Correspondence

Chernoff Bounds on Pairwise Error Probabilities of Space–Time Codes

Aleksandar Dogandžić, *Member, IEEE*

***Abstract*—We derive Chernoff bounds on pairwise error probabilities of coherent and noncoherent space–time signaling schemes. First, general Chernoff bound expressions are derived for a correlated Ricean fading channel and correlated additive Gaussian noise. Then, we specialize the obtained results to the cases of space–time-separable noise, white noise, and uncorrelated fading. We derive approximate Chernoff bounds for high and low signal-to-noise ratios (SNRs) and propose optimal signaling schemes. We also compute the optimal number of transmitter antennas for noncoherent signaling with unitary mutually orthogonal space–time codes.**

***Index Terms*—Chernoff bounds, correlated Ricean fading, multiple antennas, pairwise error probability, space–time modulation, transmitter and receiver diversity.**

I. Introduction

Chernoff bounds on pairwise error probabilities have been used to design coded modulation schemes for fading channels [1]–[6] and, more recently, to design space–time codes and analyze the performance of multiple-input–multiple-output (MIMO) wireless communication systems, see [7]–[11]. Assuming uncorrelated Ricean- and Rayleigh-fading channels and white noise, exact and approximate expressions for pairwise error probabilities of coherent space–time signaling schemes were derived in [12] and [13] (see also references therein). (For comprehensive treatment of methods for computing pairwise error probabilities in fading channels, see [14].) Recently, exact and asymptotic pairwise error probabilities have been computed for space–time signaling schemes in correlated Rayleigh fading and white noise, see [15] and [16]. In this correspondence, we derive Chernoff-bound expressions for coherent and noncoherent signaling in a correlated flat Ricean-fading channel and correlated additive Gaussian noise, generalizing the corresponding results in [7] and [8]. Approximate Chernoff bounds are derived for high and low scattering signal-to-noise ratios (SNRs) and optimal signaling schemes are proposed.

First, we introduce measurement and fading models. The $n_\mathrm{R} \times 1$ vector signal received by the receiver array at time t is modeled as

$$\boldsymbol{y}(t) = \boldsymbol{H}\boldsymbol{\phi}(t) + \boldsymbol{e}(t), \qquad t = 1, \ldots, N \tag{1.1}$$

where $\boldsymbol{H}$ is the $n_\mathrm{R} \times n_\mathrm{T}$ channel response matrix, $\boldsymbol{\phi}(t)$ is the $n_\mathrm{T} \times 1$ vector of symbols transmitted by n_T transmitter antennas and received by the receiver array at time t, and $\boldsymbol{e}(t)$ is additive noise. Stacking all N samples into a single vector, the above set of equations may be written as

$$\boldsymbol{y} = (\boldsymbol{\Phi}^T \otimes \boldsymbol{I}_{n_\mathrm{R}})\, \boldsymbol{h} + \boldsymbol{e} \tag{1.2}$$

where

$$\begin{aligned} \boldsymbol{y} &= [\boldsymbol{y}(1)^T, \boldsymbol{y}(2)^T, \ldots, \boldsymbol{y}(N)^T]^T \\ \boldsymbol{h} &= \mathrm{vec}\{\boldsymbol{H}\} \\ \boldsymbol{e} &= [\boldsymbol{e}(1)^T, \boldsymbol{e}(2)^T, \ldots, \boldsymbol{e}(N)^T]^T \end{aligned}$$

and

$$\boldsymbol{\Phi} = [\boldsymbol{\phi}(1)\, \boldsymbol{\phi}(2) \cdots \boldsymbol{\phi}(N)]$$

is the matrix of symbols received in the coherent interval $t = 1, 2, \ldots, N$. Here, the vec operator stacks the columns of a matrix one below another into a single column vector, $\boldsymbol{I}_n$ denotes the identity matrix of size n, and "T" and $\otimes$ denote transpose and Kronecker product, respectively. Furthermore, we assume that the noise $\boldsymbol{e}$ is a zero-mean complex Gaussian vector with positive-definite covariance matrix $\mathrm{E}[\boldsymbol{e}\boldsymbol{e}^H] = \boldsymbol{R}$, and that the vector of fading coefficients $\boldsymbol{h}$ is complex Gaussian with mean $\mathrm{E}[\boldsymbol{h}]$ and positive-definite covariance $\mathrm{E}[(\boldsymbol{h} - \mathrm{E}[\boldsymbol{h}])(\boldsymbol{h} - \mathrm{E}[\boldsymbol{h}])^H]$, denoted as

$$\mathrm{E}[\boldsymbol{h}] = \boldsymbol{\mu}_h, \quad \mathrm{E}[(\boldsymbol{h} - \boldsymbol{\mu}_h)(\boldsymbol{h} - \boldsymbol{\mu}_h)^H] = \boldsymbol{\Psi}_h \tag{1.3}$$

where "H" denotes Hermitian (conjugate) transpose. We will examine the following model for the mean $\boldsymbol{\mu}_h$ of the fading coefficient vector (see also [17]):

$$\boldsymbol{\mu}_h = x \cdot \boldsymbol{a}_\mathrm{T} \otimes \boldsymbol{a}_\mathrm{R} \tag{1.4}$$

where $\boldsymbol{a}_\mathrm{T}$ and $\boldsymbol{a}_\mathrm{R}$ are *line-of-sight* transmitter and receiver array responses, and x is the complex amplitude of the line-of-sight signal.

In Section II, we derive Chernoff bound expressions for coherent signaling. We specialize these expressions to the space–time separable noise scenario (Section II-A) and examine optimal signaling schemes. White noise and uncorrelated fading models are considered in Section II-A1. In Section III, we derive Chernoff bounds for noncoherent signaling. Based on approximate expressions for high and low scattering SNRs, we propose optimal code design criteria for noncoherent signaling (Sections III-A and III-B). Finally, in Section III-C we examine equal-energy orthogonal signaling and compute the optimal number of transmitter antennas for this scenario.

II. Coherent Signaling

We compute Chernoff bounds for *coherent* signaling (i.e., assuming that the channel is known to the receiver) by obtaining the Chernoff bound expression for a given channel realization, and averaging it over all possible channel realizations under a correlated flat Ricean-fading model.

Consider the measurement and fading models in Section I, where the channel $\boldsymbol{h}$ and noise covariance $\boldsymbol{R}$ are known to the receiver. Assume that we wish to decide between two space–time codes $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_0$, i.e., to test the hypothesis H_1: $\boldsymbol{\Phi}_1$ transmitted versus the alternative H_0: $\boldsymbol{\Phi}_0$ transmitted. Assume also that $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_0$ are equiprobable. Under H_1, the received measurement vector $\boldsymbol{y}$ is a complex multivariate normal with mean $\boldsymbol{Z}_1\boldsymbol{h}$ and covariance $\boldsymbol{R}$, whereas under H_0 it is a complex multivariate normal with mean $\boldsymbol{Z}_0\boldsymbol{h}$ and covariance $\boldsymbol{R}$. Denote the probability density function (pdf) of $\boldsymbol{y}$ under H_i as $p_i(\boldsymbol{y}|\boldsymbol{h})$, $i = 0, 1$, and define

$$\boldsymbol{Z}_1 = \boldsymbol{\Phi}_1^T \otimes \boldsymbol{I}_{n_\mathrm{R}}, \qquad \boldsymbol{Z}_0 = \boldsymbol{\Phi}_0^T \otimes \boldsymbol{I}_{n_\mathrm{R}}. \tag{2.1}$$

Manuscript received January 25, 2002; revised December 10, 2002. The material in this correspondence was presented in part at the Second IEEE Sensor Array and Multichannel Signal Processing Workshop, Rosslyn, VA, August 2002.

The author is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: ald@iastate.edu).

Communicated by A. Kavčić, Associate Editor for Detection and Estimation.
Digital Object Identifier 10.1109/TIT.2003.810652

Given $\boldsymbol{h}$, the Chernoff bound on pairwise error probability for deciding between H_1 and H_0 is

$$P_{\mathrm{CB}}(\boldsymbol{h}, \lambda) = \tfrac{1}{2}\ \exp[\xi(\lambda|\boldsymbol{h})], \qquad 0 \le \lambda \le 1 \tag{2.2}$$

where

$$\xi(\lambda|\boldsymbol{h}) = \ln \mathrm{E}\{\exp[\lambda \ln p_0(\boldsymbol{y}|\boldsymbol{h}) - \lambda \ln p_1(\boldsymbol{y}|\boldsymbol{h})]|\mathrm{H}_1\} \tag{2.3}$$

see [18, Ch. 2.7], [19, Ch. 3.4], and [8, Appendix B]. Here, (2.3) becomes

$$\xi(\lambda|\boldsymbol{h}) = \ln \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{|\pi \boldsymbol{R}|} \cdot \exp[-\lambda\, (\boldsymbol{y}-\boldsymbol{Z}_0\boldsymbol{h})^H \boldsymbol{R}^{-1}(\boldsymbol{y}-\boldsymbol{Z}_0\boldsymbol{h}) \\ -(1-\lambda)\, (\boldsymbol{y}-\boldsymbol{Z}_1\boldsymbol{h})^H \boldsymbol{R}^{-1}(\boldsymbol{y}-\boldsymbol{Z}_1\boldsymbol{h})]\, d\boldsymbol{y} \tag{2.4}$$

where $d\boldsymbol{y} = d(\mathrm{Re}\boldsymbol{y})d(\mathrm{Im}\{\boldsymbol{y}\})$. Let us denote complex conjugation by "$*$." To compute (2.4), we use the following lemma.[1]

Lemma 1: Let $\boldsymbol{B}$ represent an $n \times n$ positive-definite Hermitian matrix and $\boldsymbol{A}$ an $n \times n$ Hermitian matrix; let $\boldsymbol{a}$ and $\boldsymbol{b}$ represent $n \times 1$ vectors of complex constants; and let a_0 and b_0 represent complex scalars. Then

$$\begin{aligned}
&\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \tfrac{1}{2}\, [\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} + \boldsymbol{x}^H\boldsymbol{a} + \boldsymbol{a}^H\boldsymbol{x} + a_0 + a_0^*] \\
&\quad \cdot \exp[-\tfrac{1}{2}(\boldsymbol{x}^H\boldsymbol{B}\boldsymbol{x} + \boldsymbol{x}^H\boldsymbol{b} + \boldsymbol{b}^H\boldsymbol{x} + b_0 + b_0^*)]\ d\boldsymbol{x} \\
&\quad = \tfrac{1}{2}\,\pi^n |\tfrac{1}{2}\,\boldsymbol{B}|^{-1} \cdot \Big[2\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}^{-1}) - \boldsymbol{b}^H\boldsymbol{B}^{-1}\boldsymbol{a} - \boldsymbol{a}^H\boldsymbol{B}^{-1}\boldsymbol{b} \\
&\qquad + \boldsymbol{b}^H\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{b} + 2\mathrm{Re}\{a_0\}\Big] \\
&\quad \cdot \exp[\tfrac{1}{2}\boldsymbol{b}^H\boldsymbol{B}^{-1}\boldsymbol{b} - \mathrm{Re}\{b_0\}].
\end{aligned} \tag{2.5}$$

Proof: See the Appendix.

