Receive Antenna Selection for MIMO Spatial Multiplexing: Theory and Algorithms

Alexei Gorokhov, Dhananjay A. Gore, and Arogyaswami J. Paulraj, *Fellow, IEEE*

***Abstract*—This paper discusses the problem of receive antenna subset selection in multiple-input multiple-output spatial multiplexing (MIMO-SM) systems. We develop selection algorithms for maximizing the channel capacity. One algorithm in particular allows tractable statistical analysis of performance. We leverage this to prove that the capacity of the system through receive antenna selection is statistically lower bounded by the capacity of a set of parallel independent single input multiple output (SIMO) channels, each with selection diversity. This provides the crucial step in proving the next main result: The diversity order achievable through antenna selection is the same as that of the full system. The result sets up strong motivation for introducing receive selection in MIMO-SM systems. The remainder of the paper discusses selection algorithms for two popular MIMO-SM systems, namely, ordered successive interference cancellation with independently encoded layers and minimum mean square error (MMSE) receiver with joint encoding of data streams. Extensive Monte Carlo simulations are presented to validate and demonstrate performance.**

***Index Terms*—Antenna subset selection, MIMO, MMSE, spatial multiplexing, VBLAST.**

I. Introduction

MULTIPLE antenna technology significantly improves wireless link performance. The degrees of freedom afforded by the multiple antennas may be used to increase reliability through space time diversity techniques [1]–[3] and/or to increase data rate through spatial multiplexing techniques [4]–[7]. However, a major limiting factor in the deployment of MIMO systems is the cost of multiple analog chains (such as low noise amplifiers, mixers, and analog-to-digital converters at the receiver end). Antenna subset selection where transmission and/or reception is performed through a selection of the total available antennas is a powerful solution that reduces the need for multiple analog chains yet retains many diversity benefits. The core idea of antenna selection is to use a limited number of analog chains that is adaptively switched to a subset of the available antennas. An appropriate subset of antennas can be identified, e.g., within the training phase, by probing all receive antennas with the available set of receive chains. This general approach provides certain diversity benefits at a low additional cost that is mainly determined by low-cost RF switches rather than by expensive analog chains.

Early work on antenna selection focused on selection in multiple-input single-output (MISO) and single-input multiple-output (SIMO) systems. This includes the hybrid selection and maximum ratio combining (MRC) approach in [8]. Recently, there has been increasing interest [9]–[14] in applying antenna subset selection techniques to MIMO links. These efforts may be broadly classified into two main categories. The first body of work [9], [10], [12] discusses antenna subset selection for maximizing the channel capacity. While significant effort has been expended in developing selection algorithms, there is, in general, little accompanying performance analysis. One exception is [10], where the authors develop analysis based on a simplified selection rule motivated by an upper bound on capacity. However, the bound employed is rather weak, thereby limiting the applicability of the analysis. Other contributions develop selection algorithms when practical signaling schemes are used on the link. Previous work includes selection algorithms for spatial multiplexing when the zero forcing (ZF) receiver is used [11], [14] and for space-time coded [13] techniques. Finally, the authors of [15] analyze the effect of receive antenna selection for the error rate performance of space-time codes. Specifically, they show that a simple selection rule that aims at maximizing received power results in an error rate behavior similar to that of a full system (with all receive antennas present). However, much remains to be done, especially for practical MIMO-SM systems.

In this paper, we focus on antenna subset selection at the receiver where the antennas are selected to maximize channel capacity. We will assume no channel state information (CSI) available at the transmitter. We first develop a suite of near-optimal algorithms for maximizing channel capacity. The first algorithm, although suboptimal, allows for a tractable statistical analysis of selection gains. We show that MIMO capacity through antenna selection is statistically lower bounded by the capacity of a set of parallel independent SIMO channels, each with selection diversity. This observation, along with the well-known representation of a MIMO channel as a set of parallel SIMO subchannels with MRC, allows us to prove the equivalence between the diversity order of the full system and the system with antenna selection at high signal-to-noise ratio (SNR). In this paper, the diversity is understood as a relationship (slope) between the capacity and outage rate in the region of small outage rates. It is worth making the distinction between this notion of the diversity and a relationship between the error rate performance and SNR, which is often addressed in the space-time coding literature; see, e.g., [15].

Manuscript received December 15, 2002; revised June 4, 2003. The associate editor coordinating the review of this paper and approving it for publication was Dr. Rick S. Blum.

A. Gorokhov is with the Philips National Research Laboratory, on leave from the Centre National de la Recherche Scientifique (CNRS), the Laboratoire des Signaux et Systèmes, École Supérieure d'Électricité (Supélec), Paris, France (e-mail: gorokhov@philips.com).

D. A. Gore is with Qualcomm, San Diego, CA 92121 USA (e-mail: dgore@qualcomm.com).

A. J. Paulraj is with the Information Systems Laboratory, Stanford University, Stanford, CA 94305 USA (e-mail: apaulraj@stanford.edu).

Digital Object Identifier 10.1109/TSP.2003.818204

The diversity equivalence result motivates the use of antenna subset selection in practical MIMO-SM systems. We develop low complexity decremental selection algorithms for two popular architectures: a system with joint encoding of the transmitted streams and MMSE combining of these streams at the receiver, later referred to as joint MMSE (JMMSE) and the famous ordered successive interference cancellation (OSIC) system with independent encoding of streams, which is often called V-BLAST.

The paper is organized as follows. In Section II, we introduce the signal model and discuss selection in SIMO systems. Section III presents some fundamental relationships between the MIMO capacity of MIMO systems with and without receive antenna selection. Section IV contains the core result on the equivalence of diversity orders. In Section V, we develop and analyze practical selection algorithms. In Section VI, we develop decremental selection algorithms for the JMMSE and OSIC systems. Monte Carlo simulations are presented in Section VII, and we conclude the paper with a summary of results.

II. Background

In this section, we specify the data model and review well-known ideas and concepts that we use later in the paper. We discuss the capacity of SIMO channels with antenna selection at the receiver and emphasize the behavior of outage capacity achieved through antenna selection.

First, we introduce the notation convention used throughout the manuscript. Unless otherwise specified, bold lowercase (uppercase) letters denote vectors (matrices). A matrix variable indexed with a single integer denotes the corresponding row. The notation $(a\,:\,b)$ stands for indices a through b, whereas (:) stands for the whole scope of indices. Additionally, superscripts $(^T)$ and $(^H)$ denote matrix transpose and conjugate transpose, whereas $\|\cdot\|$ is the Euclidean vector norm.

A. Data Model

Assume a nonselective (flat fading) linear time-invariant MEA channel between N transmit and M receive antennas, as described in Fig. 1. The data model is

$$\boldsymbol{x}[k] = \sqrt{E_s}\boldsymbol{H}\boldsymbol{s}[k] + \boldsymbol{n}[k] \tag{1}$$

where the $M \times 1$ vector $x[k] = [x_1[k], \ldots, x_M[k]]^T$ represents the kth sample of signals collected at the outputs of M receivers and sampled at the symbol rate, $s[k] = [s_1[k], \ldots, s_N[k]]^T$ is the vector of N symbols transmitted by the transmit antennas, E_s is the average signal energy per receive antenna and per channel use, $n[k] = [n_1[k], \ldots, \boldsymbol{n}_M[k]]^T$ is the additive white Gaussian noise (AWGN) with energy $(N_0/2)$ per complex dimension, and $\boldsymbol{H} = [\boldsymbol{H}_{:,1}, \ldots, \boldsymbol{H}_{:,N}]$ is the $M \times N$ channel matrix, $\boldsymbol{H}_{:,q} = [\boldsymbol{H}_{1,q}, \ldots, \boldsymbol{H}_{M,q}]^T$, $1 \leq q \leq N$, where $\boldsymbol{H}_{p,q}$ is a scalar channel between the pth receive and the qth transmit antenna.

The capacity $C(\boldsymbol{H})$ of a fixed MIMO channel specified by $\boldsymbol{H}$ is given by (see [6], [7])

$$C(\boldsymbol{H}) = \log_2 \det \left(\boldsymbol{I}_N + \left(\frac{E_s}{N_0} \right) \boldsymbol{R}_{ss} \boldsymbol{H}^H \boldsymbol{H} \right), \quad \mathrm{tr}(\boldsymbol{R}_{ss}) = 1 \tag{2}$$

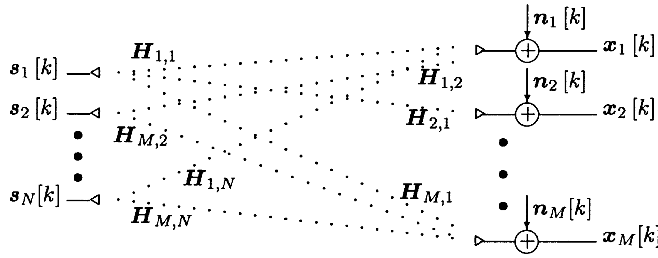

Fig. 1. MEA transmission system model.

where $\boldsymbol{R}_{ss} = \mathbb{E}\{s[k]s[k]^H\}$ is the covariance matrix of the transmitted signals, $\mathrm{tr}(\cdot)$ and $\det(\cdot)$ denote trace and determinant, respectively, and $\boldsymbol{I}_N$ is the $N \times N$ identity matrix.

