Optimal Antenna Diversity Signaling for Wide-Band Systems Utilizing Channel Side Information

Eko N. Onggosanusi, Akbar M. Sayeed, *Member, IEEE*, and Barry D. Van Veen, *Member, IEEE*

Abstract—**Optimal multi-antenna wide-band signaling schemes are derived for multipath channels assuming perfect channel state information at the transmitter. The scheme that minimizes the bit-error probability in the single-user case is a rank-one space–time beamformer which focuses the signal transmission in the direction of the most dominant channel mode. Several suboptimal variations are discussed for multiuser applications. The optimal signaling scheme given channel statistics at the transmitter is also derived. The optimal scheme in this case is a full-rank space–time beamformer that transmits on all channel modes. Analysis and simulation results are used to compare the schemes proposed in this paper. Finally, we discuss the optimal signaling scheme when a delayed version of the channel state is available at the transmitter. It is shown that in this case the optimal scheme is a rank-1 beamformer when the channel variations are sufficiently slow and is a full rank beamformer in a sufficiently fast fading channel.**

Index Terms—**Antenna arrays, diversity signaling, feedback, multipath, spread spectrum communication systems, transmit beamforming.**

I. INTRODUCTION

SPATIO-TEMPORAL diversity, combining multiple antennas and temporal signaling, has emerged as a key technology in state-of-the-art systems. For example, an antenna array is required at the base station in the third-generation WCDMA standard [1]. Receive and transmit antenna diversity are utilized for uplink and downlink applications, respectively. In receive diversity, multiple copies of the transmitted data are processed to combat channel fading. Transmit diversity, on the other hand, utilizes a predesigned signaling scheme to send multiple copies of the data for the same purpose. Use of antenna arrays at both the base station and mobile is envisioned in future wireless communication systems [2]. In this case, transmit and receive diversity techniques can be used simultaneously to enhance system performance.

The use of time-division duplexing, where uplink and downlink transmission are interleaved in time, as well as feedback channels for frequency-division duplexing, allow the transmitter to obtain channel information. This information can be used to design an efficient signaling scheme at the transmitter. Several methods for transmit signal design using channel information are found throughout the literature. For example, a filter-bank-based scheme for maximizing data rate in multicarrier systems is described in [3]. The use of channel side information (CSI) at the transmitter for narrow-band systems is discussed in [4]. Block signaling design to achieve maximum signal-to-noise ratio (SNR) and diversity gain for flat fading channels can be found in [5] and [6].

This paper focuses on the design of optimal diversity signaling schemes that minimize the bit error rate (BER) given CSI at the transmitter. We show that when perfect channel state information is available, the BER-optimal solution is characterized by a rank-1 structure: the same signature code is transmitted through each antenna and weighted according to the channel state. For a single-path channel, we show that the signature code is arbitrary and the optimal antenna weights focus the transmission at the most dominant spatial channel mode. For a multipath channel, however, we show that the optimal signature code is a discrete-time sinusoid with frequency equal to the channel frequency having maximum gain. This optimal scheme can be viewed as adaptive frequency hopping with spatial beamforming. A beamforming solution for flat fading channels was also suggested in [4] and [7] when perfect or sufficiently accurate CSI is available at the transmitter.

In multiuser scenarios, strictly enforcing optimality for each user may result in two or more users having the same maximizing frequency and hence the same signature code at a particular instant. One way to resolve this problem is to use the subdominant frequencies for other users having the same maximizing frequency as the first user. Another alternative approach is to assign distinct fixed signature codes to different users. Then, beamformer weights are chosen to minimize the BER for each user given the sub-optimal code. We term this structure the suboptimal beamforming (space-only optimization) scheme.

When only channel statistics are available at the transmitter, we show that BER is minimized by transmitting over all the channel modes. For a given statistics, distinct linearly independent signature codes are chosen for different transmit antennas to minimize the average BER. Similar results can be found in [8]. This structure is termed the multicode scheme, which is related to the block signaling scheme in [5]. We also analyze the effect of delayed channel state information at the transmitter and demonstrate that the BER-optimal solution deviates from the rank-1 scheme and is analogous to a multicode scheme when the channel fading is sufficiently fast relative to the delay. Previous work related to this issue can be found in [7], where analytical comparison between rank-one

E. N. Onggosanusi is with DSPS & R&D Center, Texas Instruments, Inc., Dallas TX 75243 USA (e-mail: eko@ti.com).

A. M. Sayeed and V. D. Van Veen are with the Department of Electrical and Computer Engineering, University of Wisconsin-Madison, Madison, WI 53706 USA (e-mail: akbar@engr.wisc.edu; vanveen@engr.wisc.edu).

beamforming (termed closed-loop transmit diversity) and a multicode (termed open-loop transmit diversity) scheme in the presence of delay was given. Similar conclusions were reached regarding the best scheme given the channel fading rate.

The rest of the paper is organized as follows. The channel model is outlined in Section II, followed by BER analysis for the coherent receiver in Section III. Diversity signaling design given perfect CSI is discussed in Section IV. Section V covers signaling design when only the channel statistics are available. The optimal scheme given delayed channel state information is discussed in Section VI. Analysis and simulation comparing different schemes are given in Section VII, followed by concluding remarks in Section VIII.

We use the following notation. Superscripts T, H, and $*$ indicate matrix transpose, matrix conjugate transpose, and complex conjugation, respectively. Uppercase boldface denotes a matrix while lowercase boldface indicates a vector. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. $\mathbf{x} \sim \mathcal{N}_C[\mathbf{m}, \mathbf{R}]$ denotes a complex circular Gaussian vector $\mathbf{x}$ with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$. Expectation is denoted as $E[\cdot]$ and the Euclidean norm of vector $\mathbf{x}$ is denoted as $\|\mathbf{x}\|$. The symbol $\otimes$ denotes the Kronecker product and $\text{vec}(\mathbf{A})$ is formed by stacking the columns of matrix $\mathbf{A}$ into a vector [9], [10]. $\mathbf{e}_n$ is the column vector $[\mathbf{0}^T \ 1 \ \mathbf{0}^T]^T$ with 1 located in the nth row. The diagonal matrix generated by the vector $\mathbf{v}$ is denoted by $\text{diag}\{\mathbf{v}\}$. The maximum eigenvalue and the corresponding eigenvector of matrix $\mathbf{A}$ is denoted by $\lambda_{\max}[\mathbf{A}]$ and $\mathbf{ev}_{\max}[\mathbf{A}]$, respectively. The following assumptions are used.

1) The systems operate in a Rayleigh fading channel [11]. We primarily consider single-user systems. The results also apply to multiuser systems in which multiuser interference can be approximated as white noise.
2) Noise-free measurements of channel coefficients and multipath delays are available at the transmitter and receiver.

II. CHANNEL AND SIGNAL MODEL

Consider a system with P transmit and Q receive antennas. The transmitted signal $\mathbf{x}(t) \in \mathbb{C}^P$ undergoes a frequency selective P-input, Q-output fading channel with delay spread of T_d. The signal $r_q(t)$ at the q-th receive antenna can be written as

$$r_q(t) = \int_0^{T_d} \mathbf{h}_q^T(\tau)\mathbf{x}(t-\tau)\,d\tau + n_q(t)$$
$$\mathbf{h}_q(\tau) = [h_{q1}(\tau)\ h_{q2}(\tau)\ \cdots\ h_{qP}(\tau)]^T \qquad (1)$$

where $x_p(t)$ is the signal transmitted via the pth antenna and $h_{qp}(\tau)$ is the channel impulse response representing the coupling between the pth transmit qth receive antenna. We assume that the additive noise process is temporally and spatially white zero mean complex Gaussian. That is, $E[n_q(t)n_{q'}^*(t')] = \sigma^2\delta(t-t')\delta_{p,p'}$.

In this paper, we focus on utilizing the available spatio-temporal degrees of freedom for diversity only—that is, a single bit stream is transmitted on all antennas. Thus, $\mathbf{x}(t) = b\,\mathbf{v}(t)$, $0 \le t < T$, where b is the data symbol transmitted

over all the P elements within 1 symbol duration T and $\mathbf{v}(t) = [v_1(t), \ldots, v_P(t)]^T$, where $v_p(t)$ is the pth transmit antenna signature waveform. For direct sequence wide-band systems, $v_p(t)$ is the waveform corresponding to the pth element length N signature code $\{s_p[i]\}_{i=0}^{N-1}$. Let B denote the (two-sided) signal bandwidth. Analogous to a direct-sequence spread spectrum system, we represent $v_p(t)$ as

$$v_p(t) = \sum_{i=0}^{N-1} s_p[i]\omega\left(t - \frac{i}{B}\right), \quad 0 \le t < T \qquad (2)$$

where $N = TB$ and $\omega(t)$ is a unit-energy waveform of duration $1/B$. We sample $r_q(t)$ at the rate $1/B$ to enable discrete-time processing without loss of information. Let

$$\mathbf{r}_q \overset{\text{def}}{=} \left[r_q(0), r_q\left(\frac{1}{B}\right), \ldots, r_q\left(\frac{(N-1)}{B}\right) \right]^T$$
$$\mathbf{s}_p \overset{\text{def}}{=} [s_P[0], s_P[1], \ldots, s_P[N-1]]^T$$
$$\mathbf{S} \overset{\text{def}}{=} [\mathbf{s}_1 \ \cdots \ \mathbf{s}_P]. \qquad (3)$$

Hence, $\mathbf{r}_q$ contains samples of the received signal at the qth antenna over one symbol duration, while $\mathbf{S}$ is an $N \times P$ matrix containing the signature codes from all transmit antennas. Now, define $\boldsymbol{\Delta}_{d_l} \in \mathbb{C}^{N \times N}$ as the time-shift matrix corresponding to the path delay d_l. We assume the delay is cyclic, so that $\boldsymbol{\Delta}_{d_l}$ is circulant. For example, when $d_l = 1$ and $N = 3$,
$\boldsymbol{\Delta}_{d_l} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. While the actual delay corresponds to a linear shift, negligible error is introduced by this assumption for sufficiently large N provided that $d_L \ll N$. Furthermore, if a *chip-level* cyclic prefix is introduced, the cyclic shift is exact [11]. Since $x_p(t)$ is essentially bandlimited to B, it suffices to look at the sampled version of the channel impulse responses at $\tau_l = d_l/B$, $d_l \in \{0, 1, \ldots, N-1\}$. By defining $\mathbf{h}_{lq}^T = [h_{q1}(\tau_l)\ h_{q2}(\tau_l)\ \cdots\ h_{qP}(\tau_l)]$, $L = \lceil T_d B \rceil$, $\boldsymbol{\Delta} = [\boldsymbol{\Delta}_{d_1}\ \cdots\ \boldsymbol{\Delta}_{d_L}] \in \mathbb{C}^{N \times NL}$, $\mathbf{s} = \text{vec}(\mathbf{S}) \in \mathbb{C}^{NP}$, $\mathbf{H}_q = [\mathbf{h}_{1q}\ \cdots\ \mathbf{h}_{Lq}] \in \mathbb{C}^{P \times L}$ and $\mathbf{h}_q = \text{vec}(\mathbf{H}_q)$, we have from (1) and (2)

$$\mathbf{r}_q = b\,\boldsymbol{\Delta}\,(\mathbf{I}_L \otimes \mathbf{S})\,\mathbf{h}_q + \mathbf{n}_q \qquad (4)$$

where $\mathbf{n}_q \sim \mathcal{N}_C[\mathbf{0}, \sigma^2\mathbf{I}_N]$.

