MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Channel Statistics-Based RF Pre-Processing with Antenna Selection

Pallav Sudarshan, Neelesh Mehta, Andreas Molisch, Jin Zhang

TR2006-101 December 2006

Abstract

We introduce two novel joint radio-frequency (RF)-baseband designs for receivers in a MIMO system with Nt transmit antennas, Nr receive antennas, but only L less-than Nr RF chains at the receiver. The joint design introduces an RF pre-processing matrix that processes the signals from the different antennas, and is followed by selection (if necessary), down-conversion, and further processing in the baseband. The schemes are similar to conventional antenna selection in that they use fewer RF chains than antenna elements, but achieve superior performance by exploiting the spatial correlation of the received signals. The first of our proposed designs uses an L x Nr RF pre-processing matrix that outputs only L streams followed by baseband signal processing, and, thus, eliminates the need for a selection switch. The second one uses an Nr x Nr RF pre-processing matrix that outputs Nr streams and is followed by a switch that selects L streams for baseband signal processing. Both spatial diversity and spatial multiplexing systems are considered and the optimum pre-processing matrices are derived for all cases. To accommodate practical RF design constraints, which prefer a variable phase-shifter-based implementation, a sub-optimal phase approximation is also introduced. Performance better than conventional antenna selection and close to the full complexity receiver is observed in both single cluster and multi-cluster wireless channels. A beam-pattern-based geometric intuition is also developed to illustrate the effectiveness of the optimal solution.

IEEE Transactions on Wireless Communications

Channel Statistics-Based RF Pre-Processing with Antenna Selection

Pallav Sudarshan, *Member, IEEE*, Neelesh B. Mehta, *Member, IEEE*, Andreas F. Molisch, *Fellow, IEEE*, and Jin Zhang *Senior Member, IEEE*

***Abstract*— We introduce two novel joint radio-frequency (RF)-baseband designs for receivers in a MIMO system with N_t transmit antennas, N_r receive antennas, but only $L < N_r$ RF chains at the receiver. The joint design introduces an RF pre-processing matrix that processes the signals from the different antennas, and is followed by selection (if necessary), down-conversion, and further processing in the baseband. The schemes are similar to conventional antenna selection in that they use fewer RF chains than antenna elements, but achieve superior performance by exploiting the spatial correlation of the received signals. The first of our proposed designs uses an $L \times N_r$ RF pre-processing matrix that outputs only L streams followed by baseband signal processing, and, thus, eliminates the need for a selection switch. The second one uses an $N_r \times N_r$ RF pre-processing matrix that outputs N_r streams and is followed by a switch that selects L streams for baseband signal processing. Both spatial diversity and spatial multiplexing systems are considered and the optimum pre-processing matrices are derived for all cases. To accommodate practical RF design constraints, which prefer a variable phase-shifter-based implementation, a sub-optimal phase approximation is also introduced. Performance better than conventional antenna selection and close to the full complexity receiver is observed in both single cluster and multi-cluster wireless channels. A beam-pattern-based geometric intuition is also developed to illustrate the effectiveness of the optimal solutions.**

***Index Terms*— MIMO systems, Diversity methods, Spatial multiplexing, Antenna arrays, Antenna selection, Channel statistics, Signal to noise ratio, Information rates, Phase shifters.**

I. INTRODUCTION

MULTIPLE input multiple output (MIMO) antenna systems deliver substantially higher bit rates and reliability using either spatial multiplexing [1], [2], where different data streams are transmitted on each antenna, or link diversity [3]–[7], where the same data stream is transmitted on all the antennas. Despite the significant gains, an important factor limiting the widespread adoption of MIMO systems is their increased hardware and signal processing complexity. The signal received (transmitted) at each antenna element requires a separate RF chain that comprises a low noise amplifier (power amplifier), demodulator (modulator), and an A/D converter (D/A converter), which are expensive.

Antenna selection, which chooses a subset of available antennas for further processing and requires fewer RF chains, has received considerable attention in the research community [8]–[17]. While antenna selection achieves the full diversity gain [14], [18], it does lead to a reduced beamforming gain.[1] This paper shows that the loss can be reduced considerably in the presence of spatial correlation. While MIMO system performance with spatial correlation [20]–[23] and antenna selection have separately received considerable attention, exploiting spatial correlation in antenna selection has not, barring a few exceptions [24]–[26].

In this paper, we introduce two novel RF pre-processing designs at the receiver that exploit spatial correlation to recover most of the beamforming gain. The design consists of a linear pre-processing matrix in the RF domain, which depends only on the channel's large-scale parameters such as mean angle of arrival (AoA), angle spread, etc., followed by selection, down-conversion, and further processing in the baseband. It is important to understand that pre-processing succeeds only because antenna selection leads to loss in performance. In the absence of this lossy step, pre-processing cannot improve the performance of an optimal baseband receiver.

Preliminary results in [27], which used a fixed (unoptimized) channel-independent Butler FFT matrix, already show a large improvement. While this scheme yields gains, it is certainly not optimal and its performance depends on the AoA and the geometry of the antenna array. The solution in [28] requires adjusting the parameters in the RF domain based on the instantaneous channel state; this can lead to tighter design constraints and require instantaneous channel state feedback when implemented at the transmitter. In this paper, we use pre-processing based on channel statistics. Using channel statistics is appealing because it varies on a much longer time scale than short-term fading [45].

The contributions of the paper are the following. First, two novel RF pre-processing designs are introduced with the pre-processing matrix depending only on the large-scale statistics of the channel. The designs are similar to antenna selection in that they use fewer RF chains, but turn out to be superior as they exploit spatial correlation. The first design outputs only

Manuscript received May 12, 2004; revised August 17, 2005 and January 17, 2006; accepted 1/27, 2006. The associate editor coordinating the review of this paper and approving it for publication was Z. Tian. Part of this work was presented at VTC 2004 (Fall) and Globecom 2004.

Pallav Sudarshan is with Motorola, 600N, US Hwy. 45, Libertyville, IL, USA (e-mail: pallav.sudarshan@motorola.com).

N. B. Mehta, A. F. Molisch, and J. Zhang are with Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, USA (e-mail: {mehta, molisch, jzhang}@merl.com).

A. F. Molisch is also at the Dept. of Electroscience, Lund University, Sweden.

Digital Object Identifier 10.1109/TWC.2006.04289.

[1]Depending upon the number of antenna elements and the subset size, a loss in spatial multiplexing gain is possible. A loss in diversity gain occurs in the presence of imperfect selection [17] or fast fading [19].

L streams and thus does not require a subsequent selection switch. The second one outputs N_r streams and is followed by a switch that then selects L streams. Both designs are optimized for either spatial diversity or spatial multiplexing. These require different analysis techniques and lead to vastly different solutions. Finally, an algorithm to implement the scheme in the RF domain using only variable phase-shifters is presented, and incurs a negligible performance loss.

The optimum solutions, derived here, turn out to be intuitively pleasing and are shown to be equivalent to a principal component analysis (PCA) [29]. However, our results are still novel as the lossy selection step fundamentally changes the behavior and, consequently, the analysis of the system. And, the covariance matrix of the vector that PCA operates upon need not be the familiar spatial covariance matrix. We also present extensive results to test the solutions for both single multipath cluster scenarios and multi-cluster scenarios. While the latter scenario often occurs in typical deployments due to high-rise buildings and mountains and can lead to significantly different performance, it has received little attention in the antenna selection literature [30].

The implementation using only variable phase-shifters is a critical issue that ultimately determines the practical viability of the proposed design [31]–[34]. While equal gain combining (EGC) at the receiver and equal gain transmission both use only phase-shifters [35]–[37], their use is limited to spatial diversity systems. They are also sub-optimal compared to maximum ratio combining (MRC) or maximum ratio transmission (MRT). For the same number of RF chains, the designs we propose outperform both EGC and MRC, as these are essentially baseband processing techniques that always succeed the selection switch.

The paper is organized as follows. Section II sets up the the system model for spatial diversity and spatial multiplexing. Sections III and IV derive the optimal solutions for the spatial diversity and spatial multiplexing systems, respectively, with the phase-only approximation being considered in Sec. V. In Sec. VI, the performance of the proposed solutions and others in the literature is compared for both single-cluster and multi-cluster channel models. The conclusions follow in Sec. VII.

II. System Model for Spatial Diversity and Spatial Multiplexing

The following matrix notation is used: $(.)^T$ stands for transpose, $(.)^\dagger$ for the Hermitian transpose, $|.|$ for the determinant, $\text{Tr}\{.\}$ for the trace, and $\mathcal{E}_X[.]$ for the expectation with respect to the random variable (RV) X. The norm of a vector is denoted by $\|.\|$. The matrix $\mathbf{I}_n$ denotes the $n \times n$ identity matrix and $\mathbf{0}$ the zero matrix.

The equivalent complex baseband notation is used to develop the analysis. However, it must be kept in mind that many of the transformations occur in the RF domain. The matrix $\mathbf{H}$ of size $N_r \times N_t$ denotes the channel, the vector $\mathbf{n}$ denotes the zero-mean additive white Gaussian noise, and ρ is the total transmitted power. The elements of $\mathbf{H}$ and $\mathbf{n}$ are normalized to have unit variance; therefore, ρ also denotes the SNR at the input of a receive antenna. We adopt the widely used Kronecker model for spatial correlation [14], [25], [38], [39].

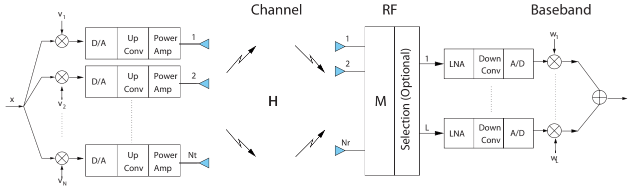


Fig. 1. Block Diagram for Diversity Transmission with RF Pre-processing.

The channel matrix is given by $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{H}_w\mathbf{T}^{\frac{1}{2}}$, where the elements of $\mathbf{H}_w$ are zero-mean unit-variance i. i. d. complex Gaussian RVs and the matrices $\mathbf{R}$ and $\mathbf{T}$ are the receive and transmit antenna covariances, respectively.

