

Spatial Multiplexing Architectures with Jointly Designed Rate-Tailoring and Ordered BLAST Decoding Part I: Diversity-Multiplexing Tradeoff Analysis

Yi Jiang Mahesh K. Varanasi

Abstract— The V-BLAST (vertical Bell Labs layered Space-Time) architecture involves independent coding/decoding per antenna (layer) with equal rate and power per antenna and a fixed order of nulling/canceling decoding but is known to suffer from poor performance; for example, in a multi-input multi-output (MIMO) Rayleigh fading channel with M_t transmit and M_r receive antennas ($M_r \geq M_t$), the diversity-multiplexing gain (D-M) tradeoff is just $(M_r - M_t + 1)(1 - r/M_t)$ for $r \in [0, M_t]$. There are two remedies available, namely, (i) channel-dependent ordered decoding at the receiver and (ii) allocation of rates and powers across the transmit antennas. However, the former doesn't improve the D-M tradeoff curve and while the latter does (with maximum diversity gain M_r and maximum multiplexing gain M_t), its tradeoff curve is still significantly inferior compared to the D-M tradeoff curve of the optimum (unconstrained) MIMO architecture. In this two-part paper, it is shown that a dramatically better D-M tradeoff and error (e.g. outage) probability can be obtained if the two remedies, i.e., ordered BLAST decoding and rate/power allocation, are judiciously combined. Indeed, a framework is developed for jointly designing channel-dependent ordered decoding at the receiver and decoding order-dependent rate/power allocation at the transmitter. The framework encompasses a large class of new spatial multiplexing architectures (SMAs). In this part, an upper bound to the D-M tradeoff for this class is obtained and found to be quite close to the optimal D-M tradeoff of the MIMO channel. Two special SMAs are proposed corresponding to two different decoding orderings. One is called the Norm ordering Rate Tailored SMA (NRT-SMA), and the other is called the Greedy ordering Rate Tailored SMA (GRT-SMA). The latter is shown to have the D-M tradeoff equal to the upper bound and is hence optimal in D-M tradeoff among the proposed class of SMAs. Compared to the traditional V-BLAST architecture, the only added complexity of these SMAs is due to (i) feeding back the decoding order information ($\leq \log_2(M_t!)$ bits) per channel realization, and (ii) applying rate and power allocation according to a pre-computed lookup table.

Keywords— MIMO, V-BLAST, diversity gain, spatial multiplexing gain, tradeoff, rate tailored, multi-access channel.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) channels have well-known advantages over their single-input single-output (SISO) counterparts, i.e., much higher spectral efficiency and greatly improved diversity gain [1][2]. A lot of work has been done in the past decade to exploit the advantages of MIMO channel, among which the V-BLAST (vertical Bell Labs layered Space-Time) architecture [3] has

Y. Jiang was with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309 USA. He is currently with the NextWave Wireless Inc, San Diego, CA 92130 (e-mail:yjiang.ee@gmail.com).

M. K. Varanasi is with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309 USA.(e-mail:varanasi@colorado.edu)

gained popularity due to its simplicity and high rate performance. V-BLAST applies independent scalar coding to multiple substreams which are then spatially multiplexed and transmitted through different transmit antennas simultaneously. Hence, V-BLAST can be regarded as a spatial multiplexing architecture (SMA). For the standard V-BLAST, the spatially multiplexed substreams (or *layers*) have the same rate and power, and at the receiver the substreams are decoded sequentially using a decision feedback equalizer (DFE) [4][5]. The substream decoded first is subject to the interference from all the other substreams, and therefore tends to have the lowest output signal-to-interference-and-noise ratio (SINR). Moreover, the erroneously decoded substream causes more errors in the subsequently decoded substreams. Hence, the error probability of the first decoded substream dominates the overall system error probability. In a Rayleigh fading channel with M_t transmit antennas and M_r receive antennas, the D-M tradeoff curve is just $(M_r - M_t + 1)(1 - r/M_t)$ for $r \in [0, M_t]$ (V-BLAST requires $M_r \geq M_t$) [6]. In particular, the diversity gain of the standard V-BLAST is hence only $M_r - M_t + 1$. Two major remedies have been proposed to improve the performance of V-BLAST.

The first remedy is to decode the substreams in an order that is determined by the channel realization; in particular, the ordering rule proposed originally in [7] was popularized in [3] known widely as the V-BLAST ordering. The V-BLAST ordering algorithm is proven in [3] to maximize the smallest layer gain among all the $M_t!$ orderings. The V-BLAST ordering can be efficiently implemented using a recursive algorithm [8][9]. A simpler sub-optimal column-norm rule was proposed in [10] which was analyzed in [11] to show no improvement in diversity gain. An optimal order was also suggested in [11] to minimize overall error probability (the probability that not all symbols transmitted are detected correctly) which is distinct from the V-BLAST ordering but was seen to give rise to a very minor improvement compared to the V-BLAST ordering. We have recently proved in [12] that no ordering, including the V-BLAST ordering, can improve the D-M tradeoff (and hence diversity gain) of the V-BLAST scheme (note that [6] provides the upper bound $(M_r - 1)(1 - r/M_t)$ which is loose), although the V-BLAST ordering does yield $10 \log_{10} M_t$ dB coding gain for zero forcing V-BLAST (ZF-VB) [12] in the high SNR regime.

The second remedy was proposed in [11], [13] which suggests that the symbols of the various transmit antennas be decoded in a pre-determined fixed order but that the

rates and powers be optimized across the transmit antennas to minimize the exact overall error probability of the system, given the constraints on total rate and total power. In the sequel, we refer to this scheme as fixed order Rate Tailored V-BLAST (fixed order RT-VB). The fixed order RT-VB has maximal diversity gain of M_r because with increasing signal-to-noise rate (SNR) for a fixed total rate, the optimum transmit powers and rates are such that no rate or power is assigned to an increasing number of transmit antennas. In the D-M tradeoff framework, rate/power allocation reduces to just multiplexing gain allocation and it was shown in [6] that an improved tradeoff curve results but which is still much poorer than the D-M tradeoff curve of the optimum (unconstrained) architecture.

Besides the above two remedies, some other work on improving the performance of V-BLAST is available in the literature. For example, in [14], the authors propose to apply turbo decoding and exploit soft information on each layer. Their method yields some coding gain, but no diversity gain improvement. More recently, the authors of [15] propose another modification of V-BLAST, namely, the so-called space-time active rotation (STAR) method. The STAR method uses only $M_t - 1$ out of M_t transmit antennas and rotates the active set over time. It is shown numerically that STAR can be quite close to the V-BLAST scheme using the maximum likelihood (ML) receiver in terms of outage probability.

With independent coding per multiplexed data substream, the MIMO system is in fact equivalent to a multiple access channel (MAC), where the M_t transmit antennas amount to M_t users with no cooperation between them [13]. For a fading MAC channel, the fundamental diversity-multiplexing (D-M) gain tradeoff is given in [13], [16], which is significantly inferior to that of a single user MIMO channel [6]. Assuming no cooperation between the transmitter and the receiver, the performance limit of any modified V-BLAST is given in [13] and [16]. In particular, note the maximal diversity gain of any modified V-BLAST can be no greater than M_r . To have diversity gain more than M_r , one may resort to applying coding across both the spatial and the temporal domain as in the D-BLAST (diagonal Bell Labs layered Space-Time) architecture [17], LAST (Lattice Space-Time) coding [18], and DLL (Diagonally Layered Lattice) schemes [19] do. However, these codes need either short and powerful error control coding (in the D-BLAST case), or a computationally expensive lattice decoder (in the LAST and DLL codes). Moreover, the concatenation of the inner space-time code and the outer error control code is much more complicated than in V-BLAST.

In this two-part paper, a different approach is proposed to achieve a dramatic D-M tradeoff and error probability improvement. We develop a framework of jointly designing channel-dependent ordered decoding at the receiver and decoding ordering-dependent rate/power allocation at the transmitter. The receiver feeds a few ($\leq \log_2(M_t!)$) bits from receiver back to the transmitter with regard to the decoding ordering. The transmitter exploits this informa-

tion to assign rates and powers to each individual transmit antenna. The two key problems then are to (a) suggest ordering rules and (b) propose for each rule a method for allocating rate/power to the antennas. Problem (b) is solved in Part II of this paper via the analysis of the error probability per layer according to a minimax optimality criterion [20]. As for problem (a), note that an ordering rule, which is a map from channel matrices into the set of permutation matrices, specifies a member of the proposed class of SMAs. There are uncountably many ordering rules but we suggest two interesting ones in this two-part paper; one is a *norm ordering rule* which corresponds to decoding the layers in the *increasing* order of their respective channel column norms (i.e., the transmit antenna with the least channel column norm is decoded first), and the second is a *greedy ordering rule* that is determined through a sequence of recursively defined Householder transformations. The SMAs based on the two orderings are referred to as Norm ordering Rate Tailored SMA (NRT-SMA) and Greedy ordering Rate Tailored SMA (GRT-SMA), respectively.