III. COHERENT RECEIVER

For simplicity, consider coherent BPSK modulation ($b \in \{\pm\sqrt{E}\}$) over a wide-band multipath channel with $\tau_L \ll T$. Here E denotes the energy in one symbol. Hence, intersymbol interference (ISI) is negligible and symbol-by-symbol detection suffices. The channel state is assumed to be constant within a symbol duration (slow fading). The maximum likelihood detector [11] in this case is

$$\hat{b} = \text{sgn}\left(\text{Re}\left\{\sum_{q=1}^{Q} Z_q\right\}\right) \qquad (5)$$

where Z_q is the coherent test statistic corresponding to the qth receive antenna obtained from matched filtering and maximum ratio combining. Using (4), Z_q can be written as

$$Z_q = \hat{\mathbf{h}}_q^H (\mathbf{I}_L \otimes \mathbf{S})^H \mathbf{\Delta}^H \mathbf{r}_q \qquad (6)$$

where $\hat{\mathbf{h}}_q$ is an estimate of the channel coefficient vector $\mathbf{h}_q$.

A. BER Analysis

Assuming perfect estimates of channel coefficients $\{\mathbf{h}_q\}_{q=1}^Q$ and multipath delays $\{\tau_l\}_{l=1}^L$ at the receiver, substitution of (4) into (6) gives

$$
\begin{aligned}
Z_q &= b\,\mathbf{h}_q^H \mathbf{R} \mathbf{h}_q + \boldsymbol{\eta}_q \\
\mathbf{R} &= (\mathbf{I}_L \otimes \mathbf{S})^H \mathbf{\Delta}^H \mathbf{\Delta} (\mathbf{I}_L \otimes \mathbf{S}) \in \mathbb{C}^{PL \times PL} \\
\boldsymbol{\eta}_q &\sim \mathcal{N}_C \left[0, \sigma^2 \mathbf{h}_q^H \mathbf{R} \mathbf{h}_q \right].
\end{aligned} \qquad (7)
$$

Given $\{\mathbf{h}_q\}_{q=1}^Q$ and assuming equally likely symbols, the *instantaneous* BER corresponding to (5) is[1]

$$\mathrm{BER}\left(\{\mathbf{h}_q\}_{q=1}^Q\right) = \mathcal{Q}\left(\sqrt{\frac{2E}{\sigma^2}\mathcal{F}}\right) \qquad (8)$$

where

$$\mathcal{F} = \sum_{q=1}^Q \mathbf{h}_q^H \mathbf{R} \mathbf{h}_q \qquad (9)$$

is the received SNR gain. The instantaneous BER assesses the system performance for a particular channel realization. The average BER reflects the system performance over a longer time scale and is defined as $\mathrm{BER}_{\mathrm{AV}} = \mathrm{E}[\mathrm{BER}(\{\mathbf{h}_q\}_{q=1}^Q)]$, where the expectation is taken over all possible channel realizations. Either the instantaneous or average BER will be used to design the diversity signaling scheme, depending upon the nature of the CSI available at the transmitter.

IV. DIVERSITY SIGNALING GIVEN CHANNEL STATES

In this section, we discuss the design of signature codes $\mathbf{S} = [\mathbf{s}_1 \; \cdots \; \mathbf{s}_P]$ (see Fig. 1) to minimize BER given the knowledge of channel states $\{\mathbf{h}_q\}_{q=1}^Q$. In practice, estimates of channel states may be available at the transmitter either via a feedback channel or indirect measurement. A feedback channel from the receiver to the transmitter is used in frequency-division duplexing (FDD) systems. Indirect channel measurement is applicable in time-division duplexing (TDD) since the uplink and downlink channels are identical. Provided that the switching time between uplink and downlink is significantly smaller than channel coherence time, accurate measurement

[1] $\mathcal{Q}(x) = (1/\sqrt{2\pi}) \int_x^\infty e^{-u^2/2} du.$
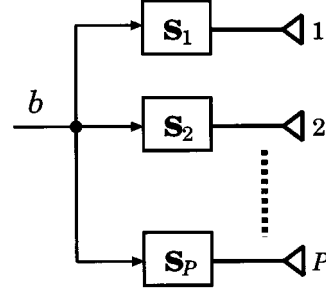

Fig. 1. Diversity signaling considered in this paper.

can be obtained. In this section, we assume that channel states are perfectly known at the transmitter. The optimal signaling scheme is derived for a general multi-antenna system in an L-path channel. A variation of the optimal scheme that provides more flexibility for multi-access scenarios is also given.

A. The Optimal Diversity Signaling Scheme: Rank-1 Space–Time Beamformer (Space-Time Optimization)

We are interested in finding the signature code matrix $\mathbf{S}$ that minimizes the instantaneous BER given $\{\mathbf{h}_q\}_{q=1}^Q$. Combining (8) and (7) and using the fact that $\mathcal{Q}(\sqrt{x})$ is decreasing with x, the optimal $\mathbf{S}$ solves

$$
\begin{aligned}
\mathbf{S} = \arg \max_{\mathbf{U} \in \mathbb{C}^{N \times P}} \sum_{q=1}^Q \mathbf{h}_q^H (\mathbf{I}_L \otimes \mathbf{U})^H \mathbf{\Delta}^H \mathbf{\Delta} (\mathbf{I}_L \otimes \mathbf{U}) \mathbf{h}_q \\
\text{s.t.} \, \mathrm{trace}\left(\mathbf{U}^H \mathbf{U}\right) = 1,
\end{aligned} \qquad (10)
$$

where the trace constraint ensures that sum of energy in all P codes is unity. The optimum solution $\mathbf{S}$ is found as follows. Let $\mathbf{H}_q \stackrel{def}{=} [\mathbf{h}_{1q} \; \cdots \; \mathbf{h}_{Lq}] \in \mathbb{C}^{P \times L}$ such that $\mathbf{h}_q = \mathrm{vec}(\mathbf{H}_q)$ and $\mathbf{u} = \mathrm{vec}(\mathbf{U})$. Note that $\mathrm{trace}(\mathbf{U}^H \mathbf{U}) = \mathrm{vec}^H(\mathbf{U})\mathrm{vec}(\mathbf{U}) = \mathbf{u}^H \mathbf{u}$. Applying the identity $\mathrm{vec}(\mathbf{A}\mathbf{X}\mathbf{B}) = (\mathbf{B}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{X})$ twice [9], we have

$$(\mathbf{I}_L \otimes \mathbf{U})\mathbf{h}_q = \mathrm{vec}(\mathbf{U}\mathbf{H}_q) = \left(\mathbf{H}_q^T \otimes \mathbf{I}_N\right)\mathbf{u}. \qquad (11)$$

Hence, defining $\mathbf{s} \stackrel{def}{=} \mathrm{vec}(\mathbf{S}) = [\mathbf{s}_1^T \; \cdots \; \mathbf{s}_P^T]^T$, we may write (10) as (12), shown at the bottom of the next page. The solution in (12) is unique up to a multiplicative factor.

Valuable insight to nature of the solution to (12) is obtained by considering the single- and multiple-path cases separately.

Theorem 1: For a single-path ($L = 1$) channel, the optimal set of signature codes is $\mathbf{s}_p = w_p \mathbf{c}$, $p = 1, \ldots, P$, where $\mathbf{c} \in \mathbb{C}^N$ is any unit-norm length-N signature code and

$$\mathbf{w} \stackrel{def}{=} [w_1 \; \cdots \; w_P]^T = \mathbf{ev}_{\max}\left[\sum_{q=1}^Q \mathbf{h}_{1q}^* \mathbf{h}_{1q}^T\right] \qquad (13)$$

$$\mathbf{s} = \arg \max_{\mathbf{u} \in \mathbb{C}^{NP}: \|\mathbf{u}\|^2 = 1} \mathbf{u}^H \left(\sum_{q=1}^Q \left(\mathbf{H}_q^T \otimes \mathbf{I}_N\right)^H \mathbf{\Delta}^H \mathbf{\Delta} \left(\mathbf{H}_q^T \otimes \mathbf{I}_N\right)\right) \mathbf{u} = \mathbf{ev}_{\max}\left[\sum_{q=1}^Q \left(\mathbf{H}_q^T \otimes \mathbf{I}_N\right)^H \mathbf{\Delta}^H \mathbf{\Delta} \left(\mathbf{H}_q^T \otimes \mathbf{I}_N\right)\right] \qquad (12)$$
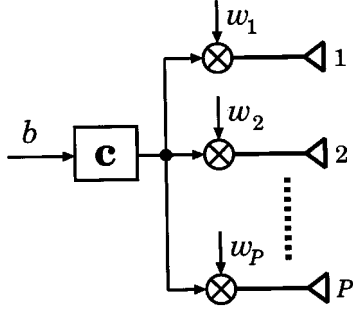
Fig. 2. The optimal diversity signaling structure given the channel states: common signature code for all antennas $\mathbf{c}$ followed by a beamformer $\mathbf{w}$.

Proof: For a single-path channel, $\boldsymbol{\Delta}^H \boldsymbol{\Delta} = \boldsymbol{\Delta}_{d_1}^H \boldsymbol{\Delta}_{d_1} = \mathbf{I}_N$. Also, $\mathbf{H}_q = \mathbf{h}_{1q}$. Hence,

$$
\begin{aligned}
\mathbf{s} &= \mathbf{ev}_{\max} \left[\sum_{q=1}^Q \left(\mathbf{h}_{1q}^T \otimes \mathbf{I}_N \right)^H \left(\mathbf{h}_{1q}^T \otimes \mathbf{I}_N \right) \right] \\
&= \mathbf{ev}_{\max} \left[\sum_{q=1}^Q \mathbf{h}_{1q}^* \mathbf{h}_{1q}^T \right] \otimes \mathbf{ev}_{\max} \left[\mathbf{I}_N \right]
\end{aligned}
$$

where the second equality follows from the eigenstructure of Kronecker product [9]. The proof is completed by noting that any length-N unit-norm vector is an eigenvector of $\mathbf{I}_N$. $\quad\square$

The instantaneous BER obtained using the signature codes in Theorem 1 is

$$
\mathcal{Q} \left(\sqrt{\frac{2E}{\sigma^2} \, \lambda_{\max} \left[\sum_{q=1}^Q \mathbf{h}_{1q}^* \mathbf{h}_{1q}^T \right]} \right).
$$

As depicted in Fig. 2, all the transmit antennas share the *same* length-N signature code $\mathbf{c}$ followed by the beamformer $\mathbf{w} \in \mathbb{C}^P$. Thus, the optimal diversity signaling scheme is decomposed into temporal and spatial elements, represented by $\mathbf{c}$ and $\mathbf{w}$, respectively. The signature code $\mathbf{c}$ is arbitrary since the channel is frequency-independent for $L = 1$. This allows different signature codes to be assigned to different users to minimize multi-acess interference and each user separately computes its optimum beamformer based upon the most recent channel states. The single path model may be applicable to a multi-antenna OFDM system in which the channel associated with each subcarrier is frequency-nonselective.

For a multipath channel ($L > 1$), the optimal solution $\mathbf{s}$ given in (12) can be further simplified by exploiting the circulant property of $\boldsymbol{\Delta}_{d_l}$, which results in the following theorem.