A. Spatial Diversity

The first system we consider is a closed-loop spatial diversity system in which replicas of one data stream are transmitted by all the antennas, as shown in Fig. 1. The transmit antenna weights are denoted by the vector $\mathbf{v} = [v_1, v_2, \ldots, v_{N_t}]$, where the i^{th} element, v_i, denotes the weight at antenna i. Perfect channel state information (CSI) is assumed at the transmitter and the receiver. The received vector $\mathbf{y}$ is given by

$$\mathbf{y} = \sqrt{\frac{\rho}{N_t}}\mathbf{H}\mathbf{v}x + \mathbf{n}, \tag{1}$$

where the scalar x is the transmitted symbol. We analyze the uncoded case in this paper. As in maximal ratio transmission (MRT), the transmitter is assumed to set $\mathbf{v} = \mathbf{v}_1$, where $\mathbf{v}_1$ is the right singular vector associated with the largest singular value of $\mathbf{H}$. While this choice is optimum in terms of maximizing the SNR when the receiver uses MRC, it is not when it uses antenna selection. However, our results show that even $\mathbf{v} = \mathbf{v}_1$ achieves nearly optimal performance. This also avoids the intractable analysis required for finding the jointly optimal $\mathbf{v}$ and RF pre-processing, and decouples the transmitter and receiver designs.

B. Spatial Multiplexing

The second system we consider is a spatial multiplexing system in which multiple data streams, as opposed to a single one in the diversity case, are transmitted simultaneously, as shown in Fig. 2. The transmitter has no CSI and the receiver has perfect CSI. When a vector, $\mathbf{x}$, of size $N_t \times 1$, is transmitted, the received vector, $\mathbf{y}$, of size $N_r \times 1$, is given by

$$\mathbf{y} = \sqrt{\frac{\rho}{N_t}}\mathbf{H}\mathbf{x} + \mathbf{n}. \tag{2}$$

III. Spatial Diversity: Optimal Time-Invariant (TI) Pre-processing

We now investigate RF-baseband design based on channel statistics. Given that we only use channel statistics to design the pre-processing matrices, they shall be referred to as time-invariant (TI) solutions. We first consider the case in which the RF pre-processing matrix (of size $L \times N_r$) takes N_r streams as input and outputs only L streams and, thus, eliminates the need for a subsequent selection switch. We then consider a receiver

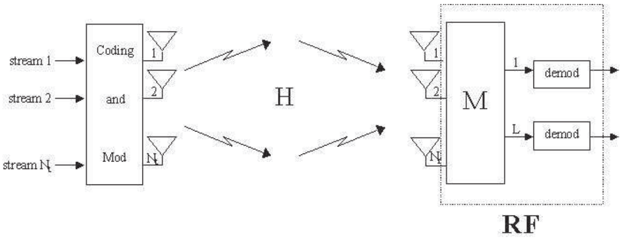


Fig. 2. Block Diagram for Spatial Multiplexing with RF Pre-processing.

design in which the pre-processing matrix (of size $N_r \times N_r$) outputs N_r streams and is followed by an instantaneous CSI-based selection switch. In both cases, we maximize the average output SNR after pre-processing (and selection, if necessary) as we consider an uncoded system. For both, we first show that semi-unitary or unitary matrices are optimal and then we find them. As we shall see, in the latter case, the presence of the selection switch makes the optimization intractable. For this case, we resort to a tractable lower bound on the SNR. We also derive the diversity order of the designs.

A. Optimal Time-Invariant (TI) Pre-Processing Without Selection Switch

Let $\mathbf{M}_L$ denote an arbitrary $L \times N_r$ pre-processing matrix. The vector $\tilde{\mathbf{y}}$ at the output of $\mathbf{M}_L$ is

$$\tilde{\mathbf{y}} = \sqrt{\frac{\rho}{N_t}}\mathbf{M}_L\mathbf{H}\mathbf{v}x + \mathbf{M}_L\mathbf{n},$$

where $\mathbf{v}$ is the transmit antenna weight vector. After pre-processing and down-conversion, the signals are combined in the baseband by the vector $\mathbf{w}^\dagger$. The average SNR, γ_{TI}, after pre-processing and combining is given by

$$\gamma_{\text{TI}} = \frac{\rho}{N_t}\mathcal{E}_{\mathbf{H}}\left[\mathbf{w}^\dagger\mathbf{W}\mathbf{M}_L\mathbf{H}\mathbf{v}\right]^2, \quad \text{s. t. } \|\mathbf{w}\| = 1, \qquad (3)$$

where the matrix $\mathbf{W}$ is the baseband noise-whitening filter such that $\mathbf{W}\mathbf{M}_L\mathbf{M}_L^\dagger\mathbf{W}^\dagger = \mathbf{I}_L$.

We now show that the optimal pre-processing matrix for maximizing the average SNR can be written in terms of a semi-unitary matrix. We then derive its form.

Theorem *1:* For a general $N_r \times N_t$ channel, $\mathbf{H}$, with singular value decomposition (SVD) $\mathbf{U\Lambda V}^\dagger$, let $\mathbf{u}_1$ denote the first column of $\mathbf{U}$ that corresponds to the largest singular value, λ_1. To maximize the average output SNR, the optimal pre-processing matrix, $\mathbf{M}_{\text{TI}}$, of size $L \times N_r$, takes the form

$$\mathbf{M}_{\text{TI}} = \mathbf{B}_L\mathbf{Z}_{\text{opt}}. \qquad (4)$$

Here, $\mathbf{B}_L$ is *any* $L \times L$ full rank matrix and $\mathbf{Z}_{\text{opt}}$ is an $L \times N_r$ semi-unitary matrix given by

$$\mathbf{Z}_{\text{opt}} = \left[\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_L\right]^\dagger,$$

where $\boldsymbol{\mu}_l$ is the singular vector corresponding to the l^{th} largest singular value of the covariance matrix $\mathbf{E}_{u_1u_1} = \mathcal{E}_{\mathbf{H}}\left[\lambda_1^2\mathbf{u}_1\mathbf{u}_1^\dagger\right]$.

Proof: The proof is given in Appendix A-1. ■

The following corollary that characterizes the optimal combining vector, $\mathbf{w}_{\text{TI}}$, trivially follows:

Corollary *1:* The optimal combining vector in the receiver for maximizing the average SNR is given by

$$\mathbf{w}_{\text{TI}} = \frac{\mathbf{Z}_{\text{opt}}\mathbf{u}_1}{\|\mathbf{Z}_{\text{opt}}\mathbf{u}_1\|}.$$

We see from Theorem 1 that $\mathbf{M}_{\text{TI}}$ is obtained by performing a PCA on $\lambda_1\mathbf{u}_1$. It is interesting to note that the covariance $\mathbf{E}_{u_1u_1}$, defined above, is not the traditional receive correlation matrix, $\mathbf{R}$, but instead quantifies the statistics of the largest eigenvector. Moreover, $\mathbf{E}_{u_1u_1}$ depends not only on the vector $\mathbf{u}_1$ but also on the eigenvalue λ_1. A statistical interpretation of the optimal matrix is that it consists of the first L eigenvectors of $\mathbf{E}_{u_1u_1}$, which contribute the most to the variance of $\lambda_1\mathbf{u}_1$. The result can also be viewed as a rigorous generalization of a result in [40], which dealt with single input single output channels.

The structure and behavior of the optimal pre-processing matrix can be intuitively understood by its beam-pattern, which is shown in Fig. 3 for $N_r = 4$ and a wavelength-relative antenna element spacing of $1/2$. The beam-patterns of $\mathbf{M}_{\text{TI}}$ for two different mean AoAs (45° and 60°) are plotted. Clearly, the beam-pattern of $\mathbf{M}_{\text{TI}}$ adapts to the mean AoA. It also adapts to the angle spread (figures not shown here). This is unlike FFT pre-processing, which has a fixed beam-pattern, and gives promising gains only for mean AoAs of 0°, 60°, 90°, 270°, and 300°.

We now show that the diversity order of any rank L RF pre-processing matrix solution, including the optimal $\mathbf{M}_{\text{TI}}$, is LN_t.

Lemma *1:* The diversity order of the TI design in a correlated channel described by the Kronecker model, in which the transmit and receive correlation matrices are full-ranked, is LN_t.

Proof: The proof is given in Appendix A-2. ■

It must be noted that the above result on diversity order measures the slope of the BER curve (uncoded) as the SNR tends to ∞. For finite SNRs, the beamforming gain is also important, as we shall see later.

B. Time-Invariant Pre-Processing With Selection (TI-S)

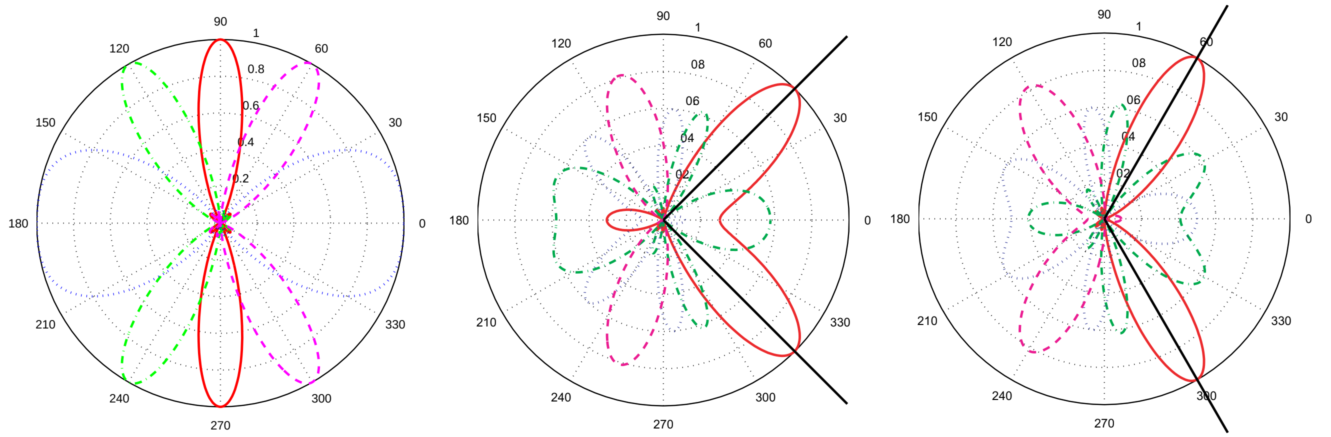
We now consider the case in which the pre-processing matrix is of size $N_r \times N_r$ and outputs the same number of streams as it receives. A subsequent switch selects L out of the N_r outputs for down-conversion. The modified received vector, $\check{\mathbf{y}}$, after pre-processing and selection is

$$\check{\mathbf{y}} = \sqrt{\frac{\rho}{N_t}}\mathbf{S}\mathbf{M}_{N_r}\mathbf{H}\mathbf{v}x + \mathbf{S}\mathbf{M}_{N_r}\mathbf{n},$$

where $\mathbf{S}$ is an $L \times N_r$ selection matrix that selects L out of N_r signals. The selection matrix $\mathbf{S}$ adapts to the instantaneous channel state. $\check{\mathbf{y}}$ is down-converted, whitened, and combined by the vector $\mathbf{w}^\dagger$ in the baseband. The average output SNR after noise-whitening by the filter $\mathbf{W}$ and combining by the receive weights $\mathbf{w}^\dagger$ is given by

$$\gamma_{\text{TI-S}} = \mathcal{E}_{\mathbf{H}}\left[\frac{\rho}{N_t}\left(\mathbf{w}^\dagger\mathbf{W}\mathbf{S}\mathbf{M}_{N_r}\mathbf{H}\mathbf{v}\right)^2\right], \text{ s. t. } \|\mathbf{w}\| = 1, \quad (5)$$

where $\mathbf{W}$, being the noise-whitening filter, satisfies the constraint $\mathbf{W}\mathbf{S}\mathbf{M}_{N_r}(\mathbf{W}\mathbf{S}\mathbf{M}_{N_r})^\dagger = \mathbf{I}_L$. The maximum SNR is



(a) Fixed FFT beam-pattern (b) $\mathbf{M}_{\text{TI-S}}$ (mean AoA $= 45°$). (c) $\mathbf{M}_{\text{TI-S}}$ (mean AoA $= 60°$).