In this paper, we will mainly consider *outage* characteristics of MIMO capacity when $\boldsymbol{H}$ is a random channel unknown to the transmitter. In other words, we assume a block fading channel model. The channel itself is modeled as follows.

As1) The entries of $\boldsymbol{H}$ are zero mean jointly circular Gaussian such that the covariance matrix of any two columns of $\boldsymbol{H}$ and the autocovariance of any column are scaled identity matrices.

This assumption requires the channel coefficients from a transmit antenna to the receive antennas to be uncorrelated. Note that we place no restrictions on transmit antenna correlation structure, and indeed, the algorithms and much of the analysis presented in the rest of the paper is independent of transmit correlation. For the sake of simplicity, we will consider uncorrelated transmitters: $\boldsymbol{R}_{ss} = (1/N)\boldsymbol{I}_N$. To summarize, the channel is assumed to be Rayleigh fading with uncorrelated receive antennas.

A subset of the total number of receive antennas is selected and used for reception, assuming that perfect CSI is available at the receiver. This assumption implies accurate channel acquisition/tracking at the receiver. As already mentioned, we assume that no CSI is available at the transmitter.

B. Antenna Selection in SIMO Systems

The SIMO scenario has been extensively studied over the past years, in particular w.r.t. antenna selection (see, e.g., [8] and [16]). In this paper, we use one important conclusion of the SIMO selection theory, namely, that antenna selection achieves the same diversity order as MRC.

Consider a $M \times 1$ SIMO channel $\boldsymbol{H}_{:,1}$ defined by the first column of $\boldsymbol{H}$. First, recall that the optimal receiver strategy for a SIMO system is to use a $1 \times M$ spatial filter $\boldsymbol{H}_{:,1}^H$ matched to the signature $\boldsymbol{H}_{:,1}$. This filtering, which is widely known as MRC, results in a scalar composite channel (e.g., between the channel input and the filter output) with capacity

$$C(\boldsymbol{H}_{:,1}) = \log_2 \left(1 + \left(\frac{E_s}{N_0} \right) \|\boldsymbol{H}_{:,1}\|^2 \right) \tag{3}$$

and is a special case of (2) with $N = 1$. In the context of antenna selection, we assume that the "best" antenna is used for reception. The capacity with optimal selection may be written as

$$C_s(\boldsymbol{H}_{:,1}) = \log_2\left(1 + \left(\frac{E_s}{N_0}\right)\max_p |\boldsymbol{H}_{p,1}|^2\right). \tag{4}$$

From (3) and (4), it is easily verified that the analytical characterization of outage capacity depends directly on the receive SNR. In both cases, the distribution laws of the respective SNR are well known. The closed-form expressions for their cumulative density functions (c.d.f.s) are as follows:

$$\begin{aligned}&\mathbb{P}\left\{\|H_{:,1}\|^2 < x\right\} = \Gamma_i(x; M)\\ &\qquad\mathbb{P}\left\{\max_p |\boldsymbol{H}_{p,1}|^2 < x\right\} = (1 - \exp(-x))^M\end{aligned} \tag{5}$$

where $\Gamma_i(x;M) = \Gamma(M)^{-1}\int_0^x u^{M-1}\exp(-u)du$ is the incomplete Gamma function with parameter M.

Diversity is preferred in random fading environments since it plays a crucial role in the relationship between the quality of service (QoS) (such as outage rate or average error rate) on one hand and system performance (such as throughput) on the other. This relationship is of particular interest at small outage rates since high QoS requirements are mandatory in most situations. Consequently, diversity order is often viewed as a slope of, e.g., outage capacity versus outage rate, or vice versa, in the region of small $(< 10^{-1})$ outage rates. Verify that the c.d.f.s in (5) stand for outage rates $\mathbb{P}\{\cdot\}$ with corresponding outage SNR x and outage capacity $\log_2(1+x)$. From (5), observe that in the region of small outage rates, we have

$$\begin{aligned}\mathbb{P}\left\{\max_p |\boldsymbol{H}_{p,1}|^2 < x\right\} &= (1 - \exp(-x))^M\\ &= (1 - (1 - x + o(x)))^M\\ &= x^M(1 + o(x))\\ \mathbb{P}\left\{\|\boldsymbol{H}_{:,1}\|^2 < x\right\} &= \Gamma(M)^{-1}\int_0^x u^{M-1}\exp(-u)du\\ &= (M!)^{-1}x^M(1 + o(x)).\end{aligned}$$

Observe that in the region of small outage rates, both c.d.f.s behave as the Mth power of SNR x. Moreover

$$\mathbb{P}\left\{\max_p |\boldsymbol{H}_{p,1}|^2 < x\right\} \to \mathbb{P}\left\{\|\boldsymbol{H}_{:,1}\|^2 < x(M!)^{\frac{1}{M}}\right\} \text{ as } x \to 0.$$

In other words, MRC and optimal selection of one antenna yield similar relationships between outage capacity and outage rate. Note, however, a constant SNR gain of $(M!)^{1/M}$ in favor of MRC. Since this gain is invariant w.r.t. the outage rate, both schemes have the same diversity order. This equivalence in diversity order is also visible through the following obvious inequality:

$$\left(\frac{1}{M}\right)\sum_p |\boldsymbol{H}_{p,1}|^2 \le \max_p |\boldsymbol{H}_{p,1}|^2 \le \sum_p |\boldsymbol{H}_{p,1}|^2. \tag{6}$$

The inequality states that the SNR through selection is bounded between two quantities, both of which have the same diversity order.

Finally, optimal selection of more than one antenna followed by any of the existing combining schemes (MRC, equal gain combining, etc. [16]) yields the same diversity order M since the performance of these schemes is upper bounded by that of MRC over M antennas and lower bounded by the performance of one antenna selection.

III. MIMO Selection of Receive Antennas

In this section, we assume that a set of M sensors is available at the receiver, whereas the number of receive chains is $N < M$. We further assume that the receiver selects a subset of N out of M available receive antennas to maximize the capacity of the resulting $N \times N$ MIMO channel.[1]

A. MIMO Capacity With Receive Selection

Denote the indices of the selected subset of receive antennas by $r = [r_1, \ldots, r_N]$ and the corresponding $N \times N$ channel matrix consisting of the rows of $\boldsymbol{H}$ (indexed by r) by $\boldsymbol{H}_r$. The capacity $C_r(\boldsymbol{H})$ associated with the selection is

$$C_r(\boldsymbol{H}) = \log_2\det\left(\boldsymbol{I}_N + \left(\frac{E_s}{N_0}\right)\boldsymbol{H}_r^H\boldsymbol{H}_r\right). \tag{7}$$

The goal of this study is to find a tractable statistical description of the capacity $C_r(\boldsymbol{H})$. Since a closed-form characterization of the optimal solution (7) is difficult, we resort to a simple tight lower bound. Define $\boldsymbol{U}$ as a $M \times N$ orthonormal basis of the column space of $\boldsymbol{H}$ and $\boldsymbol{U}_r$ as the $N \times N$ block of $\boldsymbol{U}$ corresponding to $\boldsymbol{H}_r$.

Lemma 1: The capacity $C_r(\boldsymbol{H})$ from (7) may be lower bounded as follows:

$$\begin{aligned}C_r(\boldsymbol{H}) \ge \log_2\det&\left(\boldsymbol{I}_N + \left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)\\ &+\log_2\det\left(\boldsymbol{U}_r\boldsymbol{U}_r^H\right)\end{aligned} \tag{8}$$

wherein the strict equality holds under **As1**) when either $M \to \infty$ or $(E_s/N_0) \to \infty$.

Proof: See the Appendix. ■

The derivation of this bound requires similar approximations as used by Foschini (see [5]) to derive a lower bound on capacity. The approximation is known to be rather inaccurate for $N = M$ and low and moderate (E_s/N_0) because of the exponential distribution of the minor component of $(\boldsymbol{H}^H\boldsymbol{H})$. The accuracy of this bound and conditioning of $(\boldsymbol{H}^H\boldsymbol{H})$ improves significantly when $M > N$. This is always the case in the context of receive antenna selection.

The first term on the right-hand side of (8) stands for the *full* capacity $C(\boldsymbol{H})$ of the $M \times N$ channel $\boldsymbol{H}$, i.e., when all receive antennas are used. The second term is nonpositive since

$$\det\left(\boldsymbol{U}_r\boldsymbol{U}_r^H\right) = \det\left(\boldsymbol{U}_r^H\boldsymbol{U}_r\right) \le \det(\boldsymbol{U}^H\boldsymbol{U}) = 1. \tag{9}$$

The term represents the capacity reduction caused by removing $(M-N)$ "least significant" antennas. We may write

$$C(\boldsymbol{H}) - C_r(\boldsymbol{H}) \le L_r(\boldsymbol{H}), \quad L_r(\boldsymbol{H}) = -\log_2\det\left(\boldsymbol{U}_r\boldsymbol{U}_r^H\right). \tag{10}$$

[1]Theoretical study of selection of more than N receive antennas will not be considered. Nonetheless, some important conclusions regarding the diversity of the system will be made in a more general context.

The term $L_r(\boldsymbol{H})$ may be interpreted as an upper bound on the capacity loss caused by antenna selection w.r.t. the full capacity $C(\boldsymbol{H})$ of the $M \times N$ channel. From (10), the antenna subset that minimizes the upper bound on the capacity loss corresponds to the $N \times N$ block of the $M \times N$ unitary matrix $\boldsymbol{U}$ that has the maximum determinant.