Theorem 2: For a multipath channel ($L > 1$), the optimal set of signature codes is $\mathbf{s}_p = w_p \mathbf{c}$, $p = 1, \ldots, P$, where

$$
\mathbf{w} = \begin{bmatrix} w_1 & \cdots & w_P \end{bmatrix}^T = \mathbf{ev}_{\max} \left[\Sigma_{\overline{n}} \right] \tag{14}
$$

$$
\overline{n} = \arg \max_{n=0,\ldots,N-1} \lambda_{\max} \left[\Sigma_n \right] \tag{15}
$$

$$
\mathbf{c} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1, e^{j2\pi \overline{n}/N}, \ldots, e^{j2\pi(N-1)\overline{n}/N} \end{bmatrix}^T \tag{16}
$$

$$
\Sigma_n = \sum_{q=1}^Q \mathbf{g}_{n,q} \mathbf{g}_{n,q}^H,
$$

$$
\mathbf{g}_{n,q} \stackrel{def}{=} \begin{bmatrix} \sum_{l=1}^L h_{lq1}^* e^{j2\pi n d_l/N} \\ \vdots \\ \sum_{l=1}^L h_{lqP}^* e^{j2\pi n d_l/N} \end{bmatrix}. \tag{17}
$$

Proof: See Appendix A. $\quad\square$

The techniques we apply to obtain the optimal solution are similar to those used in [12] to derive a space–time channel (block) diagonalization. Notice that the pth element of $\mathbf{g}_{n,q}$ is the complex conjugate of the frequency response of the channel between the qth receive and pth transmit antenna at frequency $2\pi n/N$. Also, for $Q = 1$, $\mathbf{w}$ simplifies to $\mathbf{g}_{\overline{n},1}/\|\mathbf{g}_{\overline{n},1}\|$, where $\overline{n} = \arg \max_n \|\mathbf{g}_{n,1}\|^2$. Lastly we observe that the N distinct signature codes $\mathbf{c}$ generated by all possible values of $\overline{n}$ are orthogonal.

The optimum solution in Theorem 2 results in the minimum instantaneous BER

$$
\mathrm{BER} \left(\{\mathbf{h}_q\}_{q=1}^Q \right) = \mathcal{Q} \left(\sqrt{\frac{2E}{\sigma^2} \, \lambda_{\max} \left[\Sigma_{\overline{n}} \right]} \right). \tag{18}
$$

with $\overline{n}$ and Σ_n given in (15) and (17), respectively. In a frequency-selective channel, the minimum BER is obtained by transmitting the signal at the (discretized) frequency corresponding to the maximum channel gain. This is reflected in the choice of signature code $\mathbf{c}$. Note that, at any particular instant, the transmission is focused to a particular frequency within the available bandwidth. However, over longer time intervals where the channel state varies significantly, the transmitted signal is likely to traverse the entire bandwidth. Once the maximizing frequency is chosen, the spatial beamformer $\mathbf{w}$ focuses the signal on the dominant spatial channel mode at that frequency to minimize instantaneous BER. We term this structure the optimal space–time beamforming scheme. In essence, focusing the transmission to the dominant spatio-temporal channel mode is a generalization of selection diversity. We note that the space–time signaling structure depicted in Fig. 2 applies to both the optimal solutions in Theorem 1 ($L = 1$) and 2 ($L > 1$). For $L = 1$, $\mathbf{c}$ is arbitrary.

B. Suboptimal Beamforming Scheme (Space-Only Optimization)

In some cases, one may be restricted to use a suboptimal temporal signature code. An example of such situation is in multi-user scenario, where strictly enforcing optimality for each user may result in two or more users having the same maximizing frequency $\overline{n}$ and hence signature code $\mathbf{c}$ at a particular instant. One way to resolve this problem is to use subdominant frequencies for other users that have the same maximizing frequency as the first user. This approach does not result in multiuser interference (MUI) even in frequency-selective channels. However, it requires coordination among users and temporal signaling for each user must be adapted based on the channel state information.

In practical CDMA systems, Gold or Walsh–Hadamard codes are used for temporal (user-specific) signature codes. Here, one

is restricted to use a fixed suboptimal code $\tilde{\mathbf{c}}$ in frequency selective channels. In this case, CSI can still be used to perform space-only optimization. Given $\tilde{\mathbf{c}}$, the spatial beamformer $\tilde{\mathbf{w}}$ is chosen to minimize the BER. In this case, $\mathbf{U}$ in (10) can be written as $\tilde{\mathbf{c}}\tilde{\mathbf{w}}^T$, where the optimization is performed over $\{\tilde{\mathbf{w}} \in \mathbb{C}^P : \|\tilde{\mathbf{w}}\|^2 = 1\}$. Analogous to (11), we have

$$
\begin{aligned}
\left(\mathbf{I}_L \otimes \tilde{\mathbf{c}}\tilde{\mathbf{w}}^T\right) \mathbf{h}_q &= \text{vec}\left(\tilde{\mathbf{c}}\tilde{\mathbf{w}}^T \mathbf{H}_q\right) \\
&= \left(\mathbf{H}_q^T \otimes \tilde{\mathbf{c}}\right) \tilde{\mathbf{w}} \\
&= \left(\mathbf{I}_L \otimes \tilde{\mathbf{c}}\right) \mathbf{H}_q^T \tilde{\mathbf{w}}
\end{aligned}
\tag{19}
$$

where the last equality follows from Identity 2 at the end of Appendix A. From (10) and (19), the optimal spatial beamformer $\tilde{\mathbf{w}}$ given the signature code $\tilde{\mathbf{c}}$ is given by (20), shown at the bottom of the page, and the resulting instantaneous BER can be written as shown at the bottom of the page. For a single-path channel, the performance of this suboptimal scheme is identical to the optimal one since the optimal signature code $\mathbf{c}$ is arbitrary, as shown in Theorem 1. Note that the above analysis and optimization ignore MUI, which in general exists when spreading codes such as Gold or Walsh–Hadamard are used in frequency-selective channels. Hence, the solution in (20) only minimizes the *single-user* performance bound of the system. This serves as a good approximation when the number of users and spreading gain are large and/or power control is used, as MUI can be approximated as additional Gaussian noise [13]. In general, when a multiuser detector is used, the optimal spatial beamformer depends on the type of multiuser detector. This problem is beyond the scope of this paper.

V. DIVERSITY SIGNALING GIVEN CHANNEL STATISTICS

When only channel statistics are available at the transmitter, the appropriate performance measure is the *average* BER. For Rayleigh fading channels, the second-order statistics completely characterize the channel. The advantage of using channel statistics is that they vary much slower in time than the channel states and thus can be measured more accurately at the transmitter in the TDD case, or require less frequent feedback in the FDD case.

To obtain an analytical expression for the average BER, it is useful to write $\sum_{q=1}^Q \mathbf{h}_q^H \mathbf{R} \mathbf{h}_q$ in (8) as $\mathbf{h}^H (\mathbf{I}_Q \otimes \mathbf{R}) \mathbf{h}$, where

$\mathbf{h} \stackrel{def}{=} [\mathbf{h}_1^T \ \cdots \ \mathbf{h}_Q^T]^T$. For Rayleigh fading, $\mathbf{h} \sim \mathcal{N}_C[\mathbf{0}, \mathbf{\Psi}]$, where $\mathbf{\Psi}$ is the channel covariance matrix. Assuming that no pair of channel coefficients are completely correlated and no channel coefficient has zero energy, $\mathbf{\Psi}$ is positive definite. It can be shown via KL expansion [14] for $\mathbf{h}$ that

$$
\begin{aligned}
\text{BER}_{AV} &= \text{E}\left[\text{BER}(\{\mathbf{h}_q\}_{q=1}^Q)\right] \\
&= \text{E}\left[\mathcal{Q}\left(\sqrt{\frac{2}{\sigma^2} \sum_{n=1}^{PQL} |\xi_n|^2}\right)\right], \\
\xi_n &\sim \mathcal{N}_C[0, \lambda_n], \quad \text{E}[\xi_n \xi_{n'}^*] = \delta_{n,n'} \lambda_n
\end{aligned}
\tag{21}
$$

where $\{\lambda_n\}_{n=1}^{PQL}$ are the eigenvalues of the $PQL \times PQL$ matrix $\mathbf{\Phi} \stackrel{def}{=} \mathbf{\Psi}(\mathbf{I}_Q \otimes \mathbf{R})$. Since $\mathbf{R}$ is nonnegative definite and Hermitian symmetric, it follows that $\mathbf{\Phi}$ is nonnegative definite. Using the identity $\mathcal{Q}(x) = (1/\pi) \int_0^{\pi/2} \exp(-x^2/(2\sin^2\theta)) \, d\theta$ [15], it can be shown that

$$
\text{BER}_{AV} = \frac{1}{\pi} \int_0^{\pi/2} \prod_{n=1}^{PQL} \left(1 + \frac{E}{\sigma^2 \sin^2\theta} \lambda_n\right)^{-1} d\theta
\tag{22}
$$

Notice that the eigenvalues of $\mathbf{\Phi}$ depend upon the channel covariance matrix $\mathbf{\Psi}$ and the signature waveform correlation structure $\mathbf{R}$. Since $\mathbf{\Psi}$ is fixed, we choose $\mathbf{S}$ such that $\mathbf{R}$ minimizes (22). The following theorem addresses this issue.

Theorem 3: Under the constraint trace $(\mathbf{\Phi}) = C$, where C is a constant, $\lambda_1 = \lambda_2 = \cdots = \lambda_{PQL}$ is a sufficient condition to minimize the average BER in (22).

Proof: See Appendix B. □

This is consistent with the fact that maximum diversity gain is obtained when all independent diversity channels have the same average energy [11]. The constraint trace $(\mathbf{\Phi}) = C$ maintains the average received power constant.

The above theorem suggests that, to minimize BER_{AV}, the set of signature codes $\mathbf{S} = [\mathbf{s}_1 \cdots \mathbf{s}_P]$ should be chosen such that

$$
\mathbf{\Phi} = \mathbf{\Psi}(\mathbf{I}_Q \otimes \mathbf{R}) = C\mathbf{I}_{PQL}.
\tag{23}
$$

Without loss of generality, we choose $C = 1/PQL$. Since $\mathbf{\Psi}$ is positive-definite and hence nonsingular, it can be inferred from (23) that the optimal code correlation matrix $\mathbf{R}$ must be

$$
\tilde{\mathbf{w}} = \arg \max_{\mathbf{u} \in \mathbb{C}^P : \|\mathbf{u}\|^2 = 1} \mathbf{u}^H \left(\sum_{q=1}^Q \mathbf{H}_q^* (\mathbf{I}_L \otimes \tilde{\mathbf{c}})^H \mathbf{\Delta}^H \mathbf{\Delta} (\mathbf{I}_L \otimes \tilde{\mathbf{c}}) \mathbf{H}_q^T\right) \mathbf{u}
$$

$$
= \mathbf{ev}_{\max}\left[\sum_{q=1}^Q \mathbf{H}_q^* (\mathbf{I}_L \otimes \tilde{\mathbf{c}})^H \mathbf{\Delta}^H \mathbf{\Delta} (\mathbf{I}_L \otimes \tilde{\mathbf{c}}) \mathbf{H}_q^T\right]
\tag{20}
$$

$$
\text{BER}\left(\{\mathbf{h}_q\}_{q=1}^Q\right) = \mathcal{Q}\left(\sqrt{\frac{2E}{\sigma^2} \lambda_{\max}\left[\sum_{q=1}^Q \mathbf{H}_q^* (\mathbf{I}_L \otimes \tilde{\mathbf{c}})^H \mathbf{\Delta}^H \mathbf{\Delta} (\mathbf{I}_L \otimes \tilde{\mathbf{c}}) \mathbf{H}_q^T\right]}\right)
$$

nonsingular. Hence, from (7), it is easy to see that $\mathbf{S}$ must be full column rank, unlike the optimal rank-one solution for $\mathbf{S}$ when CSI is available ($\mathbf{S} = \mathbf{c}\mathbf{w}^T$). This implies that, when only channel statistics are available at the transmitter, the signature codes used for different transmit antennas should be linearly independent. We refer to this particular structure as the *multicode* scheme. Similar idea of signaling design based upon the channel statistics can be found in [8].