Let us define

$$\boldsymbol{Q} = (\boldsymbol{Z}_1 - \boldsymbol{Z}_0)^H \boldsymbol{R}^{-1}(\boldsymbol{Z}_1 - \boldsymbol{Z}_0). \tag{2.6}$$

Then, applying Lemma 1 to (2.4), with $\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{Z}_1\boldsymbol{h}$, $\boldsymbol{B} = 2\boldsymbol{R}^{-1}$, $\boldsymbol{A} = \boldsymbol{0}$, $\boldsymbol{a} = \boldsymbol{0}$, $\boldsymbol{b} = 2\lambda \cdot \boldsymbol{R}^{-1}(Z_1 - Z_0)\boldsymbol{h}$, $a_0 = 1/|\pi\boldsymbol{R}|$, and $b_0 = \lambda\boldsymbol{h}^H\boldsymbol{Q}\boldsymbol{h}$, we get

$$\xi(\lambda|\boldsymbol{h}) = (\lambda^2 - \lambda) \cdot \boldsymbol{h}^H\boldsymbol{Q}\boldsymbol{h} \tag{2.7}$$

which is minimized for $\lambda = 1/2$, yielding the optimal Chernoff bound for the case of ideal channel state information

$$P_{\mathrm{CB}}(\boldsymbol{h}) = \tfrac{1}{2}\exp(-\tfrac{1}{4}\boldsymbol{h}^H\boldsymbol{Q}\boldsymbol{h}). \tag{2.8}$$

Now, following the Ricean-fading model for the channel coefficients in (1.3), we average (2.8) over $\boldsymbol{h}$

$$P_{\mathrm{CB}} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P_{\mathrm{CB}}(\boldsymbol{h})\ \frac{1}{|\pi\boldsymbol{\Psi}_h|} \\ \cdot \exp\{-(\boldsymbol{h} - \boldsymbol{\mu}_h)^H\boldsymbol{\Psi}_h^{-1}(\boldsymbol{h} - \boldsymbol{\mu}_h)\}\ d\boldsymbol{h} \tag{2.9}$$

which can be computed by applying Lemma 1 with $\boldsymbol{x} = \boldsymbol{h}$, $\boldsymbol{B} = 2\boldsymbol{\Psi}_h^{-1} + \frac{1}{2}\boldsymbol{Q}$, $\boldsymbol{A} = \boldsymbol{0}$, $\boldsymbol{a} = \boldsymbol{0}$, $\boldsymbol{b} = -2\boldsymbol{\Psi}_h^{-1}\boldsymbol{\mu}_h$, $a_0 = 1/(2|\pi\boldsymbol{\Psi}_h|)$, and $b_0 = \boldsymbol{\mu}_h^H\boldsymbol{\Psi}_h^{-1}\boldsymbol{\mu}_h$, yielding

$$P_{\mathrm{CB}} = \frac{1}{2} \cdot \frac{1}{|\boldsymbol{I}_{n_\mathrm{T} n_\mathrm{R}} + \frac{1}{4}\,\boldsymbol{Q}\boldsymbol{\Psi}_h|} \\ \cdot \exp[-\tfrac{1}{4}\,\boldsymbol{\mu}_h^H\,\left(\boldsymbol{I}_{n_\mathrm{T} n_\mathrm{R}} + \tfrac{1}{4}\,\boldsymbol{Q}\boldsymbol{\Psi}_h\right)^{-1}\boldsymbol{Q}\boldsymbol{\mu}_h]. \tag{2.10}$$

[1]The integral in Lemma 1 has a fairly general form. Throughout this correspondence, we will use its special case with $\boldsymbol{A} = \boldsymbol{0}$ and $\boldsymbol{a} = \boldsymbol{0}$.

Interestingly, the exact minimum pairwise error probability for coherent detection (averaged over channel realizations) can be readily computed using (2.10) and [14, Ch. 12.1.2.1 and eq. (4.2)]

$$\mathrm{PEP}(\boldsymbol{\Phi}_0, \boldsymbol{\Phi}_1) = \frac{1}{\pi} \cdot \int_0^{\pi/2} \frac{1}{|\boldsymbol{I}_{n_\mathrm{T} n_\mathrm{R}} + \frac{1}{4\sin^2\theta} \cdot \boldsymbol{Q}\boldsymbol{\Psi}_h|} \\ \cdot \exp\left[-\frac{1}{4\sin^2\theta} \cdot \boldsymbol{\mu}_h^H\left(\boldsymbol{I}_{n_\mathrm{T} n_\mathrm{R}} + \frac{1}{4\sin^2\theta} \cdot \boldsymbol{Q}\boldsymbol{\Psi}_h\right)^{-1}\boldsymbol{Q}\boldsymbol{\mu}_h\right] d\theta \tag{2.11}$$

see also [12], where a similar expression was derived for uncorrelated fading and white noise.

Let us now introduce some terminology and notation. A positive-semidefinite Hermitian matrix is "large" if its nonzero eigenvalues are significantly larger than 1. Similarly, a positive-semidefinite Hermitian matrix is "small" if its eigenvalues are significantly smaller than 1. Also, we will denote by $\boldsymbol{\Psi}^{1/2}$ a Hermitian square root of a Hermitian matrix $\boldsymbol{\Psi}$; then $\boldsymbol{\Psi}^{-1/2} = (\boldsymbol{\Psi}^{1/2})^{-1}$.

Large $\frac{1}{4}\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}$: If the matrix $\frac{1}{4}\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}$ is "large" then we can approximate (2.10) as

$$P_{\mathrm{CB}} \approx \frac{1}{2} \cdot \frac{1}{\left|\frac{1}{4}\,\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}\right|_{\mathrm{rank}(\boldsymbol{Q})}} \\ \cdot \exp[-\boldsymbol{\mu}_h^H\boldsymbol{\Psi}_h^{-1/2} \cdot \boldsymbol{\Pi}(\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}) \cdot \boldsymbol{\Psi}_h^{-1/2}\boldsymbol{\mu}_h] \tag{2.12}$$

where $\boldsymbol{\Pi}(\boldsymbol{X})$ denotes the projection matrix[2] onto the column space of $\boldsymbol{X}$, and $|\boldsymbol{A}|_r$ denotes the product of the r largest eigenvalues of a Hermitian matrix $\boldsymbol{A}$. Note also that $\mathrm{rank}(\boldsymbol{Q}) = n_\mathrm{R} \cdot \mathrm{rank}(\boldsymbol{\Phi}_1 - \boldsymbol{\Phi}_0)$. If $\boldsymbol{\Phi}_1 - \boldsymbol{\Phi}_0$ has full rank n_T (which is the *rank criterion* in [7]), the above expression simplifies to

$$P_{\mathrm{CB\ full\ rank}} \approx \frac{1}{2} \cdot \frac{1}{\left|\frac{1}{4}\,\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}\right|}\ \exp(-\boldsymbol{\mu}_h^H\boldsymbol{\Psi}_h^{-1}\boldsymbol{\mu}_h). \tag{2.13}$$

To achieve the full rank n_T of $\boldsymbol{\Phi}_1 - \boldsymbol{\Phi}_0$, the number of time samples needs to be equal to or larger than the number of transmitter antennas, i.e., $N \ge n_\mathrm{T}$. From (2.13), it follows that the optimal codes should maximize

$$|\boldsymbol{Q}| = |[(\boldsymbol{\Phi}_1^* - \boldsymbol{\Phi}_0^*) \otimes \boldsymbol{I}_{n_\mathrm{R}}] \cdot \boldsymbol{R}^{-1} \cdot [(\boldsymbol{\Phi}_1^T - \boldsymbol{\Phi}_0^T) \otimes \boldsymbol{I}_{n_\mathrm{R}}]| \tag{2.14}$$

which is an extension of the *determinant criterion* in [7] to this scenario.

Small $\frac{1}{4}\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}$: If the matrix $\frac{1}{4}\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}$ is "small," then we can approximate (2.10) as

$$\begin{aligned} P_{\mathrm{CB}} &\approx \tfrac{1}{2} \cdot \exp\{-\tfrac{1}{4}\ \mathrm{tr}[\boldsymbol{Q}(\boldsymbol{\mu}_h\boldsymbol{\mu}_h^H + \boldsymbol{\Psi}_h)]\} \\ &= \tfrac{1}{2} \cdot \exp\{-\tfrac{1}{4}\ \mathrm{E}(\boldsymbol{h}^H\boldsymbol{Q}\boldsymbol{h})\}. \end{aligned} \tag{2.15}$$

Clearly, the optimal codes need to maximize

$$\mathrm{tr}[\boldsymbol{Q}(\boldsymbol{\mu}_h\boldsymbol{\mu}_h^H + \boldsymbol{\Psi}_h)] = \mathrm{tr}\left\{[(\boldsymbol{\Phi}_1^* - \boldsymbol{\Phi}_0^*) \otimes \boldsymbol{I}_{n_\mathrm{R}}] \cdot \boldsymbol{R}^{-1} \\ \cdot [(\boldsymbol{\Phi}_1^T - \boldsymbol{\Phi}_0^T) \otimes \boldsymbol{I}_{n_\mathrm{R}}] \cdot (\boldsymbol{\mu}_h\boldsymbol{\mu}_h^H + \boldsymbol{\Psi}_h)\right\} \tag{2.16}$$

which can be viewed as a measure of the SNR. We can define the Ricean K-factor as the ratio between the line-of-sight and scattering SNR components

$$K = \frac{\boldsymbol{\mu}_h^H\boldsymbol{Q}\boldsymbol{\mu}_h}{\mathrm{tr}(\boldsymbol{Q}\boldsymbol{\Psi}_h)} \tag{2.17}$$

generalizing the single-input single-output (SISO) definition in, e.g., [21, Ch. 1.2.2]. If the line-of-sight SNR component is dominant

[2]For the definition of a projection matrix, see, e.g., [20, Ch. 12.3].

(i.e., large K), then the optimal code design criterion for small $\frac{1}{4}\,\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}$ reduces to maximizing $\boldsymbol{\mu}_h^H\boldsymbol{Q}\boldsymbol{\mu}_h$ with respect to $\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0$.

In the following, we specialize the above results to the case of space–time-separable additive noise.

A. *Space–Time-Separable Noise*

In certain practical applications, it is reasonable to assume that additive noise is separable with respect to space and time, i.e., its spatial covariance is constant in time and its temporal covariance is the same at all sensors (see, for example, [22] and [23]). Therefore, the covariance matrix of the space–time noise snapshot $\boldsymbol{e}$ can be written as

$$\boldsymbol{R}=\boldsymbol{C}^T\otimes\boldsymbol{\Sigma} \tag{2.18}$$

where $\boldsymbol{C}$ and $\boldsymbol{\Sigma}$ are the noise temporal and spatial covariance matrices. Then, (2.6) simplifies to

$$\boldsymbol{Q}=\boldsymbol{U}^*\otimes\boldsymbol{\Sigma}^{-1} \tag{2.19}$$

where

$$\boldsymbol{U}=(\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0)\boldsymbol{C}^{-1}(\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0)^H. \tag{2.20}$$

For large $\frac{1}{4}\,\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Q}\boldsymbol{\Psi}_h^{1/2}=\frac{1}{4}\,\boldsymbol{\Psi}_h^{1/2}(\boldsymbol{U}^*\otimes\boldsymbol{\Sigma}^{-1})\boldsymbol{\Psi}_h^{1/2}$ and $\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0$ having full rank n_{T}, an approximate Chernoff bound follows easily from (2.13)

$$P_{\mathrm{CB\ full\ rank}}\approx\frac{1}{2}\cdot\frac{|\boldsymbol{\Sigma}|^{n_{\mathrm{T}}}}{|\frac{1}{4}\,\boldsymbol{\Psi}_h|\cdot|\boldsymbol{U}|^{n_{\mathrm{R}}}}\cdot\exp(-\boldsymbol{\mu}_h^H\boldsymbol{\Psi}_h^{-1}\boldsymbol{\mu}_h) \tag{2.21}$$

and the determinant criterion simplifies to maximizing

$$|\boldsymbol{U}|=|(\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0)\boldsymbol{C}^{-1}(\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0)^H| \tag{2.22}$$

with respect to $\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0$.

Consider now the case where $\frac{1}{4}\,\boldsymbol{\Psi}_h^{1/2}(\boldsymbol{U}^*\otimes\boldsymbol{\Sigma}^{-1})\boldsymbol{\Psi}_h^{1/2}$ is "small" and the mean of the fading coefficient vector $\boldsymbol{h}$ follows (1.4). Then, an approximate Chernoff bound follows by substituting (2.19) and (1.4) into (2.15)

$$\begin{aligned}P_{\mathrm{CB}}\approx\tfrac{1}{2}\cdot\exp\Big\{&-\tfrac{1}{4}\,|x|^2\cdot\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{U}^*\boldsymbol{a}_{\mathrm{T}}\cdot\boldsymbol{a}_{\mathrm{R}}^H\boldsymbol{\Sigma}^{-1}\boldsymbol{a}_{\mathrm{R}}\\&-\tfrac{1}{4}\,\mathrm{tr}[(\boldsymbol{U}^*\otimes\boldsymbol{\Sigma}^{-1})\boldsymbol{\Psi}_h]\Big\}.\end{aligned} \tag{2.23}$$