B. Loss Characterization

Statistical properties of the full capacity $C(\boldsymbol{H})$ are well understood nowadays. We focus on the statistical properties of the *loss factor* $L_r(\boldsymbol{H})$.

Lemma 2: Assume **As1**); then, $C(\boldsymbol{H})$ and $L_r(\boldsymbol{H})$ are statistically independent.

Proof: Note that $C(\boldsymbol{H})$ may be expressed in terms of the singular values of $\boldsymbol{H}$, whereas $L_r(\boldsymbol{H})$ is defined by the left-hand singular vectors of $\boldsymbol{H}$. These quantities are statistically independent under **As1**); see e.g., [17]. ■

Remark: The importance of Lemma 2 lies in the fact that it allows the performance of MIMO selection algorithms to be quantified by a loss w.r.t. the full MIMO capacity, which is statistically independent from the full capacity. Clearly, this lemma may be extended to any $\boldsymbol{H}$ whose distribution is invariant w.r.t. a left-hand rotation. Hence, it enables the analysis of antenna selection for a fairly general fading model. In other words, under the assumption **As1**), the lemma implies that the loss term L_r is invariant w.r.t. $\boldsymbol{R}_{ss}$. This in turn implies that transmit correlation is irrelevant for the analysis of receive antenna selection and for algorithm design.

A closed-form statistical description of the optimal solution is quite difficult and in fact provides no further intuition. Instead, we propose below a suboptimal selection algorithm. The performance of this suboptimal selection policy lower bounds the optimal performance. The main feature of this algorithm is that it allows a tractable statistical analysis and helps prove the main theorem. We make no claims regarding the practicality of this algorithm. In a later section, we provide more practical selection algorithms.

The suboptimal selection rule that computes the $N \times N$ block $\underline{\boldsymbol{U}}$ of $\boldsymbol{U}$ is specified in Algorithm 1.

```
ALGORITHM I INCREMENTAL LOSS MINIMIZATION
Set U_ = U_{r1}, where r_1 = arg max_{1≤l≤M} ||U_l||^2.
  for n = 1 to (N − 1)
    compute B⊥(U_) a N × n orthogonal projector onto the row null-space of U_;
    update U_ := [U_^T, U_{r_n}^T]^T, where r_n = arg max_{1≤l≤M} ||U_l B⊥(U_)||;
  end
```

The core idea behind this approach is to minimize $L_r(\boldsymbol{H})$ incrementally w.r.t. the entries of r. Specifically, the r_nth receive antenna is appended at the nth stage to the previously selected set $\{r_l\}_{1 \le l < n}$ to maximize the determinant of $(\boldsymbol{U}_{r_{1:n}}^H \boldsymbol{U}_{r_{1:n}})$. The algorithm is initialized with the maximum norm row of $\boldsymbol{U}$. In the following statement of the algorithm, $\boldsymbol{B}_\perp(\underline{\boldsymbol{U}})$ denotes any arbitrary orthonormal matrix whose column space coincides with the orthogonal complement to the row space of matrix $\underline{\boldsymbol{U}}$.

The update in Algorithm I becomes clear from the determinant decomposition (which follows from QR factorization):

$$\det(\underline{\boldsymbol{U}}^H \underline{\boldsymbol{U}}) = \|\underline{\boldsymbol{U}}_1\|^2 \prod_{k=1}^{N-1} \left\| \underline{\boldsymbol{U}}_{k+1} \boldsymbol{B}_\perp(\underline{\boldsymbol{U}}_{1:k}) \right\|^2. \qquad (11)$$

Clearly, Algorithm I remains suboptimal, primarily because of the approximation in (10) but also because of incremental rather than simultaneous maximization w.r.t. the entries of r. Incremental selection is attractive computationally as an alternative to the search over all $\binom{M}{N}$ candidate sets r for moderately big N and/or M. The most appealing feature of this algorithm, however, is that it allows a tractable statistical decomposition of the associated loss factor. Note that

$$L_*(\boldsymbol{H}) \triangleq -\log_2 \det(\underline{U}^H \underline{U})$$

where $L_*(\boldsymbol{H})$ is the loss associated with the subset selected by Algorithm I. It may be decomposed as follows.

Theorem 1: Under the assumptions of lemma 1 and the selection rule specified in Algorithm I

$$L_*(\boldsymbol{H}) = -\sum_{k=1}^{N} \log_2 \beta_k^2$$
$$\beta_k^2 = \left\| \underline{U}_{k+1} \boldsymbol{B}_\perp(\underline{U}_{1:k}) \right\|^2, \quad 1 \le k \le N \qquad (12)$$

where $\{\beta_k^2\}_{1 \le k \le N}$ are statistically independent, and each β_k^2 is distributed as a maximum of squared row norms of a random $(M-k+1) \times (N-k+1)$ unitary matrix uniformly distributed w.r.t. Haar measure.

Proof: See [18]. ■

The result of Theorem 1 brings valuable insights into the performance of antenna selection. In fact, it allows us to extend the fundamental results regarding the diversity order of SIMO selection to the MIMO (receive) selection scenario. We analyze the diversity of MIMO selection in the following section.

IV. Diversity of MIMO Antenna Selection

Contrary to the SIMO scenario discussed in Section II-B, a precise meaning and quantification of diversity in MIMO systems depends to a great extent on the adopted analysis criteria. The criteria commonly used in the literature on MIMO systems fall into two main categories. On one hand, performance of MIMO systems is analyzed in terms of bit or block error rates, assuming some underlying space-time coding scheme. Solid foundations for this approach have been provided in the landmark paper by Tarokh *et al.* [2]. The effective diversity orders based on the outage error rate have been analyzed [19]. The analysis based on error rate performance implies, under some favorable assumptions about the channel statistics, that diversity order (MN) is achievable. A recent work by Zheng and Tse [20] offers a more general framework to study the diversity/multiplexing tradeoffs.

Another direction is provided through Foschini's celebrated work [5] and a paper by Varanasi *et al.* [21] that interprets MIMO capacity as the capacity of a set of parallel SIMO

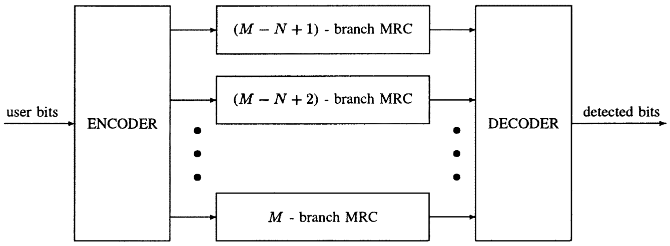

Fig. 2. MIMO capacity: Lower bound scheme.

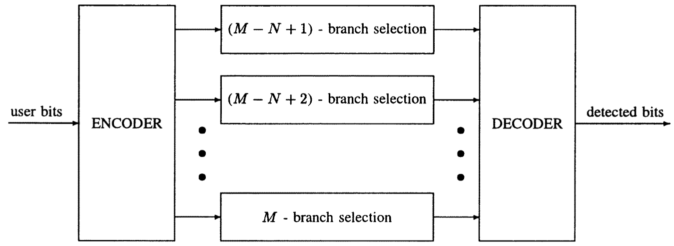

Fig. 3. MIMO capacity with receive selection: lower bound scheme.

systems with each SIMO channel characterized by a certain receive diversity order. Since our focus here is MIMO capacity, we will pursue the line of thought initiated by Foschini to interpret the effect of MIMO selection. First, we will revisit the main results by Foschini with an emphasis on diversity orders. Next, we will present the new results for the MIMO selection.

A. MIMO Diversity: Standard Results

As presented in [21], the capacity of a $M \times N$ channel $\boldsymbol{H}$ is

$$C(\boldsymbol{H}) = \sum_{k=1}^{N} \log_2 \left(1 + \left(\frac{E_s}{N_0}\right) \boldsymbol{H}_{:,k}^H \times \left(\boldsymbol{I}_M + \left(\frac{E_s}{N_0}\right) \boldsymbol{H}_{:,k+1:N} \boldsymbol{H}_{:,k+1:N}^H\right)^{-1} \boldsymbol{H}_{:,k}\right) \quad (13)$$

and is achieved with a combination of the successive minimum mean square (MMSE) cancellation of interfering signals and decision-based removal of the previously detected signals. This elegant result, however, does not allow a simple statistical description, mainly because of statistical dependence between the terms in the sum (13). A tractable statistical characterization is obtained from (13) by neglecting $\boldsymbol{I}_M$, assuming high (E_s/N_0) and/or M (see [5]). In fact, this bound corresponds to replacing the optimal MMSE cancellation with zero forcing (ZF) cancellation. More importantly, this approximation transforms the additive terms in (13) into statistically independent variables

$$C(\boldsymbol{H}) > \sum_{k=1}^{N} \log_2 \left(1 + \left(\frac{E_s}{N_0}\right) \gamma_{M-k+1}^2\right) \quad (14)$$

where γ_n^2 has *Gamma distribution* with n degrees of freedom with c.d.f. $\mathbb{P}\{\gamma_n^2 < x\} = \Gamma_i(x; n)$; see [5] and [6]. Comparing each term in the sum (14) with the MRC capacity given by (3) and the corresponding distribution described in the left-hand equation in (5), we see that the MIMO capacity $C(\boldsymbol{H})$ may be lower bounded by the capacity of N parallel SIMO channels with the respective diversity orders ranging from $(M - N + 1)$ up to M. The block diagram of this equivalent channel is shown in Fig. 2. This interpretation of MIMO capacity yields a natural diversity measure that consists of a vector of diversity orders $(M - N + 1)$ through M. Note that this definition may be extended to the case of limited transmit diversity, i.e., when the matrix of transmit correlations is rank deficient. Similar analysis applies in these cases, with the number of transmit antennas being replaced by the rank of the matrix of transmit correlations [22], [23].