In this paper, the NRT-SMA and GRT-SMA are analyzed in terms of diversity-multiplexing (D-M) gain tradeoff [6]. We show that while both can achieve the maximal diversity gain $M_r M_t^1$, a much more significant result is that the D-M gain tradeoff of the GRT-SMA is actually quite close to the optimal tradeoff of the single-user MIMO channel. We also establish an upper bound to the D-M gain tradeoff of the class of SMAs, which turns out to be exactly the D-M tradeoff of GRT-SMA. This result elevates the status of GRT-SMA from merely achieving very good performance to that of an optimal SMA among the proposed class of SMAs in terms of D-M tradeoff. However, our analysis does not preclude the existence of a SMA with the same D-M tradeoff but better coding gain performance than that of GRT-SMA.

The major advantage of the class of SMAs over D-BLAST [17], the LAST code [18], and the DLL code [19] is that the SMAs admit simple independent scalar coding for the multiplexed substreams, which reduces the implementation complexity significantly. For instance, for the SMAs the inner detector can be easily concatenated with outer error control coding. Compared to the classic V-BLAST, the only added complexity is (i) feedback of a few bits of decoding ordering information per channel realization, and (ii) allocating rates and powers across the transmit antennas according to a lookup table. The table is obtained *offline* as we will discuss in detail in Part II of this paper [20].

The remainder of this paper is organized as follows. Section II introduces the channel model and some useful preliminary results. Section III introduces the basic concept of the class of new SMAs and analyzes their D-M gain tradeoffs. Section IV studies ordering rules and ordered QR decompositions, including two special cases, i.e., the

¹this should not be not surprising since allocating the total rate and power to the transmit antenna with the largest channel norm or even to the transmit-receive antenna pair with the largest channel gain, through $\log M_t$ bits of feedback, will achieve this.

Norm QR decomposition and Greedy QR decomposition. Furthermore, we establish the upper bound to the diversity gains of the layers obtained by any ordered ZF-VB decoder, which shows that the ZF-VB decoder with greedy ordering achieves the upper bound. Because the proofs of the theorems in Section IV are rather technical, we relegate them to Section V.

II. CHANNEL MODEL AND PRELIMINARIES

A. Channel Model

We consider a communication system with M_t transmit and M_r receive antennas in a frequency flat fading channel. The sampled baseband signal is given by

$$\mathbf{y} = \mathbf{H}\mathbf{\Pi}\mathbf{W}^{\frac{1}{2}}\mathbf{s} + \mathbf{z}, \quad (1)$$

where $\mathbf{s} \in \mathbb{C}^{M_t \times 1}$ is the information symbols, \mathbf{W} is a diagonal matrix with diagonal entries $\{w_i\}_{i=1}^{M_t}$ denoting the power allocation, $\mathbf{\Pi}$ is a channel-dependent permutation matrix corresponding to the decoding ordering, and $\mathbf{y} \in \mathbb{C}^{M_r \times 1}$ is the received signal and $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ is the independent, identically distributed (iid) Rayleigh flat fading channel matrix. We assume that $\mathbf{z} \sim N(0, \sigma_z^2 \mathbf{I}_{M_r})$ is the circularly symmetric complex Gaussian noise where \mathbf{I}_{M_r} denotes an identity matrix with dimension M_r . Without loss of generality, we normalize \mathbf{s} such that $\mathbb{E}[\mathbf{s}\mathbf{s}^*] = \mathbf{I}$, and denote $W \triangleq \sum_{i=1}^{M_t} w_i$ as the total input power. Here $\mathbb{E}[\cdot]$ stands for the expectation, and $(\cdot)^*$ is the conjugate transpose. The input SNR is defined as $\rho = \frac{W}{\sigma_z^2}$. The individual input SNR of the i th substream is $\rho_i = \frac{w_i}{\sigma_z^2}$, $i = 1, 2, \dots, M_t$, and $\sum_{i=1}^{M_t} \rho_i = \rho$.

Different from the V-BLAST scheme, the class of SMAs make no assumption with regard to the numbers of transmit and receive antennas. Denote $N \triangleq \min\{M_t, M_r\}$. As we will see later, the rate tailored SMAs only select $K \leq N$ transmit antennas for use based on the information on decoding ordering.

B. Order Statistics

We reproduce some useful results on order statistics from [21].

Definition II.1: If the random variables X_1, X_2, \dots, X_n are rearranged in ascending order of magnitude and then written as $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, we call $X_{(i)}$ the i th order statistic ($i = 1, 2, \dots, n$).

Lemma II.2: Suppose X_1, X_2, \dots, X_n are n iid variates with probability density function (pdf) $f(x)$ and cumulative distribution function (cdf) $F(x)$. Then the i th order statistic $X_{(i)}$ has pdf

$$f_{X_{(i)}}(x) = \frac{1}{\beta(i, n-i+1)} F^{i-1}(x) [1-F(x)]^{n-i} f(x),$$

where $\beta(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$.

The following is an immediate corollary of Lemma II.2.

Corollary II.3: Suppose that X_1, \dots, X_{n-1} are statistically independent and are uniformly distributed over the

unit interval $(0, 1)$. Then the pdf of the i th order statistic is

$$f_{X_{(i)}}(x) = \frac{1}{\beta(i, n-i)} x^{i-1} (1-x)^{n-i-1} \quad 0 \leq x \leq 1. \quad (2)$$

C. Two Useful Theorems

The following theorem easily follows from [6].

Theorem II.4: Consider the iid Rayleigh fading channel \mathbf{H} given in (1) with ordered singular values $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N > 0$. Then

$$\lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(\lambda_i^2 < \epsilon)}{\log \epsilon} = (M_t - i + 1)(M_r - i + 1), \quad 1 \leq i \leq N. \quad (3)$$

Proof: Note that λ_i in [6, eq. (15)] is actually λ_{N-i+1}^2 here. We can rewrite [6, eq. (15)] as

$$\lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(\lambda_n^2 < \epsilon^{b_n}, \forall n)}{\log \epsilon} = \sum_{n=1}^N (|M_r - M_t| + 2(N - n + 1) - 1) b_n, \quad (4)$$

where $b_n \geq 0, \forall n$. Observe that the event $\{\lambda_i^2 < \epsilon\} = \{\lambda_n^2 < \epsilon, i \leq n \leq N\}$ corresponds to

$$b_n = \begin{cases} 0 & 1 \leq n \leq i-1 \\ 1 & i \leq n \leq N \end{cases}. \quad (5)$$

Inserting (5) into (4), we obtain (3). \blacksquare

The following theorem was established in [22] and will be used in Theorem IV.2.

Theorem II.5: Let $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ have rank N with singular values $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N > 0$. There exists an upper triangular matrix $\mathbf{R} \in \mathbb{C}^{N \times N}$ and matrices \mathbf{Q} and \mathbf{P} with orthonormal columns such that $\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^*$ if and only if the diagonal elements of \mathbf{R} satisfy

$$\prod_{i=1}^n |r_{[ii]}| \leq \prod_{i=1}^n \lambda_i, \text{ for } 1 \leq n < N, \text{ and } \prod_{i=1}^N |r_{ii}| = \prod_{i=1}^N \lambda_i, \quad (6)$$

where $|r_{[ii]}|$ is the i th largest entry among $\{|r_{ii}|\}_{i=1}^N$, the absolute values of the diagonal elements of \mathbf{R} .

III. A CLASS OF SPATIAL MULTIPLEXING ARCHITECTURES

In this section, we introduce the basic concept and features of a class of new SMAs. We also derive the D-M gain tradeoffs.

A. Basic Concept

At the receiver side of the SMAs, an ordered V-BLAST decoder is used, which applies the procedure of *successive interference nulling and decision feedback equalization*. In the *successive interference nulling* step, the decoder suppresses the interference by the minimum mean-squared-error (MMSE) or zero-forcing (ZF) criterion, which are called MMSE-VB and ZF-VB, respectively. In this paper, we constrain the receiver to be the ZF-VB which can be concisely represented by the QR decomposition

$\mathbf{H} = \mathbf{QR}$, where \mathbf{Q} is an $M_r \times N$ matrix with its orthonormal columns being the ZF nulling vectors, and \mathbf{R} is an $M_t \times M_t$ upper triangular matrix with positive diagonal if $M_t = N = \min\{M_t, M_r\}$. If $M_t > M_r = N$, then the firsts N columns of \mathbf{R} form an upper triangular matrix. In this case, the last $M_t - N$ transmit antennas are not used. Correspondingly, the ordered ZF-VB decoder can be represented by applying the QR decomposition to \mathbf{H} with its columns permuted, i.e., $\mathbf{H}\mathbf{\Pi} = \mathbf{QR}$ where $\mathbf{\Pi} \in \mathbb{R}^{M_t \times M_t}$ is a permutation matrix. If $M_t > N$, then only N transmit antennas corresponding to the first N columns of $\mathbf{H}\mathbf{\Pi}$ are used. Correspondingly, the trailing $M_t - N$ diagonal entries of the diagonal matrix \mathbf{W} in (1) are zeros. Hence the effective channel matrix is pruned to be $\tilde{\mathbf{H}} \triangleq \mathbf{H}\mathbf{\Pi} \in \mathbb{C}^{M_r \times N}$ where $\tilde{\mathbf{\Pi}}$ is the submatrix consisting of the first N columns of $\mathbf{\Pi}$. For the sake of notational simplicity, we still use $\tilde{\mathbf{H}} = \mathbf{QR}$ to denote the QR decomposition.