Assuming that the line-of-sight SNR component is dominant, the noise is temporally white (i.e., $\boldsymbol{C}=\boldsymbol{I}_N$), and signaling is antipodal (i.e., $\boldsymbol{\Phi}_1=-\boldsymbol{\Phi}_0$), the optimal codes need to maximize $\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{\Phi}_1^*\boldsymbol{\Phi}_1^T\boldsymbol{a}_{\mathrm{T}}$. Under the power constraint $\mathrm{tr}(\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H)=1$, $\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{\Phi}_1^*\boldsymbol{\Phi}_1^T\boldsymbol{a}_{\mathrm{T}}$ is maximized for

$$\boldsymbol{\Phi}_1^*\boldsymbol{\Phi}_1^T=\left(1\Big/\sqrt{\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{a}_{\mathrm{T}}}\right)\cdot\boldsymbol{a}_{\mathrm{T}}\boldsymbol{a}_{\mathrm{T}}^H \tag{2.24}$$

implying that the optimal $\boldsymbol{\Phi}_1$ is a rank-one matrix with the following structure:

$$\begin{aligned}\boldsymbol{\Phi}_1&=-\boldsymbol{\Phi}_0\\&=\frac{1}{\sqrt{\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{a}_{\mathrm{T}}\cdot\sum_{t=1}^{N}|s(t)|^2}}\cdot\boldsymbol{a}_{\mathrm{T}}^*\,[s(1),\ s(2),\ \ldots,\ s(N)]\end{aligned} \tag{2.25}$$

which corresponds to *beamforming*, i.e., "spatial matching" to the line-of-sight transmitter array response. This can be easily seen by rewriting (1.4) as $\mathrm{E}(\boldsymbol{H})=x\cdot\boldsymbol{a}_{\mathrm{R}}\boldsymbol{a}_{\mathrm{T}}^T$, and by observing that $\boldsymbol{a}_{\mathrm{T}}^*=(\boldsymbol{a}_{\mathrm{T}}^T)^H$. Note also that $\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{a}_{\mathrm{T}}=n_{\mathrm{T}}$ if the transmitter array consists of isotropic omnidirectional antennas. In the following discussion, we assume that *both* the receiver and transmitter arrays consist of isotropic omnidirectional antennas, implying that

$$\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{a}_{\mathrm{T}}=n_{\mathrm{T}}\quad\text{and}\quad\boldsymbol{a}_{\mathrm{R}}^H\boldsymbol{a}_{\mathrm{R}}=n_{\mathrm{R}}. \tag{2.26}$$

1) White Noise and Uncorrelated Fading: Assuming that the additive noise is both spatially and temporally white (i.e., $\boldsymbol{R}=\sigma^2\boldsymbol{I}_{n_{\mathrm{R}}N}$), the fading coefficients are uncorrelated with equal variances ψ_h^2 (i.e., $\boldsymbol{\Psi}_h=\psi_h^2\boldsymbol{I}_{n_{\mathrm{T}}n_{\mathrm{R}}}$), and the mean of the fading coefficients follows (1.4), (2.10) simplifies to

$$\begin{aligned}P_{\mathrm{CB}}=&\ \frac{1}{2}\cdot\frac{1}{|\boldsymbol{I}_{n_{\mathrm{T}}}+\psi_h^2/(4\sigma^2)\cdot\boldsymbol{U}|^{n_{\mathrm{R}}}}\\&\cdot\exp\left[-\frac{n_{\mathrm{R}}|x|^2}{4\sigma^2}\cdot\boldsymbol{a}_{\mathrm{T}}^H\left(\boldsymbol{I}_{n_{\mathrm{T}}}+\frac{\psi_h^2}{4\sigma^2}\boldsymbol{U}^*\right)^{-1}\boldsymbol{U}^*\boldsymbol{a}_{\mathrm{T}}\right]\end{aligned} \tag{2.27}$$

where

$$\boldsymbol{U}=(\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0)(\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0)^H. \tag{2.28}$$

Here we have used the fact that $\boldsymbol{a}_{\mathrm{R}}^H\boldsymbol{a}_{\mathrm{R}}=n_{\mathrm{R}}$, which holds if antennas at the receiver are isotropic (see also (2.26)).

For full-rank $\boldsymbol{U}$ and large $\psi_h^2/(4\sigma^2)\cdot\boldsymbol{U}$, (2.27) is approximated as

$$P_{\mathrm{CB\ full\ rank}}\approx\frac{1}{2}\cdot\frac{1}{|\psi_h^2/(4\sigma^2)\cdot\boldsymbol{U}|^{n_{\mathrm{R}}}}\cdot\exp\left[-\frac{n_{\mathrm{T}}\cdot n_{\mathrm{R}}\cdot|x|^2}{\psi_h^2}\right] \tag{2.29}$$

where we have used the fact that $\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{a}_{\mathrm{T}}=n_{\mathrm{T}}$, which holds if the transmitter array consists of isotropic antennas (see also (2.26)). Equation (2.29) can be obtained by substituting $\boldsymbol{\Sigma}=\sigma^2\boldsymbol{I}_{n_{\mathrm{R}}}$, $\boldsymbol{C}=\boldsymbol{I}_N$, $\boldsymbol{\Psi}_h=\psi_h^2\boldsymbol{I}_{n_{\mathrm{T}}n_{\mathrm{R}}}$, (1.4), and (2.26) into (2.21).

For small $\psi_h^2/(4\sigma^2)\cdot\boldsymbol{U}$, (2.27) approximately equals

$$P_{\mathrm{CB}}\approx\frac{1}{2}\cdot\exp\left[-\frac{n_{\mathrm{R}}|x|^2}{4\sigma^2}\cdot\boldsymbol{a}_{\mathrm{T}}^H\boldsymbol{U}^*\boldsymbol{a}_{\mathrm{T}}-\frac{n_{\mathrm{R}}\psi_h^2}{4\sigma^2}\cdot\mathrm{tr}(\boldsymbol{U})\right]. \tag{2.30}$$

It is interesting to examine optimal signaling schemes under the two scenarios above. For simplicity, let us concentrate on antipodal signaling

$$\boldsymbol{\Phi}_1=-\boldsymbol{\Phi}_0 \tag{2.31}$$

implying that $\boldsymbol{U}=4\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H$. We define normalized[3] *line-of-sight* and *scattering* SNRs as

$$\mathrm{SNR}_{\mathrm{LOS}}=\frac{|x|^2}{\sigma^2},\qquad\mathrm{SNR}_{\mathrm{SC}}=\frac{\psi_h^2}{\sigma^2}. \tag{2.32}$$

We also impose a power constraint on the transmitted symbols

$$\mathrm{tr}(\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H)=\mathrm{tr}(\boldsymbol{\Phi}_0\boldsymbol{\Phi}_0^H)=1. \tag{2.33}$$

Large, Full-Rank $\mathrm{SNR}_{\mathrm{SC}}\cdot\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H$: For large, full-rank $\mathrm{SNR}_{\mathrm{SC}}\cdot\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H$, the optimal antipodal codes (which minimize (2.29)) are constructed by maximizing

$$|\mathrm{SNR}_{\mathrm{SC}}\cdot\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H| \tag{2.34}$$

subject to (2.33). Clearly, the optimal $\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H$ has all eigenvalues equal to $1/n_{\mathrm{T}}$; hence,

$$|\mathrm{SNR}_{\mathrm{SC}}\cdot\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H|_{\mathrm{MAX}}=\left(\frac{\mathrm{SNR}_{\mathrm{SC}}}{n_{\mathrm{T}}}\right)^{n_{\mathrm{T}}} \tag{2.35}$$

[3]The SNRs defined in (2.32) are normalized so that they do not depend on the numbers of receiver and transmitter antennas.

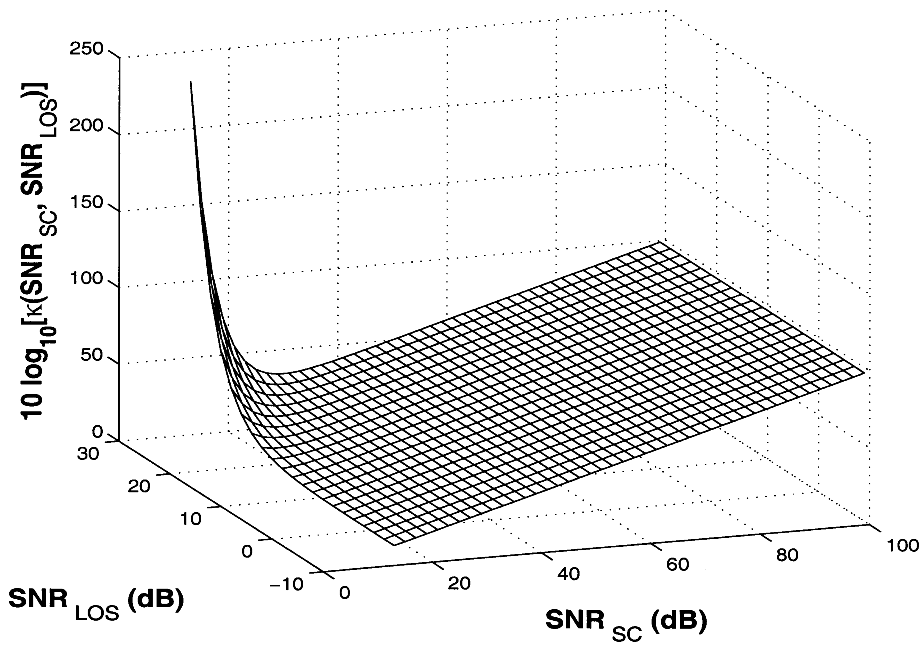

Fig. 1. The approximate single-input–single-output gain as a function of line-of-sight and scattering SNRs.

see also [24] and [25]. In addition, the condition that $\mathrm{SNR}_{\mathrm{SC}} \cdot \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_1^H$ is large simplifies to the requirement that $\mathrm{SNR}_{\mathrm{SC}}/n_{\mathrm{T}}$ is large. For example, orthogonal designs, satisfying

$$\boldsymbol{\Phi}_1 \boldsymbol{\Phi}_1^H = (1/n_{\mathrm{T}}) \cdot \boldsymbol{I}_{n_{\mathrm{T}}} \tag{2.36}$$

are optimal. Substituting (2.35) into (2.29), we obtain an approximate Chernoff bound for optimal antipodal signaling

$$\begin{aligned} &P_{\mathrm{CB,\,opt}}\Big|_{\mathrm{large}\ \mathrm{SNR}_{\mathrm{SC}}/n_{\mathrm{T}}} \\ &\approx \frac{1}{2} \cdot \left(\frac{\mathrm{SNR}_{\mathrm{SC}}}{n_{\mathrm{T}}}\right)^{-n_{\mathrm{T}} n_{\mathrm{R}}} \cdot \exp\left[-\frac{n_{\mathrm{T}} \cdot n_{\mathrm{R}} \cdot |x|^2}{\psi_h^2}\right] \\ &= \frac{1}{2} \cdot \left[\frac{1}{n_{\mathrm{T}}} \cdot \kappa(\mathrm{SNR}_{\mathrm{SC}},\, \mathrm{SNR}_{\mathrm{LOS}})\right]^{-n_{\mathrm{T}} n_{\mathrm{R}}} \end{aligned} \tag{2.37}$$

where

$$\kappa(\mathrm{SNR}_{\mathrm{SC}},\, \mathrm{SNR}_{\mathrm{LOS}}) = \mathrm{SNR}_{\mathrm{SC}} \cdot \exp\left(\frac{\mathrm{SNR}_{\mathrm{LOS}}}{\mathrm{SNR}_{\mathrm{SC}}}\right). \tag{2.38}$$

In a SISO system, (2.37) simplifies to $\frac{1}{2} \cdot 1/\kappa(\mathrm{SNR}_{\mathrm{SC}},\, \mathrm{SNR}_{\mathrm{LOS}})$. Therefore, we refer to $\kappa(\mathrm{SNR}_{\mathrm{SC}},\, \mathrm{SNR}_{\mathrm{LOS}})$ as an approximate "gain" of a SISO system. In Fig. 1, we show $\kappa(\mathrm{SNR}_{\mathrm{SC}},\, \mathrm{SNR}_{\mathrm{LOS}})$ in decibels as a function of scattering and line-of-sight SNRs, also in decibels. To meet the large $\mathrm{SNR}_{\mathrm{SC}}/n_{\mathrm{T}}$ requirement, we consider only the values of scattering $\mathrm{SNR}_{\mathrm{SC}}$ larger than 20 (13 dB). For fixed line-of-sight $\mathrm{SNR}_{\mathrm{LOS}}$ and for $\mathrm{SNR}_{\mathrm{LOS}} < \mathrm{SNR}_{\mathrm{SC}}$, $\kappa(\mathrm{SNR}_{\mathrm{SC}},\, \mathrm{SNR}_{\mathrm{LOS}})$ grows linearly with $\mathrm{SNR}_{\mathrm{SC}}$. For fixed $\mathrm{SNR}_{\mathrm{SC}}$ and $\mathrm{SNR}_{\mathrm{SC}} < \mathrm{SNR}_{\mathrm{LOS}}$, $\kappa(\mathrm{SNR}_{\mathrm{SC}},\, \mathrm{SNR}_{\mathrm{LOS}})$ grows exponentially with $\mathrm{SNR}_{\mathrm{LOS}}$. Also, note that the minimum value of $\kappa(\mathrm{SNR}_{\mathrm{SC}},\, \mathrm{SNR}_{\mathrm{LOS}})$ in Fig. 1 is 20, which corresponds to the Chernoff bound of 0.025 in a SISO system.

