# Optimal Transmitter Eigen-Beamforming and Space-Time Block Coding Based on Channel Mean Feedback

Shengli Zhou, *Student Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Optimal transmitter designs obeying the water-filling principle are well-documented; they are widely applied when the propagation channel is deterministically known and regularly updated at the transmitter. Because channel state information is impossible to be known perfectly at the transmitter in practical wireless systems, we design, in this paper, an optimal multiantenna transmitter based on the knowledge of mean values of the underlying channels. Our optimal transmitter design turns out to be an eigen-beamformer with multiple beams pointing to orthogonal directions along the eigenvectors of the correlation matrix of the estimated channel at the transmitter and with proper power loading across beams. The optimality pertains to minimizing an upper bound on the symbol error rate, which leads to better performance than maximizing the expected signal-to-noise ratio (SNR) at the receiver. Coupled with orthogonal space-time block codes, two-directional eigen-beamforming emerges as a more attractive choice than conventional one-directional beamforming with uniformly improved performance, without rate reduction, and without essential increase in complexity. With multiple receive antennas and reasonably good feedback quality, the two-directional eigen-beamformer is also capable of achieving the best possible performance in a large range of transmit-power-to-noise ratios, without a rate penalty.**

***Index Terms*—Beamforming, mean feedback, space-time block codes, transmit diversity.**

## I. Introduction

Multiantenna diversity is well motivated for wireless communications through fading channels. Although receive-antenna diversity has been widely applied in practice, in certain cases, e.g., cellular downlink, multiple receive antennas may be expensive or impractical to deploy, which endeavors transmit-diversity systems. Equipped with space-time coding at the transmitter and intelligent signal processing at the receiver, multiantenna transceivers offer significant diversity and coding advantages over single antenna systems [1], [17], [18]. Our attention in this paper is thereby focused on application scenarios dealing with multiple transmit antennas.

Manuscript received November 1, 2001; revised May 13, 2002. This work was supported by the the National Science Foundation (NSF) under Grant 0105612, the NSF under Wireless Initiative Grant 99-79443, and by the ARL/CTA under Grant DAAD19-01-2-011. This work was presented in part at the International Conference on Acoustics Speech and Signal Processing, Orlando, FL, USA, May 13–17, 2002. The associate editor coordinating the review of this paper and approving it for publication was Prof. Brian L. Hughes.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: szhou@ece.umn.edu; georgios@ece.umn.edu).

Publisher Item Identifier 10.1109/TSP.2002.803355.

Multiantenna systems can further enhance performance and capacity when perfect or partial channel state information (CSI) is made available at the transmitter [3], [12], [19]. Collect the channel coefficients from $N_t$ transmit antennas to one receive antenna into an $N_t \times 1$ vector $\mathbf{h}$. Given partial CSI at the transmitter, the spatial channel $\mathbf{h}$ can be generally modeled as a complex Gaussian random vector with nonzero mean $\bar{\boldsymbol{h}}$ and nonwhite covariance matrix $\boldsymbol{\Sigma}_h$ [19]. Two special forms of partial feedback are possible [19]: *mean feedback* and *covariance feedback*. Mean feedback assumes knowledge of the channel mean $\bar{\boldsymbol{h}}$ and models the covariance as white with $\boldsymbol{\Sigma}_h$ proportional to an identity matrix. For slowly varying wireless channels, this is achieved, for example, by feeding back to the transmitter an unquantized, or quantized, channel estimate $\hat{\mathbf{h}}$ acquired at the receiver. The transmitter's uncertainty about the channel around its mean is embodied in the nonzero vector $\mathbf{h} - \bar{\boldsymbol{h}}$, which may arise due to

- channel estimation errors at the receiver;
- quantization errors;
- errors induced by the feedback channel;
- channel variations during the feedback delay.

In time division duplex (TDD) or frequency division duplex (FDD) systems, channel mean values can be also obtained from uplink measurements by exploiting the time or frequency correlation between downlink and uplink channels [3]. Covariance feedback, on the other hand, is motivated when the channel $\mathbf{h}$ varies too rapidly for the transmitter to track its mean. In this case, the channel mean is set to zero, and the relative geometry of the propagation paths manifests itself in a nonwhite covariance matrix $\boldsymbol{\Sigma}_h$ [19].

Based on either mean feedback or covariance feedback, optimal multiantenna transmitter design has been pursued in [7], [8], [12], and [19] based on a *capacity criterion*, which specifies the theoretical maximum rate of reliable communication achievable in the absence of delay and processing constraints (see also [13], when no feedback is available at the transmitter). With covariance feedback, optimal transmitter precoding was designed in [2] to minimize the symbol error rate (SER) for differential binary phase-shift keying (BPSK) transmissions and in [6] for PSK, based on channel estimation error and conditional mutual information criteria. With a *SER upper bound as criterion*, optimal precoding with covariance feedback has been generalized in [20] to widely used constellations.

In this paper, we design SER-bound optimal multiantenna transmit precoders for widely used constellations based on channel mean feedback. The optimal precoder turns out to be a generalized beamformer with multiple beams formed using the orthogonal eigenvectors of the correlation matrix of the estimated channel at the transmitter, hence the name optimal transmitter eigen-beamforming. The optimal eigen-beams are power loaded according to a “spatial water-filling” principle. Our performance-oriented designs rely on an upper bound of the SER and outperform the designs that are based on maximizing the expected signal-to-noise ratio (SNR) at the receiver. The latter lead to conventional beamforming, which transmits all power along the strongest direction that the feedback dictates, no matter how reliable the feedback information is [3], [12].

To increase the data rate without compromising the performance, we also develop parallel transmissions equipped with orthogonal space-time block coding (STBC) [1], [5], [17] across optimally loaded eigen-beams. Wedding optimal precoding with orthogonal STBC leads to a two-directional eigen-beamforming, which turns out to enjoy uniformly better performance than the conventional one-directional beamforming without rate reduction and without complexity increase. With two transmit antennas, the two-directional eigen-beamformer achieves the best possible performance. However, even with more than two transmit antennas, if multiple receive antennas are present and the feedback quality is reasonably good, the two-directional eigen-beamformer achieves the best possible performance over a large range of transmit-power to noise ratios without a rate penalty. Thanks to its full-rate capability and superior performance, the two-directional eigen-beamformer has strong application potential in future wireless systems with multiple ($N_t \geq 2$) transmit-antennas and channel mean feedback.

The combination of orthogonal STBC with beamforming has also been studied in [9] and [11]. We detail the differences and novelties of this paper relative to [9] and [11] in Section V-C.

The rest of this paper is organized as follows. Section II describes the system model. Section III develops optimal eigen-beamformers for a single receive antenna, whereas Section III deals with multiple receive antennas. Section V is devoted to jointly exploiting orthogonal STBC and optimal eigen-beamforming. Numerical results are presented in Section VI, and conclusions are drawn in Section VII.

*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 or the determinant of a matrix and $\|\cdot\|$ for the Euclidean norm of a vector; $E\{\cdot\}$ stands for expectation, tr$\{\cdot\}$ for the trace of a matrix; Re$\{\cdot\}$ stands for the real part of a complex number and Im$\{\cdot\}$ for the imaginary part; sign$(\cdot)$ denotes the sign of a real number, and $\lfloor\cdot\rfloor$ the integer floor; $\mathbf{I}_K$ denotes the identity matrix of size $K$; $\mathbf{0}_{K\times P}$ denotes an all-zero matrix with size $K \times P$; diag$(\mathbf{x})$ stands for a diagonal matrix with $\mathbf{x}$ on its diagonal; $[\cdot]_p$ denotes the $p$th entry of a vector; and $[\cdot]_{p,q}$ denotes the $(p,q)$th entry of a matrix. The special notation $\mathbf{h} \sim \mathcal{CN}(\bar{\boldsymbol{h}}, \boldsymbol{\Sigma}_h)$ indicates that $\mathbf{h}$ is complex Gaussian distributed with mean $\bar{\boldsymbol{h}}$ and covariance $\boldsymbol{\Sigma}_h$.

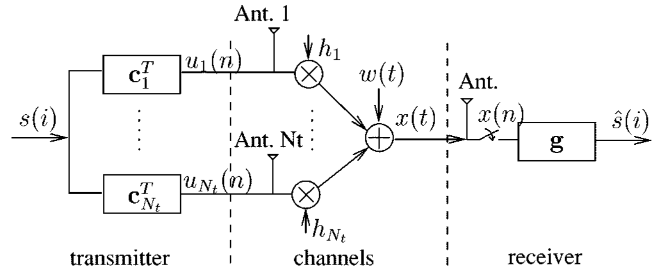


Fig. 1. Discrete-time baseband equivalent model.

## II. Modeling and Preliminaries

Fig. 1 depicts the block diagram of a transmit diversity system with a single receive and $N_t$ transmit antennas. The extension to multiple receive antennas is postponed until Section IV. In the $\mu$th ($\mu \in [1, N_t]$) transmit antenna, each information-bearing symbol $s(i)$ is first spread by the code $\mathbf{c}_\mu := [c_\mu(0), \ldots, c_\mu(P-1)]^T$ of length $P$ to obtain the chip sequence $u_\mu(n) = \sum_{i=-\infty}^{\infty} s(i)c_\mu(n-iP)$. After spectral shaping by the transmit-filter $\varphi(t)$ (which is not shown in Fig. 1), the continuous-time signal $u_\mu(t) = \sum_{n=-\infty}^{\infty} u(n)\varphi(t-nT_c)$ is transmitted through the $\mu$th antenna, where $T_c$ is the chip duration. The transmission channels are flat faded (frequency nonselective) with complex fading coefficients $\{h_\mu\}_{\mu=1}^{N_t}$. The received signal in the presence of additive white Gaussian noise $w(t)$ is thus given by $x(t) = \sum_{\mu=1}^{N_t} h_\mu u_\mu(t) + w(t)$. After receive filtering with $\bar{\varphi}(t)$, which is matched to $\varphi(t)$, the received signal $x(t)$ is sampled at $t = nT_c$ to yield the discrete time samples $x(n) := x(t)|_{t=nT_c}$. Selecting $\varphi(t)$ and $\bar{\varphi}(t)$ to possess the square root Nyquist-$T_c$ property avoids intersymbol interference and allows one to express $x(n)$ as

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

where $w(n) := (w \star \bar{\varphi})(t)|_{t=nT_c}$ with $\star$ denoting linear convolution. Because $w(t)$ is white Gaussian and $\bar{\varphi}(t)$ is square root Nyquist, the sampled noise sequence $w(n)$ is also white Gaussian.

To cast (1) into a convenient matrix-vector form, we define the $P \times 1$ vectors $\mathbf{x}(i) := [x(iP+0), \ldots, x(iP+P-1)]^T$, and $\mathbf{w}(i) := [w(iP+0), \ldots, w(iP+P-1)]^T$; the $N_t \times 1$ channel vector $\mathbf{h} := [h_1, \ldots, h_{N_t}]^T$, and the $P \times N_t$ spreading code matrix[1] $\mathbf{C} := [\mathbf{c}_1, \ldots, \mathbf{c}_{N_t}]$. The block version of (1) can be rewritten 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

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

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

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

[1]The spreading matrix $\mathbf{C}$ can be viewed (and will be invariably referred to) as a precoder or as a beamformer.

The MRC receiver is known to maximize the signal-to-noise ratio (SNR) at its output [14]. Furthermore, slicing the MRC output $y = \mathbf{g}_{\text{opt}}^{\mathcal{H}}\mathbf{x} = \mathbf{h}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{x}$, yields the desired symbol estimate $\hat{s}$, e.g., with BPSK, we obtain $\hat{s} = \text{sign}(\text{Re}\{y\})$.

### A. Problem Statement and Assumptions

For a given precoder $\mathbf{C}$, (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 $\mathbf{C}$ if perfect or imperfect channel state information is available at the transmitter. The optimal $\mathbf{C}$ based on channel covariance feedback is provided in [20]. In this paper, we look for the optimal $\mathbf{C}$ based on channel mean feedback.

We will first optimize $\mathbf{C}$ for the configuration of Fig. 1. Due to spreading, this multiantenna transmitter is not rate efficient since we transmit $(1/P)$ symbols/s/Hz with $N_t$ antennas. Such a redundant transmitter was also studied in [2] and has its own merits for "power-limited" (e.g., spread spectrum military communication) systems, where spectral resources are not at a premium, but low transmission power is desired. To enable operation in "bandwidth-limited" scenarios, we will combine our optimum low-rate precoder with orthogonal space-time coding in Section V. This combination will lead us to an important transmitter design that enjoys full rate (1 symbol/s/Hz) transmissions for any number of transmit antennas.

Throughout this paper, we will adopt the following assumptions.

**a0)** *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.,* $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_P)$.

**a1)** *The transmitter obtains an unbiased channel estimate* $\bar{\boldsymbol{h}}$ *based on partial CSI received from the receiver through the feedback channel; before updated feedback arrives, the transmitter treats* $\bar{\boldsymbol{h}}$ *as deterministic, and in order to account for CSI imperfections, it relies on an estimate*[2] *of the true channel* $\mathbf{h}$, *which is formed as*

$$\check{\mathbf{h}} := \bar{\boldsymbol{h}} + \boldsymbol{\epsilon} \tag{4}$$

*where* $\boldsymbol{\epsilon} \sim \mathcal{CN}(\mathbf{0}, \sigma_\epsilon^2\mathbf{I}_{N_t})$.

Assumption a1) corresponds to the mean feedback in [19].

