; and so on until we will find the desired $\bar{N}_t$.

The practical power loading algorithm is summarized as:

S1) For $r = 1, \ldots, N_t$, calculate $\bar{\gamma}_{th,r}$ based only on the first r channel eigenvalues as

$$\bar{\gamma}_{th,r} := (1/g)(r/\lambda_{rr} - \textstyle\sum_{\mu=1}^{r} 1/\lambda_\mu). \quad (17)$$

S2) With the given power budget E_s ensuring that E_s/N_0 falls in the interval $[\bar{\gamma}_{th,r}, \bar{\gamma}_{th,r+1}]$, set $f_{r+1}, \ldots, f_{N_t} = 0$, and obtain $f_1, \ldots, f_r$ according to (16) based only on $\lambda_1, \ldots, \lambda_r$.

Having specified the optimal f_μ^2, we have found the optimal $\mathbf{D}_f^2$ in (13). The optimal $\mathbf{C}$ can be factored from (13) as:

$$\mathbf{C} = \mathbf{\Phi}\mathbf{D}_f\mathbf{U}_h^{\mathcal{H}}, \quad (18)$$

where the columns of $\mathbf{\Phi}$ are orthonormal, and the diagonal entries of $\mathbf{D}_f$ are given by (16). We summarize our result as:

Theorem 1: *Suppose a0)-a3) hold true. The optimum receive-filter $\mathbf{g}_{opt}$ is given by* (3)*, and the optimum precoding matrix $\mathbf{C}_{opt} = \mathbf{\Phi}\mathbf{D}_f\mathbf{U}_h^{\mathcal{H}}$ has $\mathbf{U}_h$ and $\mathbf{D}_f$ formed as in* (5), (16) *and* (13) *with $\mathbf{\Phi}$ an arbitrary orthonormal $P \times N_t$ matrix. Optimality in $\mathbf{g}_{opt}$ refers to maximum-SNR, while optimality in $\mathbf{C}_{opt}$ pertains to minimizing the Chernoff bound on the average symbol error rate.*

B. Optimally Loaded Eigen-beamforming Interpretation

The $P \times N_t$ optimal precoder $\mathbf{C}$ in (18) can be interpreted as a generalized beamformer. Different from conventional beamforming that transmits all available power along the channel's strongest direction (implemented via the first row of $\mathbf{U}_h^{\mathcal{H}}$), here N_t beams are formed pointing to N_t orthogonal directions along the eigenvectors of the channel covariance matrix $\mathbf{R}_{hh}$; thus, the name eigen-beamforming. The matrix $\mathbf{D}_f$ takes care of power loading across all beams. Notice that more power is distributed to stronger channels since $f_1^2 \geq f_2^2 \cdots \geq f_{N_t}^2$. Furthermore, $f_\mu^2 + (1/\lambda_\mu)(N_0/gE_s)$ is constant $\forall \mu$; thus, the power allocation obeys the water-filling principle.

When the system operates at a prescribed power: $E_s/N_0 \in [\bar{\gamma}_{th,r}, \bar{\gamma}_{th,r+1}]$, it is clear that only $r = \text{rank}(\mathbf{D}_f)$ eigen-beams are used, and a diversity order of r is achieved. Full diversity schemes correspond to $r = N_t$. Based on (17), one can easily determine what diversity level to be used for the best performance with a given power budget E_s. We thus have:

Corollary 1: *The optimal diversity order is r, when E_s/N_0 falls in the interval: $[\bar{\gamma}_{th,r}, \bar{\gamma}_{th,r+1}]$, with $\bar{\gamma}_{th,r}$ defined in* (17).

Corollary 2: *Full diversity schemes are not SER-bound optimal across the entire SNR range; their optimality is ensured only when the SNR is sufficiently high: $E_s/N_0 > \bar{\gamma}_{th,N_t}$.*

Notice that apart from requiring it to be orthonormal, so far we left the $P \times N_t$ matrix $\mathbf{\Phi}$ unspecified. To fully exploit the diversity offered by N_t antennas, $P \geq N_t$ is required. On the other hand, the choice $P > N_t$ does not improve performance; it is thus desirable to choose P as small as possible to increase the transmission rate. When the desired diversity order is r, as in Corollary 1, we can reduce the $P \times N_t$ matrix $\mathbf{\Phi}$ to an $r \times N_t$ fat matrix $[\tilde{\mathbf{\Phi}}, \mathbf{0}_{r\times(N_t-r)}]$, where $\tilde{\mathbf{\Phi}}$ is any $r \times r$ orthonormal matrix, without loss of optimality. This way, we can achieve rate $1/r$ for a diversity transmission of order r.