Small $\mathrm{SNR}_{\mathrm{SC}} \cdot \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_1^H$: For small $\mathrm{SNR}_{\mathrm{SC}} \cdot \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_1^H$ and nonzero $\mathrm{SNR}_{\mathrm{LOS}}$, the optimal antipodal codes (which minimize (2.30)) are constructed by maximizing $\mathrm{SNR}_{\mathrm{LOS}} \cdot \boldsymbol{a}_{\mathrm{T}}^H \boldsymbol{\Phi}_1^* \boldsymbol{\Phi}_1^T \boldsymbol{a}_{\mathrm{T}} + \mathrm{SNR}_{\mathrm{SC}} \cdot \mathrm{tr}(\boldsymbol{\Phi}_1 \boldsymbol{\Phi}_1^H)$, subject to (2.33), and are given in (2.25). Then, $\boldsymbol{\Phi}_1^* \boldsymbol{\Phi}_1^T = (1/n_{\mathrm{T}}) \cdot \boldsymbol{a}_{\mathrm{T}} \boldsymbol{a}_{\mathrm{T}}^H$ and the condition that $\mathrm{SNR}_{\mathrm{SC}} \cdot \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_1^H$ is small simplifies to the requirement that $\mathrm{SNR}_{\mathrm{SC}}$ is small. Substituting these results into (2.30), we obtain an approximate Chernoff bound for optimal antipodal signaling

$$\begin{aligned} &P_{\mathrm{CB,\,opt}}\Big|_{\mathrm{small}\ \mathrm{SNR}_{\mathrm{SC}}} \\ &\approx \frac{1}{2} \cdot \exp\left[-n_{\mathrm{T}} n_{\mathrm{R}} \cdot \mathrm{SNR}_{\mathrm{LOS}} - n_{\mathrm{R}} \cdot \mathrm{SNR}_{\mathrm{SC}}\right] \\ &= \frac{1}{2} \cdot \left[\exp\left(\mathrm{SNR}_{\mathrm{LOS}} + \frac{1}{n_{\mathrm{T}}}\, \mathrm{SNR}_{\mathrm{SC}}\right)\right]^{-n_{\mathrm{T}} n_{\mathrm{R}}}. \end{aligned} \tag{2.39}$$

Chernoff Bound for Antipodal Orthogonal Designs: We compute the Chernoff bound for antipodal orthogonal designs by substituting (2.31) and (2.36) into (2.27)

$$\begin{aligned} &P_{\mathrm{CB}}\Big|_{\substack{\boldsymbol{\Phi}_1 = -\boldsymbol{\Phi}_0 \\ \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_1^H = (1/n_{\mathrm{T}}) \cdot \boldsymbol{I}_{n_{\mathrm{T}}}}} \\ &= \frac{1}{2} \cdot \left[\left(1 + \frac{\mathrm{SNR}_{\mathrm{SC}}}{n_{\mathrm{T}}}\right) \exp\left(\frac{\frac{\mathrm{SNR}_{\mathrm{LOS}}}{n_{\mathrm{T}}}}{1 + \frac{\mathrm{SNR}_{\mathrm{SC}}}{n_{\mathrm{T}}}}\right)\right]^{-n_{\mathrm{T}} n_{\mathrm{R}}}. \end{aligned} \tag{2.40}$$

Interestingly, the Ricean K-factor in (2.17) simplifies to

$$K = \frac{\mathrm{SNR}_{\mathrm{LOS}}}{\mathrm{SNR}_{\mathrm{SC}}}. \tag{2.41}$$

Observe that (2.40) decreases exponentially as the number of receiver antennas n_{R} grows. Increasing the number of transmitter antennas n_{T} results in two opposite effects. Clearly, the diversity gain (equal to $n_{\mathrm{T}} n_{\mathrm{R}}$) increases with n_{T} thereby reducing error probability. However, due to the power constraint (2.33), the signal power per transmitter antenna decreases, which results in larger error probability per diversity branch. These two effects can also be seen by observing (2.40), which decreases exponentially with n_{T}, but the argument of the exponent (in square brackets) also decreases with n_{T}. In this case, the first effect is dominant: (2.40) decreases as n_{T} grows for all possible $\mathrm{SNR}_{\mathrm{LOS}}$ and $\mathrm{SNR}_{\mathrm{SC}}$. Consequently, the corresponding pairwise error probability decreases with n_{T} as well, see (2.11). Hence, for the fading, signal, and noise models considered here, it is desirable (in terms of minimizing the pairwise error probability) to use as many transmitter antennas as possible.[4] This is not true for noncoherent signaling, see Section III-C. Information-theoretic criteria have also been used to determine the optimal number of transmitter antennas. For example, maximizing nonergodic Shannon capacity for coherent low-rank channels was proposed in [26] and [27], resulting in optimal $n_{\rm T}$ that is equal to the channel rank.

III. Noncoherent Signaling

We compute Chernoff bounds for *noncoherent* signaling, i.e., assuming that the channel is not known to the receiver. Consider the measurement model (1.2) with known noise covariance $\boldsymbol{R}$ and unknown channel coefficient vector $\boldsymbol{h}$, described by known mean $\boldsymbol{\mu}_h$ and known covariance $\boldsymbol{\Psi}_h$, see (1.3). (Efficient methods for estimating statistical properties of MIMO Ricean-fading channels have been recently proposed in [28] and [29].) As before, we consider testing the hypothesis $\rm H_1$: $\boldsymbol{\Phi}_1$ transmitted versus the alternative $\rm H_0$: $\boldsymbol{\Phi}_0$ transmitted and assume that $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_0$ are equiprobable. In this scenario, the received measurement vector $\boldsymbol{y}$ under $\rm H_1$ is a complex multivariate normal with mean $\boldsymbol{Z}_1\boldsymbol{\mu}_h$ and covariance $\boldsymbol{R}_{y,1} = \boldsymbol{Z}_1\boldsymbol{\Psi}_h\boldsymbol{Z}_1^H + \boldsymbol{R}$, whereas under $\rm H_0$ it is a complex multivariate normal with mean $\boldsymbol{Z}_0\boldsymbol{\mu}_h$ and covariance $\boldsymbol{R}_{y,0} = \boldsymbol{Z}_0\boldsymbol{\Psi}_h\boldsymbol{Z}_0^H + \boldsymbol{R}$. Note that the above likelihood functions are *marginal likelihoods*, where the unknown channel vector $\boldsymbol{h}$ has been integrated out with respect to its prior distribution. (It is also possible to construct concentrated-likelihood detectors that do not require knowledge of $\boldsymbol{\mu}_h$, $\boldsymbol{\Psi}_h$, or $\boldsymbol{R}$ at the receiver. In these detectors, the likelihood function is concentrated with respect to the unknown channel and noise parameters [15], [30], [31] or statistical channel parameters [31] in a manner similar to that used to derive deterministic and stochastic maximum-likelihood methods for sensor array processing [32]. Performance analysis of concentrated-likelihood detectors is beyond the scope of this correspondence.)

As before, denote the pdf of $\boldsymbol{y}$ under $\rm H_i$ as $p_i(\boldsymbol{y})$, $i = 0, 1$. Then, following [18, Ch. 2.7], [19, Ch. 3.4], and [8, Appendix B], the Chernoff bound on pairwise error probability for deciding between $\rm H_1$ and $\rm H_0$ is

$$P_{\rm CB}(\lambda) = \tfrac{1}{2}\exp[\xi(\lambda)], \qquad 0 \le \lambda \le 1 \tag{3.1}$$

where

$$\xi(\lambda) = \ln {\rm E}\{\exp[\lambda \ln p_0(\boldsymbol{y}) - \lambda \ln p_1(\boldsymbol{y})] \,|\, {\rm H}_1\}. \tag{3.2}$$

Here, (3.2) becomes

$$\begin{aligned}\xi(\lambda) = \ln \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} &\frac{1}{|\pi\boldsymbol{R}_{y,0}|^{\lambda}\cdot|\pi\boldsymbol{R}_{y,1}|^{1-\lambda}} \\ &\cdot \exp\Big[-\lambda(\boldsymbol{y}-\boldsymbol{Z}_0\boldsymbol{\mu}_h)^H\boldsymbol{R}_{y,0}^{-1}(\boldsymbol{y}-\boldsymbol{Z}_0\boldsymbol{\mu}_h) \\ &\quad - (1-\lambda)\,(\boldsymbol{y}-\boldsymbol{Z}_1\boldsymbol{\mu}_h)^H\boldsymbol{R}_{y,1}^{-1}(\boldsymbol{y}-\boldsymbol{Z}_1\boldsymbol{\mu}_h)\Big]\, d\boldsymbol{y}\end{aligned} \tag{3.3}$$

which can be computed by applying Lemma 1 with $\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{Z}_1\boldsymbol{\mu}_h$, $\boldsymbol{B} = 2\cdot[\lambda\boldsymbol{R}_{y,0}^{-1} + (1-\lambda)\boldsymbol{R}_{y,1}^{-1}]$, $\boldsymbol{A} = \boldsymbol{0}$, $\boldsymbol{a} = \boldsymbol{0}$, $\boldsymbol{b} = 2\lambda\cdot\boldsymbol{R}_{y,0}^{-1}\cdot(\boldsymbol{Z}_1-\boldsymbol{Z}_0)\boldsymbol{\mu}_h$, $a_0 = 1/(|\pi\boldsymbol{R}_{y,0}|^{\lambda}\cdot|\pi\boldsymbol{R}_{y,1}|^{1-\lambda})$, and $b_0 = \lambda\boldsymbol{\mu}_h^H(\boldsymbol{Z}_1-\boldsymbol{Z}_0)^H\cdot\boldsymbol{R}_{y,0}^{-1}\cdot(\boldsymbol{Z}_1-\boldsymbol{Z}_0)\boldsymbol{\mu}_h$:

$$\begin{aligned}\xi(\lambda) = &-\ln[|\lambda\boldsymbol{R}_{y,0}^{-1}+(1-\lambda)\boldsymbol{R}_{y,1}^{-1}|\cdot|\boldsymbol{R}_{y,0}|^{\lambda}\cdot|\boldsymbol{R}_{y,1}|^{1-\lambda}] \\ &- \boldsymbol{\mu}_h^H(\boldsymbol{Z}_1-\boldsymbol{Z}_0)^H\left(\frac{1}{\lambda}\boldsymbol{R}_{y,0}+\frac{1}{1-\lambda}\boldsymbol{R}_{y,1}\right)^{-1}(\boldsymbol{Z}_1-\boldsymbol{Z}_0)\boldsymbol{\mu}_h\end{aligned} \tag{3.4a}$$

$$\begin{aligned}= \ln\left\{\frac{|\boldsymbol{R}_{y,1}|^{\lambda}\cdot|\boldsymbol{R}_{y,0}|^{1-\lambda}}{|\lambda\boldsymbol{R}_{y,1}+(1-\lambda)\boldsymbol{R}_{y,0}|}\right\} &- \lambda(1-\lambda)\cdot\boldsymbol{\mu}_h^H(\boldsymbol{Z}_1-\boldsymbol{Z}_0)^H \\ &\cdot[\lambda\boldsymbol{R}_{y,1}+(1-\lambda)\boldsymbol{R}_{y,0}]^{-1}(\boldsymbol{Z}_1-\boldsymbol{Z}_0)\boldsymbol{\mu}_h\end{aligned} \tag{3.4b}$$

[4]Note, however, that $P_{\rm CB}$ in (2.40) converges to $(1/2)\cdot\exp[-n_{\rm R}\cdot({\rm SNR}_{\rm SC}+{\rm SNR}_{\rm LOS})]$ as $n_{\rm T} \to \infty$.

where, to compute (3.4a), we have used the identity

$$\begin{aligned}-&\left(\frac{1}{\lambda}\boldsymbol{R}_{y,0}+\frac{1}{1-\lambda}\boldsymbol{R}_{y,1}\right)^{-1} \\ &= \lambda^2\boldsymbol{R}_{y,0}^{-1}\cdot[\lambda\boldsymbol{R}_{y,0}^{-1}+(1-\lambda)\boldsymbol{R}_{y,1}^{-1}]^{-1}\cdot\boldsymbol{R}_{y,0}^{-1} - \lambda\boldsymbol{R}_{y,0}^{-1}\end{aligned} \tag{3.5}$$

which follows from the matrix inversion lemma in [20, eq. (2.22), p. 424]. Then, the Chernoff bound for a given λ is

$$\begin{aligned}P_{\rm CB}(\lambda) &= \frac{1}{2}\cdot\frac{|\boldsymbol{R}_{y,1}|^{\lambda}\cdot|\boldsymbol{R}_{y,0}|^{1-\lambda}}{|\lambda\boldsymbol{R}_{y,1}+(1-\lambda)\boldsymbol{R}_{y,0}|} \\ &\quad\cdot\exp\Big\{-\lambda(1-\lambda)\cdot\boldsymbol{\mu}_h^H(\boldsymbol{Z}_1-\boldsymbol{Z}_0)^H \\ &\qquad\cdot[\lambda\boldsymbol{R}_{y,1}+(1-\lambda)\boldsymbol{R}_{y,0}]^{-1}(\boldsymbol{Z}_1-\boldsymbol{Z}_0)\boldsymbol{\mu}_h\Big\} \\ &= \frac{1}{2}\cdot\frac{|\boldsymbol{R}_{y,1}|^{\lambda}\cdot|\boldsymbol{R}_{y,0}|^{1-\lambda}}{|\boldsymbol{R}+\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)\boldsymbol{Z}^H|} \\ &\quad\cdot\exp\Big\{-\lambda(1-\lambda)\cdot\boldsymbol{\mu}_h^H(\boldsymbol{Z}_1-\boldsymbol{Z}_0)^H \\ &\qquad\cdot\Big[\boldsymbol{R}+\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)\boldsymbol{Z}^H\Big]^{-1}(\boldsymbol{Z}_1-\boldsymbol{Z}_0)\boldsymbol{\mu}_h\Big\}\end{aligned} \tag{3.6}$$

where $0 \le \lambda \le 1$, and

$$\boldsymbol{Z} = [\boldsymbol{Z}_1,\ \boldsymbol{Z}_0] \tag{3.7a}$$

$$\hat{\boldsymbol{\Psi}}_h(\lambda) = \begin{bmatrix}\lambda\boldsymbol{\Psi}_h & 0 \\ 0 & (1-\lambda)\boldsymbol{\Psi}_h\end{bmatrix}. \tag{3.7b}$$