Fig. 3. Beam-pattern as a function of azimuth angle for time-invariant pre-processing and FFT pre-processing.

then

$$\gamma_{\text{TI-S}}^{\max} = \max_{\mathbf{M}_{N_r}} \mathcal{E}_{\mathbf{H}} \left[\max_{\mathbf{S},\mathbf{w},\mathbf{W}} \frac{\rho}{N_t} \left(\mathbf{w}^\dagger \mathbf{W}\mathbf{S}\mathbf{M}_{N_r}\mathbf{H}\mathbf{v} \right)^2 \right], \quad \text{s. t. } \|\mathbf{w}\| = 1. \tag{6}$$

We now show that it is sufficient to restrict the search to the set of all unitary matrices.

Lemma *2:* The combination $\mathbf{W}\mathbf{S}\mathbf{M}_{N_r}$ of RF pre-processing, $\mathbf{M}_{N_r}$ followed by selection, $\mathbf{S}$, and noise-whitening, $\mathbf{W}$, can be written as

$$\mathbf{W}\mathbf{S}\mathbf{M}_{N_r} = \mathbf{S}^{'}\mathbf{U}_{N_r}, \tag{7}$$

where $\mathbf{S}^{'}$ is a selection matrix that selects L out of N_r rows and $\mathbf{U}_{N_r}$ is a unitary matrix.

Proof: Given that $\mathbf{W}$ is a noise-whitening filter, it satisfies the $\mathbf{W}\mathbf{S}\mathbf{M}_{N_r}(\mathbf{W}\mathbf{S}\mathbf{M}_{N_r})^\dagger = \mathbf{I}_L$. Therefore, $\mathbf{W}\mathbf{S}\mathbf{M}_{N_r}$ is a semi-unitary matrix. Every $L \times N_r$ semi-unitary matrix is a sub-matrix of another $N_r \times N_r$ unitary matrix, say $\mathbf{U}_{N_r}$. Therefore, there exists a selection matrix, $\mathbf{S}^{'}$, such that $\mathbf{W}\mathbf{S}\mathbf{M}_{N_r} = \mathbf{S}^{'}\mathbf{U}_{N_r}$. ■

We can now restrict our search space to unitary matrices. It is difficult to analytically find the $\mathbf{M}_{N_r}$ that maximizes the SNR because of the presence of $\mathbf{S}$, which depends on the instantaneous channel realization $\mathbf{H}$. We therefore derive an analytically tractable lower bound and show that the solution proposed below achieves it. A successive refinement-based justification is also presented. The simulation results show that this approach works well. The lower bound is obtained by interchanging the order of $\mathcal{E}_{\mathbf{H}}$ and $\max$ as follows:

$$\gamma_{\text{TI-S}}^{\max} \geq \max_{\mathbf{S}} \max_{\mathbf{M}_{N_r}} \mathcal{E}_{\mathbf{H}} \left[\max_{\mathbf{w},\mathbf{W}} \frac{\rho}{N_t} \left(\mathbf{w}^\dagger \mathbf{W}\mathbf{S}\mathbf{M}_{N_r}\mathbf{H}\mathbf{v} \right)^2 \right]. \tag{8}$$

The following theorem derives conditions that partially determine the optimal unitary matrix that achieves this lower bound.

Theorem *2:* Let $\mathcal{U}_L = \text{span}\{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_L\}$, denote the span of $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_L$, and let $\mathcal{N}(\mathcal{U}_L)$ denote its null space. Then, the $N_r \times N_r$ matrix that achieves the lower bound in (8) must be of the form

$$\mathbf{M}_{\text{TI-S}} = \mathbf{P} \left[\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_L, \mathbf{z}_1, \ldots \mathbf{z}_{N_r-L} \right]^\dagger, \tag{9}$$

where $\mathbf{P}$ is any $N_r \times N_r$ permutation matrix and $\mathbf{z}_1, \ldots, \mathbf{z}_L$ are orthonormal vectors in $\mathcal{N}(\mathcal{U}_L)$.

Proof: From Lemma 2, we can restrict $\mathbf{M}_{\text{TI-S}}$ to be unitary. Then, for a given selection matrix $\mathbf{S}_0$, the problem in (8) is similar to that in Section III-A (with $\mathbf{B}_L = \mathbf{I}_L$). To maximize (8), the rows of $\mathbf{M}_{N_r}$ that are selected by $\mathbf{S}_0$ need to be the eigenvectors corresponding to the L largest eigenvalues of $\mathbf{E}_{u_1 u_1}$. For example, if $\mathbf{S}_0 = [\mathbf{I}_L \quad \mathbf{0}]$, then $\mathbf{M}_{\text{TI-S}} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_L, \mathbf{z}_1, \ldots, \mathbf{z}_{N_r-L}]^\dagger$. Given that $\mathbf{M}_{\text{TI-S}}$ is unitary, $\mathbf{z}_1, \ldots, \mathbf{z}_{N_r-L}$ are orthonormal vectors in $\mathcal{N}(\mathcal{U}_L)$. Any other selection matrix is a permutation of $\mathbf{S}_0$, therefore, so is $\mathbf{M}_{\text{TI-S}}$. ■

To characterize the matrix $\mathbf{M}_{\text{TI-S}}$ that maximizes average SNR, we now need to identify $\mathbf{z}_1, \ldots, \mathbf{z}_{N_r-L}$. For this we propose the following algorithm that successively improves the SNR. We conjecture that the solution that we obtain is optimal.

1) Successive Improvement Approach to Determine $\mathbf{M}_{\textit{TI-S}}$*:* The outline of the algorithm is as follows: We divide the space of all selection matrices into $N_r - L$ non-intersecting subspaces. We then find $\mathbf{z}_k$, such that for all selection matrices belonging to the k^{th} subset, the average output SNR, $\gamma_{\text{TI-S}}$, is improved over the SNR for the previously considered $k-1$ subspaces. This is successively done over all the $N_r - L$ subsets to obtain $\mathbf{z}_1, \ldots, \mathbf{z}_{N_r-L}$.

Let $\mathcal{S}$ be the set of all possible $L \times N_r$ selection matrices that select L out of the N_r rows of $\mathbf{M}_{\text{TI-S}}\mathbf{H}$. Then, $\mathcal{S}$ contains $\binom{N_r}{L}$ elements. For $p = 0, \ldots, N_r - L$, define $\mathcal{S}_p$ as the set of all permutation matrices that select the $(L+p)^{\text{th}}$ row and any $(L-1)$ from the first $(L+p-1)$ rows of $\mathbf{M}_{N_r}\mathbf{H}$. Clearly, $\mathcal{S}_p$ is a subset of $\mathcal{S}$. Furthermore, $\mathcal{S}_0, \ldots, \mathcal{S}_{N_r-L}$ cover $\mathcal{S}$, *i.e.*, $\mathcal{S} = \mathcal{S}_0 \bigcup \cdots \bigcup \mathcal{S}_{N_r-L}$, where $\bigcup$ denotes the union of sets.

$\mathcal{S}_0$ contains only one element $\mathbf{S}_0 = [\mathbf{I}_L \quad \mathbf{0}]$, which selects the first L rows of $\mathbf{M}_{N_r}\mathbf{H}$. From Theorem 2, we get that if $\mathbf{S} = \mathbf{S}_0$, then $\mathbf{M}_{N_r}^{(0)}$ (the optimal pre-processing matrix when $\mathbf{S}$ is restricted to $\mathcal{S}_0$) is

$$\mathbf{M}_{N_r}^{(0)} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_L, \mathbf{z}_1, \ldots, \mathbf{z}_{N_r-L}]^\dagger. \tag{10}$$

Irrespective of the choice of $\mathbf{z}_k$'s, setting the first L rows of $\mathbf{M}_{\text{TI-S}}$ to be $[\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_L]^\dagger$ ensures optimal performance if the selection matrix is in $\mathbf{S}_0$. Now, if the selection matrix is $\mathbf{S}_1 \in$

$\mathcal{S}_1$, then Theorem 1 and the constraint $\mathbf{z}_k \in \mathcal{N}(\mathcal{U}_L)$ imply that the average SNR is maximized when $\mathbf{z}_1 = \boldsymbol{\mu}_{L+1}$. Therefore, $\mathbf{M}_{N_r}^{(1)} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{L+1}, \mathbf{z}_2, \ldots, \mathbf{z}_{N_r-L}]^\dagger$. Fixing the first L elements to be $\boldsymbol{\mu}_i$'s ensures that subsequent selections for $\mathbf{z}_i$'s do not affect the performance of $\mathbf{S}_0$. Following a similar procedure, we get $\mathbf{z}_2 = \boldsymbol{\mu}_{L+2}, \ldots, \mathbf{z}_{N_r-L} = \boldsymbol{\mu}_{N_r}$. Therefore,

$$\mathbf{M}_{\text{TI-S}} = \left[\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{N_r}\right]^\dagger. \tag{11}$$

In effect, the successive refinement approach ensures that the above solution achieves the following lower bound, which is tighter than that in (8):

$$\gamma_{\text{TI-S}}^{\max} \geq \max\left[\max_{\mathbf{S}\in\mathcal{S}_0, \mathbf{M}_{N_r}} [\gamma_{\text{TI-S}}], \max_{\mathbf{S}\in\mathcal{S}_1,\ \mathbf{M}_{N_r}\ \text{given}\ \mathbf{M}_{N_r}^{(0)}} [\gamma_{\text{TI-S}}], \ldots\right].$$

It can also be seen that the diversity order of TI-S is $N_t N_r$. This can be reasoned as follows. The signal input to the selection switch passes through an equivalent channel $\mathbf{M}_{\text{TI-S}}\mathbf{H}$, of size $N_r \times N_t$. Moreover, $\mathbf{M}_{\text{TI-S}}$, $\mathbf{T}$ and $\mathbf{R}$ have the highest possible ranks N_r, N_t, and N_r, respectively. Using results for instantaneous channel-based selection [14], [18] or by an approach similar to Lemma 1, it can be seen that $\mathbf{M}_{\text{TI-S}}$ achieves full diversity.[2]

IV. Spatial Multiplexing: Optimal Time-Invariant Pre-Processing

For a spatial multiplexing system, in which multiple streams are transmitted simultaneously, it is the information rate that is the performance metric. As in the diversity case, we consider the cases in which the pre-processor either outputs only L streams or outputs N_r streams and is followed by a selection switch that reduces the number of streams from N_r to L. Given the very different nature of the spatial diversity and spatial multiplexing systems, the analysis turns out to be very different. However, we can again show that it is sufficient to restrict the search space to semi-unitary or unitary matrices, and then we derive their optimal forms.