### B. Channel Mean Feedback

We next highlight three specific possibilities where a1) holds true and illustrate how to obtain $(\bar{\boldsymbol{h}}, \sigma_\epsilon^2)$ from partial CSI; more realistic cases could be derived similarly.

*Case 1 (Delayed Feedback) [9], [12], [19]:* Here, we assume the following.

i) The channel coefficients are slowly time varying according to Jakes' model with Doppler frequency $f_d$.
ii) The transmit antennas are well separated, and thus, $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \sigma_h^2\mathbf{I}_{N_t})$.
iii) The channel is acquired perfectly at the receiver and is fed back to the transmitter via a noiseless channel with delay $D(PT_c)$.

Let $\hat{\mathbf{h}}_f$ denote the channel feedback. Notice that both $\mathbf{h}$ and $\hat{\mathbf{h}}_f$ are complex Gaussian vectors, drawn from the same distribution $\mathcal{CN}(\mathbf{0}, \sigma_h^2\mathbf{I}_{N_t})$. It can be shown that $E\{\mathbf{h}\hat{\mathbf{h}}_f^{\mathcal{H}}\} = \rho\sigma_h^2\mathbf{I}_{N_t}$, where the correlation coefficient $\rho := J_0(2\pi f_d DPT_c)$ determines the feedback quality. The minimum mean-square error (MMSE) estimator of $\mathbf{h}$ based on $\hat{\mathbf{h}}_f$ is given by $E\{\mathbf{h}|\hat{\mathbf{h}}_f\} = \rho\hat{\mathbf{h}}_f$, with estimation error having covariance matrix $\sigma_h^2(1 - |\rho|^2)\mathbf{I}_{N_t}$. Thus, for each realization of $\hat{\mathbf{h}}_f = \hat{\mathbf{h}}_{f,0}$, the transmitter obtains [9], [12]

$$\bar{\boldsymbol{h}} = \rho\hat{\mathbf{h}}_{f,0}, \quad \sigma_\epsilon^2 = \sigma_h^2(1 - |\rho|^2). \tag{5}$$

The transmitter treats $\bar{\boldsymbol{h}}$ as deterministic and updates its value when the next feedback becomes available.

*Case 2 (Quantized Feedback) [10], [12]:* In this case, we assume that the channel is acquired at the receiver and is quantized to $2^b$ code words $\{\mathbf{a}(j)\}_{j=1}^{2^b}$. The quantizer output is then encoded by $b$ information bits, which are fed back to the transmitter with a negligible delay over a noiseless low-speed feedback channel. We assume that the transmitter has the same code book and reconstructs the channel as $\mathbf{a}(j)$ if the index $j$ is suggested by the received $b$ bits. Although the quantization error is non-Gaussian and nonwhite in general, we assume that the quantization errors can be approximated by zero-mean and white Gaussian noise samples in order to simplify the transmitter design. With $\epsilon_Q^2$ denoting the approximate variance of the quantization error, the parameters in a1) are

$$\bar{\boldsymbol{h}} = \mathbf{a}(j), \text{ if index } j \text{ is received}, \quad \sigma_\epsilon^2 = \epsilon_Q^2. \tag{6}$$

*Case 3 (Uplink Measurements) [3]:* In TDD or FDD systems, the downlink channel can be estimated from uplink measurements.[3] This can be viewed as a form of feedback as well. The mobile can, for example, send $N_p$ pilot symbols $\{p(n)\}_{n=1}^{N_p}$ periodically for the base station to estimate the uplink channel through the received signals $\{z_\mu(n)\}_{n=1}^{N_p}$ on the $\mu$th antenna. Because the uplink and downlink channels are correlated in time or frequency, these measurements $\{z_\mu(n)\}_{\mu=1,n=1}^{N_t,N_p}$ can be also used to estimate the downlink channels through linear MMSE (a.k.a. Wiener) filtering [3]. Denote the channel estimates as $\hat{\mathbf{h}}$ and the estimation error as $\mathbf{e}$ so that $\mathbf{h} = \hat{\mathbf{h}} + \mathbf{e}$. Assume that the antennas are well separated, and thus, the channel estimation errors are uncorrelated, with zero mean and variance $\sigma_e^2$. In this case, we have

$$\bar{\boldsymbol{h}} = \hat{\mathbf{h}}, \quad \sigma_\epsilon^2 = \sigma_e^2. \tag{7}$$

The linear MMSE $\sigma_\epsilon^2$ can be calculated directly from the filter coefficients; see e.g., [3, eq. (20)].

## III. Optimal Eigen-Beamforming

Our goal in this section is to optimize the precoder $\mathbf{C}$ based on a0) and a1). Different from the optimal transmitter designs based on capacity criteria [7], [8], [12], [19], we will investigate the uncoded system (2), and our criterion will be SER. Notice

[2]We use the notation $\check{\mathbf{h}}$ to differentiate the transmitter's estimate of the channel from the receiver's channel estimate, which can be made as accurate as possible.

[3]Notice that the receiver needs to know the precoder $\mathbf{C}$ for reception. Therefore, $(\bar{\boldsymbol{h}}, \sigma_\epsilon^2)$ should be sent to the receiver before the data transmission starts. This operation is also required in Cases 1 and 2 if the feedback channel induces errors.

that error-control codes developed for single antenna transmissions (termed scalar codes in [4], [13]) can certainly be applied as outer codes in our system, and the uncoded SER will provide a good indicator for the coded bit error rate as well. In the following, we will first derive a closed-form SER expression, that will facilitate our optimal precoder design.

### A. Exact SER Expressions

For each realization of $\check{\mathbf{h}}$, the SNR at the MRC receiver output is

$$\gamma = \frac{E\left\{|\check{\mathbf{h}}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\check{\mathbf{h}}s|^2\right\}}{E\left\{\check{\mathbf{h}}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{w}\mathbf{w}^{\mathcal{H}}\mathbf{C}\check{\mathbf{h}}\right\}} = \check{\mathbf{h}}^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\check{\mathbf{h}}\frac{E_s}{N_0} \tag{8}$$

where $E_s := E\{|s|^2\}$ is the average energy of the underlying signal constellation. Since the SER depends on the constellation used, we will consider two widely used constellations: $M$-ary phase shift keying ($M$-PSK) and square $M$-ary quadrature amplitude modulation ($M$-QAM) [14] (see also, [20, Tab. 1]). Extension to $M$-ary amplitude modulation ($M$-AM) is straightforward [20].

Because $\check{\mathbf{h}}$ is random, the expected SER should be considered by averaging over all possible channel realizations. To arrive at a closed-form average SER expression, we will first simplify (8). Toward this objective, we will use the spectral decomposition of the positive semi-definite matrix $\mathbf{C}^{\mathcal{H}}\mathbf{C}$

$$\mathbf{C}^{\mathcal{H}}\mathbf{C} = \mathbf{U}_c\mathbf{D}_c\mathbf{U}_c^{\mathcal{H}}, \quad \mathbf{D}_c := \text{diag}(\delta_1, \ldots, \delta_{N_t}) \tag{9}$$

where $\mathbf{U}_c$ is unitary, and $\delta_\mu$ denotes the $\mu$th eigenvalue of $\mathbf{C}^{\mathcal{H}}\mathbf{C}$ that is non-negative: $\delta_\mu \geq 0$. Without of loss of generality, we can arrange $\delta_\mu$ in a nonincreasing order $\delta_1 \geq \cdots \geq \delta_{N_t}$ by reordering the eigenvectors in $\mathbf{U}_c$.

Using (9), we can express the SNR of (8) as

$$\begin{aligned}\gamma &= \check{\mathbf{h}}^{\mathcal{H}}\mathbf{U}_c\mathbf{D}_c\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}\frac{E_s}{N_0} = \sum_{\mu=1}^{N_t}\gamma_\mu \\ \gamma_\mu &:= \delta_\mu\left|\left[\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}\right]_\mu\right|^2\frac{E_s}{N_0}.\end{aligned} \tag{10}$$

Notice that the SNR expression (10) coincides with that of the MRC output for $N_t$ independent channels [15], with $\gamma_\mu$ denoting the $\mu$th subchannel's SNR. Since the channel coefficient on each path $|[\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}]_\mu|$ is Ricean distributed, the quality of each path is determined by two important factors. The first is the Ricean factor

$$\mathcal{K}_\mu := \frac{\left|\left[\mathbf{U}_c^{\mathcal{H}}\bar{\boldsymbol{h}}\right]_\mu\right|^2}{\sigma_\epsilon^2} \tag{11}$$

which indicates the ratio of the direct path $[\mathbf{U}_c^{\mathcal{H}}\bar{\boldsymbol{h}}]_\mu$ power over the power of the diffuse components captured by $[\mathbf{U}_c^{\mathcal{H}}\boldsymbol{\epsilon}]_\mu$. The second is the variance of each path

$$\Omega_\mu := E\left\{\left|\left[\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}\right]_\mu\right|^2\right\} = \left|\left[\mathbf{U}_c^{\mathcal{H}}\bar{\boldsymbol{h}}\right]_\mu\right|^2 + \sigma_\epsilon^2 = (1+\mathcal{K}_\mu)\sigma_\epsilon^2. \tag{12}$$

Based on $\Omega_\mu$, the expected received SNR of the $\mu$th subchannel is [cf. (10)]

$$\bar{\gamma}_\mu = E\{\gamma_\mu\} = \delta_\mu(1+\mathcal{K}_\mu)\sigma_\epsilon^2\frac{E_s}{N_0}. \tag{13}$$

Notice the simple dependence of $(\mathcal{K}_\mu, \bar{\gamma}_\mu)$ on the mean feedback parameters $(\bar{\boldsymbol{h}}, \sigma_\epsilon^2)$, and the given $\mathbf{U}_c$ and $\mathbf{D}_c$ factors of the code matrix. The SER averaged over the Ricean distributed $|[\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}]_\mu|$ will depend on $\{\mathcal{K}_\mu\}_{\mu=1}^{N_t}$ and $\{\bar{\gamma}_\mu\}_{\mu=1}^{N_t}$. In fact, [15] and [16] show that the average SER for various signal constellations can be found in closed form. For convenience, we list here the expressions for $M$-PSK and $M$-QAM:

$$P_{s,\text{PSK}} = \frac{1}{\pi}\int_0^{(M-1)\pi/M}\prod_{\mu=1}^{N_t} I_\mu(\bar{\gamma}_\mu, g_{\text{PSK}}, \theta)d\theta \tag{14}$$

$$\begin{aligned}P_{s,\text{QAM}} =& \frac{b_{\text{QAM}}}{\sqrt{M}}\int_0^{\pi/4}\prod_{\mu=1}^{N_t} I_\mu(\bar{\gamma}_\mu, g_{\text{QAM}}, \theta)d\theta \\ &+ b_{\text{QAM}}\int_{\pi/4}^{\pi/2}\prod_{\mu=1}^{N_t} I_\mu(\bar{\gamma}_\mu, g_{\text{QAM}}, \theta)d\theta\end{aligned} \tag{15}$$

where $b_{\text{QAM}} := 4(1 - 1/\sqrt{M})/\pi$, and $I_\mu(x, g, \theta)$ is the moment generating function of the probability density function (p.d.f.) of $|[\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}]_\mu|$ evaluated at $-gx/\sin^2\theta$ [15, eq. (24)]. The constellation-specific constant $g$ is given by

$$g_{\text{PSK}} = \sin^2\left(\frac{\pi}{M}\right) \text{ for } M\text{-ary PSK} \tag{16}$$

$$g_{\text{QAM}} = \frac{3}{2(M-1)} \text{ for } M\text{-ary QAM.} \tag{17}$$

For Ricean distributed $|[\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}]_\mu|$, the function $I_\mu(x, g, \theta)$ assumes the following form [15]:

$$\begin{aligned}I_\mu(x, g, \theta) = & \frac{(1+\mathcal{K}_\mu)\sin^2\theta}{(1+\mathcal{K}_\mu)\sin^2\theta + gx} \\ & \times\exp\left(-\frac{\mathcal{K}_\mu gx}{(1+\mathcal{K}_\mu)\sin^2\theta + gx}\right).\end{aligned} \tag{18}$$

For any given precoder $\mathbf{C}$, (14) or (15) provides the exact SER expected at the transmitter, based on channel mean feedback.

### B. SER-Bound Optimality Criterion

Our ultimate goal is to minimize the SER in (14), or (15), 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 rely on an upper bound on the SER to design the optimal $\mathbf{C}$ that will enable simple closed-form precoder solutions.

Based on the fact that $I_\mu(x, g, \theta)$ in (18) peaks at $\theta = \pi/2$, one can find an upper bound on the SER in (14) and (15) in a unifying form [16, p. 275], [20]

$$\begin{aligned}P_{s,\text{bound}} &= \alpha\prod_{\mu=1}^{N_t} I_\mu\left(\bar{\gamma}_\mu, g, \frac{\pi}{2}\right) \\ &= \alpha\prod_{\mu=1}^{N_t}\frac{1}{1+\delta_\mu\beta}\exp\left(\frac{-\mathcal{K}_\mu\delta_\mu\beta}{1+\delta_\mu\beta}\right)\end{aligned} \tag{19}$$

where for notational brevity, we have defined

$$\alpha := \frac{M-1}{M}, \quad \beta := g\sigma_\epsilon^2 \frac{E_s}{N_0} \tag{20}$$

with $g$ taking constellation-specific values as in (16) or (17).