When all the channel coefficients are independent and identically distributed (i.i.d.) so that $\boldsymbol{\Psi} = (1/QL)\mathbf{I}_{PQL}$, the problem is reduced to choosing $\mathbf{S}$ such that $\mathbf{R} = (1/P)\mathbf{I}_{PL}$. For a single-path channel, any set of orthogonal codes can be used, such as Walsh–Hadamard codes. However, for a multipath channel ($L > 1$), the code correlation matrix $\mathbf{R}$ resulting from Walsh–Hadamard codes is quite different from $(1/P)\mathbf{I}_{PL}$. A set of Gold sequences is one choice that results in $\mathbf{R} \approx (1/P)\mathbf{I}_{PL}$ [11]. Hence, choosing $\mathbf{S}$ to be a set of Gold sequences will result in near-optimal performance.

Since some channel coefficients can be correlated or have different average energy, the channel covariance matrix $\boldsymbol{\Psi}$ can be any $PQL \times PQL$ Hermitian positive definite matrix. In this case, using codes that result in $\mathbf{R} \approx (1/P)\mathbf{I}_{PL}$ may result in performance loss compared to the optimal codes. In many cases, there is no set of codes $\mathbf{S}$ that satisfy (23) exactly. A special case when an exact solution exists is when $Q = L = 1$ since for any $N \times P$ semi-unitary matrix $\mathbf{V}$ and $\boldsymbol{\Psi}$ Hermitian and positive definite, we may write $\boldsymbol{\Psi} = \boldsymbol{\Psi}^{H/2}\mathbf{V}^H\mathbf{V}\boldsymbol{\Psi}^{1/2}$. Hence, since $\mathbf{R} = \mathbf{S}^H\mathbf{S}$

$$\mathbf{S} = \mathbf{V}\boldsymbol{\Psi}^{-1/2} \qquad (24)$$

satisfies the optimality condition in (23).

VI. OPTIMAL SCHEME GIVEN DELAYED CHANNEL STATES

The results in Section IV assume that the CSI at the transmitter is perfect. However, in practical systems, some nonidealities may exist. For instance, both FDD and TDD provide delayed versions of the channel state at the transmitter. In addition, for FDD systems, CSI is quantized and suffers from feedback bit error. Delay is by far the most prominent nonideality since sufficiently fine quantization and low error rate feedback channels can be used. In this section, we discuss optimal signaling given a delayed version of the channel state. A comparison between multicode and beamforming schemes in the presence of delay is given in [7] using average BER. Codebook optimization for single path, single receive antenna systems based upon the average received SNR gain and mutual information given nonideal CSI is reported in [4].

We use the Markov (AR-1) model used in [7] for channel state evolution. Define $\hat{h}_{lqp}$ as the delayed version of h_{lqp} and we assume that $\{h_{lqp}\}$ evolve independently. Then, for $l = 1, \ldots, L$, $q = 1, \ldots, Q, p = 1, \ldots, P$, we have

$$h_{lqp} = \rho\,\hat{h}_{lqp} + \theta_{lqp}$$
$$\mathrm{E}\left[\hat{h}_{lqp}^*\theta_{lqp}\right] = 0, \quad \theta_{lqp} \sim \mathcal{N}_C\left[0, (1-\rho^2) \times \mathrm{E}\,|h_{qpl}|^2\right] \qquad (25)$$

where $\rho = J_0(2\pi f_d D)$, $J_0(x)$ denotes the zeroth-order Bessel function of the first kind, f_d is the channel Doppler spread, and D is the amount of delay. The parameter ρ represents the channel fading rate. As the fading rate increases, ρ decreases and the correlation between h_{lqp} and $\hat{h}_{lqp}$ decreases.

To obtain an optimal signaling scheme, we use two criteria: maximize the expected SNR gain, $\mathrm{E}[\mathcal{F}|\{\hat{h}_{qp}\}]$ with $\mathcal{F}$ defined in (9) and minimize the expected BER, $\mathrm{E}[\mathrm{BER}|\{\hat{h}_{qp}\}]$. Due to the effect of delay, maximizing $\mathrm{E}[\mathcal{F}|\{\hat{h}_{qp}\}]$ is not in general equivalent to minimizing $\mathrm{E}[\mathrm{BER}|\{\hat{h}_{qp}\}]$.

The solution to maximization of expected SNR gain is similar to the ideal case in Section IV-A. Substituting $\rho\hat{\mathbf{H}}_q + \boldsymbol{\Theta}_q$ for $\mathbf{H}_q$ into (12), it can be shown that the optimal solution is

$$\mathbf{s} = \mathbf{ev}_{\max}\Big[\sum_{q=1}^{Q} \rho^2 \left(\hat{\mathbf{H}}_q^T \otimes \mathbf{I}_N\right)^H \boldsymbol{\Delta}^H \boldsymbol{\Delta} \left(\hat{\mathbf{H}}_q^T \otimes \mathbf{I}_N\right)$$
$$+ \mathrm{E}\left[\left(\boldsymbol{\Theta}_q^T \otimes \mathbf{I}_N\right)^H \boldsymbol{\Delta}^H \boldsymbol{\Delta} \left(\boldsymbol{\Theta}_q^T \otimes \mathbf{I}_N\right)\right]\Big].$$

Using the techniques shown in Appendix A, this solution can be expressed in the form given in Theorem 2 by defining

$$\boldsymbol{\Sigma}_n = \sum_{q=1}^{Q} \rho^2 \hat{\mathbf{g}}_{n,q}\hat{\mathbf{g}}_{n,q}^H + \mathrm{E}\left[\boldsymbol{\theta}_{n,q}\boldsymbol{\theta}_{n,q}^H\right]$$

$$\hat{\mathbf{g}}_{n,q} = \frac{1}{\sqrt{N}}\begin{bmatrix} \sum_{l=1}^{L} \hat{h}_{lq1}^* e^{j2\pi(n-1)d_l/N} \\ \vdots \\ \sum_{l=1}^{L} \hat{h}_{lqP}^* e^{j2\pi(n-1)d_l/N} \end{bmatrix}$$

$$\boldsymbol{\theta}_{n,q} = \frac{1}{\sqrt{N}}\begin{bmatrix} \sum_{l=1}^{L} \theta_{lq1}^* e^{j2\pi(n-1)d_l/N} \\ \vdots \\ \sum_{l=1}^{L} \theta_{lqP}^* e^{j2\pi(n-1)d_l/N} \end{bmatrix}. \qquad (26)$$

This result demonstrates that the rank-1 beamformer structure maximizes $\mathrm{E}[\mathcal{F}|\{\hat{h}_{qp}\}]$ given the delayed channel state. However, the rank-1 beamformer structure does not generally minimize $\mathrm{E}[\mathrm{BER}|\{\hat{h}_{qP}\}]$.

Exact minimization of $\mathrm{E}[\mathrm{BER}(\mathbf{h})|\{\hat{h}_{qp}\}]$ over $\mathbf{R}$ is not tractable. Hence, we develop some insight by considering a $Q = L = 1$ system and minimize the Chernoff upper bound $\mathrm{BER}(\mathbf{h}) \leq (1/2)\exp(-\mathcal{E}\mathcal{F}/\sigma^2)$ [11]. Here $\mathcal{F} = \mathbf{h}^H\mathbf{S}^H\mathbf{S}\mathbf{h}$ and we assume i.i.d. fading across all transmit antennas: $\mathrm{E}[\mathbf{h}\mathbf{h}^H] = \sigma_h^2\mathbf{I}_P$, where σ_h^2 is the channel coefficient variance. This implies that given the delayed channel state $\hat{\mathbf{h}}$, $\mathbf{h} \sim \mathcal{N}_C[\rho\hat{\mathbf{h}}, \sigma_h^2\mathbf{I}_P]$. Minimization of the Chernoff upper bound under these conditions is discussed in detail in Appendix C. In particular, we derive a closed-form solution for $P = 2$. It is shown in Appendix C that the optimal signature code matrix for $P = 2$ is

$$\mathbf{S} = \sqrt{\lambda_{\mathrm{opt}}} \times \mathbf{v}_1 \frac{\hat{\mathbf{h}}^H}{\|\hat{\mathbf{h}}\|} + \sqrt{1 - \lambda_{\mathrm{opt}}} \times \mathbf{v}_2 \frac{\hat{\mathbf{h}}_\perp^H}{\|\hat{\mathbf{h}}_\perp\|} \quad (27)$$

$$\lambda_{\mathrm{opt}} = \min\{1, \lambda_+\} \qquad (28)$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthonormal vectors, $\hat{\mathbf{h}}^H\hat{\mathbf{h}}_\perp = 0$, and $(1/2) \leq \lambda_{\mathrm{opt}} \leq 1$. The coefficient λ_+ is given in (38) in Appendix C. When $\lambda_{\mathrm{opt}} = 1$, $\mathbf{S}$ represents the beamforming solu-
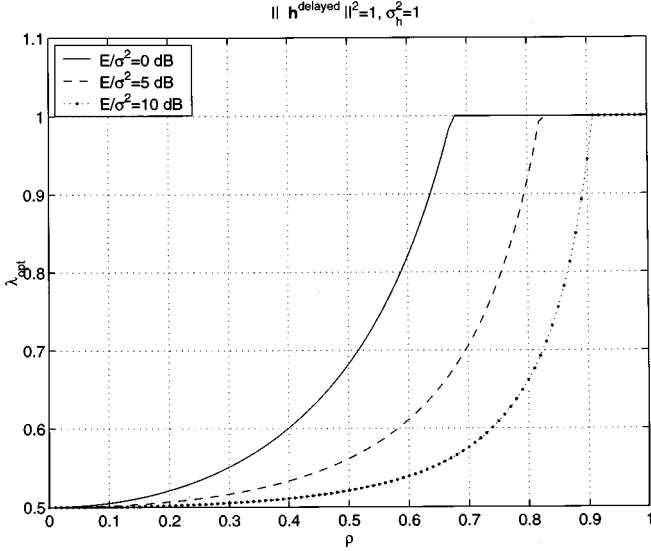
Fig. 3. The optimal λ for $P = 2$, $Q = 1$, and $L = 1$ as a function of ρ for different values of E/σ^2 with $\|\mathbf{h}\|^2 = \sigma_h^2 = 1$.

tion depicted in Fig. 2. On the other hand, $\lambda_{\text{opt}} < 1$ indicates a multicode solution as $\mathbf{S}$ is of rank 2. Notice that the rank 2 $\mathbf{S}$ is a perturbation from the beamforming solution $\mathbf{v}_1 \hat{\mathbf{h}}^H / \|\hat{\mathbf{h}}\|$. This perturbation is needed to achieve optimality due to the uncertainty in the most dominant channel mode derived from the delayed channel state $\hat{\mathbf{h}}$. The value of λ_{opt} in (28) is depicted in Fig. 3 as a function of ρ for various values of E/σ^2 while fixing $\|\hat{\mathbf{h}}\|^2 = 1$ and $\sigma_h^2 = 1$. Observe that the rank-one beamforming solution becomes optimal at smaller values of ρ (faster fading) as E/σ^2 is decreased.

VII. COMPARISON AND EXAMPLES

It is shown in Section IV-B that the suboptimal beamforming (space-only optimized) scheme is inferior to the optimal (space–time optimized) scheme when CSI is available at the transmitter. Also, for single path channel, it is shown that the optimal scheme coincides with the suboptimal scheme.