A. Optimal Time-Invariant (TI) Pre-Processing Without Selection Switch

Let $\mathbf{M}_L$ be an $L \times N_r$ pre-processing matrix as shown in Fig. 2. The modified received vector, $\tilde{\mathbf{y}}$, after RF pre-processing, becomes

$$\tilde{\mathbf{y}} = \sqrt{\frac{\rho}{N_t}}\mathbf{M}_L\mathbf{H}\mathbf{x} + \mathbf{M}_L\mathbf{n}.$$

The transmitter has no CSI and allocates the same power to all the transmit antennas. It uses i. i. d. Gaussian signaling.[3] The receiver has perfect CSI. The information rate, C_{TI}, for such a system is given by

$$C_{\text{TI}} = \mathcal{E}_{\mathbf{H}}\left[\log_2\left|\mathbf{I}_{N_t} + \frac{\rho}{N_t}\mathbf{H}^\dagger\mathbf{M}_L^\dagger(\mathbf{M}_L\mathbf{M}_L^\dagger)^{-1}\mathbf{M}_L\mathbf{H}\right|\right]. \tag{12}$$

The following two Lemmas shall come in handy.

Lemma *3:* Let $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ be diagonal matrices with non-negative elements such that each element of $\boldsymbol{\Lambda}_1$ is less than or equal to the corresponding element of $\boldsymbol{\Lambda}_2$. Then for any arbitrary complex matrix $\mathbf{A}$, we have

$$\left|\mathbf{I}_{N_t} + \mathbf{A}\boldsymbol{\Lambda}_2\mathbf{A}^\dagger\right| \geq \left|\mathbf{I}_{N_t} + \mathbf{A}\boldsymbol{\Lambda}_1\mathbf{A}^\dagger\right|.$$

Proof: The proof is given in Appendix A-3. ■

Lemma *4:* Let $\mathbf{Q}$ be an $L \times N_r$ semi-unitary matrix and $\mathbf{G}$ be a full rank matrix. For a matrix $\mathbf{A}$, let $e_i(\mathbf{A})$ denote the i^{th} largest eigenvalue of $\mathbf{A}\mathbf{A}^\dagger$. Then,

$$e_{N_r-L+i}(\mathbf{G}) \leq e_i(\mathbf{Q}\mathbf{G}) \leq e_i(\mathbf{G}). \tag{13}$$

Furthermore, $e_i(\mathbf{Q}\mathbf{G}) = e_i(\mathbf{G})$ if $\mathbf{Q}$ is the conjugate transpose of the L eigenvectors of $\mathbf{G}\mathbf{G}^\dagger$ that correspond to its L largest eigenvalues.

Proof: The proof is given in Appendix A-4. ■

Using the above two Lemmas, we now show that it is sufficient to only search over the space of semi-unitary matrices to find the optimal pre-processing matrix, $\mathbf{M}_{\text{TI}}$, that maximizes the average information rate and we then find it.

Theorem *3:* To maximize the average information rate, over a channel matrix of the form $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{H}_w\mathbf{T}^{\frac{1}{2}}$, the optimal $L \times N_r$ time-invariant pre-processing matrix, $\mathbf{M}_{\text{TI}}$, is of the form

$$\mathbf{M}_{\text{TI}} = \mathbf{B}_L\mathbf{Q}_{\text{opt}}, \tag{14}$$

where $\mathbf{B}_L$ is *any* $L \times L$ full rank matrix, and $\mathbf{Q}_{\text{opt}}$ is given by

$$\mathbf{Q}_{\text{opt}} = [\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_L]^\dagger,$$

and $\mathbf{r}_l$ is the singular vector of $\mathbf{R}$ corresponding to its l^{th} largest eigenvalue. The maximum information rate is then given by

$$C_{\text{TI}} = \mathcal{E}_{\mathbf{H}}\left[\log_2\left|\mathbf{I}_{N_t} + \frac{\rho}{N_t}\mathbf{H}^\dagger\mathbf{Q}_{\text{opt}}^\dagger\mathbf{Q}_{\text{opt}}\mathbf{H}\right|\right]. \tag{15}$$

Proof: Any $L \times N_r$ matrix with rank L can be written as $\mathbf{M}_L = \mathbf{B}_L\mathbf{Q}$, where $\mathbf{B}_L$ is any $L \times L$ matrix and $\mathbf{Q}$ is an $L \times N_r$ semi-unitary matrix.[4] Since $\mathbf{Q}\mathbf{Q}^\dagger = \mathbf{I}_L$, the maximum information rate expression in (12) can be written in terms of $\mathbf{Q}$ as

$$C_{\text{TI}} = \max_{\mathbf{Q}}\mathcal{E}_{\mathbf{H}}\left[\log_2\left|\mathbf{I}_{N_t} + \frac{\rho}{N_t}\mathbf{H}^\dagger\mathbf{Q}^\dagger\mathbf{Q}\mathbf{H}\right|\right]. \tag{16}$$

Note that $\mathbf{B}_L$ is absent in the above equation. Recall that $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{H}_w\mathbf{T}^{\frac{1}{2}}$. Let the product term $\mathbf{Q}\mathbf{R}^{\frac{1}{2}}$ be denoted by the matrix $\mathbf{X}_L$, and let its SVD be denoted by $\mathbf{X}_L = \mathbf{U}_X\boldsymbol{\Sigma}_X\mathbf{V}_X^\dagger$.

[2]Note that the result can change if space-time coding is also used. With space-time trellis codes, the diversity order depends on the underlying fading channel model while the behavior of orthogonal space-time block codes is different [41], [42].

[3]We do not use the Gaussian signaling derived in [43] for antenna selection as the power allocation parameter α defined in it needs to be determined numerically.

[4]A matrix $\mathbf{M}_L$ with rank less than L is not considered as it is sub-optimal.

The above information rate expression simplifies as follows:

$$\begin{aligned} C_{\text{TI}} &= \max_{\mathbf{X}_L} \mathcal{E}_{\mathbf{H}_w} \left[\log_2 \left| \mathbf{I}_{N_t} + \frac{\rho}{N_t} \mathbf{T}^{\frac{1}{2}} \mathbf{H}_w^\dagger \mathbf{X}_L^\dagger \mathbf{X}_L \mathbf{H}_w \mathbf{T}^{\frac{1}{2}} \right| \right], \\ &= \max_{\mathbf{U}_X, \mathbf{\Sigma}_X} \mathcal{E}_{\mathbf{H}_w} \left[\log_2 \left| \mathbf{I}_{N_t} + \frac{\rho}{N_t} \mathbf{T}^{\frac{1}{2}} \mathbf{H}_w^\dagger \mathbf{U}_X \mathbf{\Sigma}_X^2 \mathbf{U}_X^\dagger \mathbf{H}_w \mathbf{T}^{\frac{1}{2}} \right| \right], \\ &= \max_{\mathbf{\Sigma}_X} \mathcal{E}_{\mathbf{H}_w} \left[\log_2 \left| \mathbf{I}_{N_t} + \frac{\rho}{N_t} \mathbf{T}^{\frac{1}{2}} \mathbf{H}_w^\dagger \mathbf{\Sigma}_X^2 \mathbf{H}_w \mathbf{T}^{\frac{1}{2}} \right| \right], \end{aligned} \tag{17}$$

where (17) holds because the statistical properties of $\mathbf{H}_w$ are not changed by multiplication with a unitary matrix. From Lemma 3, it follows that (17) is maximized when each element of the non-negative diagonal matrix $\mathbf{\Sigma}_X$ is maximized. Since the elements of $\mathbf{\Sigma}_X^2$ correspond to the eigenvalues of $\mathbf{QRQ}^\dagger$, Lemma 4 implies that (17) is maximized when the eigenvalues of $\mathbf{QRQ}^\dagger$ reach the maximum value of $e_i(\mathbf{R})$. Thus the optimal $\mathbf{Q}$ is then the conjugate transpose of the L eigenvectors corresponding to the L largest eigenvalues of $\mathbf{R}$. ■

B. Time-Invariant Pre-Processing With Selection (TI-S)

We now consider the case in which the RF pre-processing matrix outputs N_r streams and is followed by a selection switch. Our aim is to derive the $N_r \times N_r$ time-invariant pre-processing matrix that maximizes the information rate after pre-processing and selection. The maximization problem then becomes

$$C_{\text{TI-S}} = \max_{\mathbf{M}_{N_r}} \left(\mathcal{E}_{\mathbf{H}} \left[\max_{\mathbf{S}} \log_2 \left| \mathbf{I}_{N_t} + \frac{\rho}{N_t} \mathbf{H}^\dagger \mathbf{M}_{N_r}^\dagger \mathbf{S}^\dagger (\mathbf{S} \mathbf{M}_{N_r} \mathbf{M}_{N_r}^\dagger \mathbf{S}^\dagger)^{-1} \mathbf{S} \mathbf{M}_{N_r} \mathbf{H} \right| \right] \right).$$

The above problem is difficult to solve because of the presence of $\mathbf{S}$, which depends on the instantaneous channel realization. However, as before, we can again show that it is sufficient to restrict the search space to unitary matrices. We then propose a solution that, at the very least, achieves the following lower bound, which is obtained by interchanging the order of expectation and max as follows:

$$C_{\text{TI-S}} \geq \max_{\mathbf{S}} \max_{\mathbf{M}_{N_r}} \left(\mathcal{E}_{\mathbf{H}} \left[\log_2 \left| \mathbf{I}_{N_t} + \frac{\rho}{N_t} \mathbf{H}^\dagger \mathbf{M}_{N_r}^\dagger \mathbf{S}^\dagger (\mathbf{S} \mathbf{M}_{N_r} \mathbf{M}_{N_r}^\dagger \mathbf{S}^\dagger)^{-1} \mathbf{S} \mathbf{M}_{N_r} \mathbf{H} \right| \right] \right).$$

Following the successive refinement argument similar to that in Sec. III-B, we can show that $\mathbf{M}_{\text{TI-S}}$ can be expressed in terms of L eigenvectors of $\mathbf{R}$ corresponding to its L largest eigenvalues, with the remaining $N_r - L$ vectors being orthonormal to them. We get

$$\mathbf{M}_{\text{TI-S}} = [\mathbf{r}_1, \ldots, \mathbf{r}_{N_r}]^\dagger. \tag{18}$$

In general, $M_{\text{TI-S}}$ can be written as $M_{\text{TI-S}} = \mathbf{DP}[\mathbf{r}_1, \ldots, \mathbf{r}_{N_r}]^\dagger$, where $\mathbf{D}$ is any $N_r \times N_r$ diagonal matrix and $\mathbf{P}$ is any $N_r \times N_r$ permutation matrix. We conjecture that this is the unique characterization of the optimal solution.

