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 $M_t$ and maximum multiplexing gain $M_r$), 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 curve of the optimum (unconstrained) MIMO architecture. The former doesn’t improve the performance of V-BLAST.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) channels have 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 $r \in [0, M_t]$. 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 ordering 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 $(M_t - 1)(1 - r/M_t)$ 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 sub-stream, 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 information 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 \( M_r M_t^3 \), 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

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

\[
y = H \Pi W \hat{z} + z,
\]

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

Different from the V-BLAST scheme, the class of SMAs makes no assumption with regard to the numbers of transmit and receive antennas. Denote \( N = \min\{M_t, M_r\} \). 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 \( X_1, X_2, \ldots, X_n \) are rearranged in ascending order of magnitude and then written as \( X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)} \), we call \( X_{(i)} \) the \( i \)th order statistic \( (i = 1, 2, \ldots, n) \).

**Lemma II.2:** Suppose \( X_1, X_2, \ldots, X_n \) 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 \( X_1, \ldots, X_{n-1} \) are statistically independent and are uniformly distributed over the unit interval \((0, 1)\). Then the pdf of the \( i \)th order statistic is

\[
f_{X_{(i)}}(x) = \frac{1}{\beta(i, n - i)} x^{i-1}(1 - x)^{n-i-1} \quad 0 \leq x \leq 1.
\]
\[ \mathbf{H} = \mathbf{QR}, \] where \( \mathbf{Q} \) is an \( M_t \times N \) 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 \( M_t = N = \min\{M_r, M_t\} \). If \( M_t > M_r = N \), then the first \( N \) 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., \( \mathbf{H}\Pi = \mathbf{QR} \) where \( \Pi \in \mathbb{R}^{M_t \times M_t} \) is a permutation matrix: If \( M_t > N \), then only \( N \) transmit antennas corresponding to the first \( N \) columns of \( \mathbf{H}\Pi \) 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 \( \tilde{\mathbf{H}} \) of (7), which is effectively the \( \mathbf{QR} \) decomposition.

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 encompass various new SMAs based on different decoding ordering rules.

\[ y = \mathbf{QRW}^\frac{1}{2}s + z, \] \( y = \mathbf{QRW}^\frac{1}{2}s + \tilde{z}, \]

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

\[ \tilde{y} = \mathbf{RW}^\frac{1}{2}s + \tilde{z}, \]

\[ \tilde{y} = \mathbf{Q}^*y \] and \( \tilde{z} = \mathbf{Q}^*z \).

The sequential signal decoding, which involves the decision feedback, is as follows:

\[ \begin{align*}
\hat{s}_i &= \mathbf{Q} \left[ \left( \tilde{y}_i - \sum_{j=i+1}^{N} r_{ij} \sqrt{w_j} s_j \right) / r_{ii} \right] \\
\text{end}
\end{align*} \]

where \( r_{ij} \) is the \((i, j)\)th entry of \( \mathbf{R} \) and \( \mathbf{Q}[i] \) 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

\[ \tilde{y}_i = r_{ii} \sqrt{w_i} s_i + \tilde{z}_i, \quad i = 1, 2, \ldots, N, \]

Because \( \mathbb{E}[\tilde{z}\tilde{z}^*] = \sigma^2 \mathbf{I} \), the output SNR of the \( i \)th layer is \( r_{ii}^2 w_i / \sigma^2 = r_{ii}^2 \rho_i \). 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, \( \Pi \) does not depend on \( \mathbf{H} \) and it is well-known that the diagonal elements of \( \mathbf{R} \) are statistically independent with \( \chi^2_{M_t - i + 1} \) 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

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

where \( \mathbb{P}(\mathcal{E}) \) stands for the probability of the event \( \mathcal{E} \). For ordered decoding, \( \Pi \) is a function of \( \mathbf{H} \) and the distributions of \( r_{ii}^2 \) 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 \textit{a priori} information on the statistics of \( \{r_{nn}\}_{n=1}^{N} \) to determine the rates and powers allocated to \( K \leq N \) sub-streams. 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 \( \{r_{nn}\}_{n=1}^{N} \) 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 encompass various new SMAs based on different decoding ordering rules.
which can be achieved using uncoded quadrature amplitude modulation (QAM) [26]. Suppose we have $N$ scalar layers with diversity gains $D_i$, $i = 1, \cdots, N$. The optimal D-M gain tradeoff of the $i$th layer is $D_i(R_i) = D_i(1 - R_i)$, where $R_i$ is the multiplexing gain of the $i$th layer.