We are now ready to optimize the $P_{s,\text{bound}}$ in (19) with respect to (w.r.t.) $\mathbf{C}$. Notice that $\gamma$ in (8), and subsequently the SER in (14) and (15), depend on $\mathbf{C}$ through $\mathbf{C}^{\mathcal{H}}\mathbf{C}$. Therefore, to optimize the $P_{s,\text{bound}}$ w.r.t. $\mathbf{C}$, it suffices to optimize it w.r.t. $\mathbf{U}_c$ and $\mathbf{D}_c$ that affect $\mathcal{K}_\mu$ and $\delta_\mu$ in (19). Once the optimal $\mathbf{U}_c$ and $\mathbf{D}_c$ are obtained, the precoder $\mathbf{C}$ can be expressed as

$$\mathbf{C} = \mathbf{\Phi}\mathbf{D}_c^{1/2}\mathbf{U}_c^{\mathcal{H}} \tag{21}$$

where the columns of $\mathbf{\Phi}$ are orthonormal. Notice that as long as $P \geq N_t$ and $\mathbf{\Phi}^{\mathcal{H}}\mathbf{\Phi} = \mathbf{I}_{N_t}$, all unitary matrices $\mathbf{\Phi}$ lead to the same performance. Exploiting the degrees of freedom available in $\mathbf{\Phi}$ brings other important benefits that will be discussed in Section V. For now, however, we will focus on selecting $\mathbf{U}_c$ and $\mathbf{D}_c$ that minimize the $P_{s,\text{bound}}$ in (19).

Our precoder $\mathbf{C}$ in (21) can be viewed as a beamformer. A beam toward a particular direction is formed by a set of steering weights that are nothing but coefficients multiplying the transmitted symbol $s$ in (2) per time slot. The $p$th row of $\mathbf{C}$ contains $N_t$ entries that are chips weighting the symbol $s$ across the transmit antennas during the $p$th chip period (time slot). The transmitted signal vector per time slot has correlation matrix of rank one. Our precoder $\mathbf{C}$ is thus a time-varying beamformer with the $p$th row playing the role of a beam-steering vector during the time slot $p \in [1, P]$.

It follows from (21) that the $p$th row of $\mathbf{C}$ can be decomposed as $\sum_{\mu=1}^{N_t}[\mathbf{\Phi}]_{p,\mu}\left(\sqrt{\delta_\mu}\mathbf{u}_{c,\mu}^{\mathcal{H}}\right)$, where $\mathbf{u}_{c,\mu}^{\mathcal{H}}$ is the $\mu$th row of $\mathbf{U}_c^{\mathcal{H}}$. Each beam-steering row of $\mathbf{C}$ can be projected onto any set of $N_t$ orthogonal basis vectors, and the basis may be different from slot to slot. However, the particular decomposition dictated by (21) uses as basis the $N_t$ orthogonal eigenvectors of the spectral factor $\mathbf{U}_c$ that remains invariant $\forall\, p \in [1, P]$. Likewise, the power loaded by the constants $\sqrt{\delta_\mu}$ does not depend on $p$. Hence, our beamformer $\mathbf{C}$ allocates invariant power along invariant eigen-directions (eigen-beams) that are multiplexed with different weights $[\mathbf{\Phi}]_{p,\mu}$ to yield time-varying beam-steering vectors for each time slot $p \in [1, P]$. Thus far, our transmitter is designed to send one symbol over $P$ time slots. In Section V, we will see how these time-varying multiplexing weights operating on those predefined eigen-beams can be used to transmit $K > 1$ symbols in $P$ time slots and, thus, make up for the rate loss that spreading has introduced.

The optimization of the beam directions $(\mathbf{U}_c)$, and the power loading across beams $(\mathbf{D}_c^{1/2})$, can be accomplished separately, as shown next.

### C. Optimal Beam Directions

Let us consider the eigen-decomposition of the rank-one matrix $\bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^{\mathcal{H}}$

$$\bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^{\mathcal{H}} = \mathbf{U}_h\mathbf{D}_h\mathbf{U}_h^{\mathcal{H}}, \quad \mathbf{D}_h := \text{diag}(\lambda, 0, \ldots, 0) \tag{22}$$

where $\lambda := \|\bar{\boldsymbol{h}}\|^2$, and the unitary matrix $\mathbf{U}_h = [\mathbf{u}_{h,1}, \ldots, \mathbf{u}_{h,N_t}]$ contains $N_t$ eigenvectors. Since $\bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^{\mathcal{H}}$ is a rank-one matrix, the eigenvector corresponding to the nonzero eigenvalue is $\mathbf{u}_{h,1} = \bar{\boldsymbol{h}}/\sqrt{\lambda}$. The remaining eigenvectors can be chosen arbitrarily as long as they are mutually orthogonal, as well as orthogonal to $\mathbf{u}_{h,1}$. Given a1), the eigen-decomposition in (22) determines the eigen-decomposition for the correlation matrix $\mathbf{R}_{\check{h}\check{h}} := E\{\check{\mathbf{h}}\check{\mathbf{h}}^{\mathcal{H}}\}$ as [cf. (6)]

$$\mathbf{R}_{\check{h}\check{h}} = \mathbf{U}_h\left(\mathbf{D}_h + \sigma_\epsilon^2\mathbf{I}_{N_t}\right)\mathbf{U}_h^{\mathcal{H}}. \tag{23}$$

Based on (23), we have the following result for the optimal beamformer $\mathbf{U}_c$.

*Proposition 1:* Under a0) and a1), the optimal beam directions minimizing the $P_{s,\text{bound}}$ in (19) are those along the eigenvectors of the channel correlation matrix $\mathbf{R}_{\check{h}\check{h}}$, as perceived by the transmitter, i.e., $\mathbf{U}_c = \mathbf{U}_h$.

*Proof:* The $P_{s,\text{bound}}$ in (19) can be rewritten as

$$\begin{aligned}P_{s,\text{bound}} =& \alpha|\mathbf{I}_{N_t} + \beta\mathbf{D}_c|^{-1} \\ &\times \exp\left(-\text{tr}\left\{\left[\frac{1}{\sigma_\epsilon^2}\mathbf{U}_c^{\mathcal{H}}\bar{\boldsymbol{h}}\,\bar{\boldsymbol{h}}^{\mathcal{H}}\mathbf{U}_c\right]\right.\right. \\ &\left.\left.\times\left[(\beta\mathbf{D}_c)(\mathbf{I}_{N_t}+\beta\mathbf{D}_c)^{-1}\right]\right\}\right) \\ =& \alpha|\mathbf{I}_{N_t} + \beta\mathbf{D}_c|^{-1}\exp\left(-\frac{1}{\sigma_\epsilon^2}\text{tr}\left\{\bar{\boldsymbol{h}}\,\bar{\boldsymbol{h}}^{\mathcal{H}}\right\}\right) \\ &\times \exp\left(\frac{1}{\sigma_\epsilon^2}\text{tr}\left\{\mathbf{U}_c^{\mathcal{H}}\bar{\boldsymbol{h}}\,\bar{\boldsymbol{h}}^{\mathcal{H}}\mathbf{U}_c(\mathbf{I}_{N_t}+\beta\mathbf{D}_c)^{-1}\right\}\right).\end{aligned} \tag{24}$$

For each fixed $\mathbf{D}_c$, we seek $\mathbf{U}_c$ that minimizes

$$\begin{aligned}&\text{tr}\left\{\mathbf{U}_c^{\mathcal{H}}\bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^{\mathcal{H}}\mathbf{U}_c(\mathbf{I}_{N_t}+\beta\mathbf{D}_c)^{-1}\right\} \\ &\quad = \text{tr}\left\{\mathbf{U}_c^{\mathcal{H}}\mathbf{U}_h\mathbf{D}_h\mathbf{U}_h^{\mathcal{H}}\mathbf{U}_c(\mathbf{I}_{N_t}+\beta\mathbf{D}_c)^{-1}\right\}.\end{aligned} \tag{25}$$

To proceed, we will need the following lemma, which is proven in [8, eq (19)].

*Lemma 1:* Suppose $\mathbf{\Lambda}_Q$ is an $N \times N$ positive semidefinite diagonal matrix with diagonal entries arranged in nonincreasing order, i.e., $[\mathbf{\Lambda}_Q]_{1,1} \geq \cdots \geq [\mathbf{\Lambda}_Q]_{N,N}$. Suppose $\mathbf{\Lambda}_A$ is an $N\times N$ positive definite diagonal matrix with diagonal entries arranged in nonincreasing order. For arbitrary unitary matrix $\mathbf{U}_Q$, define $\mathbf{Q} = \mathbf{U}_Q\mathbf{\Lambda}_Q\mathbf{U}_Q^{\mathcal{H}}$. It holds that

$$\begin{aligned}\text{tr}\left\{\mathbf{Q}\mathbf{\Lambda}_A^{-1}\right\} =& \text{tr}\left\{\mathbf{\Lambda}_A^{-1/2}\mathbf{Q}\mathbf{\Lambda}_A^{-1/2}\right\} \\ \geq& \text{tr}\left\{\mathbf{\Lambda}_A^{-1/2}\mathbf{\Lambda}_Q\mathbf{\Lambda}_A^{-1/2}\right\} \\ =& \text{tr}\left\{\mathbf{\Lambda}_Q\mathbf{\Lambda}_A^{-1}\right\}\end{aligned} \tag{26}$$

where the equality holds when $\mathbf{U}_Q = \mathbf{I}_N$.

Recall that for each fixed power allocation, we can arrange the diagonal entries of $\mathbf{D}_c$ in a nonincreasing order by reordering the eigenvectors in $\mathbf{U}_c$. Applying Lemma 1 to (25), we find the optimal beam directions $\mathbf{U}_c = \mathbf{U}_h$.

As suggested by Proposition 1, the optimal precoder turns out to be a beamformer multiplexing orthogonal beams that are pointing to directions along the eigenvectors of the channel correlation matrix as perceived by the transmitter, thus, the name eigen-beamformer. We next proceed to optimize power allocation across the eigen-beams.

### D. Power Loading Across Beams

With the optimal $\mathbf{U}_c = \mathbf{U}_h$, we can rewrite (24) as

$$P_{s,\text{bound}} = \alpha |\mathbf{I}_{N_t} + \beta \mathbf{D}_c|^{-1} \exp\left(-\frac{1}{\sigma_\epsilon^2}\text{tr}\{\mathbf{D}_h\}\right) \times \exp\left(\frac{1}{\sigma_\epsilon^2}\text{tr}\{\mathbf{D}_h(\mathbf{I}_{N_t}+\beta\mathbf{D}_c)^{-1}\}\right). \quad (27)$$

Without any constraint, minimizing $P_{s,\text{bound}}$ with respect to $\mathbf{D}_c$ leads to the trivial solution that requires infinite power to be transmitted ($\text{tr}\{\mathbf{C}^{\mathcal{H}}\mathbf{C}\} = \infty$). A practical constraint that takes into account limited budget resources is the average transmitted power, which is expressed as $\mathcal{P}_0 = E\{\text{tr}\{(s\mathbf{C})^{\mathcal{H}}(s\mathbf{C})\}\} = E_s \text{tr}\{\mathbf{C}^{\mathcal{H}}\mathbf{C}\}$. Without loss of generality, we assume that $\mathcal{P}_0 = E_s$, and $\text{tr}\{\mathbf{C}^{\mathcal{H}}\mathbf{C}\} = \text{tr}\{\mathbf{D}_c\} = \sum_{\mu=1}^{N_t} \delta_\mu = 1$, i.e., the total transmit-power per symbol is $E_s$.

Since $\ln(\cdot)$ is a monotonically increasing function, and $\mathbf{D}_h$ in (22) has only one nonzero element $\lambda$, our equivalent constrained optimization problem is

$$\min_{\mathbf{D}_c} \mathcal{E}_1, \text{ where } \mathcal{E}_1 := \ln P_{s,\text{bound}} = \ln\alpha - \frac{\lambda}{\sigma_\epsilon^2} - \ln|\mathbf{I}_{N_t} + \beta\mathbf{D}_c| + \frac{\lambda}{\sigma_\epsilon^2(1+\beta\delta_1)}$$
$$\text{subject to } \mathcal{C} := \sum_{\mu=1}^{N_t} \delta_\mu - 1 = 0, \text{ and } \delta_\mu \geq 0. \quad (28)$$

We adopt the special notation $[x]_+ := \max(x, 0)$. Differentiating the Lagrangian $\mathcal{E}_1 + \zeta\mathcal{C}$ with respect to $\delta_\mu$, where $\zeta$ denotes the Lagrange multiplier, and equating it to zero, we obtain

$$\delta_2 = \cdots = \delta_{N_t} = \left[\frac{1}{\zeta} - \frac{1}{\beta}\right]_+. \quad (29)$$

Recalling from (20) that $\beta > 0$, and from (9) that $\delta_\mu \geq 0, \forall\, \mu$, we deduce from (29) that $\zeta > 0$. Since $\zeta > 0$, differentiating $\mathcal{E}_1 + \zeta\mathcal{C}$ with respect to $\delta_1$ yields

$$\delta_1 = \left[\frac{1+\sqrt{1+\frac{4\lambda\zeta}{\beta\sigma_\epsilon^2}}}{2\zeta} - \frac{1}{\beta}\right]_+. \quad (30)$$

Comparing (30) with (29) reveals that $\delta_1 \geq \delta_2$. Suppose there exists a $\zeta$ such that $\delta_2 = \cdots = \delta_{N_t} > 0$, and $\delta_1 = 1 - (N_t - 1)\delta_2 > 0$. This $\zeta$ should satisfy $\beta > \zeta > N_t/(1+N_t/\beta) := \zeta_0$. Substituting (29) and (30) into the power constraint $\mathcal{C}$ of (28), we find that $\zeta$ should also satisfy the quadratic equation

$$a\zeta^2 - b\zeta + c = 0 \quad (31)$$

where

$$a := \left(1+\frac{N_t}{\beta}\right)^2, \quad c := N_t(N_t-1)$$
$$b := \left[\frac{\lambda}{\beta\sigma_\epsilon^2} + \left(1+\frac{N_t}{\beta}\right)(2N_t - 1)\right]. \quad (32)$$

If $\Delta := b^2 - 4ac > 0$, then (31) has two possible roots: $\zeta_1 = (b+\sqrt{\Delta})/(2a)$ and $\zeta_2 = (b-\sqrt{\Delta})/(2a)$. Since $\zeta_1 > \zeta_2 > 0$ and $\zeta_1\zeta_2 = c/a < \zeta_0^2$, we infer that $\zeta_2 < \zeta_0$, and thus, $\zeta_2$ can be discarded. Our final solution can then be expressed as

$$\delta_2 = \cdots = \delta_{N_t} = \left[\frac{2a}{b+\sqrt{b^2-4ac}} - \frac{1}{\beta}\right]_+$$
$$\delta_1 = 1 - (N_t - 1)\delta_2. \quad (33)$$

Equation (33) provides the optimal loading in *closed form*. The possibility of this form obtained by solving a quadratic equation, when $N_r = 1$, was also pointed out in [11]. To facilitate generalization to multiple receive antennas, we derive next a simpler (albeit suboptimum) loading solution.