V. Implementation Amenable for Phase-Shifter-Based Designs

The $\mathbf{M}_{\text{TI}}$ and $\mathbf{M}_{\text{TI-S}}$ matrices derived so far for spatial diversity and spatial multiplexing systems consist of complex elements with arbitrary amplitudes. We now look for solutions in which $\mathbf{M}_{\text{TI}}$ and $\mathbf{M}_{\text{TI-S}}$ consist only of phase-shift elements, which, as discussed before, are more practical. We shall also allow for the use of on/off switches, which control the phase-shifter output. It is known from studies in other contexts, such as [44], that multiple local maxima can exist, making an analytical determination of optimal phase-shifters difficult. This motivates the use of the sub-optimal phase approximation scheme that we consider below.

We first consider $\mathbf{M}_{\text{TI}}$, with the approach for $\mathbf{M}_{\text{TI-S}}$ being very similar. Under the constraint of phase-only pre-processing, we propose using the phase matrix $\mathbf{\Phi}_{\text{TI}}$ to replace $\mathbf{M}_{\text{TI}}$, where $\mathbf{\Phi}_{\text{TI}} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{N_r}]$ such that for each i, the column $\boldsymbol{\phi}_i$ is closest in angle to $\mathbf{m}_i$, the i^{th} column of $\mathbf{M}_{\text{TI}}$. Thus, the received signals still add coherently to the extent possible. Therefore, each element ϕ_{ji} is given, in closed-form, by

$$\phi_{ji} = a_{ji} e^{j \arg(m_{ji})}, \tag{19}$$

where the switch a_{ji} is 0 or 1 and m_{ji} denotes the $(j, i)^{\text{th}}$ element of $\mathbf{M}_{\text{TI}}$. To maximize q_i, which is the cosine of the angle between $\boldsymbol{\phi}_i$ and $\mathbf{m}_i$, we need to determine a_{ji}, $1 \leq j \leq L$, for each column i. In general, this requires $O(2^L)$ computations. The following sub-optimal algorithm requires only $O(L \log(L))$ computations:

- Sort the entries of $\mathbf{m}_i$ in the descending order of absolute values to get $\{m_{[1]i}, \ldots, m_{[L]i}\}$, where $|m_{[1]i}| \geq |m_{[2]i}| \geq \cdots \geq |m_{[L]i}|$. Let $[k]$ denote the index of the k^{th} largest entry.
- Define $q_{il} = \frac{\boldsymbol{\phi}_i^\dagger \boldsymbol{m}_i}{\|\boldsymbol{\phi}_i\|}$, such that $\boldsymbol{\phi}_i$ has exactly l non-zero entries at the positions $j = [1], \ldots, [l]$. Thus, $q_{il} = \frac{\sum_{k=1}^{l} |m_{[k]i}|}{\sqrt{l}}$, $1 \leq l \leq L$. Select $l_{\max}(i)$ such that $q_{il_{\max}(i)}$ is maximum.[5]
- Then, $a_{ji} = \begin{cases} 1, & \text{if } j = [1], \ldots, [l_{\max}(i)] \\ 0, & \text{otherwise} \end{cases}$.

The proposed algorithm outperforms the one in [28].

It must be noted that in practice, using RF elements introduces several non-idealities such as insertion loss, phase quantization and errors, calibration errors, etc., which can have a detrimental effect on the overall system performance. Their impact on RF pre-processing was investigated in [34], which showed that it is very robust to phase non-idealities. For high insertion losses, low noise amplifiers may need to be placed before the selection switch, which can increase the overall cost of the system.

VI. Simulation Results

A thorough performance evaluation and comparison of the proposed receiver designs with those in the literature is performed in this section. For the spatial diversity and spatial multiplexing systems, the time-invariant solution without

[5]It might appear that a threshold K exists such that $q_{i1} \leq q_{i2} \leq \cdots \leq q_{i(K-1)} \leq q_{iK} \geq q_{i(K+1)}$ is a sufficient condition for q_{iK} to be maximum. However, counter examples exist.

TABLE I
AVERAGE OUTPUT SNR (IN DB) COMPARISON

	FC	TI-S	TI	FFT-Sel.	Ant. Sel.
$\theta_r = 45°, \sigma_r = 6°$	15.8	15.8	15.8	13.6	10.8
$\theta_r = 60°, \sigma_r = 6°$	15.8	15.8	15.8	15.8	10.8
$\theta_r = 60°, \sigma_r = 15°$	14.8	14.2	14.1	14.1	11.4

selection (TI) and the time-invariant solution with selection (TI-S) are compared with conventional antenna selection (Ant-Sel), FFT-selection (FFT-Sel), and a full complexity (FC) receiver (which has N_r RF chains at its disposal). The performance of the phase-only approximation solution, TI-Ph, is also investigated. Another reference case is instantaneous time-variant (TV) processing, in which the matrix is adjusted perfectly to the instantaneous channel state [28].

We evaluate both the single-cluster and multi-cluster channel models. For the multi-cluster scenario, we shall assume that the multipaths from the multiple clusters arrive at the same time, *i.e.*, there is no time resolvability. The time-dispersive case is beyond the scope of this paper. We consider an equi-powered two-cluster channel, where each cluster follows the Kronecker channel model: $\mathbf{H}_1 = \mathbf{R}_1^{\frac{1}{2}}\mathbf{H}_{w_1}\mathbf{T}^{\frac{1}{2}}$ and $\mathbf{H}_2 = \mathbf{R}_2^{\frac{1}{2}}\mathbf{H}_{w_2}\mathbf{T}^{\frac{1}{2}}$. It can be shown that the effective channel, $\mathbf{H}_{\text{eq}}$, also follows the Kronecker model, albeit with a different covariance matrix. The effective channel is $\mathbf{H}_{\text{eq}} = \mathbf{R}_{\text{eq}}^{\frac{1}{2}}\mathbf{H}_w\mathbf{T}^{\frac{1}{2}}$, where $\mathbf{R}_{\text{eq}} = \mathbf{R}_1 + \mathbf{R}_2$. While the Kronecker model holds, the spatial correlation is quite different.

The results are evaluated for uniform linear arrays with isotropic transmit and receive antenna elements. Given that transmit correlation does not affect the optimum RF pre-processing at the receiver, we set $\mathbf{T} = \mathbf{I}_{N_t}$. The receive antenna elements are spaced half a wavelength apart. All receivers, except FC, use $L < N_r$ RF chains. The angular dispersion at the receiver is Gaussian distributed with mean θ_r and standard deviation (angle spread) σ_r. The broadside of the antenna array corresponds to $\theta_r = 0°$.

To study the performance comprehensively, we plot the average BER vs. SNR curves for spatial diversity and the average information rate vs. SNR curves for spatial multiplexing. In addition, we also plot the cumulative distribution function (CDF) of the SNR (for spatial diversity) and information rate (for spatial multiplexing). The CDF is of interest because it provides a complete characterization of the probability distribution, as opposed to looking at only the average. For example, the outage probability can be easily inferred from the CDF [45, Chp. 5].

A. Single Cluster Scenario: Spatial Diversity

For the spatial diversity system, the average bit error rate (BER) for BPSK as a function of the input SNR, ρ, is shown in Fig. 4 for mean $\theta_r = 45°$ and an angular spread of $\sigma_r = 15°$. The number of antenna elements is set as $N_t = N_r = 4$ with $L = 1$. In Fig. 4, we observe that, at a BER of 10^{-3} with one RF chain, TI yields up to a 1.5 dB gain over conventional antenna selection. Recall that TI does not require a selection switch. TI-S is within 1 dB of FC. As expected, the slope of the TI curve is smaller than that of TI-S and antenna selection.

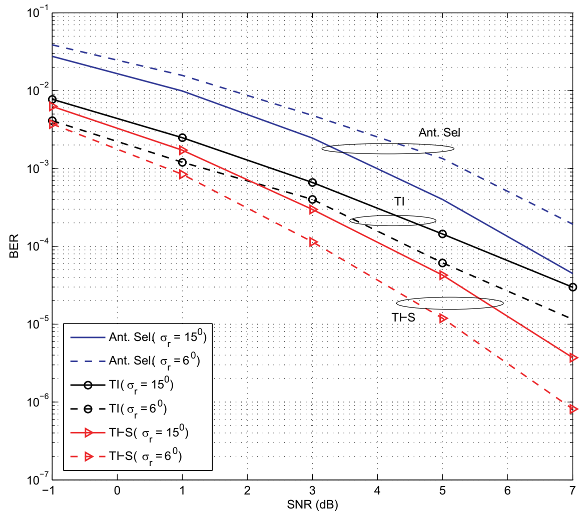


Fig. 4. Single cluster: Average BER vs. SNR for spatial diversity system with $N_t = 4$, $N_r = 4$, $L = 1$, $\theta_r = 45°$, and different angle spreads, σ_r.