Now we can rewrite (1) as

$$\mathbf{y} = \mathbf{QR}\mathbf{W}_1^{\frac{1}{2}}\mathbf{s} + \mathbf{z}. \quad (7)$$

where $\mathbf{W}_1 \in \mathbb{R}^{N \times N}$ is the leading submatrix of \mathbf{W} given in (1) ($\mathbf{W}_1 = \mathbf{W}$ if $M_t \leq M_r$). Multiplying \mathbf{Q}^* to both sides of (7), which is effectively the *nulling step*, yields

$$\tilde{\mathbf{y}} = \mathbf{R}\mathbf{W}_1^{\frac{1}{2}}\mathbf{s} + \tilde{\mathbf{z}}, \quad (8)$$

$\tilde{\mathbf{y}} = \mathbf{Q}^*\mathbf{y}$ and $\tilde{\mathbf{z}} = \mathbf{Q}^*\mathbf{z}$. The sequential signal decoding, which involves the *decision feedback*, is as follows:

for $i = N : -1 : 1$

$$\hat{s}_i = \mathcal{Q} \left[\left(\tilde{y}_i - \sum_{j=i+1}^N r_{ij} \sqrt{w_j} \hat{s}_j \right) / r_{ii} \right]$$

end

where r_{ij} is the (i, j) th entry of \mathbf{R} and $\mathcal{Q}[\cdot]$ stands for mapping to the nearest point in the symbol constellation. Ignoring the error-propagation effect for now, we see that the MIMO channel is decomposed into N parallel layers

$$\tilde{y}_i = r_{ii} \sqrt{w_i} s_i + \tilde{z}_i, \quad i = 1, 2, \dots, N. \quad (9)$$

Because $\mathbb{E}[\tilde{\mathbf{z}}\tilde{\mathbf{z}}^*] = \sigma_z^2 \mathbf{I}$, the output SNR of the i th layer is $r_{ii}^2 w_i / \sigma_z^2 = r_{ii}^2 \rho_i$. Hence given the input SNR, the output SNRs of the substreams are completely determined by the diagonal entries of the upper triangular matrix \mathbf{R} which in turn depend on the permutation matrix $\mathbf{\Pi}$.

With fixed order decoding, $\mathbf{\Pi}$ does not depend on \mathbf{H} and it is well-known that the diagonal elements of \mathbf{R} are statistically independent with $\chi_{2(M_r - i + 1)}^2$ distribution (see, e.g., [23]). The diversity gain of a SISO channel only depends on the distribution of the channel gain around zero [24]. From this fact, the diversity gain of the i th layer is

$$D_i = \lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} = M_r - i + 1, \quad \text{for } 1 \leq i \leq N, \quad (10)$$

where $\mathbb{P}(\mathcal{E})$ stands for the probability of the event \mathcal{E} . For ordered decoding, $\mathbf{\Pi}$ is a function of \mathbf{H} and the distributions of r_{ii}^2 's are much more complicated. We shall address this issue in Section IV.

Now we have seen that with the ZF-VB decoder, the MIMO channel is converted into N parallel SISO channels with different channel gains. Under the overall constraints of data rate and input power, the transmitter attempts to optimally allocate the rates and powers over the N SISO channels to minimize the system error probability. Note that the exact values of the channel gains $\{r_{nn}\}_{n=1}^N$ are unavailable to the transmitter. Instead, the transmitter uses the *a priori* information on the statistics of $\{r_{nn}^2\}_{n=1}^N$ to determine the rates and powers allocated to $K \leq N$ substreams. Given the decoding ordering information fed back from the receiver, the transmitter maps the substreams to transmit antennas. A major difficulty here is that in many cases the exact pdfs of $\{r_{nn}^2\}_{n=1}^N$ are intractable due to the channel-dependent ordering. This problem is solved in Part II of this paper.

In summary, the transmitter determines, based on the statistics of the layer gains, the rates and powers allocated to the individual substreams which are independently coded with SISO error control codes. Using the channel state information, the receiver determines the decoding ordering, and then feeds it back to transmitter. Using the ordering information, the transmitter maps the substreams to the transmit antennas. Finally, the receiver applies an ordered ZF-VB decoder to decode the substreams sequentially. This framework encompasses various new SMAs based on different decoding ordering rules.

B. Diversity-Multiplexing Tradeoff Analysis

We analyze the D-M gain tradeoff of the class of SMAs in the asymptotically high SNR regime. Denote $R(\rho)$ as the data rate of any communication scheme (a family of codes parameterized by SNR) with input SNR ρ . The diversity gain and multiplexing gain are defined as follows [6].

Definition III.1: A scheme is said to have multiplexing gain \mathcal{R} and diversity gain \mathcal{D} if the data rate $R(\rho)$ satisfies

$$\lim_{\rho \rightarrow \infty} \frac{R(\rho)}{\log \rho} = \mathcal{R}, \quad (11)$$

and the average error probability $P_e(\rho)$ satisfies

$$-\lim_{\rho \rightarrow \infty} \frac{\log P_e(\rho)}{\log \rho} = \mathcal{D}. \quad (12)$$

The tradeoff between diversity gain and multiplexing gain is denoted by $\mathcal{D}(\mathcal{R})$ which is always a non-increasing function. For the iid Rayleigh fading channel given in (1), the optimal D-M gain tradeoff is a piece-wise linear curve obtained by connecting the following $N + 1$ points [6]

$$\{(n, (M_r - n)(M_t - n))\}_{n=0}^N. \quad (13)$$

The more recent work extends the D-M tradeoff framework to the diversity-multiplexing-delay tradeoff [25]. However, this extended framework is not relevant here since our proposed architecture does not require retransmissions and hence does not incur delay.

The optimal D-M gain tradeoff of a scalar fading channel with maximal diversity gain D is

$$\mathcal{D}(\mathcal{R}) = D(1 - \mathcal{R}), \quad 0 \leq \mathcal{R} \leq 1, \quad (14)$$

which can be achieved using uncoded quadrature amplitude modulation (QAM) [26]. Suppose we have N scalar layers with diversity gains D_i , $i = 1, \dots, N$. The optimal D-M gain tradeoff of the i th layer is $\mathcal{D}_i(\mathcal{R}_i) = D_i(1 - \mathcal{R}_i)$, where \mathcal{R}_i is the multiplexing gain of the i th layer.

Because the class of SMAs apply independent scalar coding to each layer, as input SNR ρ increases, the error probability of the layer with the smallest diversity gain dominates the overall error probability of the system. To optimize the system performance, the rates, and therefore the multiplexing gains, should be tailored according to the following optimization problem:

$$\begin{aligned} \mathcal{D}(\mathcal{R}) = \max_{\mathcal{R}_i} \min_{i: \mathcal{R}_i > 0} & D_i(1 - \mathcal{R}_i) \\ \text{subject to} & 0 \leq \mathcal{R}_i \leq 1, \quad 1 \leq i \leq N, \\ & \sum_{i=1}^N \mathcal{R}_i = \mathcal{R}, \end{aligned} \quad (15)$$

Denote $\mathcal{I} \triangleq \{i : \mathcal{R}_i > 0\}$ the set of the layers in use. Then we observe that at the optimal solution to (15), the multiplexing gains \mathcal{R}_i 's must be such that $D_i(1 - \mathcal{R}_i) = \mathcal{D}$ for all $i \in \mathcal{I}$. Otherwise, if there exists $i \in \mathcal{I}$ such that $D_i(1 - \mathcal{R}_i) < D_j(1 - \mathcal{R}_j)$ for $\forall j \in \mathcal{I}$ and $j \neq i$, then we can always reduce \mathcal{R}_i to $\mathcal{R}_i - \delta$ (note that $\mathcal{R}_i > 0$) and increase \mathcal{R}_j to $\mathcal{R}_j + \delta$ for some $j \neq i$ (note that $\mathcal{R}_j < 1$) such that the constraints are still satisfied but the cost function is increased from \mathcal{D} to $\mathcal{D} + D_i\delta$, which leads to a contradiction. From the observation that $D_i(1 - \mathcal{R}_i) = \mathcal{D}$, $\forall i \in \mathcal{I}$, we have $\mathcal{R}_i = 1 - \frac{\mathcal{D}}{D_i}$, $\forall i \in \mathcal{I}$. According to the constraint $\sum_{i \in \mathcal{I}} \mathcal{R}_i = \mathcal{R}$, we have $\sum_{i \in \mathcal{I}} (1 - \frac{\mathcal{D}}{D_i}) = \mathcal{R}$. Thus

$$\mathcal{D} = \frac{K - \mathcal{R}}{\sum_{i \in \mathcal{I}} D_i^{-1}} \quad (16)$$

where $K = |\mathcal{I}|$ is the cardinality of the finite set \mathcal{I} , i.e., the number of transmit antennas in use. Hence we can simplify (16) to be

$$\mathcal{D}(\mathcal{R}) = \max_{\substack{\mathcal{I} \subseteq \{1, 2, \dots, M_t\} \\ |\mathcal{I}| = K}} \frac{K - \mathcal{R}}{\sum_{i \in \mathcal{I}} D_i^{-1}}. \quad (17)$$