In this section we first demonstrate that the suboptimal beamforming scheme in Section IV-B based on CSI is superior to the multicode scheme. This is intuitively satisfying since channel states convey more information about the channel than channel statistics. For simplicity, assume that $\boldsymbol{\Psi} = (1/QL)\mathbf{I}_{PQL}$ and

$$(\mathbf{I}_L \otimes \tilde{\mathbf{c}})^H \boldsymbol{\Delta}^H \boldsymbol{\Delta} (\mathbf{I}_L \otimes \tilde{\mathbf{c}}) = \mathbf{I}_L,$$
for the suboptimal beamforming scheme (29)

$$(\mathbf{I}_L \otimes \mathbf{S})^H \boldsymbol{\Delta}^H \boldsymbol{\Delta} (\mathbf{I}_L \otimes \mathbf{S}) = \frac{1}{P} \mathbf{I}_{PL},$$
for the multicode scheme. (30)

These conditions result in the same transmitted signal energy. For the multicode scheme, (30) results in minimum BER_{AV} as it satistifies the optimality condition given in (23). The condition (29) for the signature code $\tilde{\mathbf{c}}$ in the suboptimal beamforming scheme indicates that $\tilde{\mathbf{c}}$ has the same autocorrelation characteristics as the codes in the multicode scheme. By the convolution theorem, it also indicates that $\tilde{\mathbf{c}}$ has an all-pass frequency response, in contrast to the optimal solution implied by Theorem

1. Let $\mathcal{F}_{\text{sub}}$ and $\mathcal{F}_{\text{multi}}$ be SNR gain as defined in (9) for (29) and (30), respectively. Then, from (20), we have

$$\mathcal{F}_{\text{sub}} = \lambda_{\max} \left[\sum_{q=1}^{Q} \mathbf{H}_q^* \mathbf{H}_q^T \right]$$

$$\mathcal{F}_{\text{multi}} = \frac{1}{P} \sum_{q=1}^{Q} \|\mathbf{h}_q\|^2 = \frac{1}{P} \sum_{q=1}^{Q} \text{trace} \left(\mathbf{H}_q^T \mathbf{H}_q^* \right)$$

$$= \frac{1}{P} \sum_{q=1}^{Q} \text{trace} \left(\mathbf{H}_q^* \mathbf{H}_q^T \right)$$

$$= \frac{1}{P} \text{trace} \left(\sum_{q=1}^{Q} \mathbf{H}_q^* \mathbf{H}_q^T \right)$$

$$\leq \frac{1}{P} \times P \times \lambda_{\max} \left[\sum_{q=1}^{Q} \mathbf{H}_q^* \mathbf{H}_q^T \right] = \mathcal{F}_{\text{sub}}.$$

This demonstrates that the best multicode scheme cannot outperform the suboptimum beamforming scheme. Hence, using CSI to optimize only the spatial beamformer is still beneficial. Moreover, for a system with one receive antenna in a single-path channel ($Q = L = 1$), $\mathcal{F}_{\text{sub}} = \|\mathbf{h}_{11}\|^2$ and $\mathcal{F}_{\text{multi}} = (1/P)\|\mathbf{h}_{11}\|^2$. This indicates a $(10 \times \log_{10} P)$ dB SNR gain for the suboptimal beamforming over the multicode scheme.

We next compare the optimal beamforming, suboptimal beamforming, and multicode schemes for a $P = 2$ system with different values of Q and L and $N = 31$. Binary Gold codes are used for one of the suboptimal beamforming and multicode schemes. The correlation properties of binary Gold codes are such that (29) and (30) are closely approximated [11]. The simulated channel is Rayleigh fading with $\boldsymbol{\Psi} = (1/QL)\mathbf{I}_{PQL}$. The average BER curves are depicted in Fig. 4(a)–(d). Observe that, for $L = 1$ [Fig. 4(a) and (b)], the suboptimal beamforming scheme coincides with the optimal one. Also, the gain of the optimal beamforming scheme relative to the suboptimal scheme is more pronounced as the number of paths L increases. This gain is obtained due to the spectral focusing of energy in the frequency selective channel as discussed in Section IV-A. The suboptimal scheme only focuses spatially. For $L = 2$ and 4, we choose $d_l \in \{0, 1\}$ and $\{0, 1, 2, 3\}$, respectively. The suboptimal beamforming schemes utilize the Gold code, the second and the fourth dominant frequencies. Fig. 4(c)–(d) show that the schemes which use subdominant frequencies still outperform the one using the Gold code. The loss associated with use of subdominant frequencies seems to be negligible, yet more pronounced for a larger number of paths L. This is because frequency selectivity increases with the number of paths L.

As an example of signaling design given the channel statistics, we choose a $P = 2$, $Q = 1$ system in a singlepath channel ($L = 1$) with channel statistics

$$\boldsymbol{\Psi} = \boldsymbol{\Psi}_{(2,1)} \stackrel{def}{=} \begin{bmatrix} 1 + \kappa & c\sqrt{1 - \kappa^2} \\ c\sqrt{1 - \kappa^2} & 1 - \kappa \end{bmatrix}. \quad (31)$$

Here, $c \in [-(1/2), (1/2)]$ is the correlation between 2 channel coefficients h_{111} and h_{112} and $\kappa \in (-1, 1)$ represents the en-
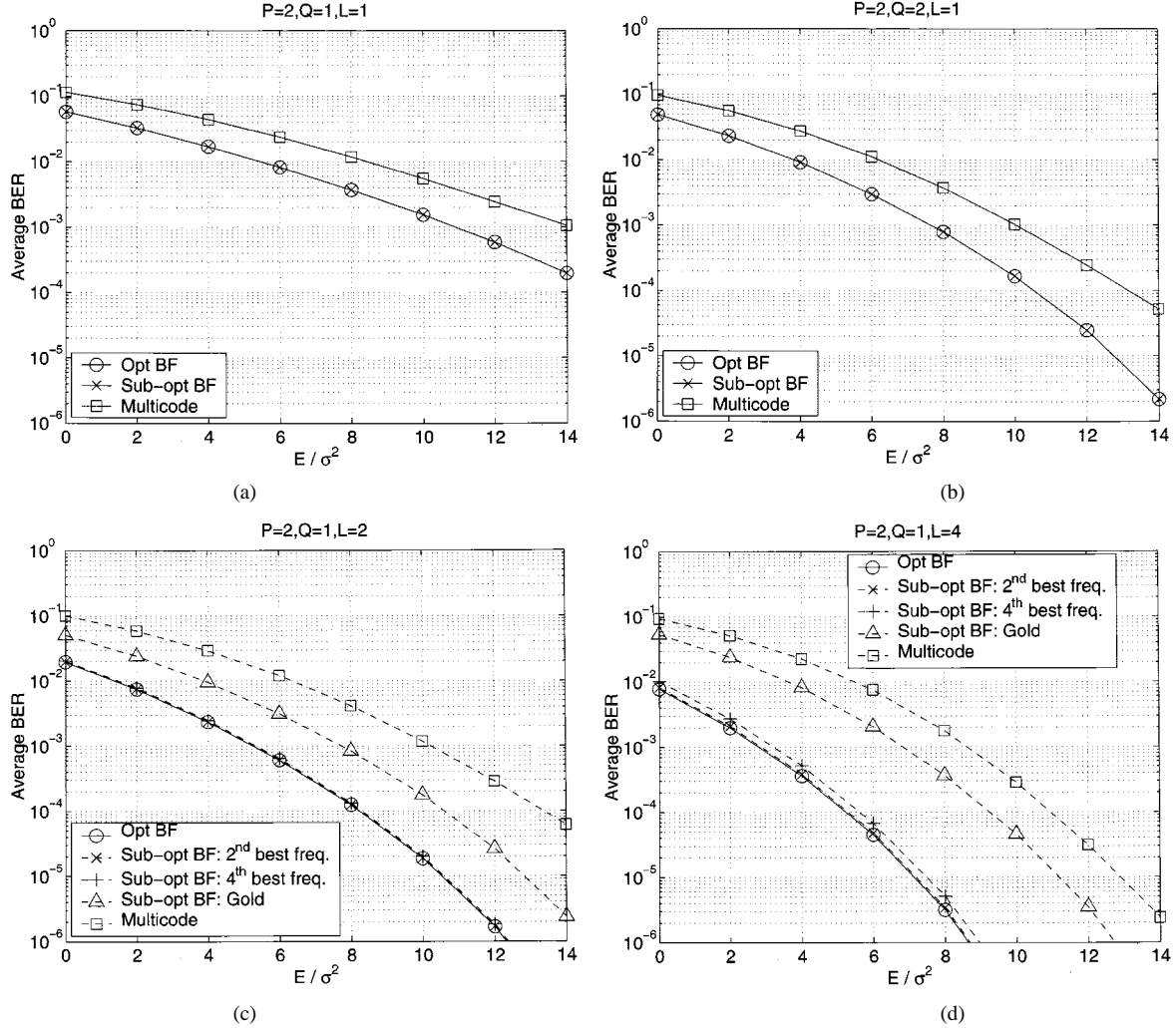
Fig. 4. Average BER comparison for optimal beamforming, sub-optimal beamforming and multicode for $P = 2$ and various values of Q and L. Gold codes are used for the multicode and one of the sub-optimal beamforming schemes. (a) $Q = 1$, $L = 1$. (b) $Q = 2$, $L = 1$. (c) $Q = 1$, $L = 2$. (d) $Q = 1$, $L = 4$.

ergy difference between h_{111} and h_{112}. Decomposing $\mathbf{\Psi}$ as the product of an upper and lower triangular matrix and applying (24) gives

$$\mathbf{s}_1 = \frac{1}{\sqrt{1+\kappa}\sqrt{1-c^2}}\mathbf{u}_1$$

$$\mathbf{s}_2 = \frac{1}{\sqrt{1-\kappa}}\mathbf{u}_2 - \frac{c}{\sqrt{1-\kappa}\sqrt{1-c^2}}\mathbf{u}_1$$

where $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{C}^N$ form an orthonormal set. Fig. 5 shows BER_{AV} of the optimal multicode and Walsh–Hadamard based multicode schemes ($\mathbf{R} = \mathbf{I}_2/2$) with $N = 32$. In this case, the performance of the optimal scheme does not depend upon κ and ρ since (23) is satisfied exactly independent of κ and ρ. As evident, the performance gain of the optimal scheme becomes more significant as c and/or κ increase.

To illustrate signature code design given the channel statistics for a multipath channel, we consider $L = 2$ with $d_1 = 0$, $d_2 = 1$. We model the channel covariance matrix as

$$\mathbf{\Psi} = \begin{bmatrix} 1+\kappa_m & c_m\sqrt{1-\kappa_m^2} \\ c_m\sqrt{1-\kappa_m^2} & 1-\kappa_m \end{bmatrix} \otimes \mathbf{\Psi}_{(2,1)}$$
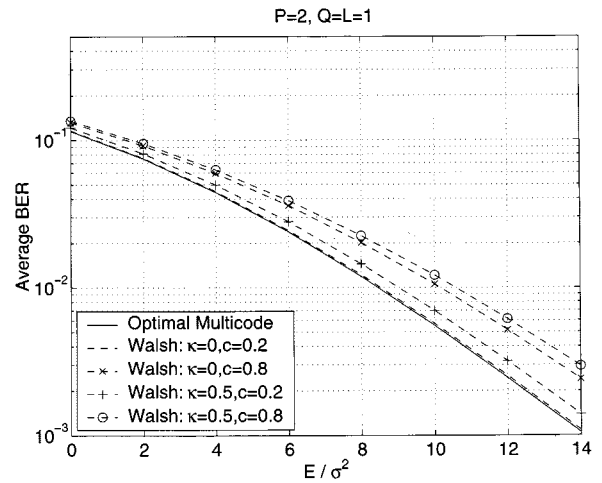

Fig. 5. Average BER comparison between the optimal and orthogonal multicode schemes for $P = 2$, $Q = L = 1$. Various values of κ and c are used.

where $\mathbf{\Psi}_{(2,1)}$ is given in (31). The parameters κ_m and c_m represent the energy difference and correlation between two paths. The code length N is set to be 31. We choose $\kappa = \kappa_m = 0.5$

and $c = c_m = 0.5$ such that the optimality condition (23) cannot be satisfied. The above model for $\boldsymbol{\Psi}$ assumes that two different paths share the same spatial channel characteristic. Since the optimal multicode scheme satisfying (23) is unknown, we employ a solution that approximately satisfies (23) by solving the following equation numerically:

$$\mathbf{S} = \arg \min_{\mathbf{X} \in \mathbb{C}^{31 \times 2}} \left\| \boldsymbol{\Psi} \mathbf{R}(\mathbf{X}) - \frac{1}{4} \mathbf{L}_4 \right\|_F$$
$$\text{s.t.} \operatorname{trace}\left(\boldsymbol{\Psi} \mathbf{R}(\mathbf{X})\right) = 1 \qquad (32)$$

where $\mathbf{R}(\mathbf{X}) = (\mathbf{I}_2 \otimes \mathbf{X})^H \boldsymbol{\Delta}^H \boldsymbol{\Delta} (\mathbf{I}_2 \otimes \mathbf{X})$ and $\| \cdot \|_F$ denotes the Frobenius norm [10]. The average BER of the optimal and Gold-sequence based multicode schemes are shown in Fig. 6. Observe that the optimal multicode outperforms the Gold-based multicode by ~ 1.5-dB at $\text{BER}_{\text{AV}} = 10^{-3}$. The results with $\boldsymbol{\Phi} = \mathbf{I}_4/4$ represents an unattainable BER lower bound, which is shown only for comparison.