B. MIMO Diversity: Receive Selection

In this section, we will show that selecting N out of M receive antennas using Algorithm I yields, at high (E_s/N_0) and/or big M, an equivalent channel that is similar to Fig. 2 in the sense that the capacity is lower bounded by N independent SIMO channels. The main difference is that instead of MRC, each SIMO link undergoes selection combining at the receiver. This "lower bound" equivalent channel is depicted by a block diagram in Fig. 3.

Before proceeding with the derivation, it is worthwhile to emphasize that this lower bound implies that selection achieves the same diversity as that of the $M \times N$ MIMO system. This conclusion follows from the following arguments. First, the equivalent channel in Fig. 2 provides an upper bound on the optimal MIMO selection since all M receive antennas are used. Second, SIMO selection and SIMO MRC yield the same diversity order, as shown in Section II-B. Hence, the respective branches of Figs. 2 and 3 have the same diversity order. Therefore, the lower bound Fig. 3 and the upper bound Fig. 2 are equivalent in terms of diversity order.

We now establish the lower bound Fig. 3 on the capacity of MIMO with receive antennas selected according to Algorithm I. Substituting (12) and (14) into (10), we may write

$$C_r(\boldsymbol{H}) \geq \sum_{k=1}^{N} \log_2 \left(\left(1 + \left(\frac{E_s}{N_0}\right) \gamma_{M-k+1}^2\right) \beta_k^2\right) \rightarrow \sum_{k=1}^{N} \log_2 \left(1 + \left(\frac{E_s}{N_0}\right) \gamma_{M-k+1}^2 \beta_k^2\right) \quad (15)$$

wherein the limit is achieved at high (E_s/N_0) or/and M. It is worth reiterating the observation, made in the previous section, that the approximation (15) is accurate even at moderate SNR when $M > N$.

Let us focus now on the right-hand side of (15). According to Theorem 1, the scalars β_k^2 are maximum row norms of a set of statistically independent $(M-k+1) \times (N-k+1)$ unitary matrix uniformly distributed w.r.t. Haar measure. These scalars are defined as $\beta_1^2 \triangleq \|\underline{\boldsymbol{U}}_1\|^2$ (the maximum row norm of $\boldsymbol{U}$), and $\beta_k^2 \triangleq \|\underline{\boldsymbol{U}}_{k+1} \boldsymbol{B}_\perp(\underline{\boldsymbol{U}}_{1:k})\|^2$ (the maximum row norm of $\boldsymbol{U}\boldsymbol{B}_\perp(\underline{\boldsymbol{U}}_{1:k})$), wherein the unitary matrices $\boldsymbol{U}\boldsymbol{B}_\perp(\underline{\boldsymbol{U}}_{1:k})$, $1 \leq k \leq N$ are derived from $\boldsymbol{U}$ according to Algorithm I.

For the purpose of diversity analysis, we replace β_k^2 with $\underline{\beta}_k^2$, where $\underline{\beta}_k^2$ is defined as the maximum absolute value of entries of a given column of the respective unitary matrix rather than its maximum row norm. Without loss of generality, we choose the first column and define

$$\underline{\beta}_1^2 \triangleq \max_l |\boldsymbol{U}_{l,1}|^2, \quad \underline{\beta}_k^2 \triangleq \max_l \left|(\boldsymbol{U}\boldsymbol{B}_\perp(\underline{\boldsymbol{U}}_{1:k}))_{l,1}\right|^2.$$

It is easy to see that, according to the latter definition, $\underline{\beta}_k^2 \leq \beta_k^2$, $1 \leq k \leq N$. This lower bound on β_k^2 induces interesting statistical properties on the capacity with antenna selection. At high (E_s/N_0) or/and M, (15) yields

$$C_r(\boldsymbol{H}) > \sum_{k=1}^{N} \log_2\left(1 + \left(\frac{E_s}{N_0}\right)\eta_k^2\right)$$
$$\eta_k^2 \triangleq \gamma_{M-k+1}^2 \underline{\beta}_k^2, \quad 1 \leq k \leq N \tag{16}$$

wherein the approximate relationship becomes a lower bound at high (E_s/N_0) or/and big M.

Lemma 3: The scalars η_k^2 are statistically independent with c.d.f.

$$\mathbb{P}\left\{\eta_k^2 < x\right\} = (1 - \exp(-x))^{M-k+1}, \quad 1 \leq k \leq N. \tag{17}$$

Proof: See the Appendix. ■

According to (16) and Lemma 3, η_k^2 is statistically equivalent to the maximum squared absolute value of a vector of $(M-k+1)$ i.i.d. zero mean unit variance complex circular Gaussian variables. In other words, η_k^2 is equal to the gain of SIMO selection with $(M-k+1)$ antennas. This observation, together with (16), validates the diagram in Fig. 3 as a lower bound for the MIMO selection capacity achieved with Algorithm I at high (E_s/N_0) or/and big M. As explained earlier in this section, this lower bound, along with the upper bound Fig. 2, implies that Algorithm I, which is applied to select N out of M receive antennas, yields the same diversity order as the optimal use of the entire $M \times N$ system.

Remark: We reiterate that the main attraction of the incremental selection algorithm is that it enables a tractable statistical analysis, and we make no claims regarding the practicality of this algorithm. Further, we note that the incremental loss minimization algorithm works only for square configurations, i.e., when the number of receive antennas selected is equal to the number of transmit antennas. It does not apply when $m > N$. However, the above arguments prove that selecting N out of M receive antennas is sufficient to provide the diversity performance of the full $M \times N$ system. We point out here that similar to the SIMO case, even though the diversity order achievable through selection is equal to that of the full system, there is a definite capacity loss since all available antennas are not used. This loss is best described as an SNR gap (or array loss). The nature of this loss is discussed again in the simulations section. This fixed loss decreases with the use of $m > N$ antennas. In the next section, we present practical algorithms that may be used to select any number of receive antennas.

V. Practical Algorithms for Antenna Selection

In this section, we develop reduced complexity algorithms to select M' out of M receive antennas.

A. General Case of Receive Antenna Selection

It is easy to see that Algorithm I for receive antenna selection is not applicable when the number M' of selected receive antennas is larger than the number of the transmit antennas N. Furthermore, as explained in Section III, this algorithm attempts to minimize a loss factor that is an approximate performance measure. Admittedly, the main value of Algorithm I is that it enables a tractable statistical characterization, particularly in terms of diversity achievable with antenna selection. Note that the main result of Section IV is that full diversity is achieved with $M' = N$. This observation suggests that using more than N receive chains would not bring a substantial improvement in outage capacity, at least in the region of small outage rates.[2] Yet, it may be desirable to have a simple selection rule that applies in the general case of $M' < M$.

Optimal selection requires maximization of (7) over all $\binom{M}{M'}$ subsets r requiring around $\binom{M}{M'}M'^3$ complex additions/multiplications. This number may be too large, particularly for big N and/or M. In the rest of this section, we will describe two simplified approaches. The first approach suggests *incremental selection*. Similarly to Algorithm I, the set r is built up incrementally. Here, we will consider the exact capacity (7) rather than the loss factor from (10).

```
ALGORITHM II INCREMENTAL SELECTION
Set A := (E_s/N_0) I_N and
  r_1 := arg max_{1≤l≤M} ||H_l||^2.
    for n = 1 to (M' − 1)
    update A := A − A H_{r_n}^H (1 + H_{r_n} A H_{r_n}^H)^{-1} H_{r_n} A;
    compute r_{n+1} = arg max_{l∉{r_1,...,r_n}} H_l A H_l^H;
  end
```

```
ALGORITHM III DECREMENTAL SELECTION
Set A := ((E_s/N_0)^{-1} I_N + H^H H)^{-1}, p =
 arg min_{1≤l≤M} H_l A H_l^H and r := {1,...,p − 1, p +
 1,...M}.
  for n = 1 to (M − M' − 1)
    update A := A + A H_p^H (1 − H_p A H_p^H)^{-1} H_p A;
    compute p = arg min_{l∈r} H_l A H_l^H,  r = p \ r;
  end
```

Let us first analyze incremental selection. For the sake of simplicity, we assume symmetric transmit antennas with no transmit correlations,[3] i.e., $\boldsymbol{R}_{ss} = (1/N)\boldsymbol{I}_N$. Suppose that after n steps, the antennas indexed with $\{r_1, \ldots, r_n\}$ have been selected and that $\boldsymbol{H}_{r_1,\ldots,r_n}$ is the $n \times N$ block of $\boldsymbol{H}$. Appending the lth antenna yields

$$C(\boldsymbol{H}_{r_1,\ldots,r_n}, \boldsymbol{H}_l) = \log_2 \det\left(\boldsymbol{I}_N\left(\frac{E_s}{N_0}\right)\right.$$
$$\left.\times\left(\boldsymbol{H}_{r_1,\ldots,r_n}^H \boldsymbol{H}_{r_1,\ldots,r_n} + \boldsymbol{H}_l^H \boldsymbol{H}_l\right)\right) \tag{18}$$

[2]Simulation results are presented in Section VII to support this observation.