If we assume, without loss of generality, that D_i 's are in non-increasing order, then the set \mathcal{I} which leads to the maximal \mathcal{D} must be $\mathcal{I} = \{1, 2, \dots, K\}$, i.e., the K layers with the highest diversity gains are used. It follows from (17) that

$$\mathcal{D}(\mathcal{R}) = \max_{\substack{\mathcal{R} \leq K \leq N \\ K \in \mathbb{Z}}} \frac{K - \mathcal{R}}{\sum_{i=1}^K D_i^{-1}}. \quad (18)$$

It is easily seen that $\mathcal{D}(N) = 0$ and $\mathcal{D}(0) = D_1$. In the open interval $\mathcal{R} \in (0, M_t)$, $\mathcal{D}(\mathcal{R})$ is a piecewise linear function and the inflections (changes of derivative) occur when the number of transmit antennas in use changes by one, which happens if

$$\frac{K - \mathcal{R}}{\sum_{i=1}^K D_i^{-1}} = \frac{K + 1 - \mathcal{R}}{\sum_{i=1}^{K+1} D_i^{-1}}, \quad 1 \leq K \leq N - 1, \quad (19)$$

or the overall multiplexing gain

$$\mathcal{R} = K - D_{K+1} \sum_{i=1}^K D_i^{-1}. \quad (20)$$

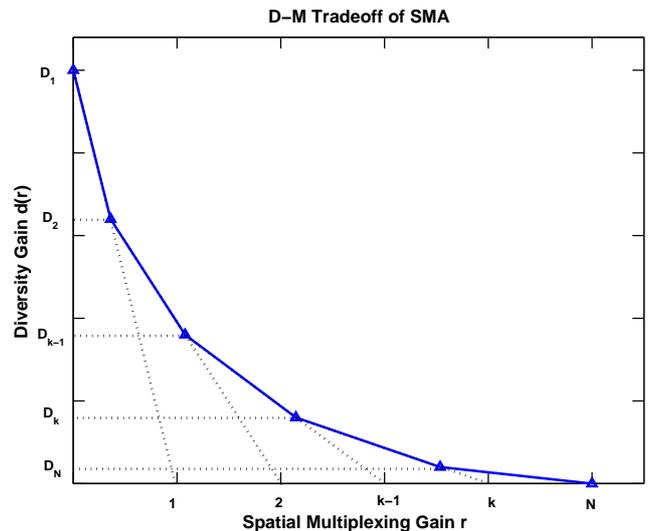


Fig. 1. Visualization of the D-M gain tradeoff of the SMAs given in (22).

Substituting (20) into (18) we get the diversity gains at the inflection points:

$$\mathcal{D}(\mathcal{R}) = D_{K+1}, \quad 1 \leq K \leq N - 1. \quad (21)$$

Hence the achievable D-M tradeoff function can be represented by a piecewise linear curve connecting the following $N + 1$ points

$$(0, D_1), \left\{ \left(K - D_{K+1} \sum_{i=1}^K D_i^{-1}, D_{K+1} \right) \right\}_{K=1}^{N-1}, \text{ and } (N, 0). \quad (22)$$

Figure 1 visualizes the D-M gain tradeoff given in (22), which also reveals the fact that the coordinate $\left(k - 1 - D_k \sum_{i=1}^{k-1} D_i^{-1}, D_k \right)$ is exactly where the line $y = D_k$ crosses the line connecting

$$\left(k - 2 - D_{k-1} \sum_{i=1}^{k-2} D_i^{-1}, D_{k-1} \right) \text{ and } (k - 1, 0).$$

This analysis suggests that to improve the overall system D-M tradeoff, the decoder should apply ordering which yields layers with large D_i 's.

There is an interesting relationship between the proposed SMAs and the system with transmit antenna selection. Indeed, a V-BLAST system with transmit antenna selection (see, e.g., [27]) can be regarded as a special SMA which allocates *equal* rate to a subset of transmit antennas and zero rate to the rest. However, such an SMA is clearly suboptimal due to the equal rate constraint.

It is worthwhile noting that the error propagation effect can be safely ignored in the D-M tradeoff analysis. The reason is as follows. By the union bound the overall probability of error is upper bounded by the sum of the error probabilities of the K individual layers, which is further upper bounded by K times of the error probability of the

worse layer. However, our proposed SMA carefully allocates rates to the layers such that all of them have the same diversity gain. Therefore, error propagation only causes some coding gain loss rather than diversity gain loss. Finally, we remark that power allocation does not appear in the above analysis because at asymptotically high SNR, it only influences the coding gain performance of the system.

IV. ORDERING RULES AND ORDERED QR DECOMPOSITIONS

We have seen in Section III that members of the proposed class of SMAs are decided by the decoding ordering rule which in turn influences the layer gains. It also follows from the D-M analysis that a good decoding ordering should yield layers with large diversity gains ($\{\mathbf{D}_i\}_{i=1}^N$). In the next we study two special ordering rules, namely, the *norm ordering rule* and *greedy ordering rule*. The QR decompositions with the two rules are referred to as Norm QR and Greedy QR decomposition, respectively.

A. Norm QR Decomposition

The procedure of the Norm QR decomposition is as follows.

- (i) Calculate the Euclidean norms $\{\|\mathbf{h}_i\|\}_{i=1}^{M_t}$ where \mathbf{h}_i is the i th column of \mathbf{H} .
- (ii) Find permutation matrix $\mathbf{\Pi}$ such that the column norms of the permuted matrix $\mathbf{H}\mathbf{\Pi}$, from the left to the right, are in non-increasing order.
- (iii) Apply the standard QR decomposition to the permuted matrix $\mathbf{H}\mathbf{\Pi} = \mathbf{Q}\mathbf{R}$.

Given the iid Rayleigh assumption of \mathbf{H} , we have derived the pdfs of all the squared diagonal elements of \mathbf{R} , $\{r_{ii}^2\}_{i=1}^N$.

Theorem IV.1: Suppose $\mathbf{H}\mathbf{\Pi} = \mathbf{Q}\mathbf{R}$ is the Norm QR decomposition of \mathbf{H} given in (1). Then the pdf of r_{11}^2 is

$$f_{r_{11}^2}(x) = \frac{1}{M_t} \cdot \frac{x^{M_r-1} e^{-x}}{(M_r-1)!} \left(1 - e^{-x} \sum_{k=0}^{M_r-1} \frac{x^k}{k!}\right)^{M_t-1}, \quad x > 0. \quad (23)$$

The pdfs of other diagonal elements are

$$f_{r_{ii}^2}(x) = \frac{x^{M_r-i} e^{-M_t x}}{\beta(M_r-i+1, i-1)\beta(M_t+1-i, i)(M_r-1)!} \times \int_0^\infty w^{i-2} e^{-M_t w} \left(\sum_{k=M_t}^\infty \frac{(x+w)^k}{k!}\right)^{M_t-i} \left(\sum_{k=0}^{M_t-1} \frac{(x+w)^k}{k!}\right)^{i-1} dw, \quad x > 0, \quad (24)$$

for $i = 2, \dots, N$. Moreover,

$$\lim_{\epsilon \rightarrow 0_+} \frac{\log \mathbb{P}(r_{11}^2 < \epsilon)}{\log \epsilon} = M_t M_r, \quad (26)$$

and

$$\lim_{\epsilon \rightarrow 0_+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} = M_r - i + 1, \quad i = 2, \dots, N. \quad (27)$$

In other words, for the ZF-VB decoder based on the norm ordering rule, the diversity gain of the i th layer is

$$\mathbf{D}_i = \begin{cases} M_r M_t & i = 1 \\ M_r - i + 1 & 2 \leq i \leq N. \end{cases} \quad (28)$$

Proof: We relegate the proof to Section V-A. \blacksquare

Given M_t and M_r , we can simplify the pdf (25) using MathematicaTM.

Recall that the i th layer of unordered V-BLAST equalizer has diversity gain of $M_r - i + 1$. Theorem IV.1 shows that Norm QR can significantly increase the diversity gain of the first layer. However, no diversity gain improvement is achieved for the other layers. To further improve layer diversity gains, we turn to the Greedy QR decomposition.

B. Greedy QR Decomposition

The Greedy QR decomposition consists of N recursive steps. We only elaborate the first step. The subsequent steps would be obvious given the first one.

In the first step, we go through the following procedures.