Observe that

$$\begin{aligned}&\boldsymbol{Z}^H[\boldsymbol{R}+\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)\boldsymbol{Z}^H]^{-1}\boldsymbol{Z} \\ &\quad= \boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z} - \boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z} \\ &\qquad\cdot\Big[\hat{\boldsymbol{\Psi}}_h(\lambda)^{-1}+\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\Big]^{-1}\cdot\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\end{aligned} \tag{3.8a}$$

$$\begin{aligned}&\quad= \boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z} - \boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda) \\ &\qquad\cdot[\boldsymbol{I}_{2n_{\rm T}n_{\rm R}}+\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)]^{-1}\cdot\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\end{aligned} \tag{3.8b}$$

$$\begin{aligned}&\quad= [\boldsymbol{I}_{2n_{\rm T}n_{\rm R}}+\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda) - \boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)] \\ &\qquad\cdot[\boldsymbol{I}_{2n_{\rm T}n_{\rm R}}+\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)]^{-1}\cdot\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\end{aligned} \tag{3.8c}$$

$$\quad= [\boldsymbol{I}_{2n_{\rm T}n_{\rm R}}+\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)]^{-1}\cdot\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}. \tag{3.8d}$$

Here, the right-hand side of (3.8a) follows by using the matrix inversion lemma in [20, eq. (2.22), p. 424]. Also,

$$\begin{aligned}&\frac{|\boldsymbol{R}_{y,1}|^{\lambda}\cdot|\boldsymbol{R}_{y,0}|^{1-\lambda}}{|\lambda\boldsymbol{R}_{y,1}+(1-\lambda)\boldsymbol{R}_{y,0}|} \\ &\quad= \frac{|\boldsymbol{R}+\boldsymbol{Z}_1\boldsymbol{\Psi}_h\boldsymbol{Z}_1^H|^{\lambda}\cdot|\boldsymbol{R}+\boldsymbol{Z}_0\boldsymbol{\Psi}_h\boldsymbol{Z}_0^H|^{1-\lambda}}{|\boldsymbol{R}+\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)\boldsymbol{Z}^H|}\end{aligned} \tag{3.9a}$$

$$\quad= \frac{|\boldsymbol{R}|\cdot|\boldsymbol{\Psi}_h|\cdot|\boldsymbol{\Psi}_h^{-1}+\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1|^{\lambda}\cdot|\boldsymbol{\Psi}_h^{-1}+\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0|^{1-\lambda}}{|\boldsymbol{R}|\cdot|\hat{\boldsymbol{\Psi}}_h(\lambda)|\cdot|\hat{\boldsymbol{\Psi}}_h(\lambda)^{-1}+\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}|} \tag{3.9b}$$

$$\begin{aligned}&\quad= |\boldsymbol{I}_{n_{\rm T}n_{\rm R}}+\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2}|^{\lambda} \\ &\qquad\cdot|\boldsymbol{I}_{n_{\rm T}n_{\rm R}}+\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}|^{1-\lambda} \\ &\qquad\cdot|\boldsymbol{I}_{2n_{\rm T}n_{\rm R}}+\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}|^{-1}\end{aligned} \tag{3.9c}$$

where (3.9b) follows by repeatedly applying the determinant formula in [20, Theorem 18.1.1, p. 416] to the numerator and denominator terms in (3.9a). Using (3.8) and (3.9), we rewrite (3.6) as (3.10), shown at the bottom of the page. (Observe also that (3.11) holds.) Equation (3.10) is the first closed-form Chernoff bound expression for noncoherent signaling in a MIMO Ricean-fading channel. A special case for orthogonal signaling in a SISO channel with independent and identically distributed (i.i.d.) fading and white noise was derived in [33, eq. (12)]. For unitary space–time codes in an i.i.d. Rayleigh-fading channel and spatially and temporally white noise, a Chernoff bound was computed in [8], see also [10].

A. Large $\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2}$, $\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}$, *and* $\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}$

If the matrices $\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2}$, $\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}$, and $\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}$ are "large" then (3.10) can be approximated using (3.12), shown at the bottom of the page. Note that

$$\mathrm{rank}(\boldsymbol{Z}_i) = n_{\mathrm{R}} \cdot \mathrm{rank}(\boldsymbol{\Phi}_i), \qquad i \in \{0, 1\}$$

and

$$\mathrm{rank}(\boldsymbol{Z}) = n_{\mathrm{R}} \cdot \mathrm{rank}([\boldsymbol{\Phi}_1^T, \boldsymbol{\Phi}_0^T]).$$

Equal Energy Signaling: Consider the case where

$$\boldsymbol{W} = \boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1 = \boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0 \tag{3.13}$$

which may be viewed as a multivariate extension of the *equal energy* condition (and is closely related to unitary space–time codes in [8]; see also Section III-C). For full-rank antipodal signaling (i.e., $\boldsymbol{\Phi}_1 = -\boldsymbol{\Phi}_0$ has full rank n_{T}), (3.12) simplifies to

$$P_{\mathrm{CB}}(\lambda)|_{\boldsymbol{\Phi}_1 = -\boldsymbol{\Phi}_0} \approx \tfrac{1}{2} \cdot \exp\left\{-4\lambda(1-\lambda) \cdot \boldsymbol{\mu}_h^H \boldsymbol{\Psi}_h^{-1}\boldsymbol{\mu}_h\right\} \tag{3.14}$$

which is minimized for $\lambda = 1/2$. As expected, antipodal signaling performs poorly in this scenario: there is no diversity gain and, additionally, this scheme breaks down if the channel coefficients have zero mean.

If, in addition to the "equal energy" condition (3.13), we assume that $[\boldsymbol{\Phi}_1^T, \boldsymbol{\Phi}_0^T]$ has full rank equal to $2n_{\mathrm{T}}$ (or, equivalently, $\boldsymbol{Z}$ has full rank equal to $2n_{\mathrm{T}}n_{\mathrm{R}}$), then (3.12) becomes

$$P_{\mathrm{CB}}(\lambda)\Bigg|_{\substack{\boldsymbol{W} = \boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1 = \boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0 \\ \mathrm{rank}\,(\boldsymbol{Z}) = 2n_{\mathrm{T}}n_{\mathrm{R}}}} \approx \frac{|\boldsymbol{W}|}{|\lambda(1-\lambda)\cdot\boldsymbol{\Psi}_h| \cdot |\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}|}\exp\left(-\boldsymbol{\mu}_h^H\boldsymbol{\Psi}_h^{-1}\boldsymbol{\mu}_h\right) \tag{3.15}$$

which is minimized for $\lambda = 1/2$. To achieve the full rank $2n_{\mathrm{T}}$ of $[\boldsymbol{\Phi}_1^T, \boldsymbol{\Phi}_0^T]$, the number of time samples must be equal to or larger than twice the number of transmitter antennas, i.e., $N \geq 2n_{\mathrm{T}}$. Based on (3.15), we formulate a determinant code optimization criterion for this scenario: maximize

$$\begin{aligned}
\mathrm{dist}(\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_0) &= \frac{1}{|\boldsymbol{W}|} \cdot \begin{vmatrix} \boldsymbol{W} & \boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0 \\ \boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1 & \boldsymbol{W} \end{vmatrix} && \text{(3.16a)}\\
&= |\boldsymbol{W} - \boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{W}^{-1}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0| && \text{(3.16b)}\\
&= |\boldsymbol{W} - \boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{W}^{-1}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1| && \text{(3.16c)}
\end{aligned}$$

where (3.16b) and (3.16c) follow by using [20, Theorem 13.3.8, p. 188]. It is clearly desirable that

$$\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0 = \boldsymbol{0} \tag{3.17}$$

which can be viewed as a condition for *orthogonality* between codes $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_0$. Then, (3.16) simplifies to

$$\mathrm{dist}(\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_0)|_{\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0 = \boldsymbol{0}} = |\boldsymbol{W}|. \tag{3.18}$$

For space–time separable noise (i.e., $\boldsymbol{R}$ satisfies (2.18)), the equal energy condition (3.13) simplifies to

$$\boldsymbol{V} = \boldsymbol{\Phi}_1\boldsymbol{C}^{-1}\boldsymbol{\Phi}_1^H = \boldsymbol{\Phi}_0\boldsymbol{C}^{-1}\boldsymbol{\Phi}_0^H \tag{3.19}$$

and the noncoherent determinant criterion in (3.16) becomes

$$\begin{aligned}
\mathrm{dist}(\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_0) &= \frac{\left|\boldsymbol{V} - \boldsymbol{\Phi}_0\boldsymbol{C}^{-1}\boldsymbol{\Phi}_1^H\boldsymbol{V}^{-1}\boldsymbol{\Phi}_1\boldsymbol{C}^{-1}\boldsymbol{\Phi}_0^H\right|^{n_{\mathrm{R}}}}{|\boldsymbol{\Sigma}|^{n_{\mathrm{T}}}}\\
&= \frac{\left|\boldsymbol{V} - \boldsymbol{\Phi}_1\boldsymbol{C}^{-1}\boldsymbol{\Phi}_0^H\boldsymbol{V}^{-1}\boldsymbol{\Phi}_0\boldsymbol{C}^{-1}\boldsymbol{\Phi}_1^H\right|^{n_{\mathrm{R}}}}{|\boldsymbol{\Sigma}|^{n_{\mathrm{T}}}}.
\end{aligned} \tag{3.20}$$

If the orthogonality condition $\boldsymbol{\Phi}_1\boldsymbol{C}^{-1}\boldsymbol{\Phi}_0^H = \boldsymbol{0}$ holds (which follows by simplifying (3.17)), the noncoherent determinant criterion reduces to maximizing $|\boldsymbol{V}| = |\boldsymbol{\Phi}_1\boldsymbol{C}^{-1}\boldsymbol{\Phi}_1^H|$.