It is well known that one can approximate well a Ricean distribution with factor $\mathcal{K}_\mu$ by a Nakagami-$m$ distribution having parameter $m_\mu$ when $m_\mu$ satisfies [16, p. 23]:

$$m_\mu = \frac{(1+\mathcal{K}_\mu)^2}{1+2\mathcal{K}_\mu}. \quad (34)$$

With the optimal beamformer $\mathbf{U}_c = \mathbf{U}_h$, we have

$$\mathcal{K}_1 = \frac{\lambda}{\sigma_\epsilon^2}, \quad \mathcal{K}_2 = \cdots = \mathcal{K}_{N_t} = 0 \quad (35)$$

which correspond [via (34)] to Nakagami parameters

$$m_1 = \frac{(\sigma_\epsilon^2+\lambda)^2}{\sigma_\epsilon^2(\sigma_\epsilon^2+2\lambda)}, \quad m_2 = \cdots = m_{N_t} = 1. \quad (36)$$

Note that a Ricean distribution with $\mathcal{K}_\mu = 0$ coincides with a Nakagami distribution having $m_\mu = 1$, and both reduce to a Rayleigh distribution. Fortunately, the moment-generating function for the Nakagami distribution has a simpler form [15], [16]:

$$\tilde{I}_\mu(x, g, \theta) = \left(1 + \frac{gx}{m_\mu \sin^2\theta}\right)^{-m_\mu}. \quad (37)$$

Approximating the Ricean distribution by the Nakagami distribution, we obtain from (19)

$$P_{s,\text{bound}} \approx \tilde{P}_{s,\text{bound}} = \alpha \prod_{\mu=1}^{N_t} \tilde{I}_\mu\left(\bar{\gamma}_\mu, g, \frac{\pi}{2}\right) = \alpha \prod_{\mu=1}^{N_t}\left(1+\frac{\delta_\mu(1+\mathcal{K}_\mu)\beta}{m_\mu}\right)^{-m_\mu}. \quad (38)$$

Taking $\ln(\cdot)$ on $\tilde{P}_{s,\text{bound}}$, our equivalent constrained optimization problem becomes

$$\min_{\mathbf{D}_c} \mathcal{E}_2, \text{ where } \mathcal{E}_2 := \ln\tilde{P}_{s,\text{bound}} = \ln\alpha - \sum_{\mu=1}^{N_t} m_\mu \ln\left(1+\frac{\delta_\mu(1+\mathcal{K}_\mu)\beta}{m_\mu}\right)$$
$$\text{subject to } \mathcal{C} := \sum_{\mu=1}^{N_t}\lambda_\mu - 1 = 0, \text{ and } \lambda_\mu \geq 0. \quad (39)$$

Differentiating the Lagrangian $\mathcal{E}_2 + \zeta\mathcal{C}$ with respect to $\delta_\mu$ and equating it to zero, we obtain

$$\delta_\mu = \left[\frac{m_\mu}{\zeta} - \frac{m_\mu}{(1+\mathcal{K}_\mu)\beta}\right]_+. \quad (40)$$

For $\delta_\mu > 0$, (40) requires $\zeta > 0$, which implies that $\delta_\mu$ is a nondecreasing function of $\mathcal{K}_\mu$. We thus have $\delta_1 \geq \delta_2 = \cdots = \delta_{N_t}$. Suppose there exists a $\zeta$ such that $\delta_2 > 0$. Plugging (40) into $\mathcal{C}$, we obtain $\zeta$, which in turn determines $\delta_2$ as

$$\delta_2^o = \frac{\sigma_\epsilon^2\left(\sigma_\epsilon^2 + 2\lambda\right)}{N_t\sigma_\epsilon^2\left(\sigma_\epsilon^2 + 2\lambda\right) + \lambda^2} \times \left[1 + \frac{1}{\beta}\left(N_t - \frac{\lambda}{\sigma_\epsilon^2 + 2\lambda}\right)\right] - \frac{1}{\beta}. \quad (41)$$

Positive $\delta_2^o$ requires the power budget $E_s$ to satisfy

$$\frac{E_s}{N_0} > \frac{\lambda}{g\sigma_\epsilon^4}\left(\frac{\sigma_\epsilon^2 + \lambda}{\sigma_\epsilon^2 + 2\lambda}\right) := \gamma_{\text{th}}. \quad (42)$$

Therefore, our simplified closed-form solution can be summarized as

$$\delta_2 = \cdots = \delta_{N_t} = \begin{cases} \delta_2^o, & \frac{E_s}{N_0} > \gamma_{\text{th}} \\ 0, & \frac{E_s}{N_0} \leq \gamma_{\text{th}} \end{cases}$$
$$\delta_1 = 1 - \delta_2(N_t - 1). \quad (43)$$

Compared with (33), the solution (43) provides additional insights. At low SNR ($E_s/N_0 \leq \gamma_{\text{th}}$), only one beam along the strongest direction $\mathbf{u}_{h,1} = \bar{\boldsymbol{h}}/\sqrt{\lambda}$ is used. When the SNR is above the threshold $\gamma_{\text{th}}$, all $N_t$ beams are used, with more power put on the strongest direction, and with the remaining power evenly distributed to the remaining $N_t - 1$ orthogonal beams. Interestingly, this observation is in agreement with [19], even though the latter relied on a capacity criterion.

Having specified the optimal $\delta_\mu$, we have found the optimal $\mathbf{D}_c$. Optimal $\mathbf{D}_c$ and $\mathbf{U}_c$ determine the optimal $\mathbf{C}$ in (21). We summarize our results so far in the following.

*Theorem 1:* Suppose a0) and a1) hold true. For the optimum receive-filter $\mathbf{g}_{\text{opt}}$ given by (3), the optimum precoding matrix is $\mathbf{C}_{\text{opt}} = \mathbf{\Phi}\mathbf{D}_c^{1/2}\mathbf{U}_h^{\mathcal{H}}$, where $\mathbf{U}_h$ and $\mathbf{D}_c$ are formed as in (9), (23), (33), or (43), with $\mathbf{\Phi}$ an arbitrary orthonormal $P \times N_t$ matrix. Optimality in $\mathbf{g}_{\text{opt}}$ refers to maximum-SNR, whereas optimality in $\mathbf{C}_{\text{opt}}$ pertains to minimizing an upper bound on the average symbol error rate.

### *E. Comparisons With Designs Maximizing the Average SNR*

Different from the SER bound used herein, [3] and [12] designed optimal transmitters to maximize the expected SNR at the receiver. With this criterion, our optimization problem becomes [cf. (8)]

$$\max_{\mathbf{C}} E\{\gamma\} \Leftrightarrow \max_{\mathbf{C}} \text{tr}\{\mathbf{R}_{\check{h}\check{h}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\}\frac{E_s}{N_0}$$
$$\text{subject to } \mathcal{C} := \text{tr}\{\mathbf{D}_c\} - 1 = 0. \quad (44)$$

Using Lemma 1, the optimal $\mathbf{U}_c$ can be easily found as $\mathbf{U}_c = \mathbf{U}_h$. It can be readily verified that the optimal $\mathbf{D}_c = \text{diag}(1, 0, 0, \ldots, 0)$. Therefore, the optimal solution reduces to an invariant beamformer (beam-steering vector) pointing to one direction along the channel mean, *no matter how reliable the channel feedback is*.

To compare these two criteria, let us first recall that maximizing the average SNR is not equivalent to minimizing the SER. A simple but illustrating example is to construct two systems as

$$x_1 = h_1 s + w, \quad x_2 = h_2 s + w. \quad (45)$$

Selecting $h_1 = 1$ corresponds to an additive white Gaussian noise (AWGN) channel. On the other hand, choosing $h_2 \sim \mathcal{CN}(0,1)$ corresponds to a flat fading channel. Obviously, both systems have the same average receiver SNR but dramatically different SER. The reason is that the average SER depends not only on the average SNR but also on higher SNR moments such as the SNR variance. Because the SER is dominated by worst errors, the SNR variance should be small to ensure that worst cases are as rare as possible. In the extreme case with perfect CSI, the SNR is deterministic. Only in this special case, where CSI is perfectly known at the transmitter, minimizing the SER is equivalent to maximizing the average SNR, and the optimal solution deploys only one, namely, the strongest eigen-beam.

In a nutshell, our proposed optimal eigen-beamformer relies on a SER-bound criterion. It does not achieve the maximum expected SNR but strikes the best compromise between the mean and the variance of the received SNR, which is a feature that we will also confirm by simulations.

## IV. Multiple Receive Antennas

In this section, we extend our results to multiple receive antennas. For simplicity, we assume that the channel mean for each receive antenna is known, whereas the variance of the channel error vectors is the same for all receive antennas. Formally, we adopt the following.

**a2)** *With $\check{\mathbf{h}}_\nu$ denoting the estimated channel at the transmitter (based on partial CSI) corresponding to the $\nu$th receive antenna, it holds that $\check{\mathbf{h}}_\nu \sim \mathcal{CN}(\bar{\boldsymbol{h}}_\nu, \sigma_\epsilon^2\mathbf{I}_{N_t})$, $\forall\, \nu \in [1, N_r]$, where $N_r$ is the number of receive antennas.*

We collect $\{\check{\mathbf{h}}_\nu\}_{\nu=1}^{N_r}$ into a matrix $\check{\mathbf{H}} := [\check{\mathbf{h}}_1, \ldots, \check{\mathbf{h}}_{N_r}]$ and likewise for the channel mean vectors $\bar{\boldsymbol{H}} := [\bar{\boldsymbol{h}}_1, \ldots, \bar{\boldsymbol{h}}_{N_r}]$. We can now relate $\check{\mathbf{H}}$ with $\bar{\boldsymbol{H}}$ via

$$\check{\mathbf{H}} = \bar{\boldsymbol{H}} + \mathbf{\Xi} \quad (46)$$

where $\mathbf{\Xi}$ is an $N_t \times N_r$ matrix with independent Gaussian entries, having zero mean and variance $\sigma_\epsilon^2$. Similar to (22), we decompose $\bar{\boldsymbol{H}}\bar{\boldsymbol{H}}^{\mathcal{H}}$ as

$$\bar{\boldsymbol{H}}\bar{\boldsymbol{H}}^{\mathcal{H}} = \mathbf{U}_H\mathbf{D}_H\mathbf{U}_H^{\mathcal{H}}, \quad \mathbf{D}_H := \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{N_t}) \quad (47)$$

where, without loss of generality, the eigenvalues are arranged in a nonincreasing order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N_t}$. Under a2), the eigen-decomposition in (47) determines the eigen-decomposition of the channel correlation matrix $\mathbf{R}_{\check{H}\check{H}} := E\{\check{\mathbf{H}}\check{\mathbf{H}}^{\mathcal{H}}\}$ as

$$\mathbf{R}_{\check{H}\check{H}} = \mathbf{U}_H\left(\mathbf{D}_H + N_r\sigma_\epsilon^2\mathbf{I}_{N_t}\right)\mathbf{U}_H^{\mathcal{H}}. \quad (48)$$

The received signal (2) at the $\nu$th antenna is now $\mathbf{x}_\nu = \mathbf{C}\check{\mathbf{h}}_\nu s + \mathbf{w}_\nu$. Let us collect the received vectors $\{\mathbf{x}_\nu\}_{\nu=1}^{N_r}$ into the $P \times N_r$ matrix $\mathbf{X} := [\mathbf{x}_1, \ldots, \mathbf{x}_{N_r}]$ and the MRC receivers $\{\mathbf{g}_\nu := \mathbf{C}\check{\mathbf{h}}_\nu\}_{\nu=1}^{N_r}$ into the $P \times N_r$ matrix $\mathbf{G} := [\mathbf{g}_1, \ldots, \mathbf{g}_{N_r}]$. The MRC output and the corresponding SNR are

$$\begin{aligned} y =& \text{tr}\{\mathbf{G}^{\mathcal{H}}\mathbf{X}\} = \sum_{\nu=1}^{N_r} \check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{x}_\nu \\ \gamma =& \frac{E_s}{N_0}\sum_{\nu=1}^{N_r} \check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}^{\mathcal{H}}\mathbf{C}\check{\mathbf{h}}_\nu \end{aligned} \tag{49}$$

where the latter includes (8) as a special case corresponding to $N_r = 1$. Following the same steps used to derive (10) and based on a2), we can decompose $\gamma$ as $\gamma = \sum_{\nu=1}^{N_r}\sum_{\mu=1}^{N_t}\gamma_{\mu,\nu}$ with $\gamma_{\mu,\nu} := \delta_\mu|[\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}_\nu]_\mu|^2 E_s/N_0$.