Because the class of SMAs apply independent scalar coding to each layer, as input SNR $\rho$ 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:

$$
D(R) = \max_{R_i, \min_i R_i > 0} D_i(1 - R_i) \quad \text{subject to} \quad 0 \leq R_i \leq 1, \quad 1 \leq i \leq N, \quad \sum_{i=1}^{N} R_i = R,
$$

(15)

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

From the observation that $D_i(1 - R_i) = D, \quad \forall i \in \mathcal{I}$, we have $R_i = 1 - \frac{D}{D_i} \quad \forall i \in \mathcal{I}$. According to the constraint $\sum_{i \in \mathcal{I}} R_i = R$, we have $\sum_{i \in \mathcal{I}} (1 - \frac{D}{D_i}) = R$. Thus

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

(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 (15) to be

$$
D(R) = \max_{x \subseteq \mathcal{I}, |x| = K} \frac{K - R}{\sum_{i \in x} D_i^{-1}}.
$$

(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 $D$ must be $\mathcal{I} = \{1, 2, \cdots, K\}$, i.e., the $K$ layers with the highest diversity gains are used. It follows from (17) that

$$
D(R) = \max_{r \leq K \leq N} \frac{K - R}{\sum_{i=1}^{K} D_i^{-1}}.
$$

(18)

It is easily seen that $D(N) = 0$ and $D(0) = D_1$. In the open interval $R \in (0, M_1)$, $D(R)$ is a piecwise linear function and the inflections (changes of derivative) occur when the number of transmit antennas in use changes by one, which happens if

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

(19)

or the overall multiplexing gain

$$
R = K - D_{K+1} \sum_{i=1}^{K} D_i^{-1}.
$$

(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:

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

(21)

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

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

(22)

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

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

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 (\(\{D_i\}_{i=1}^N\)). 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 \(\{||h_i||\}_{i=1}^{M_t}\) where \(h_i\) is the \(i\)th column of \(H\).

(ii) Find permutation matrix \(\Pi\) such that the column norms of the permuted matrix \(H\Pi\), from the left to the right, are in non-increasing order.

(iii) Apply the standard QR decomposition to the permuted matrix \(H\Pi = QR\).

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

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

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

The pdfs of other diagonal elements are

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

for \(i = 2, \ldots, N\). Moreover, we have

\[
\lim_{\epsilon \to 0} \frac{\log \mathbb{P}(r_{11}^2 < \epsilon)}{\log \epsilon} = M_t M_r,
\]

and

\[
\lim_{\epsilon \to 0} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} = M_r - i + 1, \quad i = 2, \ldots, N.
\]

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

\[
D_i = \begin{cases} M_t M_r & i = 1 \\ M_r - i + 1 & 2 \leq i \leq N. \end{cases}
\]

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

Given \(M_t\) and \(M_r\), we can simplify the pdf (25) using Mathematica™.

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 \(\{||h_i||\}_{i=1}^{M_t}\).

(ii) Permute \(h_i\) and \(h_j\) where \(j = \arg \max_{1 \leq i \leq M_t} \{||h_i||\}\). This operation can be represented by \(H_1 = H\Pi_1\) with \(\Pi_1\) being the permutation matrix. (If \(j = 1\), \(\Pi_1\) degrades to be \(I_{M_t}\).)