On the other hand, one can *a priori* force the matrix $\mathbf{C}$ (and thus $\mathbf{\Phi}$) to be fat with dimensionality $d \times N_t$, which corresponds to setting $f_{d+1}, \ldots, f_{N_t} = 0$, deterministically. Optimal power loading can then be applied to the remaining d beams. We will term this scheme (with $\mathbf{C}$ chosen beforehand to be $d \times N_t$) an d-directional eigen-beamformer. As a consequence of Theorem 1, we then have:

Corollary 3: *With $d < N_t$, the d-directional eigen-beamformer achieves the same average SER performance as an N_t-directional eigen-beamformer, when $E_s/N_0 < \gamma_{th,d+1}$.*

Two interesting special cases of Corollary 3 arise. The first is conventional one-directional (1D) eigen-beamforming with $d = 1$, [5, 7, 11]. As asserted in Corollary 3, the 1D eigen-beamformer will be optimal when $E_s/N_0 < \bar{\gamma}_{th,2} = (1/\lambda_2 - 1/\lambda_1)/g$; i.e., when the first and second eigenvalues are disparate enough, or, when E_s/N_0 is sufficiently low.

The more interesting case is 2D eigen-beamforming which corresponds to $d = 2$. The 2D eigen-beamformer is optimal when $E_s/N_0 < \bar{\gamma}_{th,3} = (2/\lambda_3 - 1/\lambda_1 - 1/\lambda_2)/g$. Notice that the optimality condition for 2D beamforming is less restrictive than that for the 1D beamforming, since $\bar{\gamma}_{th,3} \geq \bar{\gamma}_{th,2}$, and $\bar{\gamma}_{th,3}$ does not depend on λ_1 and λ_2. As we shall see in Section IV, the 2D eigen-beamforming also achieves the same rate as 1D beamforming, and subsumes the latter as a special case.

IV. Eigen-Beamforming and STBC

In the system model (2), we transmit only one symbol over P time slots (chip-periods), which essentially amounts to repetition coding (or a spread-spectrum) scheme. To overcome the associated rate loss, it is possible to send K symbols $s_1, \ldots, s_K$ simultaneously. Certainly, this would require symbol separation at the receiver. But let us suppose temporarily that the separation is indeed achievable, and each symbol is essentially going through a separate channel identical to the one we dealt with in Section III. The optimal precoder $\mathbf{C}_k$ for s_k will then be

$$\mathbf{C}_k = \mathbf{\Phi}_k \mathbf{D}_f \mathbf{U}_h^{\mathcal{H}}, \quad k = 1, 2, \ldots, K. \tag{19}$$

Because the factor $\mathbf{D}_f \mathbf{U}_h^{\mathcal{H}}$ in (19) is common $\forall k$, designing separable precoders is equivalent to choosing separable $\mathbf{\Phi}_k$'s. Fortunately, this degree of freedom can be afforded by our design in Section III because so far the $\mathbf{\Phi}_k$'s are only required to have orthonormal columns.

The desired means of data multiplexing that enables symbol separability at the receiver is possible through orthogonal space-time block coding (STBC) [1, 3, 10]. Our combining of optimal eigen-beamforming with STBC is treated next for complex symbols (see [12] for real symbols).

Let s_k^R and s_k^I denote the real and imaginary part of s_k, respectively. The following orthogonal STBC designs are available for complex symbols [3, 10]:

Definition: *For complex symbols* $\{s_k = s_k^R + j s_k^I\}_{k=1}^K$, *and* $P \times N_t$ *matrices* $\{\mathbf{\Phi}_k, \mathbf{\Psi}_k\}_{k=1}^K$ *each having entries drawn from* $\{1, 0, -1\}$, *the space time coded matrix*

$$\mathcal{O}_{N_t} = \textstyle\sum_{k=1}^K \mathbf{\Phi}_k s_k^R + j \sum_{k=1}^K \mathbf{\Psi}_k s_k^I \tag{20}$$