Mimicking the derivation of (19) and (24) and denoting with $I_{\mu,\nu}(g,x,\theta)$ the moment-generating function for the path $|[\mathbf{U}_c^{\mathcal{H}}\check{\mathbf{h}}_\nu]_\mu|$, we obtain the upper bound on the SER as

$$\begin{aligned} P_{s,\text{bound}} =& \alpha \prod_{\mu=1}^{N_t}\prod_{\nu=1}^{N_r} I_{\mu,\nu}\left(\bar{\gamma}_{\mu,\nu}, g, \frac{\pi}{2}\right) \\ =& \alpha|\mathbf{I}_{N_t} + \beta\mathbf{D}_c|^{-N_r} \exp\left(-\frac{1}{\sigma_\epsilon^2}\sum_{\nu=1}^{N_r}\text{tr}\left\{\bar{\boldsymbol{h}}_\nu\bar{\boldsymbol{h}}_\nu^{\mathcal{H}}\right\}\right) \\ &\times \exp\left(\frac{1}{\sigma_\epsilon^2}\sum_{\nu=1}^{N_r}\text{tr}\left\{\mathbf{U}_c^{\mathcal{H}}\bar{\boldsymbol{h}}_\nu\bar{\boldsymbol{h}}_\nu^{\mathcal{H}}\mathbf{U}_c(\mathbf{I}_{N_t}+\beta\mathbf{D}_c)^{-1}\right\}\right) \\ =& \alpha|\mathbf{I}_{N_t} + \beta\mathbf{D}_c|^{-N_r} \exp\left(-\frac{1}{\sigma_\epsilon^2}\text{tr}\{\mathbf{D}_H\}\right) \\ &\times \exp\left(\frac{1}{\sigma_\epsilon^2}\text{tr}\left\{\mathbf{U}_c^{\mathcal{H}}\overline{\boldsymbol{H}}\overline{\boldsymbol{H}}^{\mathcal{H}}\mathbf{U}_c(\mathbf{I}_{N_t}+\beta\mathbf{D}_c)^{-1}\right\}\right). \end{aligned} \tag{50}$$

Mimicking the proof of Proposition 1, we establish the following.

*Proposition 2:* Under a0) and a2), the optimal beam directions are along the eigenvectors of the channel correlation matrix $\mathbf{R}_{\check{H}\check{H}}$, i.e., $\mathbf{U}_c = \mathbf{U}_H$.

With the optimal $\mathbf{U}_c = \mathbf{U}_H$, we can rewrite (50) as

$$\begin{aligned} P_{s,\text{bound}} =& \alpha\left[|\mathbf{I}_{N_t}+\beta\mathbf{D}_c|^{-1}\exp\left(-\frac{1}{\sigma_\epsilon^2}\text{tr}\left\{\frac{\mathbf{D}_H}{N_r}\right\}\right)\right. \\ &\left.\times\exp\left(\frac{1}{\sigma_\epsilon^2}\text{tr}\left\{\frac{\mathbf{D}_H}{N_r}(\mathbf{I}_{N_t}+\beta\mathbf{D}_c)^{-1}\right\}\right)\right]^{N_r} \\ :=& \alpha(P_{b1})^{N_r}. \end{aligned} \tag{51}$$

Equation (51) implies that minimizing $P_{s,\text{bound}}$ is equivalent to minimizing $P_{b1}$. Notice that $\alpha P_{b1}$ can be obtained from (27) by replacing $\mathbf{D}_h$ with the matrix $\tilde{\mathbf{D}}_h := \mathbf{D}_H/N_r$. Thus, the optimization problem with multiple receive antennas can be reduced to the one with a single receive antenna. The difference is that unlike $\mathbf{D}_h$, the matrix $\tilde{\mathbf{D}}_h$ has more than one nonzero entries. We define the set of Ricean factors as

$$\tilde{\mathcal{K}}_\mu = \frac{\lambda_\mu}{N_r\sigma_\epsilon^2}, \quad \forall\, \mu \in [1, N_t]. \tag{52}$$

Because $\tilde{\mathbf{D}}_h$ has $\bar{n} := \min(N_t, N_r)$ nonzero entries, optimization based on Ricean factors turns out to be complex. A polynomial equation of order $2^{\bar{n}}$ is involved in general, which includes (31) as a special case when $\bar{n} = N_r = 1$. A one-dimensional (single-parameter) search has been proposed in [11] using numerical optimization.

By approximating Ricean distributions with Nakagami distributions, we show again that *simple closed-form solutions* are possible. Define the set of Nakagami parameters $\{\tilde{m}_\mu\}_{\mu=1}^{N_t}$ from $\{\tilde{\mathcal{K}}_\mu\}_{\mu=1}^{N_t}$ using (34). Formulating the problem similar to (39), we obtain the optimal loading as [cf. (40)]

$$\delta_\mu = \left[\frac{\tilde{m}_\mu}{\zeta} - \frac{\tilde{m}_\mu}{(1+\tilde{\mathcal{K}}_\mu)\beta}\right]_+ \tag{53}$$

which also implies that $\delta_1 \geq \cdots \geq \delta_{N_t}$, as discussed before. Since more power is allocated to stronger subchannels, this power allocation also obeys a "spatial water-filling" principle [7], [19]. If there are $\bar{N}_t$ nonzero loadings, we have $\delta_\mu = 0$, for $\mu \geq \bar{N}_t + 1$. For each $\mu = 1, \ldots, \bar{N}_t$, we solve $\zeta$ using the power constraint to obtain

$$\delta_\mu = \frac{\tilde{m}_\mu}{\sum_{l=1}^{\bar{N}_t}\tilde{m}_l}\left(1 + \sum_{l=1}^{\bar{N}_t}\frac{\tilde{m}_l}{(1+\tilde{\mathcal{K}}_l)\beta}\right) - \frac{\tilde{m}_\mu}{(1+\tilde{\mathcal{K}}_\mu)\beta}. \tag{54}$$

To ensure that $\delta_{\bar{N}_t} > 0$, the transmitted power should adhere to

$$\frac{E_s}{N_0} > \frac{1}{g\sigma_\epsilon^2}\sum_{l=1}^{\bar{N}_t-1}\frac{(\lambda_l - \lambda_{\bar{N}_t})(N_r\sigma_\epsilon^2 + \lambda_l)}{(N_r\sigma_\epsilon^2 + \lambda_{\bar{N}_t})(N_r\sigma_\epsilon^2 + 2\lambda_l)} := \bar{\gamma}_{\text{th},\bar{N}_t}. \tag{55}$$

From (54) and (55), we describe our *practical power loading algorithm* in the following steps.

Step 1) For $r = 1, \ldots, N_t$, calculate $\bar{\gamma}_{\text{th},r}$ from (55), based only on the first $r$ eigenvalues of $\mathbf{D}_H$.

Step 2) With the given power budget $E_s$ ensuring that $E_s/N_0$ falls in the interval $[\bar{\gamma}_{\text{th},r}, \bar{\gamma}_{r+1}]$, set $\delta_{r+1}, \ldots, \delta_{N_t} = 0$, and obtain $\delta_1, \ldots, \delta_r$ according to (54) with $\bar{N}_t = r$.

We summarize our results for multiple receive antennas in the following theorem.

*Theorem 2:* Suppose a0) and a2) hold true. With MRC receivers, the optimum precoding matrix $\mathbf{C}_{\text{opt}} = \mathbf{\Phi}\mathbf{D}_c^{1/2}\mathbf{U}_H^{\mathcal{H}}$ has $\mathbf{U}_H$ and $\mathbf{D}_c$ formed as in (9), (48), and (53), where $\mathbf{\Phi}$ is an arbitrary orthonormal $P \times N_t$ matrix. Optimality in $\mathbf{C}_{\text{opt}}$ pertains to minimizing an upper bound on the average SER.

When the system operates at a prescribed power $E_s/N_0 \in [\bar{\gamma}_{\text{th},r}, \bar{\gamma}_{\text{th},r+1}]$, it is clear that only $r = \text{rank}(\mathbf{D}_c)$ eigen-beams are used. Therefore, the transmit diversity order achieved is $r$. Recalling that the full diversity order with $N_t$ transmit and $N_r$ receive antennas is $N_tN_r$, we infer that full transmit diversity is achieved when $r = N_t$. Based on (55), one can easily determine what diversity order should be used to achieve the best performance for a given power budget $E_s$ under a0) and a2). Specifically, we deduce the following from Theorem 2 and (55).

*Corollary 1:* The optimal transmit diversity order is $r$ when $E_s/N_0$ falls in the interval of $[\bar{\gamma}_{\text{th},r}, \bar{\gamma}_{\text{th},r+1}]$, with $\bar{\gamma}_{\text{th},r}$ defined as in (55).

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, the spreading factor must satisfy $P \geq N_t$; otherwise, the matrix $\mathbf{C}^{\mathcal{H}}\mathbf{C}$ loses rank and is forced to have zero eigenvalues. On the other hand, the choice $P > N_t$ does not gain anything in terms of optimizing (50) relative to the minimum choice $P = N_t$. It is thus desirable to choose $P$ as small as possible in order to minimize bandwidth expansion or, equivalently, increase the transmission rate. When the desired transmit 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 enables one to achieve rate $(1/r)$ symbols/s/Hz for a transmit diversity transmission of order $r$.

Alternatively, 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 *a fortiori* setting $\delta_{d+1}, \ldots, \delta_{N_t} = 0$. Optimal power loading can then be applied to the remaining $d$ beams. We will term this scheme (with $\mathbf{C}$ and $\mathbf{\Phi}$ chosen beforehand to be $d \times N_t$) a $d$-directional $(dD)$ eigen-beamformer. As a consequence of Theorem 2 and Corollary 1, we then have the following.

*Corollary 2:* With $d < N_t$, the $d$-directional ($\mathbf{C}$: $d \times N_t$) eigen-beamformer achieves the same average SER performance as an $N_t$-directional ($\mathbf{C}$: $P \times N_t$ with $P \geq N_t$) eigen-beamformer when $E_s/N_0 < \bar{\gamma}_{\text{th},d+1}$.

Two particularly interesting special cases arise from Corollary 2. The first one is the conventional one-directional eigen-beamforming with $d = 1$, which was pursued in [3], [7], [8], [12], and [19]. As asserted by Corollary 1, the one-directional eigen-beamformer will be optimal when $E_s/N_0 < \bar{\gamma}_{\text{th},2}$. However, a more attractive case is the two-directional eigen-beamforming, which corresponds to $d = 2$. This two-directional eigen-beamformer is optimal when $E_s/N_0 < \bar{\gamma}_{\text{th},3}$. Notice that the optimality condition for two-directional eigen-beamforming is less restrictive than that for one-directional beamforming since $\bar{\gamma}_{\text{th},3} > \bar{\gamma}_{\text{th},2}$. For example, in the special case of $N_r = 2$, the matrix $\mathbf{D}_H$ has at most two nonzero eigenvalues: $\lambda_1$ and $\lambda_2$. We can thus verify from (55) that

$$\bar{\gamma}_{\text{th},2} = \frac{1}{g\sigma_\epsilon^2} \frac{(\lambda_1 - \lambda_2)\left(2\sigma_\epsilon^2 + \lambda_1\right)}{(2\sigma_\epsilon^2 + \lambda_2)(2\sigma_\epsilon^2 + 2\lambda_1)}$$
$$\bar{\gamma}_{\text{th},3} = \frac{1}{g\sigma_\epsilon^2}\left(\frac{\lambda_1\left(2\sigma_\epsilon^2 + \lambda_1\right)}{2\sigma_\epsilon^2\left(2\sigma_\epsilon^2 + 2\lambda_1\right)} + \frac{\lambda_2\left(2\sigma_\epsilon^2 + \lambda_2\right)}{2\sigma_\epsilon^2\left(2\sigma_\epsilon^2 + 2\lambda_2\right)}\right). \quad (56)$$

When the feedback suggests two good directions so that $\lambda_1 \approx \lambda_2$, it is evident that $\bar{\gamma}_{\text{th},2}$ can be much smaller than $\bar{\gamma}_{\text{th},3}$. Compared with the one-directional beamformer, the two-directional eigen-beamformer is optimal over a larger $E_s/N_0$ range or, equivalently, over a larger set of fading channels for a given $E_s/N_0$. Notice that rate loss occurs when $d > 1$. However, as we will see in Section V-B, two-directional eigen-beamforming achieves the same rate as one-directional beamforming and subsumes the latter as a special case.

## V. Eigen-Beamforming and STBC

In the system model (2), we transmit only one symbol every $P$ time slots (chip-periods), which amounts to a spread-spectrum scheme. As we mentioned in Section II-A, this is useful for "power-limited" (e.g., military) communication systems, where bandwidth is not at a premium, but low transmission power is desired [2], [6]. For "bandwidth-limited" systems on the other hand, it is possible to mitigate the rate loss by sending $K > 1$ symbols $s_1, \ldots, s_K$ simultaneously. The rate will then increase to $(K/P)$ symbols/s/Hz. Notice that our single symbol transmission in (2) achieves the best possible performance, which serves as an upper bound on the performance of multiplexed symbol transmissions. Indeed, when detecting one particular symbol $s_k$, the best scenario happens when all other symbols have been detected correctly, and their effect on $s_k$ has been perfectly cancelled.