(iii) Apply a Householder matrix \(Q_1\) to transform the first column of \(H_1\) to a scaled \(e_1\), where \(e_1\) is the first column of \(I_{M_r}\).

The procedure (i–iii) can be illustrated by

\[
\begin{pmatrix}
\times & \times & \times & \times \\
\times & \times & \times & \times \\
\times & \times & \times & \times \\
\times & \times & \times & \times 
\end{pmatrix}
Q_1 H_{1} \rightarrow
\begin{pmatrix}
r_{11} & \times & \times & \times \\
0 & \times & \times & \times \\
0 & 0 & \times & \times \\
0 & 0 & 0 & \times 
\end{pmatrix}.
\]

Note that \(r_{11} = \max \{||h_i|| : 1 \leq i \leq M_t\}\). In the next step, the same procedures are applied to the trailing \((M_t - 1) \times (M_r - 1)\) submatrix on the right hand side of (29), which yields a permutation matrix \(\Pi_2\) and a Householder matrix \(Q_2\). After \(N\) recursive steps, we obtain the desired QR decomposition \(R = Q^T \Pi_N\), or equivalently,

\[
H_{11} = QR
\]

where \(\Pi = \Pi_1\Pi_2 \cdots \Pi_N\) and \(Q = Q_1Q_2 \cdots Q_N\) \((Q_N = I\) 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 \(R\) 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 \(\{r_{11}^2\}_{i=1}^N\).

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

²The Greedy QR decomposition is not new. The built-in Matlab function is \(Q \cdot R \cdot \Pi = QR(H)\).
\( \mathbf{H} = \mathbf{QR} \) be the Greedy QR decomposition. Then
\[
\sum_{j=i}^{N} \lambda_j^2 \leq \frac{r_{ii}^2}{M_i - i + 1} \leq \lambda_i^2 \prod_{j=1}^{i-1} (M_j - j + 1), \quad i = 1, 2, \ldots, N.
\]

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

We first assume that \( M_r \geq M_t \). Then \( \mathbf{H} = \mathbf{QRI}^T \).

It follows from Theorem II.5 that
\[
\prod_{j=1}^{i} r_{jj}^2 \leq \prod_{j=1}^{i} \lambda_j^2, \quad 1 \leq i \leq N.
\]

It follows from the proven lower bound that \( r_{ii}^2 \geq \frac{\lambda_i^2}{M_i - i + 1} \).

Hence
\[
\prod_{j=1}^{i-1} r_{jj}^2 \geq \prod_{j=1}^{i-1} \frac{\lambda_j^2}{M_j - j + 1}.
\]

Combining (33) and (32) yields
\[
r_{ii}^2 \leq \lambda_i^2 \prod_{j=1}^{i-1} (M_j - j + 1), \quad i = 1, 2, \ldots, N.
\]

Now we consider the case \( M_t > M_r \). Denote \( \mathbf{R} \in \mathbb{C}^{N \times N} \) the submatrix consisting of the first \( N \) columns of \( \mathbf{R} \). Then \( \mathbf{H} \triangleq \mathbf{QR} \) is also a submatrix of \( \mathbf{H} \) with singular values \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0 \). According to the same argument which leads to (34), we can prove that
\[
r_{ii}^2 \leq \lambda_i^2 \prod_{j=1}^{i-1} (M_j - j + 1), \quad i = 1, 2, \ldots, N.
\]

Furthermore, because \( \mathbf{HH}^* \geq \mathbf{HH}^* \), \( \exists \) we have \( \lambda_2^2 \geq \lambda_N^2 \) for \( \forall \ i \) [29, Corollary 4.3.3]. Hence
\[
r_{ii}^2 \leq \lambda_i^2 \prod_{j=1}^{i-1} (M_j - j + 1), \quad i = 1, 2, \ldots, N,
\]

also hold for \( M_t > M_r \). The theorem is proven.