is termed a generalized complex orthogonal design (GCOD) in variables $\{s_k\}_{k=1}^K$ *of size* $P \times N_t$ *and rate* K/P, *if either one of two equivalent conditions holds true:*
i) $\mathcal{O}_{N_t}^{\mathcal{H}} \mathcal{O}_{N_t} = (\sum_{k=1}^K |s_k|^2) \mathbf{I}_{N_t}$ *[10], or,*
ii) The matrices $\{\mathbf{\Phi}_k, \mathbf{\Psi}_k\}_{k=1}^K$ *satisfy the conditions [3]:*

$$\begin{aligned} &\mathbf{\Phi}_k^{\mathcal{H}} \mathbf{\Phi}_k = \mathbf{I}_{N_t}, \mathbf{\Psi}_k^{\mathcal{H}} \mathbf{\Psi}_k = \mathbf{I}_{N_t}, \qquad \forall k \\ &\mathbf{\Phi}_k^{\mathcal{H}} \mathbf{\Phi}_l = -\mathbf{\Phi}_l^{\mathcal{H}} \mathbf{\Phi}_k, \mathbf{\Psi}_k^{\mathcal{H}} \mathbf{\Psi}_l = -\mathbf{\Psi}_l^{\mathcal{H}} \mathbf{\Psi}_k, \quad k \neq l \\ &\mathbf{\Phi}_k^{\mathcal{H}} \mathbf{\Psi}_l = \mathbf{\Psi}_l^{\mathcal{H}} \mathbf{\Phi}_k, \qquad \forall k, l \qquad \square \end{aligned} \tag{21}$$

For each complex symbol $s_k = s_k^R + j s_k^I$, we define two precoders corresponding to $\{\mathbf{\Phi}_k, \mathbf{\Psi}_k\}$ as: $\mathbf{C}_{k,1} = \mathbf{\Phi}_k \mathbf{D}_f \mathbf{U}_h^{\mathcal{H}}$, and $\mathbf{C}_{k,1} = \mathbf{\Psi}_k \mathbf{D}_f \mathbf{U}_h^{\mathcal{H}}$. The transmitted STBC is now

$$\mathcal{Z}_{N_t} = \textstyle\sum_{k=1}^K \mathbf{C}_{k,1} s_k^R + j \sum_{k=1}^K \mathbf{C}_{k,2} s_k^I = \mathcal{O}_{N_t} \mathbf{D}_f \mathbf{U}_h^{\mathcal{H}}. \tag{22}$$

At the kth detector output, the decision variable is formed by

$$\begin{aligned} y_k &= \mathrm{Re}\{\mathbf{h}^{\mathcal{H}} \mathbf{C}_{k,1}^{\mathcal{H}} \mathbf{x}\} + j \mathrm{Re}\{-j \mathbf{h}^{\mathcal{H}} \mathbf{C}_{k,2}^{\mathcal{H}} \mathbf{x}\} \\ &= \mathbf{h}^{\mathcal{H}} \mathbf{U}_h \mathbf{D}_f^2 \mathbf{U}_h^{\mathcal{H}} \mathbf{h} s_k + w_k, \quad \forall k \in [1, K], \end{aligned} \tag{23}$$

where w_k has variance $N_0 \mathbf{h}^{\mathcal{H}} \mathbf{U}_h \mathbf{D}_f^2 \mathbf{U}_h^{\mathcal{H}} \mathbf{h}$; and the second equality in (23) can be easily verified by using (21) [3]. Notice that (23) is nothing but the MRC output for the single symbol transmission studied in Section III; thus, the optimal loading in (16) enables space-time block coded transmissions to minimize the Chernoff bound on SER, but with rate K/P. The combination of orthogonal STBC with beamforming has also been studied in [6]. However, the focus in [6] is on channel mean feedback [11] for slowly fading channels, while our approach here is tailored for fast fading random channels.