However, note that at the SNRs considered, it still outperforms antenna selection. The performance of the sub-optimal phase-only approximation TI-Ph, described in Section V, is within 0.01 dB of ideal TI in this case, and is not shown. Similar observations also hold for the phase-only approximation to TI-S. Therefore, we no longer comment on the efficacy of the phase-only approximation. The average output SNRs of FC, TI-S, TI, and antenna selection receivers are compared in Table I for different AoAs and angle spreads. It can be seen that TI and TI-S deliver the same performance as FC for $\sigma_r = 6°$. Compared to them, FFT-Sel incurs a penalty of 2.2 dB, while antenna selection incurs a penalty of 5.0 dB. For $\sigma_r = 15°$ (lower spatial correlation), TI-S is 0.6 dB below FC, while antenna selection is 3.4 dB worse.

The CDF of the output SNR after combining is plotted in Fig. 5 for different angle spreads for $\theta_r = 60°$. Interestingly, as the correlation increases, the CDF for FC shifts to the right (increases) and its slope decreases, while that for conventional antenna selection shifts to the left (decreases). This is because the largest eigenvalue increases as the correlation increases, and spatial diversity tracks only this eigenvalue [46]. Figure 5 also shows that TI and TI-S can extract the additional beamforming gains as correlation increases. In the extreme case of a completely uncorrelated channel, TI (but not TI-S) may be worse than conventional antenna selection because of its lower diversity order.

B. Single Cluster Scenario: Spatial Multiplexing

Figure 6 compares the average information rates for $\sigma_r = 6°$ and $\theta_r = 45°$. Unlike diversity, any antenna selection scheme with $L < N_r$ RF chains results in a rate loss because it reduces the number of available spatial dimensions. We observe that for $L = 1$, TI, TI-S and TV achieve a 1 bits/s/Hz gain over conventional antenna selection. FFT-Sel performs 0.5 bits/s/Hz worse than TI since the mean AoA falls on a minimum of the FFT beam-pattern. With $L = 2$, the receivers

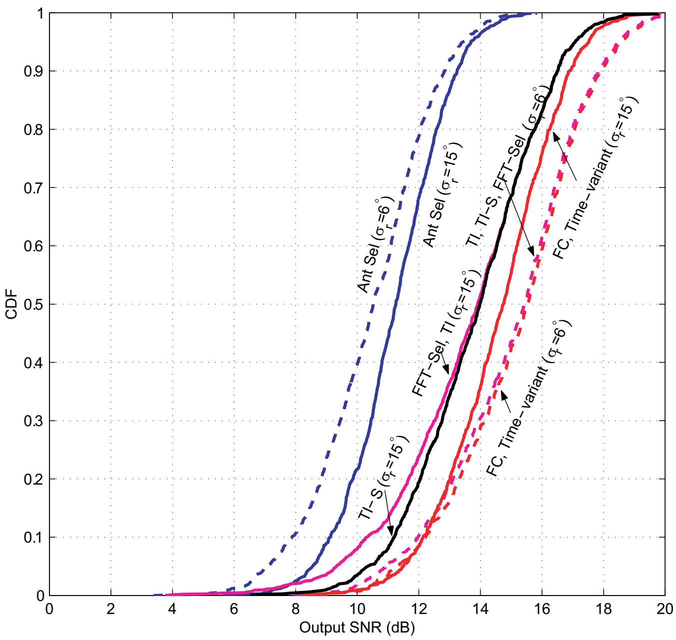


Fig. 5. Single cluster: Effect of angle spread, σ_r, on output SNR for diversity system with $N_t = 4$, $N_r = 4$, $L = 1$, $\theta_r = 60°$, and input SNR $\rho = 10$ dB.

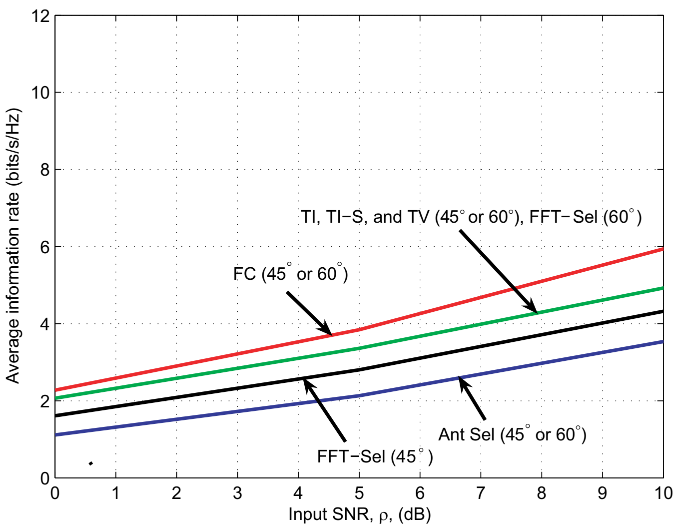


Fig. 6. Single cluster: Average information rate for spatial multiplexing system with $N_t = 2$, $N_r = 4$, $\theta_r = 45°$, and $\sigma_r = 6°$.

employing pre-processing achieve performance parity with FC, while antenna selection is 1 bits/s/Hz less than FC. Note that while FFT-Sel performs worse than TI for $L = 1$, it surprisingly achieves the same performance as TI for $L = 2$.

The CDF of the information rate for various designs for $N_t = 2$ and $N_r = 4$ for different angle spreads is shown in Fig. 7. As σ_r increases, the performance of antenna selection improves, while that of TI, TI-S and FFT-Sel degrades. In the extreme case of a completely uncorrelated channel, antenna selection, FFT-Sel and TI-S will have the same performance, while TI will be slightly worse. This is because the efficacy of the statistics-based solutions reduces as the channel correlated reduces. Note that the information rate of FC decreases as σ_r decreases [22]. This is in contrast to spatial diversity, where reducing σ_r improves the output SNR.

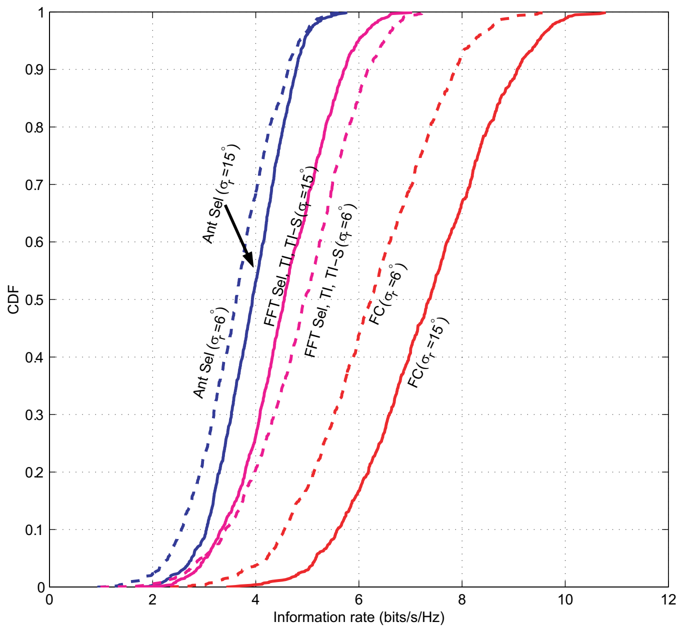


Fig. 7. Single cluster: Effect of angle spread, σ_r, on CDF of capacity for spatial multiplexing system with $N_t = 2$, $N_r = 4$, $L = 1$, $\theta_r = 45°$, and input SNR $\rho = 10$ dB.

C. Multi-Cluster Scenario

For two clusters with mean AoAs of $\theta_{r_1} = 45°$ and $\theta_{r_2} = 75°$, a common angle spread of $\sigma_r = 6°$, and $N_t = N_r = 4$, the main beam in the beam-pattern of TI for spatial diversity points at $60°$, and not at either of the mean AoAs of the two clusters. Additional investigations, results for which are not included due to space constraints, show that the beam-pattern also adapts to the angle spread and the ratio of powers of the two clusters. TI and TI-S still outperform antenna selection by 1 dB and 2 dB, respectively, when $L = 1$. It was observed that the accuracy of phase approximation is marginally worse in this case because of greater amplitude variations in the optimal RF pre-processing matrix elements.

The CDF of the average information rate for a spatial multiplexing system with a two cluster wireless channel with $N_t = 2$, $N_r = 4$, and $L = 2$ is plotted in Fig. 8. The mean AoAs are $45°$ and $75°$, and the angle spread of each cluster is $6°$. We see that both TI and TI-S outperform conventional antenna selection by 1.5 bits/s/Hz and FFT-Sel by 0.5 bits/s/Hz.

VII. Conclusions

In this paper, two novel joint RF-baseband pre-processing designs were proposed for receivers of spatial diversity and spatial multiplexing systems and optimized. The designs exploit spatial correlation to recover the beamforming gain that is lost by conventional antenna selection techniques. They have the potential to reduce the hardware complexity and the signal processing power required by MIMO systems. The first design (TI) reduces the number of output streams to the number of RF chains, and thus eliminates the selection switch. The second design (TI-S) is more sophisticated as the pre-processing matrix outputs the same number of streams input to it and is followed by a selection switch. At the expense of

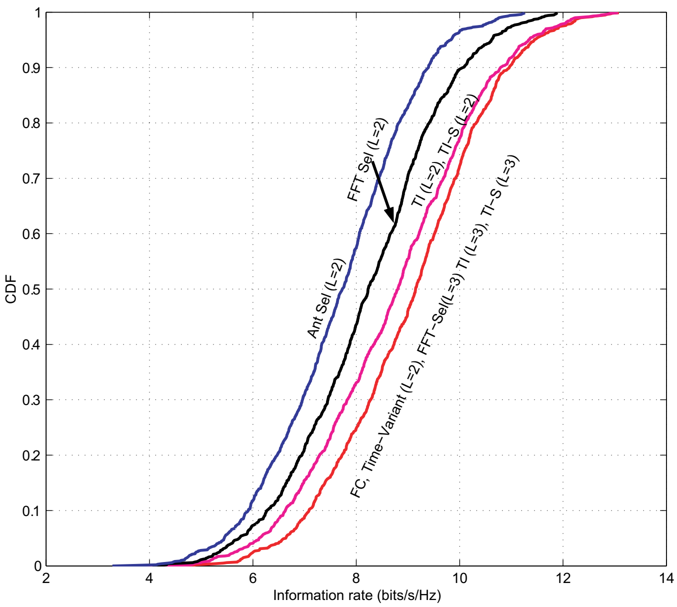


Fig. 8. Multi-cluster: CDF of information rate for spatial multiplexing system with $N_t = 2$, $N_r = 4$, and input SNR $\rho = 10$ dB for two clusters with $\theta_{r_1} = 45°$, $\theta_{r_2} = 75°$, and $\sigma_r = 6°$.

additional computations, this design always outperforms TI. Analogous algorithms hold for the transmitter as well. The transmitter-side implementation consists of up-conversion of L streams, a multiplexing switch (if necessary) that routes the L streams to N_t outputs, followed by RF post-processing that outputs to N_t transmit antennas. The transmitter-side implementation differs in one important manner: depending on the spatial correlation, the transmitter, unlike the receiver, can improve performance by using fewer transmit antennas.