To illustrate the results obtained in Section VI when delayed channel states are used, we compare the performance of the optimal solution in (27) to the beamforming (rank-1) and multicode schemes with $P = 2$, $Q = L = 1$. The multicode scheme consists of the two length-32 Walsh-Hadamard codes for two transmit antennas. The simulated average BER for $\rho = 0.95, 0.8, 0.6$ are depicted in Fig. 7(a)–(c)[2]. Observe that for all fading rates, the performance of the optimal solution tends to coincide with that of beamforming for small E/σ^2 and with multicode for higher E/σ^2. At moderate E/σ^2, beamforming is near-optimal for low Doppler spread and multicode is near optimal for high Doppler spread.

These results indicate that the beamforming solution is optimal only when the delay is small relative to the channel fading rate or the fading is sufficiently slow, i.e., $2\pi D f_d \ll 1$. As the fading rate increases, the optimality of the multicode scheme is attributed to the increase of uncertainty of the most dominant channel mode. However, the rank-1 scheme is always better when E/σ^2 is small. It is reasonable to expect that this claim also holds for $P > 2$ and $Q, L > 1$. Furthermore, for other types of nonidealities, beamforming solution is perceived to be optimal when the nonidealities are sufficiently mild. This conclusion is similar to the results in [4].

VIII. CONCLUSION

The design of diversity signaling schemes that use CSI at the transmitter is investigated. The channel side information is either the channel state or the channel statistics. It is shown that when perfect channel state information is available at the transmitter, the BER-minimizing scheme consists of a common signature code for all transmit antennas, followed by a beamformer which focuses the transmission to the most dominant spatio-temporal channel mode at any particular instant (the space-time optimized scheme). The optimal space-time beamformer scheme can be efficiently implemented
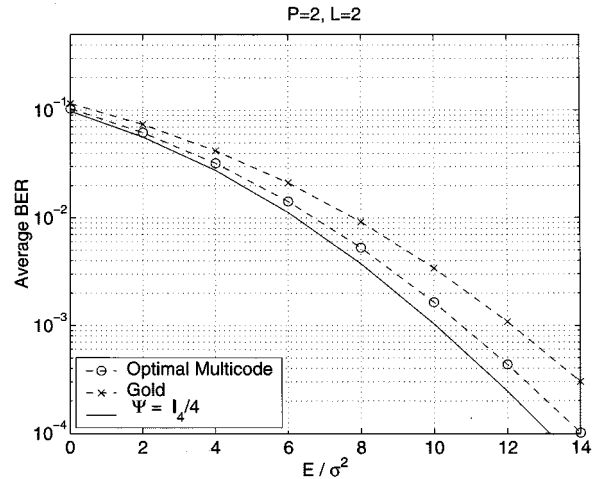
Fig. 6. Average BER comparison between the optimal and orthogonal multicode schemes for $P = L = 2$, $Q = 1$. We use $\kappa = \kappa_m = 0.5$ and $c = c_m = 0.5$. The results with $\boldsymbol{\Phi} = \mathbf{I}_4/4$ represent an unattainable BER lower bound, which is shown only for comparison.

using IDFT-DFT bank as in multicarrier systems, followed by an adaptive array. The beamformer computation involves a DFT and finding the most dominant eigenvector of a $P \times P$ matrix. When the temporal signature code is fixed, the CSI can still be used to perform space-only optimization. When only the channel statistics are available at the transmitter, the BER-minimizing solution suggests that the signal should be transmitted throughout all the channel modes to combat the effects of fading. This is done by utilizing a set of linearly independent signature codes for the transmitter array that match the channel statistics so as to provide independent and identical fading subchannels (the multicode scheme). Finally, we discuss signaling design based upon delayed channel state information. It is demonstrated that the BER-minimizing scheme is a rank-1 space-time beamformer for lower fading rates and a multicode scheme for high fading rates.

When CSI is available at the transmitter, the solution presented in this paper (space–time or space-only optimization) requires the computation of the dominant eigenvector of $NP \times P$ matrices. When $P = 2$, a closed-form solution is possible. For a relatively small P, numerical techniques such as power method and its variants [14], which compute only the dominant eigenvector can be utilized to provide the solution in a reasonable amount of time. When P is large, more sophisticated algorithms may be needed.

APPENDIX A
PROOF OF THEOREM 2

Since $\boldsymbol{\Delta}_{d_l}$ is circulant, $\boldsymbol{\Delta}_{d_l} = \mathbf{F}^H \boldsymbol{\Lambda}_{d_l} \mathbf{F}$, where $\mathbf{F} \in \mathbb{C}^{N \times N}$ is the N-DFT matrix and [14]

$$\boldsymbol{\Lambda}_{d_l} = \sqrt{N} \operatorname{diag}\{\mathbf{F} \mathbf{e}_{d_l+1}\}$$
$$= \operatorname{diag}\left\{1, e^{-j2\pi d_l/N}, \ldots, e^{-j2\pi(N-1)d_l/N}\right\}$$

since $\mathbf{e}_{d_l+1}$ is the first column of $\boldsymbol{\Delta}_{d_l}$. Define the matrix

---

[2]For 0.667-ms delay (1 slot in WCDMA [1]), $\rho = 0.95, 0.8, 0.6$ correspond to Doppler spreads of 107, 219 319 Hz ($\sim 58, 118, 172$ kmph at 2 GHz center frequency), respectively.
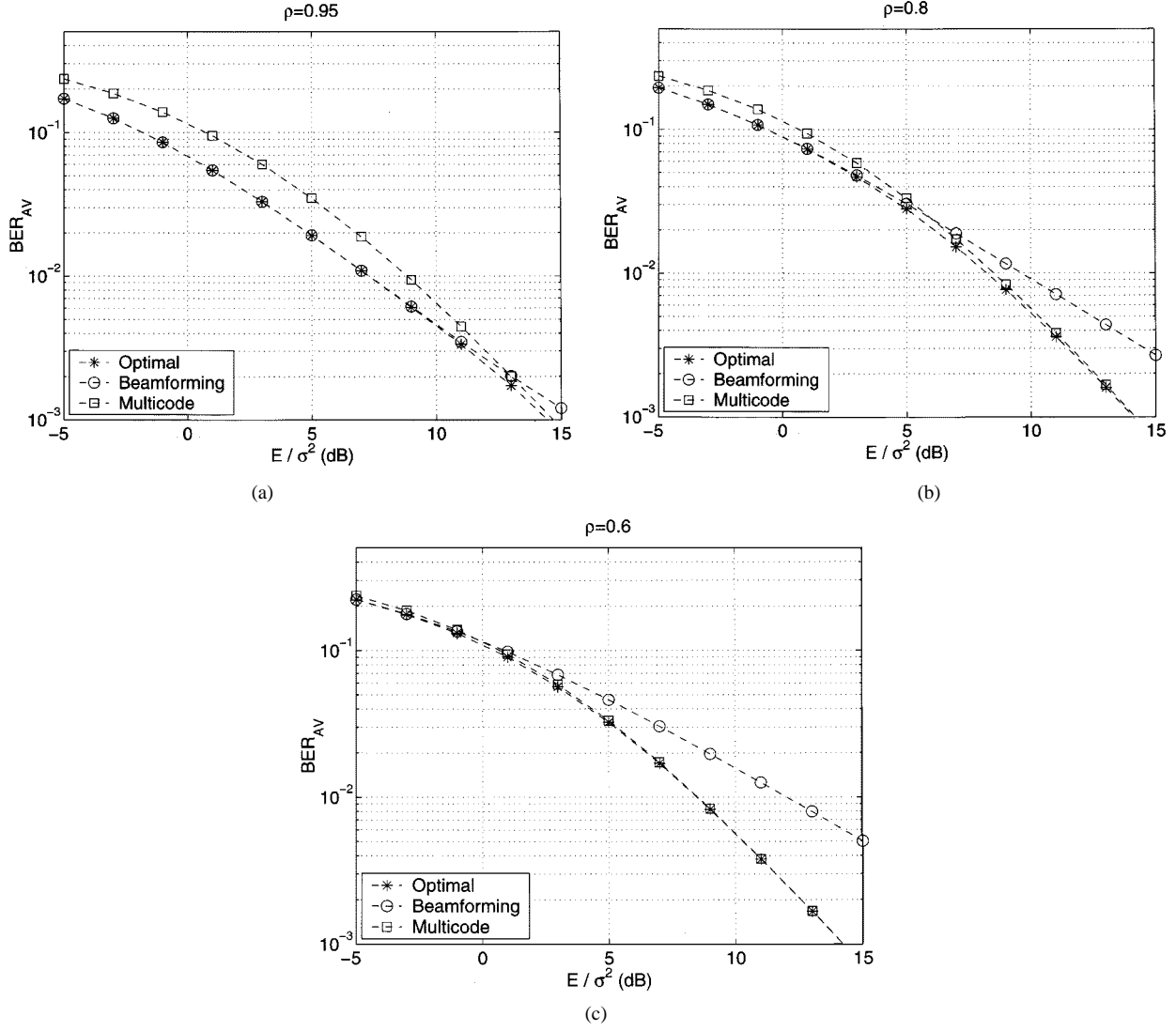
Fig. 7.   Average BER comparison for optimal diversity scheme, beamforming and multicode schemes for $P = 2$, $Q = L = 1$ when the delayed channel state is used at the transmitter. (a) $\rho = 0.95$. (b) $\rho = 0.8$. (c) $\rho = 0.6$.