B. Small $\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2}$, $\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}$, *and* $\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}$

If the matrices $\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2}$, $\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}$, and $\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}$ are "small" then we may approximate (3.10) as

$$\begin{aligned}
P_{\mathrm{CB}}(\lambda) \approx \tfrac{1}{2} \cdot \exp\Big(-\lambda(1-\lambda) \cdot \Big\{&\boldsymbol{\mu}_h^H\boldsymbol{Q}\boldsymbol{\mu}_h\\
&+ \tfrac{1}{2}\mathrm{tr}[(\boldsymbol{Z}_1\boldsymbol{\Psi}_h\boldsymbol{Z}_1^H\boldsymbol{R}^{-1} - \boldsymbol{Z}_0\boldsymbol{\Psi}_h\boldsymbol{Z}_0^H\boldsymbol{R}^{-1})^2]\Big\}\Big)
\end{aligned} \tag{3.21}$$

where $\boldsymbol{Q}$ was defined in (2.6). To derive (3.21), we have used the following approximation:

$$\ln|\boldsymbol{I}_n + \boldsymbol{A}| \approx \mathrm{tr}(\boldsymbol{A}) - \tfrac{1}{2}\mathrm{tr}(\boldsymbol{A}^2) \tag{3.22}$$

$$\begin{aligned}
P_{\mathrm{CB}}(\lambda) = \frac{1}{2} \cdot &\frac{|\boldsymbol{I}_{n_{\mathrm{T}}n_{\mathrm{R}}} + \boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2}|^{\lambda} \cdot |\boldsymbol{I}_{n_{\mathrm{T}}n_{\mathrm{R}}} + \boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}|^{1-\lambda}}{\left|\boldsymbol{I}_{2n_{\mathrm{T}}n_{\mathrm{R}}} + \hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\right|}\\
&\cdot \exp\left\{-\lambda(1-\lambda)\cdot\left[\boldsymbol{\mu}_h^H, -\boldsymbol{\mu}_h^H\right]\cdot\left[\boldsymbol{I}_{2n_{\mathrm{T}}n_{\mathrm{R}}} + \boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)\right]^{-1}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\cdot\begin{bmatrix}\boldsymbol{\mu}_h\\ -\boldsymbol{\mu}_h\end{bmatrix}\right\}
\end{aligned} \tag{3.10}$$

$$\begin{aligned}
&\boldsymbol{I}_{2n_{\mathrm{T}}n_{\mathrm{R}}} + \hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\\
&= \begin{bmatrix} \boldsymbol{I}_{n_{\mathrm{T}}n_{\mathrm{R}}} + \lambda\cdot\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2} & \sqrt{\lambda(1-\lambda)}\cdot\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}\\ \sqrt{\lambda(1-\lambda)}\cdot\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2} & \boldsymbol{I}_{n_{\mathrm{T}}n_{\mathrm{R}}} + (1-\lambda)\cdot\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}\end{bmatrix}
\end{aligned} \tag{3.11}$$

$$\begin{aligned}
P_{\mathrm{CB}}(\lambda) \approx \frac{1}{2} \cdot &\frac{\left|\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1\boldsymbol{\Psi}_h^{1/2}\right|_{\mathrm{rank}(\boldsymbol{Z}_1)}^{\lambda} \cdot \left|\boldsymbol{\Psi}_h^{1/2}\boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\boldsymbol{\Psi}_h^{1/2}\right|_{\mathrm{rank}(\boldsymbol{Z}_0)}^{1-\lambda}}{\left|\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\right|_{\mathrm{rank}(\boldsymbol{Z})}}\\
&\cdot \exp\left\{-\lambda(1-\lambda)\cdot\left[\boldsymbol{\mu}_h^H, -\boldsymbol{\mu}_h^H\right]\hat{\boldsymbol{\Psi}}_h(\lambda)^{-1/2}\cdot\boldsymbol{\Pi}\left(\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\boldsymbol{Z}^H\boldsymbol{R}^{-1}\boldsymbol{Z}\hat{\boldsymbol{\Psi}}_h(\lambda)^{1/2}\right)\cdot\hat{\boldsymbol{\Psi}}_h(\lambda)^{-1/2}\begin{bmatrix}\boldsymbol{\mu}_h\\ -\boldsymbol{\mu}_h\end{bmatrix}\right\}.
\end{aligned} \tag{3.12}$$

which holds for a "small" $n \times n$ positive-semidefinite Hermitian matrix $\boldsymbol{A}$. The approximation (3.22) was applied to (logarithms of) the three determinant terms in (3.10). Note that (3.21) is minimized for $\lambda = 1/2$. Clearly, the optimal codes need to maximize

$$\mathrm{tr}[\boldsymbol{Q}\boldsymbol{\mu}_h\boldsymbol{\mu}_h^H + \tfrac{1}{2}(\boldsymbol{Z}_1\boldsymbol{\Psi}_h\boldsymbol{Z}_1^H\boldsymbol{R}^{-1} - \boldsymbol{Z}_0\boldsymbol{\Psi}_h\boldsymbol{Z}_0^H\boldsymbol{R}^{-1})^2] \tag{3.23}$$

which further simplifies to

$$\boldsymbol{\mu}_h^H\boldsymbol{Q}\boldsymbol{\mu}_h$$

if the line-of-sight component is dominant in (3.23), i.e.,

$$\boldsymbol{\mu}_h^H\boldsymbol{Q}\boldsymbol{\mu}_h \gg \tfrac{1}{2}\,\mathrm{tr}[(\boldsymbol{Z}_1\boldsymbol{\Psi}_h\boldsymbol{Z}_1^H\boldsymbol{R}^{-1} - \boldsymbol{Z}_0\boldsymbol{\Psi}_h\boldsymbol{Z}_0^H\boldsymbol{R}^{-1})^2].$$

In the following, we simplify (3.21) to the case of white noise and uncorrelated fading.

1) White Noise and Uncorrelated Fading: Using the optimal $\lambda = 1/2$, and assuming that the additive noise is both spatially and temporally white (i.e., $\boldsymbol{R} = \sigma^2\boldsymbol{I}_{n_\mathrm{R}N}$), the fading coefficients are uncorrelated with equal variances (i.e., $\boldsymbol{\Psi}_h = \psi_h^2\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}}$), and the mean of the fading coefficients follows (1.4) (see also (2.26)), then (3.21) simplifies to

$$P_\mathrm{CB} \approx \frac{1}{2}\cdot\exp\left(-\frac{n_\mathrm{R}}{4}\cdot\left\{\mathrm{SNR}_\mathrm{LOS}\cdot\boldsymbol{a}_\mathrm{T}^H(\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0)^*(\boldsymbol{\Phi}_1-\boldsymbol{\Phi}_0)^T\boldsymbol{a}_\mathrm{T} + \tfrac{1}{2}\,\mathrm{SNR}_\mathrm{SC}^2\cdot\mathrm{tr}\left[\left(\boldsymbol{\Phi}_1^H\boldsymbol{\Phi}_1 - \boldsymbol{\Phi}_0^H\boldsymbol{\Phi}_0\right)^2\right]\right\}\right). \tag{3.24}$$

This approximation is valid if the matrices $\mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H$, $\mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{\Phi}_0\boldsymbol{\Phi}_0^H$, and $\frac{1}{2}\mathrm{SNR}_\mathrm{SC}\cdot[\boldsymbol{\Phi}_1^T,\ \boldsymbol{\Phi}_0^T]^T[\boldsymbol{\Phi}_1^H,\ \boldsymbol{\Phi}_0^H]$ are small. If the line-of-sight component in the exponent of (3.24) is dominant, then the beamforming scheme in (2.25) is optimal.

Rayleigh Fading: If $\mathrm{SNR}_\mathrm{LOS} = 0$ (Rayleigh fading), it follows from (3.24) that the optimal codes need to maximize

$$\mathrm{tr}[(\boldsymbol{\Phi}_1^H\boldsymbol{\Phi}_1 - \boldsymbol{\Phi}_0^H\boldsymbol{\Phi}_0)^2]. \tag{3.25}$$

(This scenario has been recently investigated in [34], where a cutoff-rate-based design criterion was proposed.) Since

$$\mathrm{tr}(\boldsymbol{\Phi}_1^H\boldsymbol{\Phi}_1\boldsymbol{\Phi}_0^H\boldsymbol{\Phi}_0) = \mathrm{tr}[(\boldsymbol{\Phi}_1\boldsymbol{\Phi}_0^H)\cdot(\boldsymbol{\Phi}_1\boldsymbol{\Phi}_0^H)^H] \geq 0$$

orthogonal signaling (i.e., $\boldsymbol{\Phi}_1\boldsymbol{\Phi}_0^H = \boldsymbol{0}$) is clearly optimal in this case. For orthogonal signaling, the code design criterion (3.25) further simplifies to

$$\mathrm{tr}[(\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H)^2] + \mathrm{tr}[(\boldsymbol{\Phi}_0\boldsymbol{\Phi}_0^H)^2]. \tag{3.26}$$

Subject to the power constraint (2.33), this criterion is maximized for $\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H = \boldsymbol{u}_1\boldsymbol{u}_1^H$ and $\boldsymbol{\Phi}_0\boldsymbol{\Phi}_0^H = \boldsymbol{u}_0\boldsymbol{u}_0{}^H$, where $\boldsymbol{u}_1^H\boldsymbol{u}_1 = \boldsymbol{u}_0{}^H\boldsymbol{u}_0 = 1$ and $\boldsymbol{u}_1^H\boldsymbol{u}_0 = 0$, which we refer to as *orthogonal-subspace beamforming.* Orthogonal-subspace beamforming is an example of a signaling scheme in which the "equal energy" condition (3.13) does not hold.

To summarize, orthogonal-subspace beamforming is optimal when SNR_SC is small, $\mathrm{SNR}_\mathrm{LOS} = 0$, and the power constraint (2.33) is imposed. The approximate Chernoff bound then simplifies to

$$P_\mathrm{CB} \approx \tfrac{1}{2}\cdot\exp\left[-n_\mathrm{R}\cdot(\tfrac{1}{2}\mathrm{SNR}_\mathrm{SC})^2\right] \tag{3.27}$$

and reasonably good performance can be achieved only if the number of receiver antennas n_R is very large.

C. Equal Energy Orthogonal Signaling

We derive the optimal Chernoff bound for the case where the "equal energy" and orthogonality conditions in (3.13) and (3.17) hold. Substituting (3.13) and (3.17) into (3.10), we obtain

$$P_\mathrm{CB} = \frac{1}{2}\cdot\frac{|\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}} + \boldsymbol{W}\boldsymbol{\Psi}_h|}{|\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}} + \lambda\boldsymbol{W}\boldsymbol{\Psi}_h|\cdot|\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}} + (1-\lambda)\boldsymbol{W}\boldsymbol{\Psi}_h|}\cdot\exp\left[-\lambda(1-\lambda)\cdot\boldsymbol{\mu}_h^H\cdot\left\{(\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}} + \lambda\cdot\boldsymbol{W}\boldsymbol{\Psi}_h)^{-1} + [\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}} + (1-\lambda)\cdot\boldsymbol{W}\boldsymbol{\Psi}_h]^{-1}\right\}\cdot\boldsymbol{W}\boldsymbol{\mu}_h\right] \tag{3.28}$$

where $\boldsymbol{W}$ was defined in (3.13). Note that (3.28) remains the same if λ is replaced with $1-\lambda$. Differentiating the logarithm of the Chernoff bound expression in (3.28) with respect to λ shows that (3.28) is minimized for $\lambda = 1/2$. Therefore, the optimal Chernoff bound is

$$P_\mathrm{CB}\Big|_{\substack{\boldsymbol{W} = \boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_1 = \boldsymbol{Z}_0^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0\\ \boldsymbol{Z}_1^H\boldsymbol{R}^{-1}\boldsymbol{Z}_0 = \boldsymbol{0}}} = \frac{1}{2}\cdot\frac{|\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}} + \boldsymbol{W}\boldsymbol{\Psi}_h|}{|\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}} + \frac{1}{2}\boldsymbol{W}\boldsymbol{\Psi}_h|^2}\cdot\exp\left[-\tfrac{1}{2}\cdot\boldsymbol{\mu}_h^H[\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}} + \tfrac{1}{2}\cdot\boldsymbol{W}\boldsymbol{\Psi}_h]^{-1}\boldsymbol{W}\boldsymbol{\mu}_h\right]. \tag{3.29}$$

In the following, we specialize (3.29) to the case of white noise and uncorrelated fading, and use the obtained result to compute the optimal number of transmitter antennas for unitary mutually orthogonal space–time codes.

1) White Noise and Uncorrelated Fading: Assume that the additive noise is both spatially and temporally white (i.e., $\boldsymbol{R} = \sigma^2\boldsymbol{I}_{n_\mathrm{R}N}$), the fading coefficients are uncorrelated with equal variances (i.e., $\boldsymbol{\Psi}_h = \psi_h^2\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}}$), and the mean of the fading coefficients follows (1.4). Then, (3.29) simplifies to

$$P_\mathrm{CB} = \frac{1}{2}\cdot\frac{|\boldsymbol{I}_{n_\mathrm{T}} + \mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{V}|^{n_\mathrm{R}}}{|\boldsymbol{I}_{n_\mathrm{T}} + \frac{1}{2}\,\mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{V}|^{2n_\mathrm{R}}}\cdot\exp\left[-\tfrac{1}{2}\,n_\mathrm{R}\,\mathrm{SNR}_\mathrm{LOS}\cdot\boldsymbol{a}_\mathrm{T}^H\left(\boldsymbol{I}_{n_\mathrm{T}} + \tfrac{1}{2}\,\mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{V}^*\right)^{-1}\boldsymbol{V}^*\boldsymbol{a}_\mathrm{T}\right] \tag{3.30}$$

where

$$\boldsymbol{V} = \boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H = \boldsymbol{\Phi}_0\boldsymbol{\Phi}_0^H \tag{3.31}$$

and $\mathrm{SNR}_\mathrm{LOS}$ and SNR_SC are defined in (2.32). For full-rank $\boldsymbol{V}$ and large $\frac{1}{2}\,\mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{V}$, the above expression simplifies to

$$P_{\mathrm{CB\ full\ rank}} \approx \frac{1}{2}\cdot\frac{1}{|\frac{1}{4}\,\mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{V}|^{n_\mathrm{R}}}\cdot\exp\left(-n_\mathrm{T}n_\mathrm{R}\cdot\frac{\mathrm{SNR}_\mathrm{LOS}}{\mathrm{SNR}_\mathrm{SC}}\right) \tag{3.32}$$

which can also be obtained by substituting $\boldsymbol{R} = \sigma^2\boldsymbol{I}_{n_\mathrm{R}N}$, $\boldsymbol{\Psi}_h = \psi_h^2\boldsymbol{I}_{n_\mathrm{T}n_\mathrm{R}}$, and $\lambda = 1/2$ into (3.15).