- (i) Calculate Euclidean norms $\{\|\mathbf{h}_i\|\}_{i=1}^{M_t}$.
- (ii) Permute \mathbf{h}_1 and \mathbf{h}_j where $j = \arg \max_{1 \leq i \leq M_t} \{\|\mathbf{h}_i\|\}$. This operation can be represented by $\mathbf{H}_1 = \mathbf{H}\mathbf{\Pi}_1$ with $\mathbf{\Pi}_1$ being the permutation matrix. (If $j = 1$, $\mathbf{\Pi}_1$ degrades to be \mathbf{I}_{M_t})
- (iii) Apply a Householder matrix \mathbf{Q}_1 to transform the first column of \mathbf{H}_1 to a scaled \mathbf{e}_1 , where \mathbf{e}_1 is the first column of \mathbf{I}_{M_r} .

The procedure (i-iii) can be illustrated by

$$\begin{pmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix} \xrightarrow{\mathbf{Q}_1^* \mathbf{H}\mathbf{\Pi}_1} \begin{pmatrix} r_{11} & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{pmatrix}. \quad (29)$$

Note that $r_{11} = \max\{\|\mathbf{h}_i\|, 1 \leq i \leq M_t\}$. In the next step, the same procedures are applied to the trailing $(M_r - 1) \times (M_t - 1)$ submatrix on the right hand side of (29), which yields a permutation matrix $\mathbf{\Pi}_2$ and a Householder matrix \mathbf{Q}_2 . After N recursive steps, we obtain the desired QR decomposition $\mathbf{R} = \mathbf{Q}^* \mathbf{H}\mathbf{\Pi}$, or equivalently,

$$\mathbf{H}\mathbf{\Pi} = \mathbf{Q}\mathbf{R} \quad (30)$$

where $\mathbf{\Pi} = \mathbf{\Pi}_1 \mathbf{\Pi}_2 \dots \mathbf{\Pi}_N$ and $\mathbf{Q} = \mathbf{Q}_1 \mathbf{Q}_2 \dots \mathbf{Q}_N$ ($\mathbf{Q}_N = \mathbf{I}$ if $M_t \geq M_r = N$). In summary, at the i th step this ordering algorithm ‘‘greedily’’ attempts to make the i th diagonal element of \mathbf{R} as large as possible.²

Note that Norm and Greedy QR decompositions yield the same r_{11}^2 whose pdf is given in (23). For Greedy QR decomposition, the pdfs of $\{r_{ii}^2\}_{i=2}^N$ are unavailable. However, we have informative bounds on $\{r_{ii}^2\}_{i=1}^N$.

Theorem IV.2: Consider a matrix $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ with nonzero singular values $\lambda_1 \geq \lambda_2 \dots \geq \lambda_N > 0$. Let

²The Greedy QR decomposition is not new. The built-in Matlab function is $[\mathbf{Q}, \mathbf{R}, \mathbf{\Pi}] = \text{QR}(\mathbf{H})$.

$\mathbf{H}\mathbf{\Pi} = \mathbf{Q}\mathbf{R}$ be the Greedy QR decomposition. Then

$$\frac{\sum_{j=i}^N \lambda_j^2}{M_t - i + 1} \leq r_{ii}^2 \leq \lambda_i^2 \prod_{j=1}^{i-1} (M_t - j + 1), \quad i = 1, 2, \dots, N. \quad (31)$$

Proof: The lower bound was proven in [28], and we omit it here. We prove the upper bound.

We first assume that $M_r \geq M_t = N$. Then $\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{\Pi}^T$. It follows from Theorem II.5 that

$$\prod_{j=1}^i r_{jj}^2 \leq \prod_{j=1}^i \lambda_j^2, \quad 1 \leq i \leq N. \quad (32)$$

It follows from the proven lower bound that $r_{ii}^2 \geq \frac{\lambda_i^2}{M_t - i + 1}$. Hence

$$\prod_{j=1}^{i-1} r_{jj}^2 \geq \prod_{j=1}^{i-1} \frac{\lambda_j^2}{M_t - j + 1}. \quad (33)$$

Combining (33) and (32) yields

$$r_{ii}^2 \leq \lambda_i^2 \prod_{j=1}^{i-1} (M_t - j + 1), \quad i = 1, 2, \dots, N. \quad (34)$$

Now we consider the case $M_t > M_r$. Denote $\tilde{\mathbf{R}} \in \mathbb{C}^{N \times N}$ the submatrix consisting of the first N columns of \mathbf{R} . Then $\tilde{\mathbf{H}} \triangleq \mathbf{Q}\tilde{\mathbf{R}}$ is also a submatrix of \mathbf{H} with singular values $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \dots \geq \tilde{\lambda}_N \geq 0$. According to the same argument which leads to (34), we can prove that

$$r_{ii}^2 \leq \tilde{\lambda}_i^2 \prod_{j=1}^{i-1} (M_t - j + 1), \quad i = 1, 2, \dots, N. \quad (35)$$

Furthermore, because $\mathbf{H}\mathbf{H}^* \succeq \tilde{\mathbf{H}}\tilde{\mathbf{H}}^*$,³ we have $\lambda_i^2 \geq \tilde{\lambda}_i^2$ for $\forall i$ [29, Corollary 4.3.3]. Hence

$$r_{ii}^2 \leq \lambda_i^2 \prod_{j=1}^{i-1} (M_t - j + 1), \quad i = 1, 2, \dots, N, \quad (36)$$

also hold for $M_t > M_r$. The theorem is proven. \blacksquare

It follows from the lower bound in (31) that

$$\begin{aligned} \lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} &\geq \lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(\lambda_i^2 < (M_t - i + 1)\epsilon)}{\log \epsilon} \\ &= \lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(\lambda_i^2 < \epsilon)}{\log \epsilon} \\ (\text{see Theorem II.4}) &= (M_t - i + 1)(M_r - i + 1). \end{aligned} \quad (37)$$

On the other hand, it follows from the upper bound in (31) that

$$\lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} \leq (M_t - i + 1)(M_r - i + 1). \quad (38)$$

Therefore $\lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} = (M_t - i + 1)(M_r - i + 1)$. Now we have proven the following theorem.

³We write $\mathbf{A} \succeq 0$ if \mathbf{A} is a positive semi-definite (p.s.d.) matrix, and $\mathbf{A} \succeq \mathbf{B}$ or $\mathbf{B} \preceq \mathbf{A}$ if $\mathbf{A} - \mathbf{B} \succeq 0$.

Theorem IV.3: The i th layer of ZF-VB decoder based on the greedy ordering rule, has diversity gain

$$\mathbf{D}_i = (M_t - i + 1)(M_r - i + 1), \quad 1 \leq i \leq N. \quad (39)$$

Compared to the norm ordering, the greedy decoding ordering yields layers with significantly higher diversity gains. Now it is interesting to find out the highest diversity-multiplexing gain tradeoff an ordered decoding can achieve. This problem is addressed in the next.

C. Upper Bound of $\{\mathbf{D}_i\}_{i=1}^N$

Rewrite the ordered QR decomposition $\mathbf{H}\mathbf{\Pi} = \mathbf{Q}\mathbf{R}$. Let us write a permuted channel matrix in its column form: $\mathbf{H}\mathbf{\Pi} = [\mathbf{h}_{\pi(1)}, \dots, \mathbf{h}_{\pi(i-1)}, \mathbf{h}_{\pi(i)}, \dots, \mathbf{h}_{\pi(M_t)}] = \mathbf{Q}\mathbf{R}$. Denote $\mathbf{H}_i \triangleq [\mathbf{h}_{\pi(1)}, \dots, \mathbf{h}_{\pi(i-1)}]$. Then $r_{ii}^2 = \mathbf{h}_{\pi(i)}^* \mathbf{P}_{\mathbf{H}_i}^\perp \mathbf{h}_{\pi(i)}$, where $\mathbf{P}_{\mathbf{H}_i}^\perp = \mathbf{I} - \mathbf{H}_i(\mathbf{H}_i^* \mathbf{H}_i)^{-1} \mathbf{H}_i^*$. Thus r_{ii}^2 is a function of \mathbf{H}_i and $\mathbf{h}_{\pi(i)}$, and it is invariant to the column permutation of \mathbf{H}_i . Hence out of the $M_t!$ decoding orderings, one may have up to $\binom{M_t}{i-1} \cdot (M_t - i + 1) = \frac{M_t!}{(i-1)!(M_t-i)!}$ different values of r_{ii}^2 ($1 \leq i \leq N$). As we have mentioned before, a good ordering rule should maximize the diversity gains \mathbf{D}_i , $i = 1, \dots, N$. Theorem IV.4 presents an upper bound to the diversity gains of the M_t layers for any ordering rule.