$\mathbf{D}_{qP} \overset{def}{=} \sum_{l=1}^{L} h_{lqP} \mathbf{\Lambda}_{d_l}$. Then

$$\mathbf{\Delta} \left(\mathbf{H}_q^T \otimes \mathbf{I}_N \right) = \left[\sum_{l=1}^{L} h_{lq1} \mathbf{\Delta}_{d_l} \ \cdots \ \sum_{l=1}^{L} h_{lqP} \mathbf{\Delta}_{d_l} \right]$$
$$= \left[\mathbf{F}^H \mathbf{D}_{q1} \mathbf{F} \ \cdots \ \mathbf{F}^H \mathbf{D}_{qP} \mathbf{F} \right]$$
$$= \mathbf{F}^H \left[\mathbf{D}_{q1} \ \cdots \ \mathbf{D}_{qP} \right] \left(\mathbf{I}_P \otimes \mathbf{F} \right) \quad (33)$$

where the last equality in (33) follows from the identity $\left(\mathbf{X}_1 \otimes \mathbf{X}_2 \right)^H = \mathbf{X}_1^H \otimes \mathbf{X}_2^H$. Hence,

$$\left(\mathbf{H}_q^T \otimes \mathbf{I}_N \right)^H \mathbf{\Delta}^H \mathbf{\Delta} \left(\mathbf{H}_q^T \otimes \mathbf{I}_N \right)$$
$$= \left(\mathbf{I}_P \otimes \mathbf{F} \right)^H \underbrace{\begin{bmatrix} \mathbf{D}_{q1}^H \mathbf{D}_{q1} & \cdots & \mathbf{D}_{q1}^H \mathbf{D}_{qP} \\ \vdots & \ddots & \vdots \\ \mathbf{D}_{qP}^H \mathbf{D}_{q1} & \cdots & \mathbf{D}_{qP}^H \mathbf{D}_{qP} \end{bmatrix}}_{\mathbf{\Gamma}_q} \left(\mathbf{I}_P \otimes \mathbf{F} \right).$$

The $PN \times PN$ matrix $\mathbf{\Gamma}_q$ can be rearranged into a block diagonal matrix $\mathbf{G}_q$ as follows:

$$\mathbf{G}_q = \begin{bmatrix} \mathbf{G}_{0,q} & & \\ & \ddots & \\ & & \mathbf{G}_{N-1,q} \end{bmatrix} = \mathbf{\Pi}_{(P,N)}^H \mathbf{\Gamma}_q \mathbf{\Pi}_{(P,N)},$$

$$\mathbf{G}_{n,q} = \mathbf{g}_{n,q} \mathbf{g}_{n,q}^H$$

where $\mathbf{\Pi}_{(P,N)} \overset{def}{=} \sum_{i=1}^{P} \mathbf{e}_i \otimes \mathbf{I}_N \otimes \mathbf{e}_i^T$, $\mathbf{e}_i \in \mathbb{C}^P$ is the (P, N) unitary permutation matrix defined in [9] and $\mathbf{g}_{n,q}$ is given in (17). Similar block diagonalization technique has been used in [12]. Thus,

$$\sum_{q=1}^{Q} \left(\mathbf{H}_q^T \otimes \mathbf{I}_N \right)^H \mathbf{\Delta}^H \mathbf{\Delta} \left(\mathbf{H}_q^T \otimes \mathbf{I}_N \right)$$

$$\times \left(\mathbf{I}_P \otimes \mathbf{F} \right)^H \mathbf{\Pi}_{(P,N)} \left(\sum_{q=1}^{Q} \mathbf{G}_q \right) \mathbf{\Pi}_{(P,N)}^H \left(\mathbf{I}_P \otimes \mathbf{F} \right).$$

To complete the proof, we need the following two basic lemmas from linear algebra.

Lemma 1: Let $\mathbf{A}$, $\mathbf{V}$ be $M \times M$ matrices and $\mathbf{V}$ be unitary. Let $\tilde{\mathbf{A}} = \mathbf{V}\mathbf{A}\mathbf{V}^H$. Then $\mathbf{ev}_{\max}[\tilde{\mathbf{A}}] = \mathbf{V}\mathbf{ev}_{\max}[\mathbf{A}]$.

Lemma 2: Let $\mathbf{A} = \begin{bmatrix} \mathbf{A}_0 & & \\ & \ddots & \\ & & \mathbf{A}_{N-1} \end{bmatrix}$ be a block diagonal matrix where $\mathbf{A}_n \in \mathbb{C}^{P \times P}$. Then, $\mathbf{ev}_{\max}[\mathbf{A}] = \mathbf{e}_{\overline{n}} \otimes \mathbf{ev}_{\max}[\mathbf{A}_{\overline{n}}]$, where $\overline{n} = \arg\max_n \lambda_{\max}[\mathbf{A}_n]$.

Applying Lemma 1, 2 and the fact that $(\mathbf{I}_P \otimes \mathbf{F})\mathbf{\Pi}_{(P,N)}$ is unitary, we have

$$
\mathbf{ev}_{\max}\left[(\mathbf{I}_P \otimes \mathbf{F})^H \mathbf{\Pi}_{(P,N)}\left(\sum_{q=1}^{Q}\mathbf{G}_q\right)\mathbf{\Pi}_{(P,N)}^H(\mathbf{I}_P \otimes \mathbf{F})\right]
$$
$$
=(\mathbf{I}_P \otimes \mathbf{F})^H \mathbf{\Pi}_{(P,N)}\left(\mathbf{e}_{\overline{n}} \otimes \mathbf{ev}_{\max}\left[\sum_{q=1}^{Q}\mathbf{G}_{\overline{n},q}\right]\right)
$$
$$
=(\mathbf{I}_P \otimes \mathbf{F})^H\left(\mathbf{ev}_{\max}\left[\sum_{q=1}^{Q}\mathbf{G}_{\overline{n},q}\right] \otimes \mathbf{e}_{\overline{n}}\right)
$$
$$
=\underbrace{\mathbf{ev}_{\max}\left[\sum_{q=1}^{Q}\mathbf{G}_{\overline{n},q}\right]}_{\mathbf{w}} \otimes \underbrace{(\mathbf{F}^H\mathbf{e}_{\overline{n}})}_{\mathbf{c}}
$$

where the second and third equalities follow from the following identities [9].

1) For any $\mathbf{X}_1 \in \mathbb{C}^{p \times q}$, $\mathbf{X}_2 \in \mathbb{C}^{s \times t}$, $\mathbf{Y}_1 \in \mathbb{C}^{q \times r}$, $\mathbf{Y}_2 \in \mathbb{C}^{t \times u}$, $\mathbf{X}_2 \otimes \mathbf{X}_1 = \mathbf{\Pi}_{(s,p)}(\mathbf{X}_1 \otimes \mathbf{X}_2)\mathbf{\Pi}_{(q,t)}$.
2) $(\mathbf{X}_1 \otimes \mathbf{X}_2)(\mathbf{Y}_1 \otimes \mathbf{Y}_2) = \mathbf{X}_1\mathbf{Y}_1 \otimes \mathbf{X}_2\mathbf{Y}_2$.

APPENDIX B
PROOF OF THEOREM 3

Let $\gamma(\theta) \overset{\text{def}}{=} E/\sigma^2 \sin^2 \theta > 0$, $\boldsymbol{\lambda} \overset{\text{def}}{=} [\lambda_1, \ldots, \lambda_{PQL}]^T$ and $f(\gamma(\theta), \boldsymbol{\lambda}) \overset{\text{def}}{=} \prod_{n=1}^{PQL}(1 + \gamma(\theta)\lambda_n)$. The constant trace constraint trace $(\boldsymbol{\Phi}) = C$ is equivalent to $\sum_{n=1}^{PQL} \lambda_n = C$. We seek to solve the following optimization problem:

$$
\boldsymbol{\lambda}_{\text{opt}} = \arg\max_{\lambda_1 \geq 0, \ldots, \lambda_{PQL} \geq 0} f(\boldsymbol{\lambda}, \gamma(\theta)) \quad \text{s.t.} \quad \sum_{n=1}^{PQL}\lambda_n = C.
$$

Note that $\text{BER}_{AV} = (1/\pi)\int_0^{\pi/2}[f(\gamma(\theta), \boldsymbol{\lambda})]^{-1} d\theta$. If the optimizer $\boldsymbol{\lambda}_{\text{opt}}$ does not depend on $\gamma(\theta)$, it follows that BER_{AV} is minimized at $\boldsymbol{\lambda}_{\text{opt}}$. Notice that $f(\gamma(\theta), \boldsymbol{\lambda})$ is trivially concave in $\boldsymbol{\lambda}$.

This problem can be solved using Lagrange multiplier techniques. We form the Lagrangian and its partial derivatives

$$
\mathcal{L}(\boldsymbol{\lambda}) = f(\boldsymbol{\lambda}, \gamma(\theta)) - \mu\left(\sum_{n=1}^{PQL}\lambda_n - C\right)
$$
$$
\frac{\partial \mathcal{L}}{\partial \lambda_{n'}} = \gamma(\theta)\prod_{n \neq n'}(1 + \gamma(\theta)\lambda_n) - \mu
$$
$$
= \gamma(\theta)\frac{\prod_{n=1}^{PQL}(1 + \gamma(\theta)\lambda_n)}{1 + \gamma(\theta)\lambda_{n'}} - \mu, \quad n' = 1, \ldots, PQL. \tag{34}
$$

Setting all the partial derivates in (34) to 0, we have

$$
\lambda_{n',\text{opt}} = \frac{\prod_{n=1}^{PQL}(1 + \gamma(\theta)\lambda_{n,\text{opt}})}{\mu} - \frac{1}{\gamma(\theta)},
$$
$$
n' = 1, \ldots, PQL. \tag{35}
$$

This implies $\lambda_{n',\text{opt}}$ is independent of n' and thus $\lambda_{1,\text{opt}} = \lambda_{2,\text{opt}} = \cdots = \lambda_{PQL,\text{opt}}$. Enforcing the constraint we have $\lambda_{n',\text{opt}} = C/PQL$ for $n' = 1, \ldots, PQL$. It is easy to check that this solution satisfies Kuhn–Tucker optimality condition.

APPENDIX C
MINIMIZING $\text{E}[\exp(-(E/\sigma^2)\mathbf{h}^H\mathbf{S}^H\mathbf{S}\mathbf{h}) \mid \hat{\mathbf{h}}]$

Applying the eigendecomposition $\mathbf{S}^H\mathbf{S} = \mathbf{U}\boldsymbol{\Sigma}^2\mathbf{U}^H$ where $\boldsymbol{\Sigma}^2 = \text{diag}\{\lambda_1, \ldots, \lambda_P\}$ and defining $\mathbf{f} = [f_1, f_2, \ldots, f_P]^T = \boldsymbol{\Sigma}\mathbf{U}^H\mathbf{h}$, it can be shown that

$$
\text{E}\left[\exp\left(-\frac{E}{\sigma^2}\mathbf{h}^H\mathbf{S}^H\mathbf{S}\mathbf{h}\right) \Big| \hat{\mathbf{h}}\right]
$$
$$
= \prod_{p=1}^{P}\underbrace{\text{E}\left[\exp\left(-\frac{E}{\sigma^2}|f_p|^2\right)\right]}_{\phi_p}
$$
$$
|f_p|^2 \sim \frac{\lambda_p\sigma_h^2(1-\rho^2)}{2}\chi_2^2\left(\frac{2\rho^2}{\lambda_p\sigma_h^2(1-\rho^2)}\left|\left[\boldsymbol{\Sigma}\mathbf{U}^H\hat{\mathbf{h}}\right]_p\right|^2\right)
$$

where $\chi_M^2(\delta^2)$ denotes a noncentral chi-squared distributed random variable with M degrees of freedom and noncentrality parameter δ^2. Note that for any $P \times P$ unitary matrix $\mathbf{U}$, there exists parameters $\{\mu_p\}_{p=1}^{P}$, $0 \leq \mu_p \leq 1$, $\sum_{p=1}^{P}\mu_p = 1$ and angles $\{\varphi_p\}_{p=1}^{P}$ such that

$$
\mathbf{U}^H\hat{\mathbf{h}} = \left\|\hat{\mathbf{h}}\right\| \times \left[\sqrt{\mu_1}e^{j\varphi_1} \sqrt{\mu_2}e^{j\varphi_2} \cdots \sqrt{\mu_P}e^{j\varphi_P}\right]^T. \tag{36}
$$

Using the above parametrization and the generating function of the noncentral chi-squared random variable [11], we have

$$
\phi_p = \text{E}\left[\exp\left(-\frac{E}{\sigma^2}|f_p|^2\right)\right]
$$
$$
= \frac{1}{1 + \frac{E}{\sigma^2}\sigma_h^2(1-\rho^2)\lambda_p}
$$
$$
\times \exp\left(-\frac{E}{\sigma^2}\rho^2\|\hat{\mathbf{h}}\|^2\frac{\lambda_p\mu_p}{1 + \frac{E}{\sigma^2}\sigma_h^2(1-\rho^2)\lambda_p}\right).
$$

Notice that the phase angles $\{\varphi_p\}$ are immaterial. Hence, the optimization problem is reduced to finding nonnegative parameters $\{(\lambda_p, \mu_p)\}_{p=1}^{P}$ such that $\sum_{p=1}^{P}\lambda_p = 1$, $\sum_{p=1}^{P}\mu_p = 1$ and $\prod_{p=1}^{P}\phi_p$ is minimized. Note that once the optimal $\{(\lambda_p, \mu_p)\}_{p=1}^{P}$ are obtained, the optimum $N \times P$ signature code matrix is $\mathbf{S} = \mathbf{V}\boldsymbol{\Sigma}\mathbf{U}^H$, where $\mathbf{V}$ is any $N \times P$ semi-unitary matrix, $\boldsymbol{\Sigma} = \text{diag}\{\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_P}\}$ and $\mathbf{U}$ is chosen to satisfy (36).