[3]The case of transmit correlations may be handled by setting $\boldsymbol{H} := \boldsymbol{H}\boldsymbol{R}_{ss}^{1/2}$.

which we may rewrite as

$$C(\boldsymbol{H}_{r_1,\ldots,r_n},\boldsymbol{H}_l)=\log_2\det\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H_{r_1,\ldots,r_n}\boldsymbol{H}_{r_1,\ldots,r_n}\right)$$
$$+\log_2\left(1+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}_l\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H_{r_1,\ldots,r_n}\boldsymbol{H}_{r_1,\ldots,r_n}\right)^{-1}\boldsymbol{H}^H_l\right)$$

According to the last expression, maximization of $C(\boldsymbol{H}_{r_1,\ldots,r_n},\boldsymbol{H}_l)$ w.r.t. l given $r_1,\ldots,r_n$ yields

$$r_{n+1}=\arg\max_{l\notin\{r_1,\ldots,r_n\}} \boldsymbol{H}_l\left(\left(\frac{E_s}{N_0}\right)^{-1}\boldsymbol{I}_N+\boldsymbol{H}^H_{r_1,\ldots,r_n}\boldsymbol{H}_{r_1,\ldots,r_n}\right)^{-1}\boldsymbol{H}^H_l. \quad (19)$$

One can see that successive application of the selection rule (19) requires a series of matrix inverses. These inverses may be efficiently computed through a recursive update based on the matrix inversion lemma [24]. Denote $\boldsymbol{A}$ as a $n\times n$ positive definite matrix and $n\times 1$ as a vector $\boldsymbol{a}$. Then

$$(\boldsymbol{A}+\boldsymbol{a}\boldsymbol{a}^H)^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{a}(1+\boldsymbol{a}^H\boldsymbol{A}^{-1}\boldsymbol{a})^{-1}\boldsymbol{a}^H\boldsymbol{A}^{-1}. \quad (20)$$

We make use of the selection rule (19) and the update (20) to define Algorithm II. One can check that the complexity of Algorithm II is upper bounded by $(MM'N^3)$ complex additions/multiplications.

This algorithm is numerically attractive when M is large with relatively small M'. Another possible scenario is large M and M' with small $(M-M')$. In this case, *decremental selection* may be applied. The idea of decremental selection is to start with the whole set of M antennas and to successively remove $(M-M')$ antennas so that at every stage, one antenna is removed, which yields minimum reduction of the capacity. A numerically efficient implementation of decremental selection is similar to Algorithm II. Indeed, the effect of removing the lth antenna from a given set $\{r_1,\ldots,r_n\}$ amounts to replacing the term $(\boldsymbol{H}^H_l\boldsymbol{H}_l)$ in (18) by $(-\boldsymbol{H}^H_l\boldsymbol{H}_l)$. Check that this change implies replacing the maximization in (18) by minimization over the set $\{r_1,\ldots,r_n\}$ and inverting $(+/-)$ signs in (20). These observations lead us to Algorithm III. Decremental selection has been introduced in [12], wherein it is also extended to the case frequency selective channels.

Remark: Note that because of the initial matrix inversion step, Algorithm III is slightly more complex than Algorithm II when the two are compared for the same number of iterations (selection steps). Conversely, Algorithm III is expected to perform better than Algorithm II. This is due to the fact that the antenna removal rule takes into account join contributions of all (remaining) antennas, whereas the incremental selection is based on individual contributions of the appended antennas. In addition, note that Algorithm III is strictly optimal when $M=M'+1$. Check that the complexity of Algorithm III is upper bounded by $M(M-M')N^3$ complex additions/multiplications.

VI. Practical Receivers

The diversity equivalence result states that with optimal MIMO signaling and with optimal selection, the diversity performance of the selected antennas is the same as that with all receive antennas. This is a powerful result and provides strong motivation for introducing receive selection in practical MIMO-SM architectures. In this section, we discuss receive selection for two well-known systems. The first one is a MIMO system with joint encoding of the transmitted streams and MMSE combining of these streams at the receiver prior to decoding. This system is later referred to as joint MMSE (JMMSE). The second system is a vertically layered architecture (V-BLAST) with ordered successive interference cancellation (OSIC). First of all, we will derive antenna selection criteria for both systems. Next, we adapt the ideas developed in Section V to come up with reduced complexity decremental selection algorithms.

A. JMMSE Architecture

In the JMMSE architecture, we assume that a sequence of encoded symbols is interleaved over N streams and transmitted by the respective antennas. At the receiver, the streams are extracted using the MMSE receiver, i.e.,

$$\hat{s}[k]=\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H x[k] \quad (21)$$

where $\hat{s}[k]$ is the MMSE estimate of the transmit signal vector. The N extracted streams are then multiplexed in to a single stream, deinterleaved, and decoded. We assume that an infinitely long capacity achieving code is used to encode the data sequence, and an optimal (ML) decoding procedure is employed at the receiver. Furthermore, the decoder deliberately ignores correlation of noise terms at the output of the MMSE combiner.[4] It is easy to show that the capacity $C_J(\boldsymbol{H})$ achievable with the JMMSE receiver is given by

$$C_J(\boldsymbol{H})=\sum_{k=1}^{N}\log_2\left(1+\rho_k^2\right),$$
$$\rho_k^2=\left[\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\right]^{-1}_{k,k}-1$$
$$1\le k\le N \quad (22)$$

where ρ_k^2 is the SINR of the k^{th} output stream of the MMSE. The proof of (22) follows from the standard typical set decoding argument. Specifically, one assumes that a typical set decoder is applied to retrieve N sets of codewords, wherein each set is jointly typical with the respective output of the MMSE filter. The final decoding stage consists of taking the intersection of N typical sets. The capacity $C_J(\boldsymbol{H})$ is an upper bound on the rate that allows for a successful decoding. The details of the proof are omitted.

[4]It is worthwhile to note that a proper use of noise correlation would amount to ML reception with a prohibitive complexity.

The JMMSE architecture essentially reduces the MIMO channel into an equivalent SISO link with achievable throughput, per (22). Clearly, the throughput that can be supported by this scheme is just the outage capacity of the equivalent SISO link.

For receive antenna subset selection, we follow a decremental strategy that is similar to that used in Section V to maximize channel capacity. The computation of $C_J(\boldsymbol{H})$ in (22) requires matrix inversion. A recursive update of $C_J(\boldsymbol{H})$ based on removing the antennas (rows of $\boldsymbol{H}$) may be simplified using (20). The resulting decremental selection rule is captured in Algorithm IV.

It is worthwhile to note that the incremental selection approach introduced in Section V is not suitable to the JMMSE architecture. Indeed, the number of receive antennas has to be at least as large as the number of transmitted streams to have reasonably good SINR values $\rho_1^2, \ldots \rho_N^2$ in (22). This condition is violated in the case of incremental selection for the first $(N-1)$ selection steps.

The benefit of the proposed selection algorithm is similar to the benefit of the algorithm presented in Section V. Specifically, a search over $\binom{M}{M'}$ subsets is replaced by a recursion including $M(M-M')$ steps for the decremental selection.

```
ALGORITHM IV DECREMENTAL SELECTION FOR JMMSE RECEIVER
Set A := ((E_s/N_0)^{-1} I_N + H^H H)^{-1}, r := {1,...,N}.
 for n = 1 to (M - M')
   compute p = arg max_{l∈r} Σ_{k=1}^{N} log_2([A + A H_l^H (1 - H_l A H_l^H)^{-1} H_l A]^{-1}_{k,k});
   update A := A + A H_p^H (1 - H_p A H_p^H)^{-1} H_p A, r := p \ r;
 end
```

B. OSIC Architecture

A vertically layered architecture with OSIC reception has been proposed in [25]. The transmitter forms N parallel data streams that are encoded and modulated independently and then sent over N antennas. At the receiver, a linear filter is used to extract the transmitted data streams successively by peeling out the contribution of the previously decoded streams (layers) from the received mixture. The OSIC principle suggests that the signal with the highest signal to interference and noise ratio (SINR) is extracted at each stage, thereby ensuring the highest worst-case SINR over the whole set of layers; see [26] for more details. We reiterate that the N data layers are encoded independently with ideal FEC codes. The nature of the decoding process and the absence of channel knowledge at the transmitter implies that the order of processing of the transmitted streams at the receiver is unknown to the transmitter. Symmetry arguments suggest equal data rates for all streams. Since the outage rate is linked directly to the statistics of the worst stream, the code rate on any antenna cannot exceed the throughput on the worst layer at the specified outage. The maximum achievable throughput of the OSIC scheme is therefore N times the throughput of the worst layer. To characterize the throughput of the OSIC, we consider optimal linear MMSE filtering at each layer. Define $\pi : \{1, \ldots, N\} \mapsto \{1, \ldots, N\}$ as a permutation that specifies the ordering. According to the OSIC principle, this permutation is defined by the following recursion:

$$\pi[n] = \arg \max_{l \notin \{\pi[1], \ldots, \pi[n-1]\}} \boldsymbol{H}_l^H \left(\boldsymbol{I}_N + \left(\frac{E_s}{N_0}\right) \boldsymbol{H}_{:,\mathcal{I}(n,l)} \times \boldsymbol{H}_{:,\mathcal{I}(n,l)}^H \right)^{-1} \boldsymbol{H}_l, \quad 1 \le n \le N \tag{23}$$

where $\mathcal{I}(n,l)$ is a subset of $\{1, \ldots, N\}$ that excludes the indices $\{l, \pi[1], \ldots, \pi[n-1]\}$. In other words, $\mathcal{I}(n,l)$ is the set of column indices of the channel matrix $\boldsymbol{H}$ that represents interference to the lth transmitted signal at the nth layer (e.g., after peeling out the signals $\pi[1]$ through $\pi[n-1]$). The corresponding SINR γ_n^2 at the nth layer is given by

$$\gamma_n^2 = \left(\frac{E_s}{N_0}\right) \boldsymbol{H}_{\pi[n]}^H \left(\boldsymbol{I}_N + \left(\frac{E_s}{N_0}\right) \boldsymbol{H}_{:,\mathcal{I}(n,\pi[n+1])} \times \boldsymbol{H}_{:,\mathcal{I}(n,\pi[n+1])}^H \right)^{-1} \boldsymbol{H}_{\pi[n]}, \quad 1 \le n \le N. \tag{24}$$