Theorem IV.4: Consider the ordered QR decomposition $\mathbf{H}\mathbf{\Pi} = \mathbf{Q}\mathbf{R}$ where $\mathbf{\Pi}$ is a permutation matrix dependent on \mathbf{H} . Let r_{ii} be the i th diagonal of \mathbf{R} . The inequality

$$\lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} \leq (M_t - i + 1)(M_r - i + 1), \quad 1 \leq i \leq N, \quad (40)$$

holds for any ordering rule. In other words, the diversity gain of the i th layer

$$\mathbf{D}_i \leq (M_t - i + 1)(M_r - i + 1), \quad 1 \leq i \leq N. \quad (41)$$

Proof: The proof is relegated to Section V-B. \blacksquare

It follows from (41) that if $N = M_t$, then $\mathbf{D}_{M_t} \leq M_r - M_t + 1$, i.e., the first decoded layer has diversity gain no higher than $M_r - M_t + 1$. It is well-known that if equal rates are allocated across the layers, the overall system performance of V-BLAST is limited by the first decoded layer. Hence an interesting corollary of Theorem IV.4 is that the V-BLAST system with any decoding ordering has diversity gain $M_r - M_t + 1$.

We see from Theorem IV.3 and IV.4 is that *the greedy ordering rule is D-M tradeoff-optimal among all ordering rules.*

Since we have the \mathbf{D}_i 's, we can obtain the D-M gain tradeoffs of the GRT-SMA and NRT-SMA according to (22). Figure 2 shows the D-M tradeoff of various schemes in a channel with dimensionality $M_r = M_t = N$. Besides GRT-SMA and NRT-SMA, Figure 2 also shows the optimal tradeoff (see (13)), the fixed order RT-VB scheme [11], the standard V-BLAST either with or without BLAST ordering [3] whose tradeoff is $\mathcal{D}_{\text{vb}}(\mathcal{R}) = (1 - \frac{\mathcal{R}}{N})$ (see [12]), and the V-BLAST architecture with an ML receiver whose tradeoff is $\mathcal{D}_{\text{mac}}(\mathcal{R}) = N(1 - \frac{\mathcal{R}}{N})$ (see [13][16]). It is further proven in [26] that the system with uncoded QAM

input and ML decoder (sphere decoder) achieves such a tradeoff. As shown in Figure 2, both NRT-SMA and GRT-SMA can achieve the two end points of the optimal D-M gain tradeoff and can have significantly larger diversity gain than the V-BLAST with an ML receiver, especially in the low/medium multiplexing gain regime. It also follows from Theorems IV.4 and IV.3 that the D-M tradeoff of GRT-SMA is an upper bound to the class of SMAs. The moderate performance loss of GRT-SMA compared to the optimal is due to using independent SISO coding/decoding for each substream.

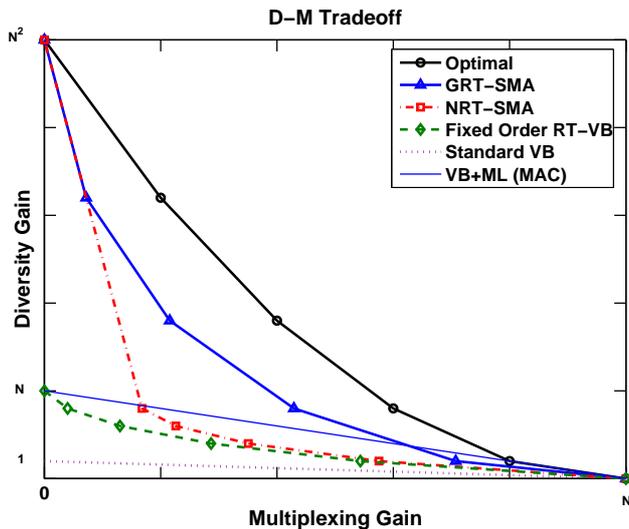


Fig. 2. Comparison of the D-M tradeoffs.

V. PROOFS OF THEOREM IV.1 AND THEOREM IV.4

A. Proof of Theorem IV.1

We first establish the following definition and lemmas.

Definition V.1: A vector \mathbf{x} is isotropic if for any unitary matrix \mathbf{U} , $\mathbf{x} \sim \mathbf{U}\mathbf{x}$, i.e., \mathbf{x} has a distribution invariant under rotations and reflections.

Lemma V.2: (See [23, Theorem 1.5.5]) Suppose $\mathbf{x} \in \mathbb{C}^n$ is a circular symmetric Gaussian vector $\mathbf{x} \sim N(0, \mathbf{I})$. The Euclidean norm $\|\mathbf{x}\|$ and the direction $\mathbf{v} \triangleq \frac{\mathbf{x}}{\|\mathbf{x}\|}$ are independent. Moreover, the direction vector \mathbf{v} has a uniform distribution over the unit sphere $\mathcal{S} = \{\mathbf{u} \in \mathbb{C}^n : \|\mathbf{u}\| = 1\}$. An immediate inference of Lemma V.2 is that the Gaussian vector $\mathbf{x} \sim N(0, \mathbf{I})$ is isotropic even if its norm is under some constraint.

Lemma V.3: For any vector $\mathbf{x} \in \mathbb{C}^n$, the associated Householder matrix \mathbf{Q} satisfying $\mathbf{Q}\mathbf{x} = \|\mathbf{x}\|\mathbf{e}$, where \mathbf{e} is the first column of an $n \times n$ identity matrix, is a function of the direction of \mathbf{x} , i.e., $\mathbf{x}/\|\mathbf{x}\|$, and is independent of $\|\mathbf{x}\|$.

Proof: The associated Householder matrix of \mathbf{x} is $\mathbf{Q} = \mathbf{I} - 2\mathbf{w}\mathbf{w}^*$, where [30]

$$\mathbf{w} = \frac{\mathbf{x} - \|\mathbf{x}\|\mathbf{e}}{\|\mathbf{x} - \|\mathbf{x}\|\mathbf{e}\|} = \frac{\mathbf{x}/\|\mathbf{x}\| - \mathbf{e}}{\|\mathbf{x}/\|\mathbf{x}\| - \mathbf{e}\|}. \quad (42)$$

Combining (42) and Lemma V.2, we have proven the lemma. \blacksquare

The Norm QR decomposition algorithm applies the QR decomposition to a matrix $\mathbf{G} \triangleq \mathbf{H}\mathbf{\Pi}$, where $\mathbf{\Pi}$ is the permutation matrix such that $\|\mathbf{g}_i\|$, $i = 1, \dots, M_t$ are in a non-decreasing ordering. The squared norms of the unordered columns are iid $\chi_{2M_r}^2$ random variables with pdf $f(x) = \frac{x^{M_r-1}e^{-x}}{(M_r-1)!}$, $x \geq 0$, and cdf $F(x) = 1 - e^{-x} \sum_{k=0}^{M_r-1} \frac{x^k}{k!}$, $x \geq 0$. According to Lemma II.2, $\|\mathbf{g}_i\|^2$ has distribution

$$f_{\|\mathbf{g}_i\|^2}(x) = \frac{1}{\beta(M_t + 1 - i, i)} \frac{x^{M_r-1}e^{-x}}{(M_r-1)!} \times \left(1 - e^{-x} \sum_{k=0}^{M_r-1} \frac{x^k}{k!}\right)^{M_t-i} \left(e^{-x} \sum_{k=0}^{M_t-1} \frac{x^k}{k!}\right)^{i-1} \quad \text{for } x > 0. \quad (43)$$

Hence $r_{11}^2 = \|\mathbf{g}_1\|^2$ has distribution given in (23).

The standard QR decomposition of \mathbf{G} is obtained by left multiplying \mathbf{G} with M_t Householder matrices successively [30, Section 5.2.1]. After left multiplying one householder matrix to \mathbf{G} , we obtain

$$\mathbf{Q}_1^* \mathbf{G} = \begin{pmatrix} r_{1,1} & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & * & \cdots & * \\ 0 & * & \cdots & * \end{pmatrix}. \quad (44)$$

The column vectors of \mathbf{G} are neither Gaussian nor mutually independent since their norms are constrained and are in non-increasing order. But they are still isotropic by Lemma V.2, and the direction vectors of all the columns are still independent because of the independence between direction vector and vector length. According to Lemma V.3, the householder matrix \mathbf{Q}_1 is independent of $\{\mathbf{g}_i\}_{i=2}^{M_t}$. Hence the trailing $(M_r) \times (M_t - 1)$ submatrix in the right hand side of (44) (represented by $*$'s) has the same distribution as its counterpart in \mathbf{G} . Hence we see that r_{22}^2 has the same distribution as $\|\mathbf{g}_2(2 : M_r)\|^2$, where $\mathbf{g}_i(i : M_r)$ consists of the last $M_r - i + 1$ elements of \mathbf{g}_i . Following the same argument, we can further show that r_{ii}^2 has the same distribution as $\|\mathbf{g}_i(i : M_r)\|^2$. To further derive the pdfs of the diagonal $\{r_{ii}^2\}_{i=2}^{M_t}$, we establish the following Lemma.