We now examine the performance of unitary orthogonal codes and discuss the optimal choice of the number of transmitter antennas n_T. Under the power constraint (2.33), the optimal $\boldsymbol{V}$ (which minimizes (3.32)) has all eigenvalues equal to $1/n_\mathrm{T}$, and therefore,

$$|\mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{V}|_\mathrm{MAX} = \left(\frac{\mathrm{SNR}_\mathrm{SC}}{n_\mathrm{T}}\right)^{n_\mathrm{T}} \tag{3.33}$$

which is the same as (2.35) obtained for antipodal coherent signaling in white noise and uncorrelated fading. The condition that $\frac{1}{2}\,\mathrm{SNR}_\mathrm{SC}\cdot\boldsymbol{V}$ is large simplifies to the requirement that $\mathrm{SNR}_\mathrm{SC}/(2n_\mathrm{T})$ is large. Let us choose

$$\boldsymbol{V} = (1/n_\mathrm{T})\cdot\boldsymbol{I}_{n_\mathrm{T}} \tag{3.34}$$

as in the unitary space–time codes [8], [9]. The optimal Chernoff bound (3.30) then simplifies to

$$P_\mathrm{CB}\Big|_{\boldsymbol{V}=(1/n_\mathrm{T})\cdot\boldsymbol{I}_{n_\mathrm{T}}} = \frac{1}{2}\cdot\left[\frac{\left(1+\frac{\mathrm{SNR}_\mathrm{SC}}{2n_\mathrm{T}}\right)^2}{1+\frac{\mathrm{SNR}_\mathrm{SC}}{n_\mathrm{T}}}\cdot\exp\left(\frac{\frac{\mathrm{SNR}_\mathrm{LOS}}{2n_\mathrm{T}}}{1+\frac{\mathrm{SNR}_\mathrm{SC}}{2n_\mathrm{T}}}\right)\right]^{-n_\mathrm{T}n_\mathrm{R}}. \tag{3.35}$$

In Fig. 2 we show the Chernoff bound in (3.35) as a function of the number of transmitter antennas n_T and the line-of-sight $\mathrm{SNR}_\mathrm{LOS}$,

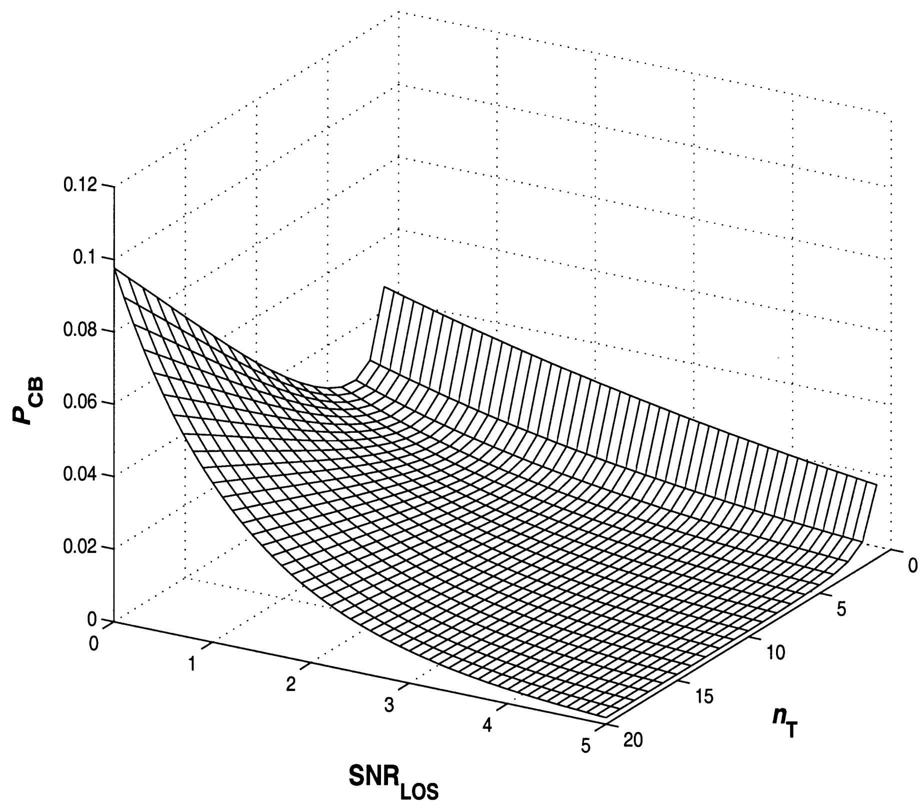

Fig. 2. Chernoff bound on error probability for equal energy orthogonal signaling as a function of number of transmitter antennas $n_{\rm T}$ and line-of-sight $\rm SNR_{LOS}$, with $n_{\rm R} = 2$ receiver antennas and scattering $\rm SNR_{SC} = 10$.

where the scattering SNR and number of receiver antennas are chosen to be $\rm SNR_{SC} = 10$ and $n_{\rm R} = 2$. For larger values of $\rm SNR_{LOS}$, the Chernoff bound decreases with $n_{\rm T}$. However, for smaller values of $\rm SNR_{LOS}$, there exists an optimal number of transmitter antennas $n_{\rm TOPT}$ for which the Chernoff bound is minimized. Hence, if too many transmitter antennas are used, the signal power per transmitter may become so small that the resulting degradation in the performance of each diversity branch cannot be compensated by the diversity gain, see also the discussion in Section II-A1 and [35]. This is consistent with early results in [35] and [36] where optimal numbers of links for Rayleigh-faded noncoherent diversity systems were obtained using criteria based on Bhattacharyya bounds and error probabilities, respectively.

Differentiating the logarithm of (3.35) with respect to $n_{\rm T}$, it can be shown that (3.35) is maximized when

$$\ln\left[\frac{\left(1+\frac{\rm SNR_{SC}}{2n_{\rm T}}\right)^2}{1+\frac{\rm SNR_{SC}}{n_{\rm T}}}\right] + \frac{\frac{\rm SNR_{SC}}{n_{\rm T}}}{1+\frac{\rm SNR_{SC}}{n_{\rm T}}} - \frac{\frac{\rm SNR_{SC}}{n_{\rm T}}}{1+\frac{\rm SNR_{SC}}{2n_{\rm T}}} + \frac{\frac{\rm SNR_{LOS}}{2n_{\rm T}}\cdot\frac{\rm SNR_{SC}}{2n_{\rm T}}}{\left(1+\frac{\rm SNR_{SC}}{2n_{\rm T}}\right)^2} = 0. \quad (3.36)$$

Solving the preceding equation gives the optimal number of transmitter antennas $n_{\rm TOPT}$ that minimizes the Chernoff bound. In Fig. 3, we show $\rm SNR_{SC}/(2n_{\rm TOPT})$ as a function of $\rm SNR_{LOS}/(2n_{\rm TOPT})$, computed using (3.36). From Fig. 3, we can easily find the optimal number of transmitter antennas for given line-of-sight and scattering SNRs. For example, assume a Rayleigh-fading scenario (i.e., $\rm SNR_{LOS} = 0$) with $\rm SNR_{SC} = 10$. Then, we read from Fig. 3 that $\rm SNR_{SC}/(2n_{\rm TOPT}) \approx 1.5$ for $\rm SNR_{LOS}/(2n_{\rm TOPT}) = 0$, and, therefore, $n_{\rm TOPT} \approx 10/(2\cdot 1.5) \approx 3$. It may be easily verified in Fig. 2 that $n_{\rm T} = 3$ is indeed the optimal number of transmitter antennas in this scenario.

IV. Concluding Remarks

We derived Chernoff-bound expressions on pairwise error probabilities for coherent and noncoherent space–time signaling schemes. First, general Chernoff-bound expressions were derived for correlated Ricean fading and correlated additive Gaussian noise, extending the corresponding results in [7] and [8]. (We also used our general Chernoff-bound expression for coherent signaling to find a simple closed-form expression for the exact pairwise error probability under this scenario, see (2.11).) Then, we specialized our results to the cases of space–time separable and white noise, and uncorrelated fading. Approximate Chernoff bounds for high and low SNRs were derived and optimal signaling schemes were proposed. We computed the optimal number of transmitter antennas (minimizing the Chernoff bound) for noncoherent signaling with unitary mutually orthogonal space–time codes.

Further research will include analyzing the accuracy of the proposed bounds and computing simple expressions for pairwise error probabilities of noncoherent and concentrated-likelihood detection schemes.

Appendix
Proof of Lemma 1

For an $n\times 1$ complex vector $\boldsymbol{a} = \mathrm{Re}\{\boldsymbol{a}\} + j\mathrm{Im}\{\boldsymbol{a}\} = \boldsymbol{a}_{\rm r} + j\boldsymbol{a}_{\rm i}$ and an $n\times n$ complex Hermitian matrix $\boldsymbol{A} = \mathrm{Re}\{\boldsymbol{A}\} + j\mathrm{Im}\{\boldsymbol{A}\} = \boldsymbol{A}_{\rm r} + j\boldsymbol{A}_{\rm i}$, define

$$\tilde{\boldsymbol{a}} = \begin{bmatrix}\boldsymbol{a}_{\rm r}\\ \boldsymbol{a}_{\rm i}\end{bmatrix}, \qquad \tilde{\boldsymbol{A}} = \frac{1}{2}\begin{bmatrix}\boldsymbol{A}_{\rm r} & -\boldsymbol{A}_{\rm i}\\ \boldsymbol{A}_{\rm i} & \boldsymbol{A}_{\rm r}\end{bmatrix}. \quad \text{(A1)}$$

Note that, since $\boldsymbol{A}$ is Hermitian, $\boldsymbol{A}_{\rm r}$ is symmetric (i.e., $\boldsymbol{A}_{\rm r} = \boldsymbol{A}_{\rm r}^T$) and $\boldsymbol{A}_{\rm i}$ is skew-symmetric (i.e., $\boldsymbol{A}_{\rm i} = -\boldsymbol{A}_{\rm i}^T$), and, therefore, $\tilde{\boldsymbol{A}}$ is sym-

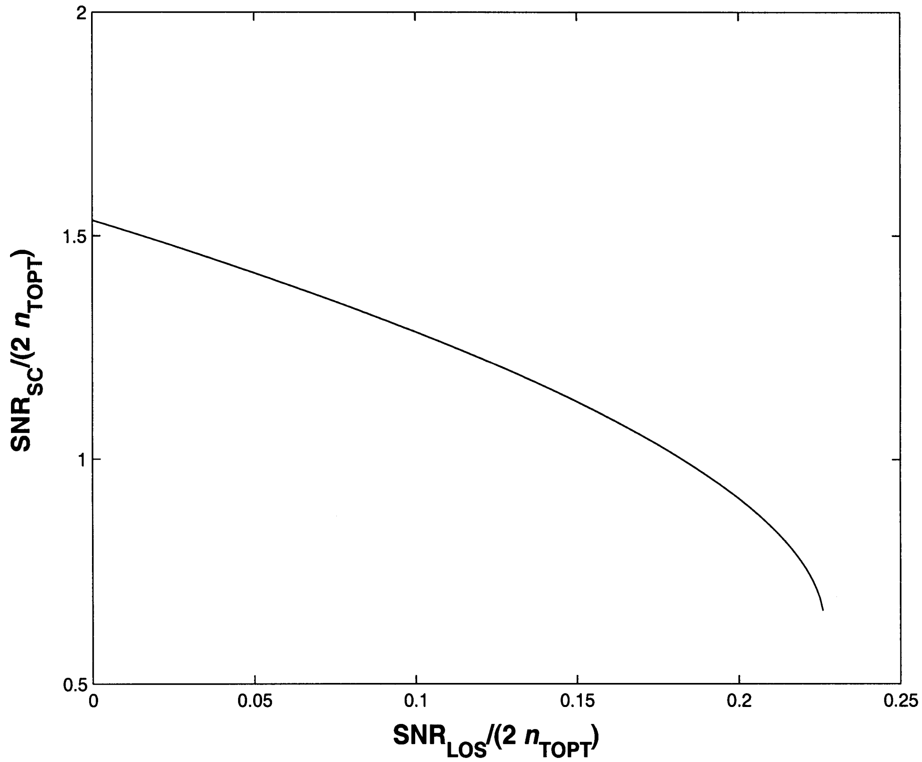

Fig. 3. Scattering $\mathrm{SNR}_{\mathrm{SC}}/(2n_{\mathrm{TOPT}})$ as a function of line-of-sight $\mathrm{SNR}_{\mathrm{LOS}}/(2n_{\mathrm{TOPT}})$, assuming that the optimal number of transmitter antennas n_{TOPT} is deployed.