It follows from the lower bound in (31) that
\[
\lim_{\epsilon \to 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} \geq \lim_{\epsilon \to 0^+} \frac{\log \mathbb{P}(\lambda_i^2 < (M_t - i + 1)\epsilon)}{\log \epsilon} = \lim_{\epsilon \to 0^+} \frac{\log \mathbb{P}(\lambda_i^2 < \epsilon)}{\log \epsilon}.
\]

(see Theorem II.4) \( = (M_t - i + 1)(M_t - i + 1). \)

On the other hand, it follows from the upper bound in (31) that
\[
\lim_{\epsilon \to 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} \leq \lim_{\epsilon \to 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} (M_t - i + 1).
\]

Therefore \( \lim_{\epsilon \to 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} = (M_t - i + 1)(M_t - i + 1) \).

Now we have proven the following theorem.

**Theorem IV.3:** The \( i \)th layer of ZF-VB decoder based on the greedy ordering rule, has diversity gain
\[
D_i = (M_t - i + 1)(M_t - i + 1), \quad 1 \leq i \leq N.
\]

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 \( \{D_i\}_{i=1}^N \)**

Rewrite the ordered QR decomposition \( \mathbf{H} = \mathbf{QR} \). Let us write a permuted channel matrix in its column form: \( \mathbf{H} = (h_{\pi(1)}, \cdots, h_{\pi(i-1)}, h_{\pi(i)}, \cdots, h_{\pi(M_t)}) \). Denote \( \mathbf{H}_i \triangleq [h_{\pi(1)}, \cdots, h_{\pi(i-1)}] \). Then \( \mathbf{H}_i^\dagger = \mathbf{H}_i^\dagger \mathbf{H}_i \mathbf{H}_i^\dagger \), where \( \mathbf{H}_i^\dagger = \mathbf{I} - \mathbf{H}_i (\mathbf{H}_i^\dagger \mathbf{H}_i)^{-1} \mathbf{H}_i^\dagger \). Thus \( \mathbf{H}_i^\dagger \) is a function of \( \mathbf{H}_i \) and \( \mathbf{H}_i \). and it is invariant to the column permutation of \( \mathbf{H}_i \). Hence out of the \( M_t \) ! ordering rule, one may have up to \( \left(M_t - i + 1 \right) \left(M_t - i + 1 \right) \) 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 \( D_i \), \( i = 1, \ldots, N \). 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 \( \mathbf{H} = \mathbf{QRI}^T \) where \( \Pi \) is a permutation matrix dependent on \( \mathbf{H} \). Let \( r_{ii} \) be the \( i \)th diagonal of \( \mathbf{R} \). The inequality
\[
\lim_{\epsilon \to 0^+} \frac{\log \mathbb{P}(r_{ii}^2 < \epsilon)}{\log \epsilon} \leq (M_t - i + 1)(M_t - i + 1), \quad 1 \leq i \leq N.
\]

holds for any ordering rule. In other words, the diversity gain of the \( i \)th layer
\[
D_i \leq (M_t - i + 1)(M_t - i + 1), \quad 1 \leq i \leq N.
\]

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

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 $x$ is isotropic if for any unitary matrix $U$, $x \sim Ux$, i.e., $x$ has a distribution invariant under rotations and reflections.

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

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

Proof: The associated Householder matrix of $x$ is $Q = I - 2ww^*$, where [30]

$$ w = \frac{x - \|x\|e}{\|x - \|x\|e\|} = \frac{x/\|x\| - e}{\|x/\|x\| - e\|}. \tag{42} $$

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

The Norm QR decomposition algorithm applies the QR decomposition to a matrix $G \triangleq HH^*$, where $H$ is the permutation matrix such that $\|g_i\|, i = 1, \ldots, M$, are in a non-decreasing ordering. The squared norms of the unordered columns are $\chi^2_{N,M_r}$ random variables with pdf $f(x) = \frac{x^{M_r-i-1}}{M_r-i}e^{-x}$, $x \geq 0$, 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, $\|g_i\|^2$ has distribution

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

Hence $r^2_{1i} = \|g_i\|^2$ has distribution given in (23).