Lemma V.4: For a complex-valued Gaussian vector $\mathbf{h} \sim N(0, \mathbf{I}_{M_r})$, the pdf of the $X \triangleq \|\mathbf{h}(i : M_r)\|^2$ conditioned on $Y \triangleq \|\mathbf{h}\|^2$ is

$$f_{X|Y}(x|y) = \frac{x^{M_r-i}(y-x)^{i-2}}{\beta(M_r-i+1, i-1)y^{M_r-1}} \quad 0 < x < y. \quad (45)$$

Proof: For the Gaussian vector $\mathbf{h} \sim N(0, \mathbf{I}_{M_r})$, the squared absolute value of each element is iid of exponential distribution. Hence we can regard $\|\mathbf{h}(M_r - i + 1 : M_r)\|^2$, $i = 1, 2, \dots, N$ as the epoch of the i th event in a Poisson process with unit intensity. For a Poisson process, given y as the epoch of the M_r th event, i.e., $\|\mathbf{h}\|^2 = y$, the arriving times of the previous $M_r - 1$ events considered as unordered random variables are independent and uniformly distributed on the interval $(0, y)$. If $M_r - 1$ points

are randomly dropped on an interval $(0, y)$ and the positions of the points from the left to the right are indexed as $x_1, x_2, \dots, x_{M_r-1}$, then according to Corollary II.3, the pdf of x_i is

$$f_{X_i}(x) = \frac{x^{i-1}(y-x)^{M_r-i-1}}{\beta(i, M_r-i)y^{M_r-1}}, \quad 0 \leq x \leq y. \quad (46)$$

As $\|\mathbf{h}(i : M_r)\|^2 = x_{M_r-i+1}$, replacing i in (46) by $M_r - i + 1$, we see that the pdf of X_{M_r-i+1} is

$$f_{X_{M_r-i+1}}(x) = \frac{x^{M_r-i}(y-x)^{i-2}}{\beta(M_r-i+1, i-1)y^{M_r-1}}, \quad 0 \leq x \leq y. \quad (47)$$

The lemma is proven. \blacksquare

Based on Lemma V.4 and the observation that $\{r_{ii}^2\}_{i=1}^{M_t}$ have the same distribution as $\|\mathbf{g}_i(i : M_r)\|^2$, we can calculate the distribution of r_{ii}^2 ($i \geq 2$) as (48) (see the top of the next page). Denote the integral in (49) as $J(x)$. It is easy to see that

$$\begin{aligned} 0 < J(0) &= \int_0^\infty w^{i-2} e^{-M_t w} \left(e^w - \sum_{k=0}^{M_t-1} \frac{w^k}{k!} \right)^{M_t-i} \\ &\quad \times \left(\sum_{k=0}^{M_t-1} \frac{w^k}{k!} \right)^{i-1} dw \\ &< \int_0^\infty w^{i-2} e^{-w} dw = (i-2)!. \end{aligned} \quad (50)$$

On the other hand, by letting $y = x + w$, we have

$$J(x) = \int_x^\infty (y-x)^{i-1} e^{-M_t y} \left(\sum_{k=M_t}^\infty \frac{y^k}{k!} \right)^{M_t-i} \left(\sum_{k=0}^{M_t-1} \frac{y^k}{k!} \right)^{i-1} dy$$

which is straight forward to show that the derivative $J'(x)$ is finite. Hence, $J(x)$ is a continuous function and by Taylor Expansion, $J(x) = J(0) + J'(\epsilon)x$ for some $\epsilon \in (0, x)$. Now we have proven that in the neighborhood of the origin

$$f_{r_{ii}^2}(x) = \tau_i \cdot x^{M_r-i} + o(x^{M_r-i}) \quad (51)$$

for some $\tau_i > 0$. Thus we conclude that

$$\lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} = M_r - i + 1, \quad \text{for } 2 \leq i \leq M_t. \quad (52)$$

Hence the i th ($i \geq 2$) layer has diversity gain of only $M_r - i + 1$. As for $i = 1$, it can be seen from (23) that

$$\begin{aligned} f_{r_{11}^2}(x) &= \frac{1}{M_t} \cdot \frac{x^{M_r-1} e^{-x}}{(M_r-1)!} \left(e^{-x} \sum_{k=M_r}^\infty \frac{x^k}{k!} \right)^{M_t-1} \\ &= \eta \cdot x^{M_r M_t-1} + o(x^{M_r M_t-1}), \quad 0 \leq x \ll 1 \end{aligned}$$

for some $\eta > 0$. Hence $\lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(r_{11}^2 < \epsilon)}{\log \epsilon} = M_t M_r$, i.e., the first layer has diversity gain of $M_t M_r$,

B. Proof of Theorem IV.4

The following lemma will be used in the proof.

Lemma V.5 (See, e.g., [31]) Let \mathbf{H} be a Gaussian matrix, whose entries are iid complex Gaussian random variables with zero-mean and unit variance. Denote its singular value decomposition (SVD) by $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$. Both \mathbf{U} and \mathbf{V} are statistically independent of the diagonal matrix $\mathbf{\Lambda}$.

Let $\mathbf{H}\mathbf{\Pi} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$ be the SVD of the permuted channel matrix, where the diagonal entries of $\mathbf{\Lambda}$ are in non-increasing order. An ordered QR decomposition is denoted by $\mathbf{H}\mathbf{\Pi} = \mathbf{Q}\mathbf{R}$. Let $\mathbf{H}_1 \in \mathbb{C}^{M_r \times i}$ and $\mathbf{V}_1 \in \mathbb{C}^{M_t \times i}$ be the submatrices consisting of the first i columns of $\mathbf{H}\mathbf{\Pi}$ and \mathbf{V}^* , respectively. The i th diagonal entry of \mathbf{R} is (see, e.g., [32])

$$r_{ii}^2 = \frac{1}{[(\mathbf{H}_1^* \mathbf{H}_1)^{-1}]_{ii}} = \frac{1}{[(\mathbf{V}_1^* \mathbf{\Lambda}^2 \mathbf{V}_1)^{-1}]_{ii}}, \quad 1 \leq i \leq N. \quad (53)$$

Let us partition the matrices:

$$\mathbf{V}_1 = \begin{pmatrix} \mathbf{V}_{11} \\ \mathbf{V}_{12} \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{pmatrix}, \quad (54)$$

where $\mathbf{V}_{11} \in \mathbb{C}^{i \times i}$, $\mathbf{V}_{12} \in \mathbb{C}^{(N-i) \times i}$, $\mathbf{\Lambda}_1 \in \mathbb{C}^{i \times i}$, and $\mathbf{\Lambda}_2 \in \mathbb{C}^{(N-i) \times (N-i)}$. Then

$$\mathbf{V}_1^* \mathbf{\Lambda}^2 \mathbf{V}_1 = \mathbf{V}_{11}^* \mathbf{\Lambda}_1^2 \mathbf{V}_{11} + \mathbf{V}_{12}^* \mathbf{\Lambda}_2^2 \mathbf{V}_{12}. \quad (55)$$

Let α be the minimal number such that $\alpha \mathbf{V}_{11}^* \mathbf{V}_{11} \succeq \mathbf{V}_{12}^* \mathbf{V}_{12}$. Such an α exists and is finite with probability one (w.p.1) (Proof: Observe that \mathbf{V}_{11} is nonsingular with probability one. Let $\alpha = \frac{1}{\lambda_{\min}(\mathbf{V}_{11}^* \mathbf{V}_{11})}$ which is finite w.p.1 and we have $\alpha \mathbf{V}_{11}^* \mathbf{V}_{11} \succeq \mathbf{I} \succeq \mathbf{V}_{12}^* \mathbf{V}_{12}$). Now, α is a function of \mathbf{V}_1 and hence is independent of $\mathbf{\Lambda}$ according to Lemma V.5. Because the diagonal of $\mathbf{\Lambda}$ is in non-increasing order,

$$\mathbf{V}_{12}^* \mathbf{\Lambda}_2^2 \mathbf{V}_{12} \preceq \lambda_i^2 \mathbf{V}_{12}^* \mathbf{V}_{12} \preceq \alpha \lambda_i^2 \mathbf{V}_{11}^* \mathbf{V}_{11} \preceq \alpha \mathbf{V}_{11}^* \mathbf{\Lambda}_1^2 \mathbf{V}_{11}. \quad (56)$$

It follows from (55) and (56) that $\mathbf{V}_1^* \mathbf{\Lambda}^2 \mathbf{V}_1 \preceq (1 + \alpha) \mathbf{V}_{11}^* \mathbf{\Lambda}_1^2 \mathbf{V}_{11}$. Invoking the fact that $\mathbf{A}^{-1} \succeq \mathbf{B}^{-1}$ if $\mathbf{A} \preceq \mathbf{B}$, we have

$$(\mathbf{V}_1^* \mathbf{\Lambda}^2 \mathbf{V}_1)^{-1} \succeq \frac{1}{1 + \alpha} (\mathbf{V}_{11}^* \mathbf{\Lambda}_1^2 \mathbf{V}_{11})^{-1} = \frac{1}{1 + \alpha} \mathbf{V}_{11}^{-1} \mathbf{\Lambda}_1^{-2} \mathbf{V}_{11}^{-*}. \quad (57)$$