### A. Increasing the Rate Without Compromising Performance

Our objective is to pursue optimal multiplexing that increases the data rate without compromising the performance. Certainly, this would require a symbol separator at the receiver that does not incur optimality loss, but let us suppose temporarily that such a separator indeed exists, and each symbol is essentially going through separate channels identical to those we dealt with in Section IV. The optimum precoder $\mathbf{C}_k$ for $s_k$ will then be

$$\mathbf{C}_k = \mathbf{\Phi}_k \mathbf{D}_c^{1/2} \mathbf{U}_H^{\mathcal{H}}, \quad k = 1, 2, \ldots, K \quad (57)$$

where the optimal $\mathbf{D}_c$ is determined as in (53). Because the factor $\mathbf{D}_c^{1/2}\mathbf{U}_H^{\mathcal{H}}$ in (57) is common $\forall\ k$, designing separable precoders is equivalent to selecting separable $\{\mathbf{\Phi}_k\}_{k=1}^K$ matrices. Fortunately, this degree of freedom can be afforded by our design because so far, our $\mathbf{\Phi}_k$s are only required to have orthonormal columns.

Specifically, we can select $\mathbf{\Phi}_k$ as an orthogonal STBC matrix [1], [5], [17]. With this choice, our transmitter implements a combination of STBC followed by optimal eigen-beamforming. Based on *covariance feedback*, we also combined optimal eigen-beamforming with STBC in [20], where $\mathbf{\Phi}_k$ designs for real and complex constellations are detailed separately. Here, we focus on complex constellations for brevity; the real constellations can be treated similar to [20].

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 [5], [17].

*Definition 1:* For complex symbols $\{s_k = s_k^R + js_k^I\}_{k=1}^K$ 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

$$\mathcal{O}_{N_t} = \sum_{k=1}^{K} \mathbf{\Phi}_k s_k^R + j \sum_{k=1}^{K} \mathbf{\Psi}_k s_k^I \quad (58)$$

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) $\mathcal{O}_{N_t}^{\mathcal{H}}\mathcal{O}_{N_t} = (\sum_{k=1}^K |s_k|^2)\mathbf{I}_{N_t}$ [17].

ii) The matrices $\{\mathbf{\Phi}_k, \mathbf{\Psi}_k\}_{k=1}^K$ satisfy the conditions [5]

$$\begin{aligned}\mathbf{\Phi}_k^{\mathcal{H}}\mathbf{\Phi}_k &= \mathbf{I}_{N_t}, \quad \mathbf{\Psi}_k^{\mathcal{H}}\mathbf{\Psi}_k = \mathbf{I}_{N_t}, \quad \forall\ k\\ \mathbf{\Phi}_k^{\mathcal{H}}\mathbf{\Phi}_l &= -\mathbf{\Phi}_l^{\mathcal{H}}\mathbf{\Phi}_k, \quad \mathbf{\Psi}_k^{\mathcal{H}}\mathbf{\Psi}_l = -\mathbf{\Psi}_l^{\mathcal{H}}\mathbf{\Psi}_k, \quad k \neq l\\ \mathbf{\Phi}_k^{\mathcal{H}}\mathbf{\Psi}_l &= \mathbf{\Psi}_l^{\mathcal{H}}\mathbf{\Phi}_k, \quad \forall\ k, l.\end{aligned} \quad (59)$$

□

For complex symbols $s_k = s_k^R + js_k^I$, we define two precoders corresponding to $\{\mathbf{\Phi}_k, \mathbf{\Psi}_k\}$ as $\mathbf{C}_{k,1} = \mathbf{\Phi}_k\mathbf{D}_c^{1/2}\mathbf{U}_H^{\mathcal{H}}$, and $\mathbf{C}_{k,2} = \mathbf{\Psi}_k\mathbf{D}_c^{1/2}\mathbf{U}_H^{\mathcal{H}}$. The combined STBC-beamforming matrix is now

$$\mathcal{Z}_{N_t} = \sum_{k=1}^{K}\mathbf{C}_{k,1}s_k^R + j\sum_{k=1}^{K}\mathbf{C}_{k,2}s_k^I = \mathcal{O}_{N_t}\mathbf{D}_c^{1/2}\mathbf{U}_H^{\mathcal{H}}. \quad (60)$$

The received vector at the $\nu$th antenna becomes $\mathbf{x}_\nu = \mathcal{Z}_{N_t}\mathbf{h}_\nu + \mathbf{w}_\nu = \sum_{k=1}^{K}(\mathbf{C}_{k,1}s_k^R + j\mathbf{C}_{k,2}s_k^I)\mathbf{h}_\nu + \mathbf{w}_\nu$. The receiver consists of $K$ parallel detectors corresponding to $K$ transmitted symbols. For the $l$th detector, the decision variable is formed by [cf. (49)]

$$\begin{aligned} y_l =& y_l^R + jy_l^I \\ =& \mathrm{Re}\left\{\sum_{\nu=1}^{N_r}\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}_{l,1}^{\mathcal{H}}\mathbf{x}_\nu\right\} + j\mathrm{Re}\left\{-j\sum_{\nu=1}^{N_r}\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}_{l,2}^{\mathcal{H}}\mathbf{x}_\nu\right\} \\ =& \mathrm{Re}\left\{\sum_{\nu=1}^{N_r}\sum_{k=1}^{K}\left(\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}_{l,1}^{\mathcal{H}}\mathbf{C}_{k,1}\mathbf{h}_\nu s_k^R + j\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}_{l,1}^{\mathcal{H}}\mathbf{C}_{k,2}\mathbf{h}_\nu s_k^I\right)\right\} \\ &+ j\mathrm{Re}\left\{\sum_{\nu=1}^{N_r}\sum_{k=1}^{K}\left(-j\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}_{l,2}^{\mathcal{H}}\mathbf{C}_{k,1}\mathbf{h}_\nu s_k^R \right.\right. \\ &\left.\left. \qquad + \check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}_{l,2}^{\mathcal{H}}\mathbf{C}_{k,2}\mathbf{h}_\nu s_k^I\right)\right\} \\ &+ \mathrm{Re}\left\{\sum_{\nu=1}^{N_r}\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}_{l,1}^{\mathcal{H}}\mathbf{w}_\nu\right\} + j\mathrm{Re}\left\{-j\sum_{\nu=1}^{N_r}\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{C}_{l,2}^{\mathcal{H}}\mathbf{w}_\nu\right\} \\ =& \sum_{\nu=1}^{N_r}\left(\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{U}_H\mathbf{D}_c\mathbf{U}_H^{\mathcal{H}}\check{\mathbf{h}}_\nu\right)s_l + w_l, \quad \forall\, l \in [1, K] \end{aligned} \quad (61)$$

where $w_l$ has variance $N_0\sum_{\nu=1}^{N_r}(\check{\mathbf{h}}_\nu^{\mathcal{H}}\mathbf{U}_H\mathbf{D}_c\mathbf{U}_H^{\mathcal{H}}\check{\mathbf{h}}_\nu)$. The last equality in (61) can be easily verified since for each $k \neq l, p \neq q$ (where $p, q \in \{1, 2\}$), the interference terms $\mathbf{h}_\nu^{\mathcal{H}}\mathbf{C}_{l,p}^{\mathcal{H}}\mathbf{C}_{k,p}\mathbf{h}_\nu a_k$ and $j\mathbf{h}_\nu^{\mathcal{H}}\mathbf{C}_{l,p}^{\mathcal{H}}\mathbf{C}_{k,q}\mathbf{h}_\nu a_k$ (where $a_k = s_k^R$ or $a_k = s_k^I$) are imaginary numbers that can be suppressed by the $\mathrm{Re}\{\cdot\}$ operation because $\mathbf{C}_{l,p}^{\mathcal{H}}\mathbf{C}_{k,p} = -\mathbf{C}_{k,p}^{\mathcal{H}}\mathbf{C}_{l,p}$ and $j\mathbf{C}_{l,p}^{\mathcal{H}}\mathbf{C}_{k,q} = -j\mathbf{C}_{k,q}^{\mathcal{H}}\mathbf{C}_{l,p}$ by design [cf. (59)]. The self interference $j\mathbf{h}_\nu^{\mathcal{H}}\mathbf{C}_{l,p}^{\mathcal{H}}\mathbf{C}_{l,q}\mathbf{h}_\nu a_l$ is suppressed for the same reason (see also [4] and [5]).

Notice that the SNR from (61) is the same as the MRC output for the single symbol transmission studied in Section IV; thus, the optimal loading in (53) enables space-time block coded transmissions to achieve the performance of single symbol transmission but with symbol rate $K/P$. Relative to single symbol transmission, (61) requires two MRC combiners per receive antenna. Since this complexity increase is negligible relative to the complexity associated with decoding the error-correcting outer codes, which are always present in practical systems, the STBC transmission of (60) entails comparable complexity to the single symbol transmission of (2).

Utilizing channel mean information at the transmitter, our optimal transmissions implement a combination of STBC and eigen-beamforming (60). Orthogonal space-time block coded transmissions are sent using $N_t$ eigen-directions, along the eigenvectors of the correlation matrix of the estimated channels at the transmitter, and are optimally power loaded. We summarize our result as follows.

*Theorem 3:* Under assumptions a0) and a2), the optimal transmission consists of orthogonal STBC across the power loaded beams formed along the eigenvectors of the correlation matrix of the estimated channels at the transmitter. The STBC is constructed as $\mathcal{Z}_{N_t}$ in (60) for complex symbols, with the power loading $\mathbf{D}_c$ as in (53); the optimality pertains to minimizing an upper bound on the symbol error rate.

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

$$\mathcal{O}_2 = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} \begin{matrix} \rightarrow \text{space} \\ \downarrow \text{time} \end{matrix}. \quad (62)$$

For $N_t = 3, 4$, rate 3/4 orthogonal STBC exist, whereas for $N_t > 4$, only rate 1/2 codes have been constructed [5], [17]. Therefore, for complex symbols, the $N_t$-directional eigen-beamformer of (60) achieves optimal performance with no rate loss only when $N_t = 2$ and pays a rate penalty up to 50% when $N_t > 2$ and complex constellations are used. To make up for this loss, the $N_t$-directional beamformer has to enlarge the constellation size, which, for the same performance, necessitates more transmit-power.

### B. Two-Directional Eigen-Beamformer With STBC

To trade off the optimal performance for a constant rate of 1 symbol/s/Hz, it is possible to combine our two-directional eigen-beamformer (see Corollary 2) with the Alamouti code applied to the strongest two eigen-beams. Setting $K = 2$ and forcing *a priori* the matrices $\{\mathbf{\Phi}_k, \mathbf{\Psi}_k\}_{k=1}^2$ to be fat with dimension $2 \times N_t$, we construct the $2 \times N_t$ space-time coded matrix for the two-directional eigen-beamformer

$$\mathcal{Z}_{2-\mathrm{d}} := [\mathcal{O}_2, \mathbf{0}_{2\times(N_t-2)}]\mathbf{D}_c^{1/2}\mathbf{U}_H^{\mathcal{H}}. \quad (63)$$

If $E_s/N_0 < \bar{\gamma}_{\mathrm{th},3}$, then according to Corollary 2, this two-directional STBC eigen-beamformer is optimal in terms of achieving the same SER bound as the $N_t$-directional design of (60). The implementation of this two-directional eigen-beamformer is depicted in Fig. 2.

The optimal scenario for one-directional beamforming, with $\mathcal{Z}_{1-\mathrm{d}} := [s_1, 0, \ldots, 0]\mathbf{D}_c^{1/2}\mathbf{U}_H^{\mathcal{H}}$ was specified in [7] and [8] from a capacity perspective. The interest in one-directional beamforming stems primarily from the fact that it allows for scalar coding with linear preprocessing at the transmit antenna array and thus relieves the receiver from the complexity burden required for decoding the capacity-achieving vector coded transmissions [7], [8], [12], [19]. Because each symbol with two-directional eigen-beamforming goes through a separate but better-conditioned channel offering diversity order 2, the same capacity-achieving scalar code applied to an one-directional beamformer can be applied also to our two-directional eigen-beamformer but for two parallel streams; see also [4] on how to achieve the maximum achievable coded diversity using scalar codes instead of vector codes. Therefore, two-directional eigen-beamforming outperforms one-directional beamforming even from a capacity perspective since it can achieve the same coded BER with less power. Notice that if $\mathbf{D}_c$ has only one nonzero entry $\delta_1 = 1$, the two-directional eigen-beamformer

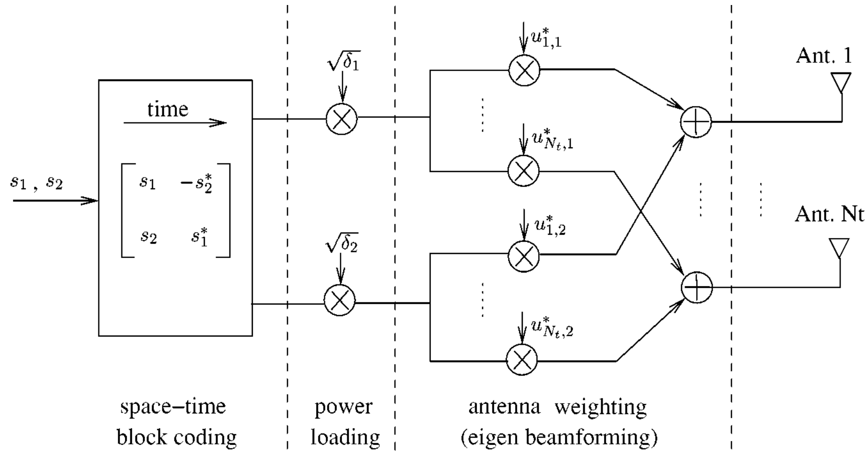


Fig. 2. Two-directional eigen-beamformer $u_{p,q} := [\mathbf{U}_H]_{p,q}$.

reduces to the one-directional beamformer, with $s_1$ and $-s_2^*$ transmitted during two consecutive time slots, as confirmed by (62) and Fig. 2. This leads to following conclusion.