A sub-optimal phase-only approximation was also proposed to address the practical design constraints imposed by today's RF integrated circuit technology. The proposed solutions outperformed conventional antenna selection and FFT Butler matrix pre-processing, which cannot adapt to the variations in the mean angle of arrival. However, the slope of the uncoded BER curve for TI is smaller at finite SNRs as it does not use a selection switch. For high SNR, TI-S is therefore preferable.

Appendix

Appendix A-1: Proof of Theorem 1

We first state a result from [47] which will come in handy in the proof.

Lemma *5:* For any $L \times N_r$ semi-unitary matrix $\mathbf{Z}$ and a positive semi-definite matrix $\mathbf{A}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{N_r}$, the following inequality holds: $\sum_{k=L+1}^{N} \lambda_k \leq \operatorname{Tr}\left\{\mathbf{Z}\mathbf{A}\mathbf{Z}^\dagger\right\} \leq \sum_{k=1}^{L} \lambda_k$. Furthermore, the upper bound is achieved if and only if the columns of $\mathbf{Z}^\dagger$ are the eigenvectors associated with the L largest eigenvalues of $\mathbf{A}$.

The constraint $\mathbf{W}\mathbf{M}_L\mathbf{M}_L^\dagger\mathbf{W}^\dagger = \mathbf{I}_L$ is equivalent to saying that $\mathbf{Z} = \mathbf{W}\mathbf{M}_L$ is semi-unitary. For MRT ($\mathbf{v} = \mathbf{v}_1$), the SNR at the receiver output is given by

$$\gamma_{\text{TI}} = \frac{\rho}{N_t}\|\lambda_1 \mathbf{w}^\dagger \mathbf{Z}\mathbf{u}_1\|^2 \leq \frac{\rho}{N_t}\lambda_1^2\|\mathbf{Z}\mathbf{u}_1\|^2. \quad (20)$$

The above inequality follows from the Cauchy-Schwartz inequality; equality is achieved if and only if $\mathbf{w} = \alpha\mathbf{Z}\mathbf{u}_1$, for some scalar α. Therefore,

$$\begin{aligned}\gamma_{\text{TI}} &= \frac{\rho}{N_t}\mathcal{E}_{\mathbf{H}}\left(\lambda_1^2\mathbf{u}_1^\dagger\mathbf{Z}^\dagger\mathbf{Z}\mathbf{u}_1\right), && (21)\\ &= \frac{\rho}{N_t}\operatorname{Tr}\left\{\mathbf{Z}\left(\mathcal{E}_{\mathbf{H}}\left(\lambda_1^2\mathbf{u}_1\mathbf{u}_1^\dagger\right)\right)\mathbf{Z}^\dagger\right\}, && (22)\\ &= \frac{\rho}{N_t}\operatorname{Tr}\left\{\mathbf{Z}\mathbf{E}_{u_1u_1}\mathbf{Z}^\dagger\right\}. && (23)\end{aligned}$$

Eqn. (22) follows from (21) because $\|\mathbf{c}\|^2 = \operatorname{Tr}\left\{\mathbf{c}\mathbf{c}^\dagger\right\}$, for any column vector $\mathbf{c}$. From Lemma 5, it follows that the optimal $\mathbf{Z}$, denoted by $\mathbf{Z}_{\text{opt}}$, is obtained by choosing the L largest eigenvectors of $\mathbf{E}_{u_1u_1}$, *i.e.*, $\mathbf{Z}_{\text{opt}} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_L]^\dagger$. Therefore, the optimal $\mathbf{M}_L$ takes the form $\mathbf{M}_L = \mathbf{W}^{-1}\mathbf{Z}_{\text{opt}}$, where $\mathbf{W}^{-1}$ is an arbitrary $L \times L$ full rank matrix.

We also get from the above proof that the optimal $\mathbf{w}_{\text{TI}}$ is given by $\mathbf{w}_{\text{TI}} = \frac{\mathbf{Z}_{\text{opt}}\mathbf{u}_1}{\|\mathbf{Z}_{\text{opt}}\mathbf{u}_1\|}$.

Appendix A-2: Proof of Lemma 1

The SNR at the output of the pre-processor, $\mathbf{M}_L$, is given by $\lambda_1^2(\mathbf{M}_L\mathbf{H})$, where $\lambda_1^2(\mathbf{M}_L\mathbf{H})$ is the largest singular value of $\mathbf{M}_L\mathbf{H}\mathbf{H}^\dagger\mathbf{M}_L^\dagger$. It satisfies the following inequality: $\frac{1}{\min(N_t,L)}\|\mathbf{M}_L\mathbf{H}\|_F^2 \leq \lambda_1^2(\mathbf{M}_L\mathbf{H}) \leq \|\mathbf{M}_L\mathbf{H}\|_F^2$, where $\|.\|_F$ denotes the Frobenius norm. Therefore, the diversity order is the same as that of a system with SNR $\|\mathbf{M}_L\mathbf{H}\|_F^2$ [14].

The diversity order is related to the moment generating function (MGF), $\mathcal{M}(s)$, of the SNR as follows [48]: If $\lim_{s\to\infty}\mathcal{M}(s) = b|s|^{-d} + O\left(|s|^{-d}\right)$, then the diversity order is d. It turns out that the MGF of $\|\mathbf{M}_L\mathbf{H}\|_F^2$ for a correlated channel can be calculated exactly using [49] for our case, in which $\mathbf{M}_L$, $\mathbf{T}$, and $\mathbf{R}$ have ranks L, N_t, and N_r, respectively, and $\|\mathbf{M}_L\mathbf{H}\|_F^2$ is the sum of LN_t complex Gaussian RVs. From [49], it can be shown that as $s \to \infty$ we have $\mathcal{M}(s) = b|s|^{-N_tL}$, where b is a constant. Hence, the result.

Appendix A-3: Proof of Lemma 3

Let $\boldsymbol{\Lambda}_1 = \operatorname{diag}(\lambda_{11}, \ldots, \lambda_{1N_r})$ and $\boldsymbol{\Lambda}_\alpha = \operatorname{diag}(\lambda_{11} + \alpha, \lambda_{12}, \ldots, \lambda_{1N_r})$, where α is a positive constant. We first show that $|\mathbf{I}_{N_t} + \mathbf{A}\boldsymbol{\Lambda}_\alpha\mathbf{A}^\dagger|$ increases monotonically with α. Define $\boldsymbol{\Sigma}_1 = \operatorname{diag}(1, 0, \ldots, 0)$, and $\mathbf{D}_1(\alpha) = \mathbf{I}_{N_t} + \mathbf{A}\boldsymbol{\Lambda}_\alpha\mathbf{A}^\dagger$. Using the formula for derivative of the determinant, we can show that

$$\begin{aligned}\frac{d|\mathbf{D}_1(\alpha)|}{d\alpha} &= |\mathbf{D}_1(\alpha)|\operatorname{Tr}\left\{\mathbf{D}_1^{-1}(\alpha)\frac{d\mathbf{D}_1(\alpha)}{d\alpha}\right\},\\ &= |\mathbf{D}_1(\alpha)|\operatorname{Tr}\left\{\mathbf{D}_1^{-1}(\alpha)\mathbf{a}_1\mathbf{a}_1^\dagger\right\},\\ &= |\mathbf{D}_1(\alpha)|\left(\mathbf{a}_1^\dagger\mathbf{D}_1^{-1}(\alpha)\mathbf{a}_1\right), && (24)\end{aligned}$$

where $\mathbf{a}_1$ is the first column of $\mathbf{A}$. Since $\mathbf{D}_1(\alpha)$ is positive-definite, $\frac{d|\mathbf{D}_1(\alpha)|}{d\alpha} > 0$. Thus, $|\mathbf{D}_1(\alpha)|$ increases with α. Similarly, we can show that $\left|\mathbf{I}_{N_t} + \mathbf{A}\boldsymbol{\Lambda}\mathbf{A}^\dagger\right|$ increases as each element of the diagonal matrix $\boldsymbol{\Lambda}$ increases.

Appendix A-4: Proof of Lemma 4

The proof for the right-hand inequality is given in [50, Lemma 3.3.1]. It can be verified by direct substitution that if $\mathbf{Q}$ is the conjugate transpose of the L eigenvectors of $\mathbf{G}$

corresponding to its L largest eigenvalues, then equality is achieved.

To prove the lower bound on $e_i(\mathbf{QG})$, let the SVD of $\mathbf{QG}$ be $\mathbf{QG} = \mathbf{U}_{QG}\mathbf{\Sigma}_{QG}\mathbf{V}_{QG}^{\dagger}$ and the SVD of $\mathbf{G}$ be $\mathbf{G} = \mathbf{U}_{G}\Sigma_{G}\mathbf{V}_{G}^{\dagger}$. Then $\mathbf{QQ}^{\dagger} = \mathbf{I}_L$ leads to

$$\mathbf{\Sigma}_{QG}\mathbf{V}_{QG}^{\dagger}\mathbf{V}_{G}\mathbf{\Sigma}_{G}^{-2}\mathbf{V}_{G}^{\dagger}\mathbf{V}_{QG}\mathbf{\Sigma}_{QG}^{\dagger} = \mathbf{I}_L,$$
$$\Rightarrow \quad (\mathbf{V}_{G}^{\dagger}\mathbf{V}_{QG})^{(L)}\Sigma_{QG}^{(L)^{-2}}(\mathbf{V}_{G}^{\dagger}\mathbf{V}_{QG})^{(L)^{\dagger}} = \mathbf{\Sigma}_{G}^{(L)^{-2}}, \quad (25)$$

where $(.)^{(L)}$ denotes the first L rows of a matrix. Notice that the matrix $\mathbf{V}' = (\mathbf{V}_{G}^{\dagger}\mathbf{V}_{QG})^{(L)}$ is unitary. Applying the right side of the inequality in (13) on $\mathbf{V}'\mathbf{\Sigma}_{QG}^{(L)^{-2}}\mathbf{V}'^{\dagger}$, we get the lower bound on $e_i(\mathbf{QG})$.

REFERENCES

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs. Tech. J.*, vol. 1, pp. 41–59, 1996.

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

[3] 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, 1998.