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

$$ Q_1^*G = \begin{pmatrix} r_{11} & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{pmatrix}. \tag{44} $$

The column vectors of $G$ 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 $Q_1$ is independent of $\{g_i\}_{i=2}^{M_r}$. Hence the trailing $(M_r) \times (M_r - 1)$ submatrix in the right hand side of (44) (represented by $*$’s) has the same distribution as its counterpart in $G$. Hence we see that $r^2_{1i}$ has the same distribution as $\|g_1(2: M_r, i)\|^2$, where $g_1(i: M_r)$ consists of the last $M_r - i + 1$ elements of $g_1$. Following the same argument, we can further show that $r^2_{1i}$ has the same distribution as $\|g_1(i: M_r)\|^2$. To further derive the pdfs of the diagonal $\{r^2_{1i}\}$, we establish the following Lemma.

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

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

Proof: For the Gaussian vector $h \sim N(0, I_{M_r})$, the squared absolute value of each element is a variable of exponential distribution. Hence we can regard $\|h(M_r - i + 1: M_r)\|^2, i = 1, 2, \cdots, N$ 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., $\|h\|^2 = y$, 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

Fig. 2. Comparison of the D-M tradeoffs.
are randomly dropped on an interval $(0, y)$ and the positions of the points from the left to the right are indexed as $x_1, x_2, \ldots, x_{M_r - 1}$, then according to Corollary II.3, the pdf of $x_i$ is
\[
f_{x_i}(x) = \frac{x^{M_r-i-1}(y-x)^{M_r-i-1}}{\beta(i, M_r - i)y^{M_r-1}}, \quad 0 \leq x \leq y.
\]
As $\|h(i : M_r)\|^2 = x_{M_r-i+1}$, replacing $i$ in (46) by $M_r - i + 1$, we see that the pdf of $x_{M_r-i+1}$ is
\[
f_{x_{M_r-i+1}}(x) = \frac{x^{M_r-i}(y-x)^{M_r-i-2}}{\beta(M_r - i + 1, i - 1)y^{M_r-1}}, \quad 0 \leq x \leq y.
\]

The lemma is proven.

Based on Lemma V.4 and the observation that $\{r_{ii}^{M_r}\}_{i=1}^{M_r}$ have the same distribution as $\|g_i(i : M_r)\|^2$, we can calculate the distribution of $r_{ii}^{M_r}$ ($i \geq 2$) as (48) (see the top of the next page). Denote the integral in (49) as $J(x)$. It is easy to see that
\[
0 < J(0) = \int_0^\infty u^{i-2}e^{-u}du - \sum_{k=0}^{M_r-1} \frac{u^k}{k!}^{M_r-i} \left( \sum_{k=0}^{M_r-1} \frac{u^k}{k!} \right)^{i-1} du
\]
\[
< \int_0^\infty u^{i-2}e^{-u}du = (i-2)!
\]
On the other hand, letting $y = x + w$, we have
\[
J(x) = \int_x^\infty (y-x)^{i-1}e^{-M_ry} \left( \sum_{k=M_t}^{M_t} \frac{y^k}{k!} \right)^{M_t-i} \left( \sum_{k=0}^{M_t-1} \frac{y^k}{k!} \right)^{i-1} du
\]
which is straightforward to show that the derivative $J'(x)$ is finite. Hence, $J(x)$ is a continuous function and by Taylor Expansion, $J(x) = J(0) + J'(0)x$ for some $x \in (0, x)$. Now we prove that in the neighborhood of the origin
\[
J_{\tau_1}(x) = \tau_1 x^{M_t-i} + o(x^{M_t-i})
\]
for some $\tau_1 > 0$. Thus we conclude that
\[
\lim_{\epsilon \to 0^+} \frac{\log P(r_{ii}^2 < \epsilon)}{\log \epsilon} = M_r - i + 1, \quad \text{for } 2 \leq i \leq M_t.
\]
Hence the ith ($i \geq 2$) layer has diversity gain of only $M_r - i + 1$. As for $i = 1$, it can be seen from (23) that
\[
f_{r_{1i}^2}(x) = \frac{1}{M_t} x^{M_t-1} \left( \frac{x}{M_t} \right)^{M_t-1} \left( \frac{1}{M_t} \right)^{M_t}, \quad 0 \leq x \ll 1
\]
for some $\eta > 0$. Hence $\lim_{\epsilon \to 0^+} \frac{\log P(r_{1i}^2 < \epsilon)}{\log \epsilon} = M_tM_r$, i.e., the first layer has diversity gain of $M_tM_r$.