*Corollary 3:* The two-directional eigen-beamformer includes the one-directional-beamformer as a special case and outperforms it uniformly, without rate reduction, and without essential increase in complexity.

Corollary 3 suggests that the two-directional eigen-beamformer is more attractive than the one-directional beamformer. It is also worthwhile recalling that the two-directional eigen-beamformer is overall optimal for systems employing $N_t = 2$ transmit antennas, but even with more than two transmit antennas, if multiple receive antennas are present and the feedback quality is reasonably good, the two-directional eigen-beamformer achieves the best possible performance of $N_t$-directional eigen-beamformer in a large range of transmit-power-to-noise ratios, which is a feature that we will also verify by simulations. Thanks to its full-rate capability and superior performance, the two-directional eigen-beamformer is expected to have major impact in practical systems.

### C. Comparisons With [9] and [11]

The combination of orthogonal STBC with beamforming has also been studied in [9] and [11]. This paper is distinct from [9] and [11] in various aspects.

*1) Coverage:* The formulation of [9] and [11] allows for more general CSI feedback, with $\mathbf{h}$ having nonzero mean and a nonwhite covariance matrix. For simplicity and tractability, closed-form algorithms in [9] and [11] are restricted to mean feedback. We focus on mean feedback at the outset. However, we come up with novel results that are not available in [9] and [11].

Arbitrary signal constellations can be deployed in [9] and [11]. We limit ourselves to commonly used PSK and square QAM constellations.

*2) Approaches:* Our approach of combining beamforming with orthogonal STBCs is different from [9] and [11]. The basic difference is epitomized in our two-directional eigen-beamformer, which maintains full-rate even with more than two transmit antennas. Specifically, the approach in [9] and [11] starts with a given fixed STBC and optimizes a *square* beamformer weight matrix to minimize the *worst-case pairwise error probability*. We start from the spread spectrum scheme of [2] and [6], which is useful in a "power-limited" scenario. We provide exact SER expressions and derive the optimum beamformer based on an upper bound on SER. To *increase the rate without sacrificing performance*, we subsequently combine our already-derived optimum beamformer with STBC, which is a combination that leads to our practically attractive two-directional eigen-beamformer. When the square beamformer of [9] and [11] is combined with orthogonal STBCs, it is destined to sacrifice up to 50% rate when more than two transmit antennas are deployed with spectrally efficient complex constellations. This is not the case with our two-directional eigen-beamformer.

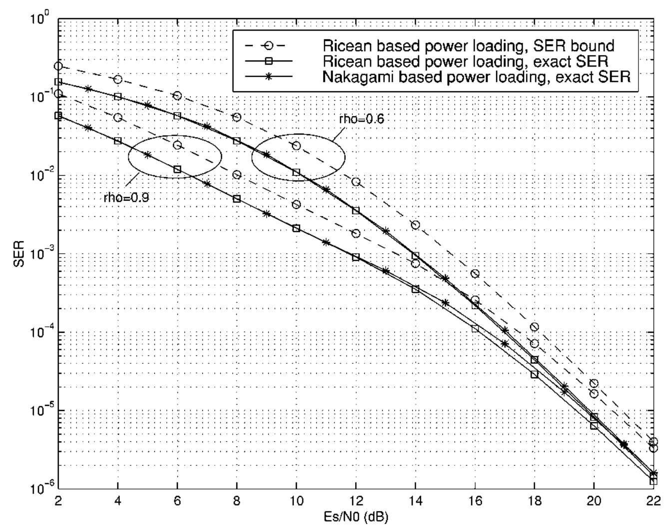


Fig. 3. Power loading based on Ricean and Nakagami distributions (QPSK).

To further appreciate our novel two-directional eigen-beamformer with STBC in "bandwidth-limited" settings, let us consider an example system with $N_t = 6$ transmit antennas, signaling with QPSK modulation. With $N_t = 6$, the approach in [9] and [11] will have to rely on a rate 1/2 orthogonal STBC that incurs 50% rate loss. To make up for this loss, [9] and [11] will have to work with a larger size (16-QAM) constellation. This will entail a considerable power loss of approximately $g_{\text{QPSK}}/g_{16\text{QAM}} = 7$ dB. Notice that in the same setting, our two-directional eigen-beamformer retains full rate of 1 symbol/s/Hz and is capable of achieving the optimal performance under the conditions we specified in Corollary 2.

It is well appreciated that Alamouti's code in (62) is neat in its simplicity. It is optimal in many aspects, and due to its practical merits, it has been introduced to the standards. Alamouti's code suffers up to 50% rate loss when extended to more than two transmit antennas with spectrally efficient complex constellations. Our two-directional eigen-beamformer shows how with a simple $2 \times N_t$ matrix $\boldsymbol{\mathcal{Z}}_{2-\text{d}}$ (whose entries we find in closed form), one can take advantage of channel mean feedback to improve on Alamouti's performance and enable full-rate operation, even with more than two transmit antennas. The two-directional eigen-beamformer is an easy-to-deploy design with very strong application potential. It is indeed interesting to know that by utilizing channel mean feedback, orthogonal STBC designs can enjoy full-rate with more than two transmit antennas.

*3) Optimality Criterion:* The criterion in [9] and [11] is the worst-case pairwise error probability (PEP); it corresponds to the Chernoff bound on the codeword error probability formed by the dominant terms in the union bound. We optimize the beamformer relying on an upper bound on the SER. However, as the optimality criteria used in [9] and [11] and in this paper become proportional, the resulting transmitters become *equivalent* when orthogonal STBCs are adopted. This optimality link was pointed out in our companion paper [20] but was not recognized in [9] and [11]. We also provide the *exact* SER expressions, which are useful to calculate the expected SER based on channel mean feedback.

In addition, we provide comparisons disclosing that our SER-bound optimal designs outperform the maximum-SNR optimal designs in [3] and [12] and provide links with transmitter designs based on capacity criterion [7], [8], [12], [19].

*4) Power Allocation:* With mean feedback, the semi-analytical solution of [11] for optimal power allocation requires an one-dimensional (single parameter) numerical search. This search may have to be performed as many as $N_t$ times. Different from [9] and [11], we here derive a *simple*, albeit *suboptimal*, closed-form solution. The closed-form solution provides interesting theoretical insights and is certainly faster than the numerical search. The overall transmitter complexity includes the eigen-decomposition of an $N_t \times N_t$ channel correlation matrix in addition to the optimal power allocation. When $N_t$ is small and eigen-decomposition is performed efficiently, the overall transmitter complexity can be considerably reduced by our closed-form power allocation.

*5) Rate-Performance Tradeoffs:* The constellation-specific thresholds provided in (55), as well as exact SER expressions, are useful for systems with adaptive modulation, where various system parameters, including constellation size, beamformer size, and transmission power, can be adjusted to strike the best tradeoff between rate and performance.

## VI. Numerical Results

We first consider an uniform linear array with $N_t = 4$ antennas at the transmitter and a single antenna at the receiver. We assume that the transmit antennas are well separated and consider the delayed channel feedback scenario outlined in Case 1 of Section III, with $\sigma_h^2 = 1$, and a given correlation coefficient $\rho$. We will present simulation results for two constellations: QPSK (4-PSK) and 16-QAM. Simulation results are averaged over 10 000 Monte Carlo feedback realizations.

We first compare optimal power loading based on the Ricean distribution (33) with that based on the Nakagami distribution (43). Fig. 3 verifies that both approaches have almost identical performance. For this reason, we subsequently plot only the performance of power loading based on (43). Fig. 3 also confirms that the SER bound is tight and has a constant difference with the exact SER across the $E_s/N_0$ range considered. This justifies well our approach of pushing down the bound to decrease the SER.

Figs. 4 and 5 compare optimal power loading, equal power loading (that has the same performance as plain STBC without beamforming), and one-directional and two-directional beamforming for both QPSK and 16–QAM. When the feedback quality is low $(\rho = 0.6)$, Fig. 4 shows that optimal power loading performs close to equal power loading, whereas it considerably outperforms conventional one-directional beamforming. On the other hand, when the feedback quality improves to $\rho = 0.9$, equal power loading is highly suboptimum. The conventional beamforming performs close to the optimal power loading at low SNR, whereas it becomes inferior at sufficiently high SNR. Notice that the two-directional beamformer outperforms the one-directional beamformer uniformly. When $E_s/N_0 > \gamma_{\text{th}}$ for each feedback realization, although both two-directional and one-directional beamformer become suboptimal, the two-directional beamformer benefits from the order-2 diversity. Since $g_{\text{QPSK}}/g_{16\text{QAM}} = 5$, we observe that 7.0 dB higher power is required for 16-QAM than QPSK to adopt $N_t$ directions.

Figs. 6 and 7 depict the probability density function (p.d.f.) of the SNR at the MRC output when the channel feedback is $\hat{\mathbf{h}}_{f,0} = [1, 1, 1, 1]^T$ with $\rho = 0.6$, and $E_s/N_0 = 15$, 20 dB, respectively. The channel uncertainty is embodied in $\boldsymbol{\epsilon}$. The p.d.f. is calculated from 50 000 realizations of $\boldsymbol{\epsilon}$. It verifies that one-directional beamforming is indeed optimal in terms of maximizing the expected SNR. However, to achieve better SER, the optimal power allocation strives for the optimal tradeoff between high SNR mean and low SNR variance. The optimal tradeoff is, of course, dependent on the chosen signal constellation, as confirmed by Figs. 6 and 7.

We next test our results with multiple receive antennas. Figs. 8 and 9 are the counterparts of Figs. 4 and 5 but with $N_r = 2$ receive antennas. It can be seen that the performance of the two-directional beamformer coincides with the optimal beamformer for a larger range of $E_s/N_0$ than that of the one-directional beamformer. This is different from the single receive antenna case, where two-directional and one-directional beamformers deviate from the optimal beamformer at the same time since there is only one dominant direction.

With multiple receive antennas and reasonably good feedback quality, the two-directional beamformer is capable of achieving the same performance as the $N_t$-directional beamformer with high probability and without rate reduction. We next verify this

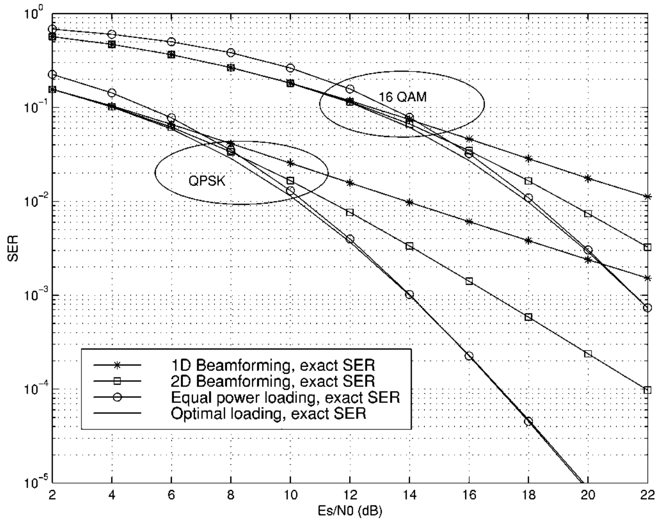


Fig. 4. SER versus $E_s/N_0$ ($\rho = 0.6$, $N_t = 4$, $N_r = 1$).

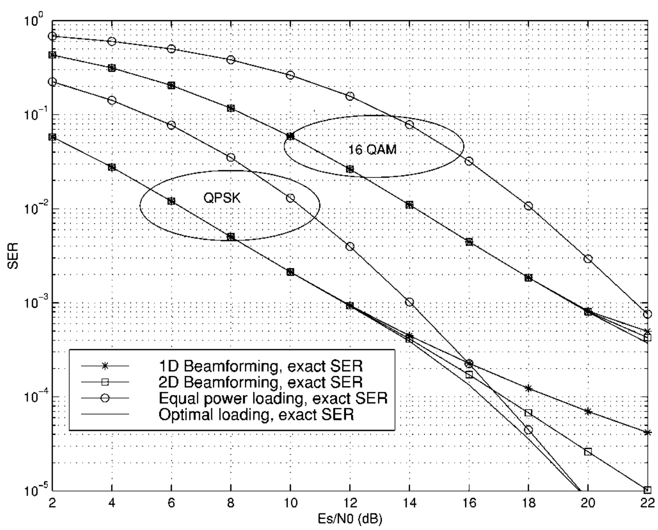


Fig. 5. SER versus $E_s/N_0$ ($\rho = 0.9$, $N_t = 4$, $N_r = 1$).