In the special case where $i = N = M_t$, we have $\mathbf{V}_{11} = \mathbf{V}_1$ and hence $\alpha = 0$. Hence it follows from (53) and (57) that

$$\begin{aligned} r_{ii}^2 &\leq \frac{1 + \alpha}{[\mathbf{V}_{11}^{-1} \mathbf{\Lambda}_1^{-2} \mathbf{V}_{11}^{-*}]_{ii}} = \frac{1 + \alpha}{\sum_{j=1}^i |v_{ij}|^2 \lambda_j^{-2}} \leq \frac{(1 + \alpha) \lambda_i^2}{|v_{ii}|^2}, \\ &1 \leq i \leq N, \end{aligned} \quad (58)$$

where v_{ij} is the (i, j) th entry of \mathbf{V}_{11}^{-1} . As both v_{ii} and α are independent of λ_i , so is $\zeta \triangleq \frac{1 + \alpha}{|v_{ii}|^2}$. Out of the $M_t!$ decoding orderings, we have $\frac{M_t!}{(i-1)!(M_t-i)!}$ different r_{ii}^2 's whose associated ζ 's are indexed as ζ_k , $k = 1, 2, \dots, \frac{M_t!}{(i-1)!(M_t-i)!}$. Denote $r_{ii, \max}^2$ the maximal among the $\frac{M_t!}{(i-1)!(M_t-i)!}$ different r_{ii}^2 's, and $\zeta_{\max} = \max_{1 \leq k \leq \frac{M_t!}{(i-1)!(M_t-i)!}} \{\zeta_k\}$. Then

$$f_{r_{ii}^2}(x) = \int_x^\infty f_{r_{ii}^2, \|\mathbf{g}_i\|^2}(x|y) f_{\|\mathbf{g}_i\|^2}(y) dy \quad (48)$$

$$= \int_x^\infty \frac{x^{M_r-i} (y-x)^{i-2} e^y \left(1 - e^{-y} \sum_{k=0}^{M_t-1} \frac{y^k}{k!}\right)^{M_t-i} \left(e^{-y} \sum_{k=0}^{M_t-1} \frac{y^k}{k!}\right)^{i-1}}{\beta(M_r-i+1, i-1) \beta(M_t+1-i, i) (M_r-1)!} dy$$

$$\text{(denote } w = y - x) = \int_0^\infty \frac{x^{M_r-i} w^{i-2} e^{(x+w)} \left(1 - e^{-(x+w)} \sum_{k=0}^{M_t-1} \frac{(x+w)^k}{k!}\right)^{M_t-i} \left(e^{-(x+w)} \sum_{k=0}^{M_t-1} \frac{(x+w)^k}{k!}\right)^{i-1}}{\beta(M_r-i+1, i-1) \beta(M_t+1-i, i) (M_r-1)!} dw$$

$$= \frac{x^{M_r-i}}{\beta(M_r-i+1, i-1) \beta(M_t+1-i, i) (M_r-1)!} \times$$

$$\int_0^\infty w^{i-2} e^{-M_t(x+w)} \left(\sum_{k=M_t}^\infty \frac{(x+w)^k}{k!}\right)^{M_t-i} \left(\sum_{k=0}^{M_t-1} \frac{(x+w)^k}{k!}\right)^{i-1} dw. \quad (49)$$

$r_{ii, \max}^2 \leq \zeta_{\max} \lambda_i^2$ with ζ_{\max} and λ_i^2 independent of each other. Using this property, we have

$$\mathbb{P}(r_{ii, \max}^2 < \epsilon) \geq \mathbb{P}(\zeta_{\max} \lambda_i^2 < \epsilon) \geq \mathbb{P}(\zeta_{\max} < c) \mathbb{P}(\lambda_i^2 < \epsilon/c), \quad (59)$$

for any positive c . We can find some finite constant c such that $\mathbb{P}(\zeta_{\max} < c)$ is a strictly positive number. Hence

$$\lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(r_{ii, \max}^2 < \epsilon)}{\log \epsilon}$$

$$\leq \lim_{\epsilon \rightarrow 0^+} \frac{\log \mathbb{P}(\lambda_i^2 < \epsilon/c)}{\log \epsilon}$$

$$= (M_t - i + 1)(M_r - i + 1) \quad \text{for } 1 \leq i \leq M_t. \quad (60)$$

where (60) follows from Theorem II.4. Theorem IV.4 is proven.

VI. CONCLUSION

The conclusion is given at the end of [20].

REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 1841–1852, November 1999.
- [3] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multiple-element arrays," *Wireless Personal Communications*, vol. 6, pp. 311–335, March 1999.
- [4] M. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel," *Proceedings of the Thirty-First Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1405–1409, Nov 2-5 1997.
- [5] G. Ginis and J. M. Cioffi, "On the relationship between V-BLAST and the GDFE," *IEEE Communications Letters*, vol. 5, pp. 364–366, September 2001.
- [6] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 1073–1096, May 2003.
- [7] M. Varanasi, "Group detection for synchronous gaussian code-division multiple-access channels," *IEEE Transactions on Information Theory*, vol. 41, pp. 1083–1096, July 1995.
- [8] B. Hassibi, "A fast square-root implementation for BLAST," *Thirty-Fourth Asilomar Conf. Signals, Systems and Computers*, pp. 1255–1259, November 2000.
- [9] J. Benesty, Y. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1722–1730, July 2003.
- [10] M. O. Damen, K. Abed-Meriam, and S. Burykh, "Iterative QR detection for an uncoded space-time communication architecture," *Proc. Allerton Conf. Communications, Control, and Computing*, October 2000.
- [11] N. Prasad and M. K. Varanasi, "Analysis of decision feedback detection for MIMO rayleigh fading channels and the optimization of rate and power allocations," *IEEE Transactions on Information Theory*, vol. 50, pp. 1009–1025, June 2004.
- [12] Y. Jiang, M. Varanasi, and J. Li, "Performance analysis of ZF and MMSE equalizers for MIMO systems: A closer study in high SNR regime," *IEEE Transactions on Information Theory*, Submitted, August, 2006.
- [13] N. Prasad and M. K. Varanasi, "Outage analysis and optimization for multiaccess/V-BLAST architecture over MIMO Rayleigh fading channels," *41st Annual Allerton Conf. on Comm. Control, and Comput., Monticello, IL*, October 2003.
- [14] S. Baro, G. Bauch, A. Pavlic, and A. Semmler, "Improving BLAST performance using space-time block codes and turbo decoding," *Proc. IEEE Globecom*, vol. 2, 2000.
- [15] A. Kannan and J. Barry, "Space-time active rotation (STAR): a new layered space-time architecture," *Proceedings of IEEE GlobeCom*, St. Louis, MO, November 2005.
- [16] D. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple access channel," *IEEE Transactions on Information Theory*, vol. 50, pp. 1859–1874, September 2004.
- [17] G. J. Foschini, Jr., "Layered space-time architecture for wireless communication in a fading environment when using multiple-element antennas," *Bell Labs Tech. Journal*, vol. 1, pp. 41–59, Autumn 1996.
- [18] H. E. Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 968–985, June 2004.
- [19] N. Prasad and M. K. Varanasi, "Analysis and optimization of diagonally layered lattice codes for mimo rayleigh fading channels," *IEEE Transactions on Information Theory*, to appear 2008.
- [20] Y. Jiang and M. Varanasi, "A class of spatial multiplexing architecture combining rate-tailored transmission and ordered BLAST detection Part II: On rates and power allocation," *IEEE Transactions on Wireless Communication*, Submitted, 2006.
- [21] H. A. David, *Order Statistics*. John Wiley & Sons, Inc, 1970.
- [22] Y. Jiang, W. Hager, and J. Li, "The generalized triangular decomposition," Available online: <http://www.sal.ufl.edu/yjiang/papers/gtd.pdf> Submitted.
- [23] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley Series in Probability and Statistics, 1982.

- [24] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge Press, 2005.
- [25] H. E. Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ channel: Diversity-multiplexing-delay tradeoff," *IEEE Transactions on Information Theory*, vol. 52, August 2006.
- [26] S. Tavildar and P. Viswanath, "Approximately universal codes over slow fading channels," *IEEE Transactions on Information Theory*, vol. 52, pp. 3233–3258, July 2006.
- [27] H. Zhang, H. Dai, Q. Zhou, and B. L. Hughes, "Transmit antenna selection for spatial multiplexing systems: A geometrical approach," *IEEE Transactions on Information Theory*, vol. 52, pp. 5297–5311, Dec. 2006.
- [28] Y. Jiang and M. Varanasi, "Diversity-multiplexing tradeoff of MIMO systems with antenna selection," *IEEE Transactions on Information Theory*, submitted.
- [29] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1985.
- [30] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins University Press, 1983.
- [31] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Hanover, MA 02339, USA: now Publishers Inc., 2004.
- [32] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Transactions on Communications*, vol. 42, pp. 1740–1750, February/March/April 1994.