B. Proof of Theorem IV.4

The following lemma will be used in the proof.

Lemma V.5 (See, e.g., [31]) Let $H$ 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 $H = U \Sigma V^*$. Both $U$ and $V$ are statistically independent of the diagonal matrix $\Lambda$.

Let $H' = U' \Sigma V^*$ be the SVD of the permuted channel matrix, where the diagonal entries of $\Lambda$ are in non-increasing order. An ordered QR decomposition is denoted by $H' = QR$. Let $H_1 \in C^{M_t \times i}$ and $V_1 \in C^{M_t \times i}$ be the submatrices consisting of the first $i$ columns of $H_1$ and $V^*$, respectively. The $i$th diagonal entry of $R$ is (see, e.g., [32])
\[
r_{1i}^2 = \frac{1}{|H_1^*H_1|^{-1}_{ii}} = \frac{1}{|(V_1^*\Lambda_1^2V_1^{-1})|^{-1}_{ii}}, \quad 1 \leq i \leq N.
\]
Let us partition the matrices:
\[
V_1 = \begin{pmatrix} V_{11} & V_{12} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix},
\]
where $V_{11} \in C^{i \times i}$, $V_{12} \in C^{(N-i) \times i}$, $\Lambda_1 \in C^{i \times i}$, and $\Lambda_2 \in C^{(N-i) \times (N-i)}$. Then
\[
V_1^*\Lambda_1^2V_1 = V_{11}^*\Lambda_1^2V_{11} + V_{12}^*\Lambda_2^2V_{12}.
\]
Let $\alpha$ be the minimal number such that $\alpha V_{11}^*V_{11} \preceq V_{12}^*V_{12}$. Such an $\alpha$ exists and is finite with probability one (w.p.1) (Proof: Observe that $V_{11}$ is nonsingular with probability one. Let $\alpha = \frac{\max(V_{11}^*V_{11})}{\min(V_{11}^*V_{11})}$ which is finite w.p.1 and we have $\alpha V_{11}^*V_{11} \succeq I \succeq V_{12}^*V_{12}$). Now, $\alpha$ is a function of $V_1$ and hence is independent of $\Lambda$ according to Lemma IV.5. Because the diagonal of $\Lambda$ is in non-increasing order,
\[
V_{11}^*\Lambda_1^2V_{11} \preceq \alpha \Lambda_1^2V_{11} \preceq \alpha V_{11}^*V_{11} \preceq \alpha \Lambda_1^2V_{11}. \tag{56}
\]
It follows from (55) and (56) that $V_1^*\Lambda_1^2V_1 \preceq (1 + \alpha)V_1^*V_1$. Invoking the fact that $\Lambda_1^{-1} \succeq B^{-1}$ if $\alpha \leq B$, we have
\[
(V_1^*\Lambda_1^2V_1)^{-1} \geq \frac{1}{1 + \alpha} (V_1^*V_1)^{-1} = \frac{1}{1 + \alpha} V_1^*\Lambda_1^{-2}V_1^{-1}. \tag{57}
\]
In the special case where $i = N = M_t$, we have $V_{11} = V_1$ and hence $\alpha = 0$. Hence it follows from (53) and (57) that
\[
r_{1i}^2 \leq \frac{1 + \alpha}{|V_{11}^*\Lambda_1^{-2}V_{11}^{-1}|} = \frac{1 + \alpha}{\sum_{j=1}^{M_t} |v_{ij}|^2 \lambda_j^{-2}} \leq \frac{1}{|v_{ii}|^2}, \quad 1 \leq i \leq N, \tag{58}
\]
where $v_{ij}$ is the $(i, j)$th entry of $V_{11}^{-1}$, i.e., $v_{ij} = \zeta_i \zeta_j$. As both $v_{ii}$ and $\alpha$ are independent of $\Lambda$, $\zeta_i$ is independent of $\Lambda$, so is $\zeta_i \zeta_i = \frac{1}{|v_{ii}|^2}$, $i \neq j$. Out of the $M_t!$ different $r_{ii}^2$’s whose associated $\zeta$’s are indexed as $\zeta_k$, $k = 1, 2, \ldots, \frac{M_t!}{(i-1)!(M_t-i)!}$. Denote $r_{ii,\text{max}}$ the maximal among the $\frac{M_t!}{(i-1)!(M_t-i)!}$ different $r_{ii}^2$’s, and $\zeta_{\text{max}} = \max_{1 \leq k \leq \frac{M_t!}{(i-1)!(M_t-i)!}} |\zeta_k|$. Then
\[ f_{z_i^2}(x) = \int_x^{\infty} \frac{1}{y} f_{\|z\|^2}(x/y) f_{\|z\|^2}(y) dy \]
= \int_x^{\infty} x^{M_i-1-y} (y-x)^{i-2} e^{y} \left( 1 - e^{-y} \sum_{k=0}^{M_i-1} \frac{y^k}{k!} \right) \frac{1}{y} \left( 1 - e^{-y} \sum_{k=0}^{M_i-1} \frac{y^k}{k!} \right) \frac{1}{y} \left( 1 - e^{-y} \sum_{k=0}^{M_i-1} \frac{y^k}{k!} \right) dy \]