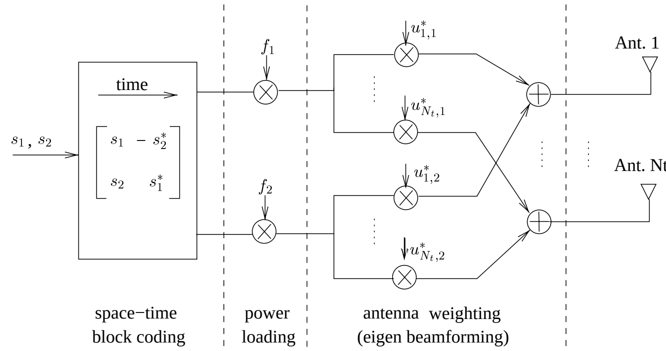

Fig. 2. The two-directional (2D) eigen-beamformer, $u_{p,q} := [\mathbf{U}_h]_{p,q}$

For complex symbols, a rate 1 GCOD only exists for $N_t = 2$. It corresponds to the well-known Alamouti code [1]:

$$\mathcal{O}_2 = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} \begin{matrix} \rightarrow \text{ space} \\ \downarrow \text{ time} \end{matrix}. \tag{24}$$

For $N_t = 3, 4$, rate $3/4$ orthogonal STBCs exist, while for $N_t > 4$, only rate $1/2$ codes have been constructed [10], [3].

Therefore, for complex symbols, the N_t-directional eigen-beamformer of (22) achieves optimal performance with no rate loss only when $N_t = 2$, and pays a rate penalty when $N_t > 2$. To tradeoff the optimal performance for a constant rate 1 transmission, it is possible to construct a 2D eigen-beamformer with the Alamouti code applied on the strongest two-directional eigen-beams. Specifically, we can construct a $2 \times N_t$ matrix $\mathcal{Z}_{\text{2-d}} := [\mathcal{O}_2, \mathbf{0}_{2\times(N_t-2)}] \mathbf{D}_f \mathbf{U}_h^{\mathcal{H}}$, which achieves the optimal performance as the N_t-directional eigen-beamformer when $E_s/N_0 < \bar{\gamma}_{th,3}$, as specified in Corollary 3. The implementation of the 2D eigen-beamformer is depicted in Fig. 2.

Notice that the optimal scenario for 1D beamforming was specified in [5] from a capacity perspective. The interest in 1D beamforming stems primarily from the fact that it allows scalar coding with linear pre-processing at the transmit-antenna array, and thus relieves the receiver from the complexity required for decoding the capacity-achieving vector coded transmissions [5, 7, 11]. Because each symbol with 2D eigen-beamforming goes through a separate but better conditioned channel, the same capacity-achieving scalar code applied to an 1D beamformer can be applied to a 2D eigen-beamformer as well. Therefore, 2D eigen-beamforming outperforms 1D beamforming even from a capacity perspective, since it can achieve the same coded BER with less power. Notice that if $\mathbf{D}_f$ has only one nonzero entry f_1, the 2D eigen-beamformer reduces to the 1D beamformer, with s_1 and $-s_2^*$ transmitted during consecu-