[4] P. Dighe, R. Mallik, and S. Jamuar, "Analysis of transmit-receive diversity in rayleigh fading," *IEEE Trans. Commun.*, pp. 694–703, 2003.

[5] N. Sollenberger, "Diversity and automatic link transfer for a TDMA wireless access link," in *Proc. Globecom*, pp. 532–536, 1993.

[6] S. Thoen *et al.*, "Performance analysis of combined transmit-SC/receive-MRC," *IEEE Trans. Commun.*, vol. 49, pp. 5–8, Jan. 2001.

[7] C. Leung and X. Feng, "Jointly optimal transmit and receive diversity," *Electron. Lett.*, vol. 38, pp. 594–596, 2002.

[8] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microwave Mag.*, pp. 46–56, Mar. 2004.

[9] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, pp. 68–73, Oct. 2004.

[10] N. B. Mehta and A. F. Molisch, "Antenna selection in MIMO systems," in *MIMO System Technology for Wireless Communications* (G. Tsulos, ed.), ch. 6, CRC Press, 2006.

[11] I. Bahceci, T. Duman, and Y. Altunbasak, "Multiple-antenna transmission systems: Performance analysis and code construction," *IEEE Trans. Inform. Theory*, vol. 10, pp. 2669–2681, 2003.

[12] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, pp. 322–324, Aug. 2002.

[13] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, pp. 1773–1776, 1999.

[14] D. A. Gore and A. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Trans. Signal Processing*, vol. 50, pp. 2580–2588, Oct. 2002.

[15] R. W. Heath, S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, pp. 142–144, Apr. 2001.

[16] A. Gorokhov, D. Gore, and A. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: theory and algorithms," *IEEE Trans. Signal Processing*, vol. 51, pp. 2796–2807, 2003.

[17] Z. Chen, "Asymptotic performance of transmit antenna selection with maximal-ratio combining for generalized selection criterion," *IEEE Commun. Lett.*, vol. 8, pp. 247–249, Apr. 2004.

[18] A. Ghrayeb and T. M. Duman, "Performance analysis of MIMO systems with antenna selection over quasi-static fading channels," *IEEE Trans. Veh. Technol.*, vol. 52, pp. 281–288, Mar. 2003.

[19] A. Ghrayeb, A. Sanei, and Y. Shayan, "Space-time trellis codes with receive antenna selection in fast fading," *Electron. Lett.*, vol. 40, 2004.

[20] C. N. Chuah *et al.*, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, 2002.

[21] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2363–2371, 2003.

[22] D.-S. Shiu *et al.*, "Fading correlation and its effect on the quality of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, May 2000.

[23] S. A. Jafar and A. Goldsmith, "Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 990–997, May 2005.

[24] D. A. Gore, R. W. Heath, and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 11, pp. 491–493, 2002.

[25] R. Narasimhan, "Spatial multiplexing with transmit antenna and constellation selection for correlated MIMO fading channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 2829–2838, Nov. 2003.

[26] H. Zhang and H. Dai, "Fast transmit antenna selection algorithms for MIMO systems with fading correlation," in *Proc. VTC*, pp. 1638–1642, 2004.

[27] A. Molisch and X. Zhang, "FFT-based hybrid antenna selection schemes for spatially correlated MIMO channels," *IEEE Commun. Lett.*, vol. 8, pp. 36–38, 2004.

[28] X. Zhang, A. F. Molisch, and S. Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Sig. Proc.*, vol. 53, pp. 4091–4103, 2005.

[29] D. Hand, H. Mannila, and P. Smyth, *Principles of Data Mining*. MIT Press, 2001.

[30] A. F. Molisch, "Effect of far scatter clusters in MIMO outdoor channel models," in *Proc. VTC*, pp. 534–538, 2003.

[31] T. Ohira, "Analog smart antennas: An overview," in *Proc. PIMRC*, pp. 1502–1506, 2002.

[32] S. Denno and T. Ohira, "Modified constant modulus algorithm for digital signal processing adaptive antennas with microwave analog beamforming," *IEEE Trans. Antennas and Propagation*, vol. 50, pp. 850–857, 2002.

[33] G. M. Rebeiz, G.-L. Tan, and J. S. Hayden, "RF MEMS phase shifters: Design and applications," *IEEE Microwave Mag.*, pp. 72–81, 2002.

[34] P. Sudarshan *et al.*, "Antenna selection with RF pre-processing: Robustness to RF and selection non-idealities," in *Proc. IEEE Radio & Wireless Conf. (RAWCON)*, 2004.

[35] A. Annamalai, C. Tellambura, and V. Bhargava, "Equal-gain diversity receiver performance in wireless channels," *IEEE Trans. Commun.*, vol. 48, pp. 1732–1745, 2000.

[36] R. Mallik, M. Win, and J. Winters, "Performance of dual-diversity predetection EGC in correlated rayleigh fading with unequal branch SNRs," *IEEE Trans. Commun.*, vol. 50, pp. 1041–1044, 2002.

[37] D. Love and R. Heath, "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, pp. 1102–1110, 2003.

[38] J. P. Kermoal *et al.*, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1211–1226, Aug. 2002.

[39] D. Asztely, "On antenna arrays in mobile communication systems: Fast fading and GSM base station receiver algorithms," Tech. Rep. IR-S3-SB-9611, Royal Institute of Technology, Mar. 1996.

[40] M. S. Alouini, A. Scaglione, and G. B. Giannakis, "PCC: Principal components combining for dense correlated multipath fading environments," in *Proc. VTC*, pp. 2510–2517, 2000.

[41] X. Zeng and A. Ghrayeb, "Performance analysis of combined convolutional coding and space-time block coding with antenna selection," in *Proc. Globecom*, pp. 795–799, 2004.

[42] X. N. Zeng and A. Ghrayeb, "Performance bounds for space-time block codes with receive antenna selection," *IEEE Trans. Inform. Theory*, vol. 50, pp. 2130–2137, 2004.

[43] P. J. Voltz, "Characterization of the optimum transmitter correlation matrix for MIMO with antenna subset selection," *IEEE Trans. Commun.*, vol. 51, pp. 1779–1782, Nov. 2003.

[44] S. T. Smith, "Phase-only adaptive nulling," *IEEE Trans. Signal Processing*, vol. 47, pp. 1835–1843, 1999.

[45] A. F. Molisch, *Wireless Communications*. Wiley-IEEE Press, 2005.

[46] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2172–2178, 2000.

[47] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Prentice-Hall, 1997.

[48] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, pp. 1389–1398, 2003.

[49] R. K. Mallik, "On multivariate rayleigh and exponential distributions," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1499–1515, 2003.

[50] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, 1994.

Pallav Sudarshan received his Bachelor of Technology degree in Electrical Engineering from Indian Institute of Technology, Delhi in 1999, and PhD degree in Electrical and Computer Engineering from North Carolina State University, Raleigh in 2004. He interned at Mitsubishi Electric Reserach Labs, Cambridge, MA, in 2003-2004, where he worked on HSDPA systems and antenna selection methods for MIMO systems. Since 2005, he is a senior electrical engineer at Mobile Devices, Motorola, Libertyville, IL. He is involved in receiver design for IEEE 802.16 and 4G standards. His research interests include OFDM systems, MIMO systems, channel coding, and system design aspects of wireless systems.

Neelesh B. Mehta received his Bachelor of Technology degree in Electronics and Communications engineering from the Indian Institute of Technology, Madras in 1996, and his M.S. and PhD degrees in Electrical engineering from the California Institute of Technology, Pasadena, CA in 1997 and 2001, respectively. He was a visiting graduate student researcher at Stanford University in 1999 and 2000.

He was a research scientist at the Wireless Systems Research group in AT&T Laboratories, Middletown, NJ until 2002. In 2002 and 2003, he was a Staff Scientist in Broadcom Corp., and was involved in GPRS and EDGE cellular handset development. Since Sept. 2003, he has been a Member of Technical Staff at the Mitsubishi Electric Research Laboratories, Cambridge, MA. His research includes work on link adaptation, multiple access protocols, WCDMA downlinks, system-level analysis and simulation of cellular systems, MIMO and antenna selection, and cooperative communications. He is also actively involved in the radio access network physical layer (RAN1) standardization activities in 3GPP.

Andreas F. Molisch (S'89, M'95, SM00, F'05) received the Dipl. Ing., Dr. techn., and habilitation degrees from the Technical University Vienna (Austria) in 1990, 1994, and 1999, respectively. From 1991 to 2000, he was with the TU Vienna, becoming an associate professor there in 1999. From 2000-2002, he was with the Wireless Systems Research Department at AT&T (Bell) Laboratories Research in Middletown, NJ. Since then, he has been a Senior Principal Member of Technical Staff with Mitsubishi Electric Research Labs, Cambridge, MA. He is also professor and chairholder for radio systems at Lund University, Sweden.

Dr. Molisch has done research in the areas of SAW filters, radiative transfer in atomic vapors, atomic line filters, smart antennas, and wideband systems. His current research interests are MIMO systems, measurement and modeling of mobile radio channels, and UWB. Dr. Molisch has authored, co-authored or edited four books (among them the recent textbook "Wireless Communications," Wiley-IEEE Press), eleven book chapters, some 95 journal papers, and numerous conference contributions.

Dr. Molisch is an editor of the IEEE Trans. Wireless Comm., co-editor of a special issue on MIMO and smart antennas in J. Wireless Comm. Mob. Comp., and co-editor of a recent IEEE JSAC special issue on UWB. He has been member of numerous TPCs, vice chair of the TPC of VTC 2005 spring, and will be general chair of ICUWB 2006. He has participated in the European research initiatives "COST 231", "COST 259", and "COST273", where he was chairman of the MIMO channel working group, he was chairman of the IEEE 802.15.4a channel model standardization group, and is also chairman of Commission C (signals and systems) of URSI (International Union of Radio Scientists). Dr. Molisch is a Fellow of the IEEE and recipient of several awards

Jin Zhang received her Ph.D. in electrical engineering from University of Ottawa in 1991. Since 2001, Dr. Zhang has been the head and Senior Principal Technical Staff of digital communications and networking group at Mitsubishi Electric Research Laboratories (MERL). She is leading many new wireless communications and networking research projects that include UWB, ZigBee, wireless sensor network, MIMO, high speed WLAN, and next generation mobile communications. Prior to joining MERL she worked for Nortel Networks for more than 10 years, where she held engineering and management positions in the areas of digital signal processing, advanced wireless technology development and optical networks. Dr. Zhang is a Senior Member of IEEE and an associate editor of IEEE Transactions on Broadcasting.