As explained in the previous paragraphs, the overall throughput of the OSIC system is defined as the N-fold throughput of the worst layer:

$$C_O(\boldsymbol{H}) = N \log_2 \left(1 + \min_{1 \le k \le N} \gamma_k^2 \right). \tag{25}$$

The problem of optimal receive antenna selection for OSIC architecture may be now formulated as maximization of (25) over all possible $\binom{M}{M'}$ subsets. We propose to apply the decremental strategy to reduce the complexity of such an exhaustive search. Specifically, $(M-M')$ antennas are removed successively so that the loss in capacity (25) is minimized at each step. The incremental strategy is not applicable here for reasons identical to those presented earlier in the JMMSE context.

Although the complexity of the decremental selection is polynomial in the number of antennas, rather than exponential (as in the case of exhaustive search), it is still rather high. Unfortunately, the very nature of the OSIC prevents any substantial simplification.

C. Discussion

The JMMSE approach is often preferred to OSIC because of lower computational complexity and smaller decoding latency. Surprisingly, the JMMSE scheme may outperform the OSIC system at low and medium SNR levels. To appreciate this fact, we state a simple upper bound on the performance of the OSIC architecture.

Lemma 4: The capacity (22) of OSIC is upper bounded as follows:

$$C_O(\boldsymbol{H}) \le N \log_2 \left(1 + \min_{1 \le k \le N} \left(\frac{E_s}{N_0}\right) \|\boldsymbol{H}_{:,k}\|^2 \right). \tag{26}$$

Proof: The result follows from the fact that the worst-case SINR over all layers may not exceed the worst-case matched filter bound or, in other words, $\min_{1\le k\le N}\gamma_k^2 \le \min_{1\le k\le N}(E_s/N_0)\|\boldsymbol{H}_{:,k}\|^2$. ■

As a matter of fact, the MMSE filter tends to the matched filter at low SNR. The SINR in turn tends to the respective matched filter bound so that the upper bound (26) becomes quite accurate. Furthermore, the SINR values of the MMSE tend to the respective matched filter bounds:

$$\rho_k^2 \to \left(\frac{E_s}{N_0}\right)\|\boldsymbol{H}_{:,k}\|^2, \quad 1\le k\le N \quad \text{as} \quad \left(\frac{E_s}{N_0}\right)\to 0. \tag{27}$$

By observing (22), (26), and (27), we conclude that the JMMSE outperforms the OSIC at low SNR limit. The numerical study in the following section shows that this relationship holds within a rather wide range of SNR.

The advantage of JMMSE reduces with increasing SNR, until eventually, the OSIC outperforms the JMMSE. To see this, we recall that the first layer of OSIC appears to be the worst at high SNR. Whenever this relationship holds, the OSIC capacity in (22) satisfies

$$C_O(\boldsymbol{H}) \approx N\log_2\left(1+\gamma_1^2\right) = N\log_2\left(1+\max_{1\le k\le N}\rho_k^2\right) \tag{28}$$

where the second equality in (28) follows from the OSIC principle. The advantage of OSIC over JMMSE at high SNR follows immediately from (22) and (28).

VII. Numerical Study

We study outage capacities that can be achieved through antenna selection for a number of different antenna configurations (varying N, M' and M).

In the following figures, we study the outage capacity of the two practical selection schemes presented in Section V. In Fig. 4, the outage capacity is plotted versus the total SNR per receive antenna (NE_s/N_0) for $M = 6$ and $N = \{2, 4\}$ for outage rate 10% (upper panel) and 1% (lower panel). The capacities without antenna selection $M = N(-\square-)$ and with the use of all M antennas $(-\star-)$ provide boundaries for the capacity achievable with antenna selection. The results of optimal selection and practical selection rules, namely Algorithms II and III, are plotted by $(-\diamond-)$, $(-\nabla-)$, and $(-\circ-)$, respectively. The algorithms studied in this paper are compared with an *ad hoc* selection $(-\times-)$, which maximizes the total received power; see, e.g., [15]. We draw attention to the fact that in the high SNR region there is a *fixed* gap in outage capacity between the system with all receive antennas and the system with antenna selection. We also note that the proposed selection algorithms exhibit quasioptimal performance, contary to the *ad hoc* selection, which yields up to 20% capacity loss.

Remark: The fixed nature of the gap shows the array loss incurred due to the use of fewer antennas (N selected) than those available (M). However, the fixed nature of this loss also points to the fact that the outage performance of the two systems is similar from a diversity perspective.

Fig. 5 shows outage capacities versus the total number M of receive antennas. Finally, Fig. 6 represents outage capacity versus the number of selected antennas (M'), with the total

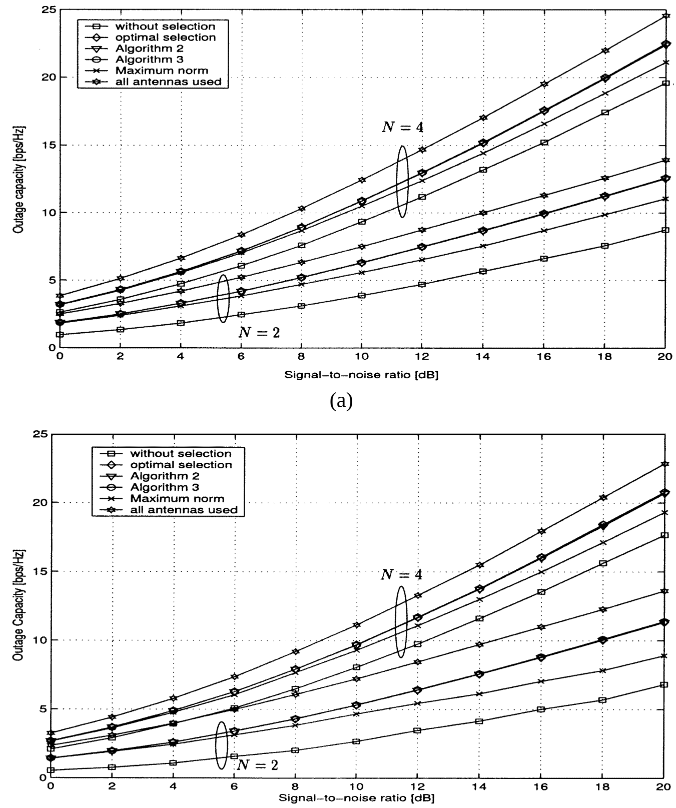

Fig. 4. Outage capacity of antenna selection versus SNR(NE_s/N_0), $M = 6$. (a) 10% outage rate. (b) 1% outage rate.

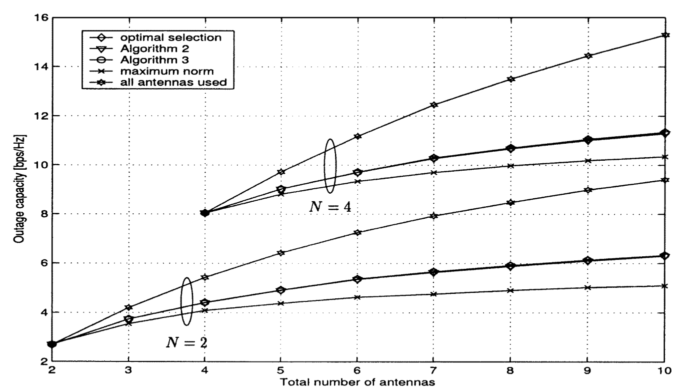

Fig. 5. Outage capacity of antenna selection (1% outage rate) versus M $(NE_s/N_0) = 10\ dB$.

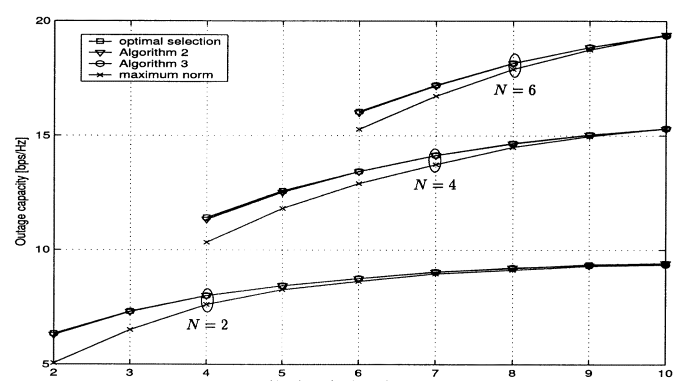

Fig. 6. Outage capacity of antenna selection (1% outage rate) versus M', $M = 10$ $(NE_s/N_0) = 10\ dB$.