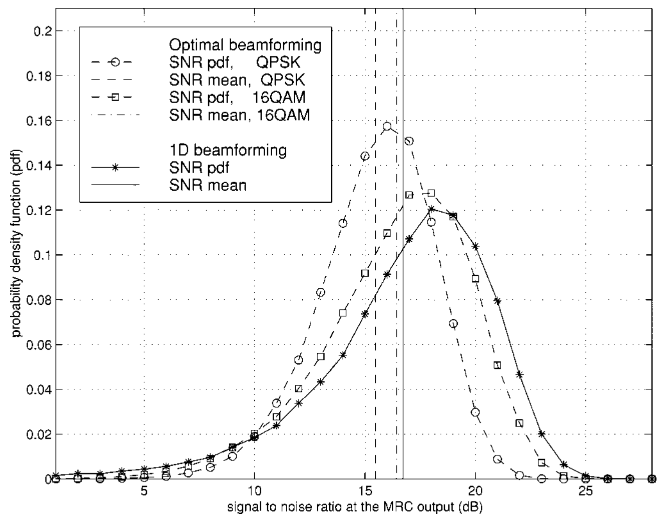


Fig. 6. P.D.F. of SNR at the MRC output ($\rho = 0.6$, $E_s/N_0 = 15$dB).

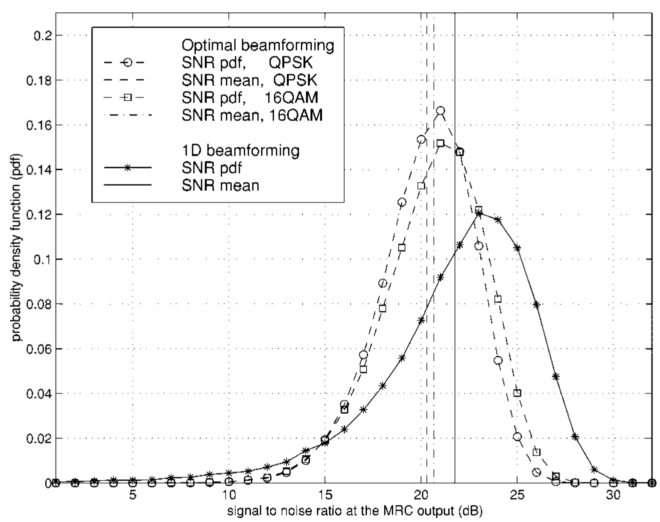


Fig. 7. P.D.F. of SNR at the MRC output ($\rho = 0.6$, $E_s/N_0 = 20$dB).

point by simulations. From 10 000 feedback realizations, we collect the percentages for which one-directional or two-directional beamforming is optimal. Let $P_1$ denote the probability that one-directional beamforming is optimal and $P_2$ the probability that the two-directional beamforming is optimal. We plot the minimal $\rho$ for each $E_s/N_0$ point that leads to $P_1 \geq 99\%$ and $P_2 \geq 99\%$ in Figs. 10 and 11 for $N_r = 2$ and $N_r = 4$, respectively. Since the mobile is unlikely to deploy more than two receive antennas in practice, Fig. 10 is practically more important. As confirmed by Fig. 10 for two receive antennas, when the quality of channel feedback improves to a level that $\rho \geq 0.95$ ($\sigma_\epsilon^2 \leq -10.1$ dB), the two-directional beamformer is optimal with $P_2 \geq 99\%$ for QPSK over the considered $E_s/N_0$ range. Even with $\rho = 0.9$ ($\sigma_\epsilon^2 = -7.3$ dB), the two-directional eigen-beamformer is optimal for QPSK when $E_s/N_0 <$ 16 dB. Notice that when $E_s/N_0 > 16$ dB, the SER for the two-directional beamformer is already extremely low (around $10^{-7}$), as shown in Fig. 9. When the number of receive antennas increases, the requirement for the feedback quality increases for the two-directional beamformer to be optimal for a given $E_s/N_0$, as shown in Fig. 11. However, in such cases, the $E_s/N_0$ range of interest should be in the lower end since multiple received copies enhance the received SNR and decrease the SER considerably. As the constellation size increases, the requirement for feedback quality also decreases for the two-directional beamformer to be optimal. Finally, Figs. 10 and 11 also confirm that the two-directional beamformer outperforms the one-directional beamformer considerably.

## VII. Conclusions

In this paper, we have derived an optimal transmitter design for multiple-antenna systems based on channel mean feedback. The optimal precoder turns out to be a generalized beamformer with multiple beams pointing to orthogonal directions along the eigenvectors of the correlation matrix of the estimated channel at the transmitter and with power loading across the multiple beams obeying a "spatial water-filling" principle. To increase the data rate without compromising the performance, orthogonal space-time block codes were naturally coupled with the proposed transmitter eigen-beamformers. A

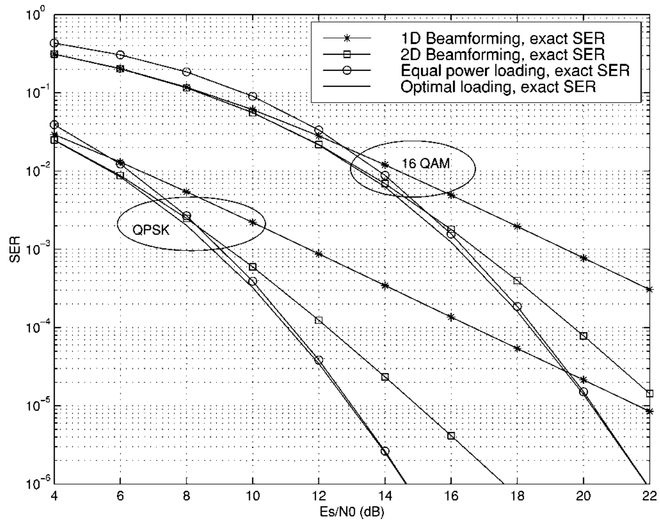


Fig. 8. SER versus $E_s/N_0$ ($\rho = 0.6$, $N_t = 4$, $N_r = 2$).

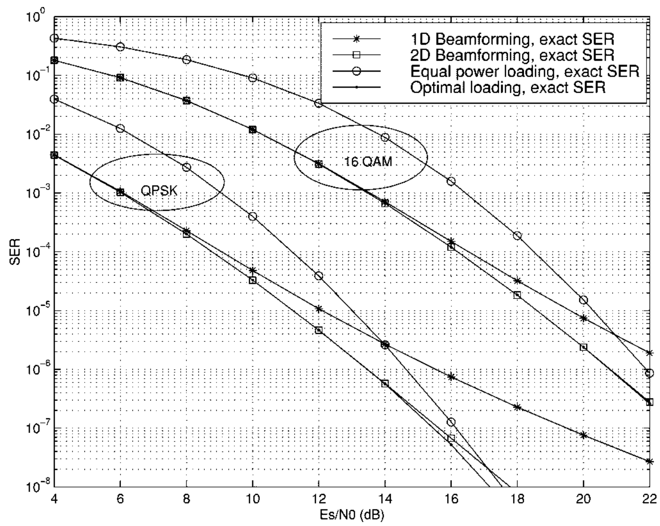


Fig. 9. SER versus $E_s/N_0$ ($\rho = 0.9$, $N_t = 4$, $N_r = 2$).

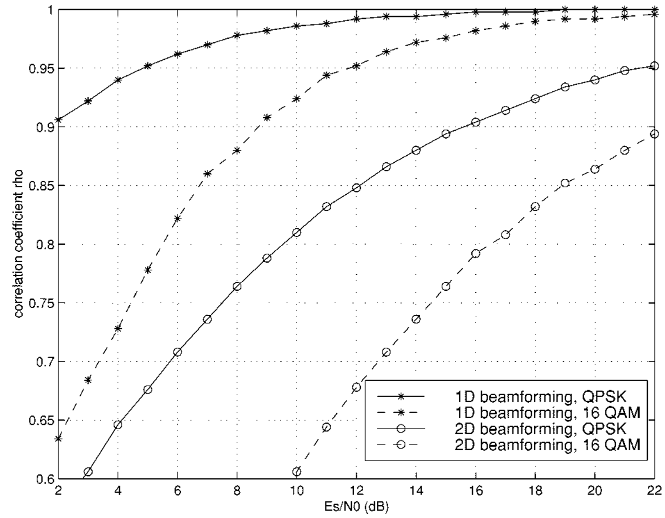


Fig. 10. Minimal $\rho$ ensuring $P_1, P_2 \geq 99\%$ ($N_r = 2$).

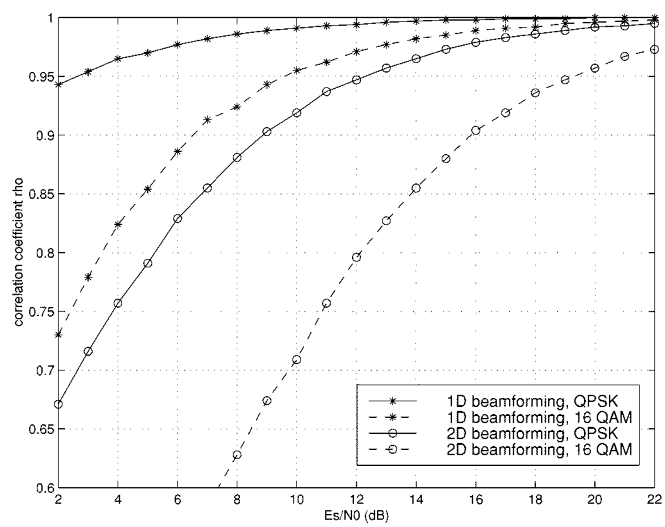


Fig. 11. Minimal $\rho$ ensuring $P_1, P_2 \geq 99\%$ ($N_r = 4$).

two-directional eigen-beamformer subsumed the conventional one-directional beamformer as a special case and was shown to outperform it uniformly without rate reduction and without essential increase in complexity. More important, with multiple receive antennas and reasonably good channel feedback quality, the two-directional eigen-beamformer is capable of achieving the best possible performance over a large range of transmitted-power-to-noise ratios, yet avoiding rate reduction. Thanks to its full-rate capability and superior performance, the two-directional eigen-beamformer has strong application potential for future wireless systems with multiple transmit antennas and channel mean feedback.

## References

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

[2] J. K. Cavers, "Optimized use of diversity modes in transmitter diversity systems," in *Proc. Veh.Technol. Conf.*, vol. 3, Amsterdam, The Netherlands, 1999, pp. 1768–1773.

[3] ——, "Single-user and multiuser adaptive maximal ratio transmission for Rayleigh channels," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 2043–2050, Nov. 2000.

[4] G. Ganesan and P. Stoica, "Achieving optimum coded diversity with scalar codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2078–2080, July 2001.

[5] ——, "Space-time block codes: A maximum SNR approach," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1650–1656, May 2001.

[6] G. B. Giannakis and S. Zhou, "Optimal transmit-diversity precoders for random fading channels," in *Proc. Globecom Conf.*, vol. 3, San Francisco, CA, Nov.–Dec. 27–1, 2000, pp. 1839–1843.

[7] S. A. Jafar and A. Goldsmith, "On optimality of beamforming for multiple antenna systems with imperfect feedback," in *Proc. Int. Symp. Inform. Theory*, Washington, DC, June 2001, pp. 321–321.

[8] 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.*, vol. 7, Helsinki, Finland, June 2001, pp. 2266–2270.

[9] G. Jöngren and B. Ottersten, "Combining transmit antenna weights and orthogonal space-time block codes by utilizing side information," in *Proc. 33rd Asilomar Conf. Signals, Syst., Comput.*, vol. 2, Pacific Grove, CA, 1999, pp. 1562–1566.

[10] G. Jöngren and M. Skoglund, "Utilizing quantized feedback information in orthogonal space-time block coding," in *Proc. Global Telecommun. Conf.*, vol. 2, 2000, pp. 995–999.

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

[12] 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. Select. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[13] A. Narula, M. D. Trott, and G. W. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2418–2433, Nov. 1999.

[14] J. Proakis, *Digital Communications*, 4th ed: McGraw-Hill, 2000.

[15] M. K. Simon and M.-S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. IEEE*, vol. 86, pp. 1860–1877, Sept. 1998.

[16] ——, *Digital Communication Over Generalized Fading Channels: A Unified Approach to the Performance Analysis*. New York: Wiley, 2000.

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

[18] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

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

[20] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inform. Theory*, Sept. 2001, submitted for publication.

**Shengli Zhou** (S'99) received the B.S. degree in 1995 and the M.Sc. degree in 1998 from the University of Science and Technology of China (USTC), Hefei, all in electrical engineering and information science. He received the Ph.D. degree in 2002 from the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis.

His broad interests lie in the areas of communications and signal processing, including transceiver optimization, blind channel estimation and equalization algorithms, wireless, multi-carrier, space-time coded, and spread-spectrum communication systems.

**Georgios B. Giannakis** (F'97) received the Diploma degree in electrical engineering from the National Technical University of Athens, Athens, Greece in 1981. From September 1982 to July 1986, he was with the University of Southern California (USC), Los Angeles, where he received the M.Sc. degree in electrical engineering in 1983, the M.Sc. degree in mathematics in 1986, and the Ph.D. in electrical engineering in 1986.

After lecturing for one year at USC, he joined the University of Virginia, Charlotteville, in 1987, where he became a Professor of electrical engineering in 1997. Since 1999, he has been a Professor with the Department of Electrical and Computer Engineering, University of Minnesota, Mineapolis, where he now 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 140 journal papers, 270 conference papers, and two edited books. Current research topics focus on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, precoding and space-time coding for block transmissions, multicarrier, and wideband wireless communication systems. He is a frequent consultant to the telecommunications industry.

Dr. Giannakis was the (co-) recipient of four best paper awards from the IEEE Signal Processing (SP) Society in 1992, 1998, 2000, and 2001. 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 LETTERS, 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 is 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 of the IEEE-SP Society's Board of Governors.