For $P = 2$, let $\lambda_1 = \lambda$ and $\mu_1 = \mu$. Let $\gamma_h \overset{def}{=} (E/\sigma^2)\sigma_h^2$. Then, we have

$$
\phi_1\phi_2 = \frac{\exp\left(-\rho^2\frac{E}{\sigma^2}\|\mathbf{h}\|^2\left[\frac{\lambda\mu}{1+\gamma_h(1-\rho^2)\lambda} + \frac{(1-\lambda)(1-\mu)}{1+\gamma_h(1-\rho^2)(1-\lambda)}\right]\right)}{(1+\gamma_h(1-\rho^2)\lambda)(1+\gamma_h(1-\rho^2)(1-\lambda))}. \tag{37}
$$

$$
\exp\left(-\rho^2\frac{E}{\sigma^2}\left\|\hat{\mathbf{h}}\right\|^2\left[\underbrace{\frac{2\lambda-1}{\left(1+\gamma_h\left(1-\rho^2\right)\lambda\right)\left(1+\gamma_h\left(1-\rho^2\right)\left(1-\lambda\right)\right)}}_{C_1}\mu+\frac{1-\lambda}{1+\gamma_h\left(1-\rho^2\right)\left(1-\lambda\right)}\right]\right)
$$

$$
\lambda_+ = \frac{\gamma_h\left(1-\rho^2\right)^2-2\left(1-\rho^2\right)-\rho^2\frac{\left\|\hat{\mathbf{h}}\right\|^2}{\sigma_h^2}}{4\gamma_h\left(1-\rho^2\right)^2}+\frac{\sqrt{\left[\gamma_h\left(1-\rho^2\right)^2+3\rho^2\frac{\left\|\hat{\mathbf{h}}\right\|^2}{\sigma_h^2}+2\left(1-\rho^2\right)\right]^2-8\rho^4\frac{\left\|\hat{\mathbf{h}}\right\|^4}{\sigma_h^4}}}{4\gamma_h\left(1-\rho^2\right)^2} \tag{38}
$$

By symmetry, we only need to consider $\lambda \in [(1/2), 1]$, $\mu \in [(1/2), 1]$. The denominator in (37) is independent of μ and the numerator can be written as shown at the top of the page. Observe that $C_1 \geq 0$ for $\lambda \in [(1/2), 1]$. This shows that the minimum of $\phi_1 \phi_2$ occurs at $\mu = 1$. Then, it follows from (36) that

$$
\mathbf{U} = \left[\frac{\hat{\mathbf{h}}}{\left\|\hat{\mathbf{h}}\right\|}\quad\frac{\hat{\mathbf{h}}_\perp}{\left\|\hat{\mathbf{h}}_\perp\right\|}\right]
$$

where $\hat{\mathbf{h}}^H \hat{\mathbf{h}}_\perp = 0$.

Now, let $J(\lambda) \stackrel{\text{def}}{=} 1/(\phi_1\phi_2)$ with $\mu = 1$. That is,

$$
J(\lambda) = \left(1+\gamma_h\left(1-\rho^2\right)\lambda\right)\left(1+\gamma_h\left(1-\rho^2\right)\left(1-\lambda\right)\right)
$$
$$
\times\exp\left(\rho^2\frac{E}{\sigma^2}\|\hat{\mathbf{h}}\|^2\frac{\lambda}{1+\gamma_h\left(1-\rho^2\right)\lambda}\right)
$$

Then, the optimization problem is reduced to finding $\lambda_{opt} = \arg\max_{\lambda\in[(1/2),1]}J(\lambda)$. The first derivative condition implies that the critical points of $J(\lambda)$ satisfy $A\lambda^2 + B\lambda + C = 0$, where

$$
A = -2\gamma_h^2\left(1-\rho^2\right)^3,
$$

$$
B = \gamma_h^2\left(1-\rho^2\right)^3-2\gamma_h\left(1-\rho^2\right)^2-\rho^2\gamma_h\frac{\left\|\hat{\mathbf{h}}\right\|^2}{\sigma_h^2}\left(1-\rho^2\right),
$$

$$
C = \gamma_h\left(1-\rho^2\right)+\rho^2\frac{\left\|\hat{\mathbf{h}}\right\|^2}{\sigma_h^2}\left(1+\gamma_h\left(1-\rho^2\right)\right).
$$

It is easy to show that the larger root of this equation is defined as in (38), shown at the top of the page. It can be shown (by the second derivative test) that $J(\lambda)$ is concave and λ_+ is a global maximum. Since

$$
\left[\gamma_h\left(1-\rho^2\right)^2+3\rho^2\frac{\left\|\hat{\mathbf{h}}\right\|^2}{\sigma_h^2}+2\left(1-\rho^2\right)\right]^2-8\rho^4\frac{\left\|\hat{\mathbf{h}}\right\|^4}{\sigma_h^4}
$$
$$
= \left[\gamma_h\left(1-\rho^2\right)^2+\rho^2\frac{\left\|\hat{\mathbf{h}}\right\|^2}{\sigma_h^2}+2\left(1-\rho^2\right)\right]^2
$$
$$
+4\rho^2\frac{\left\|\hat{\mathbf{h}}\right\|^2}{\sigma_h^2}\left[\gamma_h\left(1-\rho^2\right)^2+2\left(1-\rho^2\right)\right]
$$
$$
\geq \left[\gamma_h\left(1-\rho^2\right)^2+\rho^2\frac{\left\|\hat{\mathbf{h}}\right\|^2}{\sigma_h^2}+2\left(1-\rho^2\right)\right]^2
$$

it follows that $\lambda_+ \geq (1/2)$. The solution λ_{opt} in (28) follows from enforcing the constraint $\lambda \in [(1/2), 1]$ and using the fact that $J(\lambda)$ is strictly increasing within $\lambda \in [(1/2), \lambda_+]$.

REFERENCES

[1] [Online]. Available: http://www.3gpp.org
[2] J. H. Winters, "Smart antennas for wireless systems," *IEEE Personal Commun.*, pp. 23–27, Feb. 1998.
[3] A. Scaglione, S. Barbarossa, and G. Giannakis, "Filterbank transceiver optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1019–1032, Apr. 1999.
[4] 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.
[5] G. Ganesan and P. Stoica, "Space-time diversity using orthogonal and amicable orthogonal designs," in *Proc. ICASSP 2000*, vol. 5, 2000, pp. 2561–2564.
[6] G. Giannakis, Y. Hua, and P. Stoica, *Signal Processing Advances in Wireless Communications*. Englewood Cliffs, NJ: Prentice-Hall, 2000, vol. 2.
[7] E. N. Onggosanusi, A. Gatherer, A. G. Dabak, and S. Hosur, "Performance analysis of closed-loop transmit diversity in the presence of feedback delay," *IEEE Trans. Commun.*, vol. 49, pp. 1618–1630, Sept. 2001.
[8] M. S. Alouini, A. Scaglione, and G. B. Giannakis, "PCC—principal components combining for dense correlated multipath fading environments," in *Proc. VTC Fall 2000*, Boston, September 2000.
[9] J. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 772–781, Sept. 1978.
[10] R. Horn and C. Johnson, *Topics in Matrix Analysis*. Cambridge, MA: Cambridge Univ. Press, 1991.
[11] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw Hill, 1995.
[12] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.
[13] S. Verdu, *Multiuser Detection*. Cambridge, MA: Cambridge Univ. Press, 1998.
[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: John Hopkins Univ. Press, 1996.
[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, September 1998.

Eko N. Onggosanusi received the B.Sc. (with highest distinction), M.Sc., and Ph.D. degrees in 1996, 1998 and 2000, respectively, from the University of Wisconsin-Madison, all in electrical and computer engineering.

During the summer of 1999 and 2000, he was a visiting student at the DSPS R & D Texas Instruments, Dallas, TX, working on antenna array diversity for wideband CDMA and adaptive modulation and coding for wireless personal area networks, respectively. Since January 2001, he has been with the DSPS R & D Texas Instruments, where he is currently a member of technical staff in the mobile wireless branch of communication systems laboratory. He is currently working on the third-generation wireless communication systems, focusing on high data rate techniques and applications. His research interests include signal processing, information theory and channel coding for digital communications.

Akbar M. Sayeed (S'89–M'97) received the B.S. degree from the University of Wisconsin-Madison in 1991 and the M.S. and Ph.D. degrees from the University of Illinois at Urbana-Champaign, in 1993 and 1996, respectively, all in electrical and computer engineering.

While at the University of Illinois, he was a research assistant in the Coordinated Science Laboratory and was also the Schlumberger Fellow in signal processing from 1992to1995. During 1996–1997, he was a Post-Doctoral Fellow at Rice University, Houston, TX. Since August 1997, he has been with the University of Wisconsin-Madison, where he is currently an Assistant Professor in Electrical and Computer Engineering. His research interests are in wireless communications, statistical signal processing, time-frequency and wavelet analysis.

Dr. Sayeed received the NSF CAREER Award in 1999 and the ONR Young Investigator Award in 2001. He is currently serving as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS.

Barry D. Van Veen (S'81–M'86) was born in Green Bay, WI. He received the B.S. degree from Michigan Technological University in 1983 and the Ph.D. degree from the University of Colorado in 1986, both in electrical engineering.

He was an ONR Fellow while working on the Ph.D. degree. In the spring of 1987, he was with the Department of Electrical and Computer Engineering at the University of Colorado–Boulder. Since August of 1987, he has been with the Department of Electrical and Computer Engineering at the University of Wisconsin–Madison and currently holds the rank of Professor. His research interests include signal processing for sensor arrays, nonlinear systems, adaptive filtering, wireless communications and biomedical applications of signal processing. He coauthored *Signals and Systems* (New York: Wiley, 1999) with Simon Haykin.

Dr. Van Veen was a recipient of a 1989 Presidential Young Investigator Award from the National Science Foundation and a 1990 IEEE Signal Processing Society Paper Award. He served as an associate editor for the IEEE Transactions on Signal Processing and on the IEEE Signal Processing Society's Technical Committee on Statistical Signal and Array Processing from 1991 through 1997 and is currently a member of the Sensor Array and Multichannel Technical Committee. He received the Holdridge Teaching Excellence Award from the ECE Department at the University of Wisconsin in 1997.