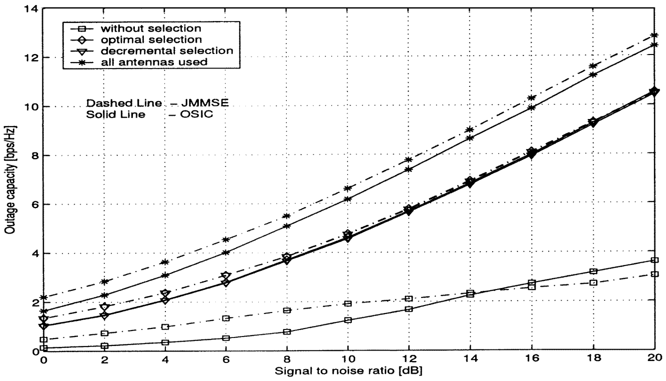

Fig. 7. Performance of receive selection algorithms for OSIC and JMMSE, $N = 2$, and $M = 6$. 1% outage capacity versus SNR. Decremental selection is near optimal. JMMSE outperforms OSIC at low and medium SNR.

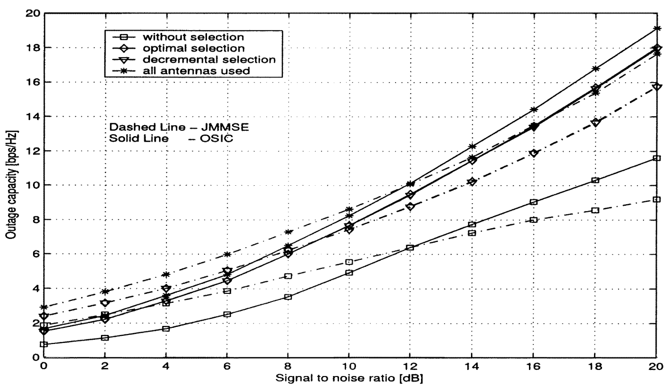

Fig. 8. Performance of receive selection algorithms for OSIC and JMMSE, $N = 4$, and $M = 6$. 1% outage capacity versus SNR. Decremental selection is near optimal. JMMSE outperforms OSIC at low and medium SNR.

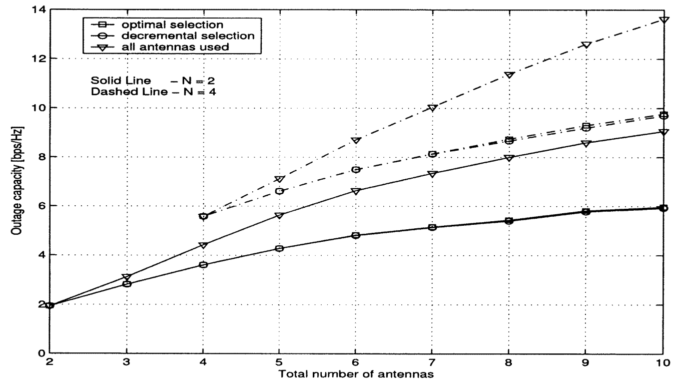

Fig. 9. Performance of receive selection algorithms for JMMSE, $N = [2, 4]$, and $NE_s/N_0 = 10\ dB$. 1% outage capacity versus M.

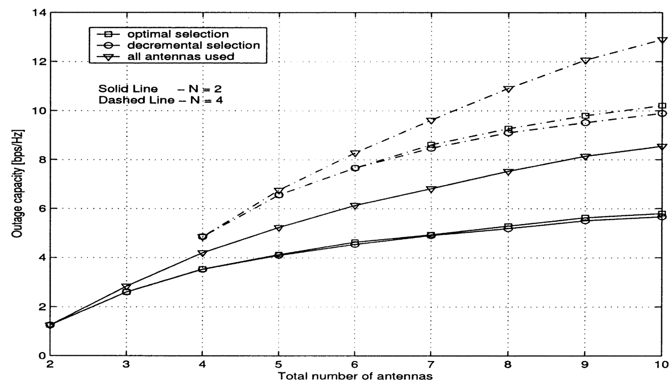

Fig. 10. Performance of receive selection algorithms for OSIC, $N = [2, 4]$, and $NE_s/N_0 = 10\ dB$. 1% outage capacity versus M.

number of antennas fixed $M = 10$ and for different number of transmit antennas. Note that the gap between the optimal selection and simplified algorithms remains small in all cases.[5] The latter figure suggests that increasing the number M' of selected antennas above the number N of transmit antennas yields a modest gain in outage capacity. In the case $N = 6$, a gain less than 20% is achieved with $M' = 10$ when the number of antennas is almost doubled. This observation agrees with the main conclusion of this paper, namely, that an appropriate selection of N antennas achieves full diversity.

Remark: An indication of possible complexity reduction may be obtained using the upper bounds on the respective complexities of the optimal selection and Algorithms II and III. Check that for $N = M' = 2$, $M = 10$ using Algorithm II yields at least a two-fold reduction, whereas $N = M' = 4$, $M = 10$ yields more than a five-fold reduction. Similarly, for $N = M' = 6$, $M = 10$ using Algorithm III yields at least a five-fold reduction. The reduction factor grows exponentially along with M; this factor is especially important when either M' or $(M - M')$ is small. To reiterate, incremental selection is attractive when M' is small. Decremental selection is motivated when $M - M'$ is small.

Next, we study the performance of receive selection algorithms in JMMSE and V-BLAST systems. Figs. 7 and 8 depict the 1% outage capacity versus the total receive SNR per receive antenna (NE_s/N_0) for $M = 6$ and $N = \{2, 4\}$, respectively. The curves corresponding to the case of no selection $(M = N)$ and the case when all antennas are used are also depicted and serve as boundaries. The performance of optimal selection and decremental selection algorithms is also presented. Note that the decremental selection algorithm works remarkably well for both JMMSE and OSIC systems. With selection, there is almost 100% improvement in outage capacity for the case of $N = 2$, $M' = 2$, $M = 6$, and $NE_s/N_0 = 10$ dB when compared with the system with no selection capability. The improvement is about 50% for the case of $N = 4$, $M' = 4$, $M = 6$, and $NE_s/N_0 = 10$ dB.

Furthermore, one can see that there is a definite crossover in the performance curves for JMMSE and OSIC systems. The reason for this effect has been explained previously in Section VI-C. The region to the left of the crossover point corresponds to the SNR region for which JMMSE is superior to OSIC in performance as well as computational complexity.

Figs. 9 and 10 depict the 1% outage capacity versus total available receive antennas (M) for JMMSE and OSIC systems, respectively, with $NE_s/N_0 = 10$ dB. We depict the curves for $N = 2$ and $M = 2$ through 10 with two antennas selected for reception and for $N = 4$ and $M = 4$ through 10 with four antennas selected for reception. The selection algorithms are quite accurate, even when the number M of antennas available for selection is very large.

[5]A more critical comparison of incremental versus decremental selection reveals a microscopic advantage of the latter one.

VIII. SUMMARY

In this paper, we provide a comprehensive study of antenna selection algorithms in MIMO-SM systems. We show that the diversity order achievable through receive antenna subset selection is the same as that with all receive antennas in place provided that the antennas are selected "sensibly." This powerful result provides us with strong motivation for introducing antenna selection schemes in practical MIMO-SM systems. Toward this end, we developed decremental selection algorithms for two popular MIMO-SM architectures. Simulations demonstrate that these algorithms have near-optimal performance with considerable reduction in computational complexity.

APPENDIX

A. Proof of Lemma 1

Denote $\underline{\boldsymbol{H}}_r$ as a $(M-N)\times N$ block of $\boldsymbol{H}$ that consists of the rows of $\boldsymbol{H}$ that are not in r.

$$\begin{aligned}
C_r(\boldsymbol{H}) &= \log_2\det\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}-\left(\frac{E_s}{N_0}\right)\underline{\boldsymbol{H}}_r^H\underline{\boldsymbol{H}}_r\right)\\
&\overset{(*)}{=} \log_2\det\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)\\
&\quad+\log_2\det\left(\boldsymbol{I}_N-\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)^{-\frac{1}{2}}\right.\\
&\quad\quad\left.\times\left(\frac{E_s}{N_0}\right)\underline{\boldsymbol{H}}_r^H\underline{\boldsymbol{H}}_r\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)^{-\frac{1}{2}}\right)\\
&= \log_2\det\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)\\
&\quad+\log_2\det\left(\boldsymbol{I}_{M-N}-\left(\frac{E_s}{N_0}\right)\underline{\boldsymbol{H}}_r\right.\\
&\quad\quad\left.\times\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\underline{\boldsymbol{H}}_r^H\right)\\
&\geq \log_2\det\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)\\
&\quad\quad+\log_2\det\left(\boldsymbol{I}_{M-N}-\underline{\boldsymbol{H}}_r(\boldsymbol{H}^H\boldsymbol{H})^{-1}\underline{\boldsymbol{H}}_r^H\right)\\
&= \log_2\det\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)\\
&\quad+\log_2\det\left(\boldsymbol{U}_r\boldsymbol{U}_r^H\right)
\end{aligned}\tag{29}$$

wherein the equality $(*)$ follows from

$$\begin{aligned}
&\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}-\left(\frac{E_s}{N_0}\right)\underline{\boldsymbol{H}}_r^H\underline{\boldsymbol{H}}_r\\
&\quad=\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)\\
&\quad\quad\cdot\left(\boldsymbol{I}_N-\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)^{-\frac{1}{2}}\left(\frac{E_s}{N_0}\right)\underline{\boldsymbol{H}}_r^H\underline{\boldsymbol{H}}_r\right.\\
&\quad\quad\quad\left.\times\left(\boldsymbol{I}_N+\left(\frac{E_s}{N_0}\right)\boldsymbol{H}^H\boldsymbol{H}\right)^{-\frac{1}{2}}\right).
\end{aligned}$$