metric (see also [37, Ch. 2.9], [38, Ch. 15]). The integral in (2.5) can be computed as follows:

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\tfrac{1}{2}\left[\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}+\boldsymbol{x}^H\boldsymbol{a}+\boldsymbol{a}^H\boldsymbol{x}+a_0+a_0^*\right] \cdot\exp[-\tfrac{1}{2}(\boldsymbol{x}^H\boldsymbol{B}\boldsymbol{x}+\boldsymbol{x}^H\boldsymbol{b}+\boldsymbol{b}^H\boldsymbol{x}+b_0+b_0^*)]\;d\boldsymbol{x} \tag{A2a}$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[\tilde{\boldsymbol{x}}^T\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}+\tilde{\boldsymbol{x}}^T\tilde{\boldsymbol{a}}+\mathrm{Re}\{a_0\}\right] \cdot\exp\left[-\left(\tilde{\boldsymbol{x}}^T\tilde{\boldsymbol{B}}\tilde{\boldsymbol{x}}+\tilde{\boldsymbol{x}}^H\tilde{\boldsymbol{b}}+\mathrm{Re}\{b_0\}\right)\right]\;d\tilde{\boldsymbol{x}} \tag{A3a}$$

$$= \tfrac{1}{2}\,\pi^n|\tilde{\boldsymbol{B}}|^{-1/2}\cdot\left[\mathrm{tr}\left(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{B}}^{-1}\right)-\tilde{\boldsymbol{b}}^T\tilde{\boldsymbol{B}}^{-1}\tilde{\boldsymbol{a}}+\tfrac{1}{2}\tilde{\boldsymbol{b}}^T\tilde{\boldsymbol{B}}^{-1}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{B}}^{-1}\tilde{\boldsymbol{b}} + 2\mathrm{Re}\{a_0\}\right]\cdot\exp\left(\tfrac{1}{4}\tilde{\boldsymbol{b}}^T\tilde{\boldsymbol{B}}^{-1}\tilde{\boldsymbol{b}}-\mathrm{Re}\{b_0\}\right) \tag{A3b}$$

$$= \tfrac{1}{2}\,\pi^n|\tfrac{1}{2}\,\boldsymbol{B}|^{-1}\cdot\left[2\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{B}^{-1}\right)-\boldsymbol{b}^H\boldsymbol{B}^{-1}\boldsymbol{a}-\boldsymbol{a}^H\boldsymbol{B}^{-1}\boldsymbol{b} + \boldsymbol{b}^H\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{b}+2\mathrm{Re}\{a_0\}\right]\exp(\tfrac{1}{2}\boldsymbol{b}^H\boldsymbol{B}^{-1}\boldsymbol{b}-\mathrm{Re}\{b_0\}) \tag{A3c}$$

where (A3b) follows by using [20, Theorem 15.12.1, p. 322] or [39, Theorem 10.5.1, p. 342]. Note also that $\mathrm{Re}\{b_0\}=(b_0+b_0^*)/2$ and $\mathrm{Re}\{a_0\}=(a_0+a_0^*)/2$. To derive (A3c) we have used the following identities:

$$|\tilde{\boldsymbol{B}}| = |\tfrac{1}{2}\boldsymbol{B}|^2 \tag{A4a}$$

$$\mathrm{tr}(\tilde{\boldsymbol{A}}\tilde{\boldsymbol{B}}^{-1}) = 2\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}^{-1}) \tag{A4b}$$

$$\tilde{\boldsymbol{b}}^T\tilde{\boldsymbol{B}}^{-1}\tilde{\boldsymbol{a}} = \boldsymbol{b}^H\boldsymbol{B}^{-1}\boldsymbol{a}+\boldsymbol{a}^H\boldsymbol{B}^{-1}\boldsymbol{b} \tag{A4c}$$

$$\tilde{\boldsymbol{a}}^T\tilde{\boldsymbol{A}}\tilde{\boldsymbol{a}} = \tfrac{1}{2}\boldsymbol{a}^H\boldsymbol{A}\boldsymbol{a} \tag{A4d}$$

$$\tilde{\boldsymbol{b}}^T\tilde{\boldsymbol{a}} = \tfrac{1}{2}\,\boldsymbol{b}^H\boldsymbol{a}+\tfrac{1}{2}\,\boldsymbol{a}^H\boldsymbol{b} \tag{A4e}$$

$$\tilde{\boldsymbol{b}}^T\tilde{\boldsymbol{B}}^{-1}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{B}}^{-1}\tilde{\boldsymbol{b}} = 2\boldsymbol{b}^H\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{b} \tag{A4f}$$

which hold for arbitrary $n\times 1$ complex vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, and $n\times n$ complex Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, where $\boldsymbol{B}$ is nonsingular. The identities (A4a) and (A4d) can also be found, for example, in [38, Ch. 15] and [37, Ch. 2.9].

ACKNOWLEDGMENT

The author is grateful to the anonymous reviewers for their helpful comments.

REFERENCES

[1] S. Benedetto and E. Biglieri, *Principles of Digital Transmission With Wireless Applications*. New York: Kluwer, 1999.

[2] S. H. Jamali and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*. Norwell, MA: Kluwer, 1994.

[3] D. Divsalar and M. K. Simon, “Trellis coded modulation for 4800–9600 bits/s transmission over a fading mobile satellite channel,” *IEEE J. Select. Areas Commun.*, vol. 5, pp. 162–175, Feb. 1987.

[4] ——, “The design of trellis coded MPSK for fading channels: Performance criteria,” *IEEE Trans. Commun.*, vol. 36, pp. 1004–1012, Sept. 1988.

[5] C. Schlegel and D. J. Costello, Jr., “Bandwidth efficient coding for fading channels: Code construction and performance analysis,” *IEEE J. Select. Areas Commun.*, vol. 7, pp. 1356–1368, Dec. 1989.

[6] C.-E. W. Sundberg and N. Seshadri, “Coded modulation for fading channels: An overview,” *Euro. Trans. Telecommun.*, vol. 4, no. 3, pp. 309–324, 1993.

[7] 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.

[8] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.
[9] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1962–1973, Sept. 2000.
[10] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.
[11] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2335–2367, Sept. 2001.
[12] M. K. Simon, "Evaluation of average bit error probability for space-time coding based on a simpler exact evaluation of pairwise error probability," *J. Commun. Networks*, vol. 3, pp. 257–264, Sept. 2001.
[13] G. Taricco and E. Biglieri, "Exact pairwise error probability of space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 510–513, Feb. 2002.
[14] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*. New York: Wiley, 2000.
[15] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2383–2399, Sept. 2001.
[16] S. Siwamogsatham, M. P. Fitz, and J. H. Grimm, "A new view of performance analysis of transmit diversity schemes in correlated Rayleigh fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 950–956, Apr. 2002.
[17] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, pp. 85–87, Mar. 2001.
[18] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, pt. 1.
[19] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. San Diego, CA: Academic, 1990.
[20] D. A. Harville, *Matrix Algebra From a Statistician's Perspective*. New York: Springer-Verlag, 1997.
[21] G. L. Stüber, *Principles of Mobile Communication*, 2nd ed. Norwell, MA: KLuwer, 2001.
[22] J. J. Blanz, A. Papathanassiou, M. Haardt, I. Furió, and P. W. Baier, "Smart antennas for combined DOA and joint channel estimation in time-slotted CDMA mobile radio systems with joint detection," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 293–306, Mar. 2000.
[23] J. Vidal, M. Cabrera, and A. Agustin, "Full exploitation of diversity in space-time MMSE receivers," in *Proc. 52nd Vehicular Technology Conf.*, Boston, MA, Sept. 2000, pp. 2497–2502.
[24] K. K. Mukkavilli, D. M. Ionescu, and B. Aazhang, "Design of space-time codes with optimal coding gain," in *Proc. 11th Int. Symp. Personal, Indoor and Mobile Radio Communications*, London, U.K., Sept. 2000, pp. 495–499.
[25] D. M. Ionescu, K. K. Mukkavilli, Z. Yan, and J. Lilleberg, "Improved 8- and 16-state space-time codes for 4 PSK with two transmit antennas," *IEEE Commun. Lett.*, vol. 5, pp. 301–303, July 2001.
[26] D. A. Gore, R. U. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, Istanbul, Turkey, June 2000, pp. 2785–2788.
[27] S. Sandhu, R. U. Nabar, D. A. Gore, and A. Paulraj, "Near-optimal selection of transmit antennas for a MIMO channel based on Shannon capacity," in *Proc. 34th Asilomar Conf. Signals, Systems and Computing*, Pacific Grove, CA, Oct. 2000, pp. 567–571.
[28] A. Dogandžić and J. Jin, "Estimating statistical properties of MIMO fading channels," *IEEE Trans. Signal Processing*, submitted for publication.
[29] ——, "Estimating statistical properties of MIMO Ricean fading channels," in *Proc. 2nd IEEE Sensor Array Multichannel Signal Process Workshop*, Rosslyn, VA, Aug. 2002, pp. 149–153.
[30] A. Dogandžić and A. Nehorai, "Space-time fading channel estimation and symbol detection in unknown spatially correlated noise," *IEEE Trans. Signal Processing*, vol. 50, pp. 457–474, Mar. 2002.
[31] E. G. Larsson, P. Stoica, and J. Li, "On maximum-likelihood detection and decoding for space-time coding systems," *IEEE Trans. Signal Processing*, vol. 50, pp. 937–944, Apr. 2002.
[32] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramér–Rao bound," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, pp. 720–741, May 1989.
[33] P. Y. Kam, "Binary orthogonal signaling over the Gaussian channel with unknown phase/fading; new results and interpretations," *IEEE Trans. Commun.*, vol. 38, pp. 1686–1692, Oct. 1990.
[34] A. O. Hero and T. L. Marzetta, "Cutoff rate and signal design for the quasistatic Rayleigh-fading space-time channel," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2400–2416, Sept. 2001.
[35] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun. Technol.*, vol. COM–15, pp. 52–60, Feb. 1967.
[36] J. N. Pierce, "Theoretical limitations on frequency and time diversity for fading binary transmissions," *IRE Trans. Commun. Syst.*, vol. COM–9, pp. 186–187, June 1961.
[37] M. S. Srivastava and C. G. Khatri, *An Introduction to Multivariate Statistics*. New York: North-Holland, 1979.
[38] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[39] F. A. Graybill, *Matrices With Applications in Statistics*, 2nd ed. Belmont, CA: Wadsworth, 1983.

Product-Function Frames in $l^2(\mathbb{Z})$

Sanjay M. Joshi, *Member, IEEE*, and
Joel M. Morris, *Senior Member, IEEE*

***Abstract*—Frames are the most widely used tool for signal representation in different domains. In this correspondence, we introduce the concept of product-function frames for $l^2(\mathbb{Z})$. The frame elements of these frames are represented as products of two or more sequences. This forms a generalized structure for many currently existing transforms. We define necessary and sufficient conditions on the frame elements so that they form a frame for $l^2(\mathbb{Z})$. We obtain windowed transforms as a special case and derive the biorthogonal-like condition. Finally, we introduce a new family of transforms for finite-dimensional subspaces of $l^2(\mathbb{Z})$, which we call the "scale-modulation transforms." The frame elements of these transforms can be obtained via scaling and modulating a "mother" window. This transform, thus, complements the shift-modulation structure of the discrete-time Gabor transform and the shift-scale structure of the discrete wavelet transform.**

***Index Terms*—Frames, Gabor transform, scale-modulation transform, signal reconstruction, windowed transforms.**

I. Introduction

Representation of a signal in joint domains has been a very active area in signal processing. Prime examples of such representation are the short-time Fourier transform (STFT) and the discrete-time Gabor transform (DTGT) in the joint time–frequency domain, and filter banks and wavelets in the joint time–scale domain. The varieties of joint-domain transforms are unified in this correspondence in the form of

Manuscript received November 30,1999; revised January 7, 2003. This work was supported in part by Maryland Industrial Partnerships (MIPS) with Nortel Networks under Contract 2218.24. The material in this correspondence was presented in part at the 35th Conference on Information Science and Systems (CISS 2001), Johns Hopkins University, Baltimore, MD, March 2001.

S. M. Joshi is with Lucent Technologies, Alameda, CA 94502 USA (e-mail: joshi@ieee.org).

J. M. Morris is with the Communications and Signal Processing Laboratory, Computer Science and Electrical Engineering Department, University of Maryland Baltimore County (UMBC), Catonsville, MD 21250 USA (e-mail: morris@umbc.edu).

Communicated by J. A. O'Sullivan, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2003.810656