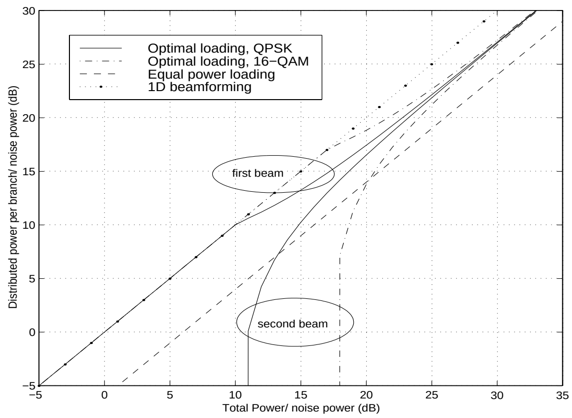

Fig. 3. Optimal vs. equal power loading

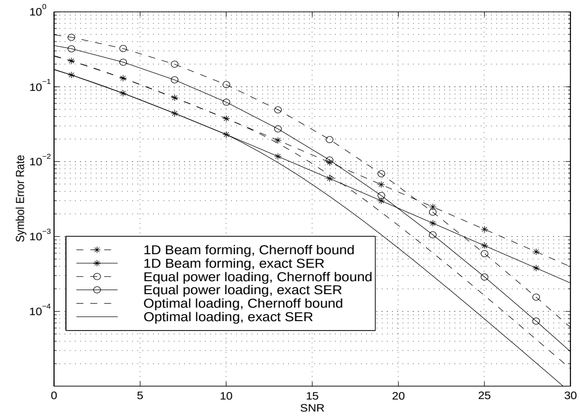

Fig. 4. SER vs E_s/N_0: QPSK

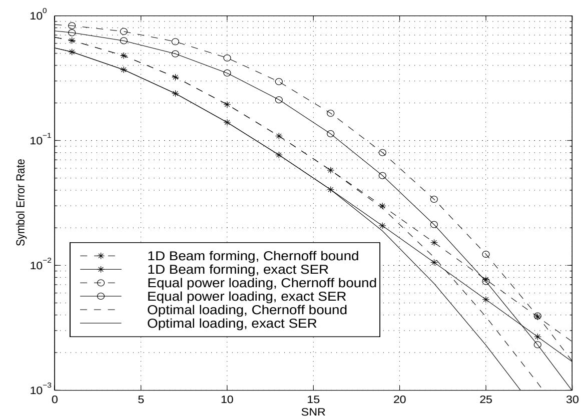

Fig. 5. SER vs E_s/N_0: 16-QAM

tive time-slots, as confirmed by (24). This leads to the following conclusion:

Corollary 4: *The 2D eigen-beamformer includes 1D beamformer as a special case and outperforms it uniformly, without rate reduction and without complexity increase.*

Corollary 4 suggests that 2D eigen-beamformer is more attractive than 1D beamformer, and deserves more attention.

V. Numerical Results

We consider a uniform linear array with $N_t = 4$ antennas at the transmitter, and a single antenna at the receiver. We assume that the side information including the distance between the transmitter and the receiver, the angle of arrival, and the angle spread are all available at the transmitter. Let λ be the wavelength of a narrowband signal, d_t the antenna spacing, and Δ the angle spread. We assume that the angle of arrival is perpendicular to the transmitter antenna array ("broadside" as in [8]), and $d_t = 0.5\lambda$, and $\Delta = 5^\circ$ (see [12] for additional setups). The correlation coefficients among the antennas are then calculated by [8, eq. (6)].

Fig. 3 shows the optimal power allocation among different beams, for both QPSK and QAM constellations. Notice that the choice of how many beams are retained depends on the constellation-specific SNR thresholds. For QPSK, we can verify that $\bar{\gamma}_{th,2} = 10.2$dB, and $\bar{\gamma}_{th,3} = 37.5$dB. Since $g_{QPSK}/g_{16QAM} = 5$, the threshold $\bar{\gamma}_{th,r}$ for 16-QAM is $10\log_{10}(5) = 7.0$dB higher for QPSK; we observe that 7.0dB higher power is required for 16-QAM before switching to the same number of beams as for QPSK.

Figs. 4 and 5 depict the exact SER, and the Chernoff bound for: optimal loading, equal power loading, and 1D beamforming. Since the considered channel is highly correlated, only $r = 2$ beams are used in the considered SNR range for optimal loading. Therefore, the 2D eigen-beamformer is overall optimal for this channel in the considered SNR range, and its performance curves coincide with those of the optimal loading. Figs. 4 and 5 confirm that the optimal allocation outperforms both the equal power allocation, and the 1D beamforming. The small gap between the Chernoff bound and the exact SER in Figs. 4 and 5 justifies our approach that pushes down the Chernoff bound to minimize the exact SER.

References

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE JSAC*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[2] J. K. Cavers, "Optimized use of diversity modes in transmitter diversity systems," in *Proc. of VTC*, 1999, vol. 3, pp. 1768–1773.

[3] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," *IEEE Trans. on Info. Theory*, pp. 1650–1656, May 2001.

[4] G. B. Giannakis and S. Zhou, "Optimal transmit-diversity precoders for random fading channels," in *Proc. of Globecom*, 2000, pp. 1839–1843.

[5] S. A. Jafar and A. Goldsmith, "On optimality of beamforming for multiple antenna systems with imperfect feedback," in *Proc. of IEEE ISIT*, June 2001, pp. 321–321.

[6] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining transmit beamforming and orthogonal space-time block codes by utilizing side information," *Proc. of 1st IEEE Sensor Array and Multichannel. Signal Proc. Workshop*, Boston, MA, Mar. 2000, pp. 153–157.

[7] 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 JSAC*, vol. 16, pp. 1423–1436, Oct. 1998.

[8] 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. on Communications*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[9] M. K. Simon and M.-S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. of the IEEE*, vol. 86, no. 9, pp. 1860–1877, Sept. 1998.

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

[11] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. on Info. Theory*, Sept. 2001.

[12] S. Zhou, and G. B. Giannakis, "Optimal transmitter eigen-Beamforming and space-time block coding based on channel correlations," *IEEE Trans. on Info. Theory*, submitted September 2001.