Under the conditions of the lemma, all eigenvalues of $(M^{-1}\boldsymbol{H}^H\boldsymbol{H})$ are uniformly bounded from 0 w.p.l. Hence, $(E_s/N_0)(\boldsymbol{I}_N+(E_s/N_0)\boldsymbol{H}^H\boldsymbol{H})^{-1}\rightarrow(\boldsymbol{H}^H\boldsymbol{H})^{-1}$ as $(E_s/N_0)\rightarrow\infty$. Similarly, $(E_s/N_0)(\boldsymbol{I}_N+(E_s/N_0\boldsymbol{H}^H\boldsymbol{H})^{-1}=(E_s/N_0)(\boldsymbol{I}_N+M(E_s/N_0)(M^{-1}\boldsymbol{H}^H\boldsymbol{H}))^{-1}\rightarrow(\boldsymbol{H}^H\boldsymbol{H})^{-1}$ as $M\rightarrow\infty$.

B. Proof of Lemma 3

Recall that η_k^2 is a product of γ_{M-k+1}^2, having Gamma distribution with $(M-k+1)$ degrees of freedom and $\underline{\beta}_k^2$ defined as the maximum squared absolute entry of a given column of a $(M-k+1)\times(N-k+1)$ unitary matrix uniformly distributed w.r.t. the Haar measure. Note that this column is uniformly distributed on the complex unit sphere. Let $\xi=[\xi_1,\ldots\xi_{M-k+1}]$ be a complex circular Gaussian i.i.d. vector: $\xi\sim\mathcal{N}_c(0,\boldsymbol{I}_{M-k+1})$. According to (5), the distribution of $\max_{1\leq l\leq M-k+1}|\xi_l|^2$ is characterized by the c.d.f. specified in (17). Hence, it remains to be shown that η_k^2 has the same distribution as $\max_{1\leq l\leq M-k+1}|\xi_l|^2$. To this end, we rewrite the latter quantity as

$$\max_{1\leq l\leq M-k+1}|\xi_l|^2=\|\xi\|^2\max_{1\leq l\leq M-k+1}\left|\left(\frac{\xi_l}{\|\xi\|}\right)\right|^2\tag{30}$$

wherein $\|\xi\|^2$ has Gamma distribution with $(M-k+1)$ degrees of freedom. The second term on the right-hand side of (30) is a maximum squared absolute entry of $(\xi/\|\xi\|)$, which is an $(M-k+1)$-dimensional vector. By the properties of $\mathcal{N}_c(0,\boldsymbol{I}_{M-k+1})$, the distribution of ξ and, therefore, that of $(\xi/\|\xi\|)$ is invariant w.r.t. a unitary matrix transform. Additionally, $(\xi/\|\xi\|)$ has a unit norm. Hence, $(\xi/\|\xi\|)$ is uniformly distributed on the complex unit sphere, whose distribution is independent of $\|\xi\|^2$. The proof is completed by noting that $\max_{1\leq l\leq M-k+1}|\xi_l|^2$ may be represented as a product of two statistically independent variables, namely, $\|\xi\|^2$ and $\max_{1\leq l\leq M-k+1}|(\xi_l/\|\xi\|)|^2$, whose respective distributions coincide with those of γ_{M-k+1}^2 and $\underline{\beta}_k^2$.

REFERENCES

[1] N. Seshadri and J. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Int. J. Wireless Inform. Networks*, vol. 1, pp. 49–60, Jan. 1994.

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

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

[4] A. Paulraj and T. Kailath, "Increasing Capacity in wireless broadcast systems using distributed transmission/directional reception," U.S. Patent 5 345 599, 1994.

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

[6] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[7] I. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Tel.*, vol. 10, no. 6, pp. 585–595, Nov/Dec. 1999.

[8] M. Win and J. Winters, "Virtual branch analysis of symbol error probability for hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 49, pp. 1926–1934, Nov. 2001.

[9] D. A. Gore, R. U. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. IEEE ICASSP*, vol. 5, Istanbul, Turkey, May 2000, pp. 2785–2788.

[10] A. Molisch, M. Win, and J. Winters, "Capacity of MIMO systems with antenna selection," in *Proc. IEEE Int. Contr. Conf.*, vol. 2, Helsinki, Finland, June 2001, pp. 570–574.

[11] R. Heath and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE Int. Contr. Conf.*, vol. 7, Helsinki, Finland, June 2001, pp. 2276–2280.

[12] A. Gorokhov, "Antenna selection algorithms for MEA transmision systems," in *Proc. IEEE ICASSP*, vol. 3, Orlando, FL, May 2002, pp. 2857–2860.

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

[14] D. Gore, R. Heath, and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, pp. 491–493, Nov. 2002.

[15] A. Grhayeb and T. Duman, "Performance analysis of MIMO systems with antenna selection over quasistatic fading channels," in *Proc. IEEE ISIT*, Lausanne, Switzerland, June 2002, p. 333.

[16] T. Eng, N. Kong, and L. Milstein, "Comparison of diversity combining techiques for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 44, pp. 1117–1129, Sept. 1996.

[17] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.

[18] A. Gorokhov, D. Gore, and A. Paulraj, "Performance of antenna subset selection in MIMO systems," *IEEE Trans. Inform. Theory*, [Online]. Available http://www.stanford.edu/dagore/IT03.ps, submitted for publication.

[19] A. Gorokhov, "Outage error probability for space-time codes over Rayleigh fading," in *Proc. IEEE ISIT*, Washington, DC, June 2001, p. 242.

[20] L. Zheng and D. Tse, "Optimal diversity-multiplexing trade-off in multi-antenna channels," in *Proc. Allerton Conf. Commun., Contr., Comput.*, Monticello, IL, Oct. 2001.

[21] M. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple access channel," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 2, Pacific Grove, CA, Nov. 1997, pp. 1405–1409.

[22] D. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[23] A. Gorokhov, "Capacity of multi-antenna Rayleigh channel with a limited transmit diversity," in *Proc. IEEE ISIT*, Sorrento, Italy, June 2000, p. 411.

[24] *Linear Systems*. Englewood Cliffs, NJ, 1980.

[25] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI ISSSE*, Sept. 1998, pp. 295–300.

[26] G. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.

Alexei Gorokhov received the Ph.D. degree in electrical engineering in 1997 from the École Nationale Supérieure des Télécommunications (Télécom Paris), Paris, France.

Since October 1997, he has been with the Centre National de la Recherche Scientifique (CNRS), namely, with the Laboratoire des Signaux et Systèmes at École Supérieure d'Électricité (Supélec), Paris, where he holds a research position. Since January 2000, he has been Senior Scientist with Philips Research Labs, Eindhoven, The Netherlands, on long-term leave from CNRS. His research interests cover various areas of digital communications, estimation theory, and spectral analysis. His present research at Philips is focused on theory, design, and implementation of multielement antenna (MIMO) systems for broadband wireless communications.

Dhananjay A. Gore received the B.Tech degree from the Indian Institute of Technology, Bombay (IITB), India, in 1998 and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 2000 and 2003, respectively, all in electrical engineering.

Since May 2003, he has been with Qualcomm Inc., San Diego, CA. His research interests are in signal processing for high-speed wireless communications, smart antennas, and MIMO.

Dr. Gore is a recipient of the Institute Silver Medal from IITB and the School of Engineering Fellowship from Stanford University.

Arogyaswami J. Paulraj (F'91) received the Ph.D. degree from the Naval Engineering College and the Indian Institute of Technology, Delhi, India.

He is a Professor with the Department of Electrical Engineering, Stanford University, where he supervises the Smart Antennas Research Group, working on applications of space-time techniques for wireless communications. His nonacademic positions included Head, Sonar Division, Naval Oceanographic Laboratory, Cochin, India; Director, Center for Artificial Intelligence and Robotics, Bangalore, India; Director, Center for Development of Advanced Computing, Pune, India; Chief Scientist, Bharat Electronics, Bangalore, India; Chief Technical Officer and Founder, Iospan Wireless Inc., Palo Alto, CA. His research has spanned several disciplines, emphasizing estimation theory, sensor signal processing, parallel computer architectures/algorithms and space-time wireless communications. His engineering experience included development of sonar systems, massively parallel computers, and more recently broadband wireless systems.

Dr. Paulraj has won several awards for his engineering and research contributions. He is the author of over 280 research papers and holds eight patents. He is a Member of the Indian National Academy of Engineering.