\[ \text{(denote } w = y - x) \]
= \int_0^{x^{M_i-1}} \frac{x^{M_i-1}}{\beta(M_i - i + 1, i - 1)} \frac{x^{M_i-1}}{\beta(M_i + 1 - i, i)(M_i - 1)!} \times \int_0^{\infty} (x + w)^k \left( \sum_{k=0}^{M_i-1} \frac{(x + w)^k}{k!} \right) \frac{1}{y} \left( 1 - e^{-y} \sum_{k=0}^{M_i-1} \frac{y^k}{k!} \right) \frac{1}{y} \left( 1 - e^{-y} \sum_{k=0}^{M_i-1} \frac{y^k}{k!} \right) \frac{1}{y} \left( 1 - e^{-y} \sum_{k=0}^{M_i-1} \frac{y^k}{k!} \right) dw. \]  

(48)

\[ r_{i_{\text{max}}}^2 \leq \zeta_{\text{max}} \lambda_i^2 \] with \( \zeta_{\text{max}} \) and \( \lambda_i^2 \) independent of each other. Using this property, we have:
\[ P\left( r_{i_{\text{max}}}^2 < \epsilon \right) \geq P\left( \zeta_{\text{max}} \lambda_i^2 < \epsilon \right) \geq P\left( \zeta_{\text{max}} < \epsilon / c \lambda_i^2 \right). \] (59)

for any positive \( c \). We can find some finite constant \( c \) such that \( P(\zeta_{\text{max}} < c) \) is a strictly positive number. Hence
\[ \lim_{c \to 0^+} \frac{\log P\left( r_{i_{\text{max}}}^2 < \epsilon \right)}{\log \epsilon} \leq \lim_{c \to 0^+} \frac{\log P\left( \zeta_{\text{max}} \lambda_i^2 < \epsilon / c \right)}{\log \epsilon} = (M_i - i + 1)(M_i - i + 1) \text{ for } 1 \leq i \leq M_i. \] (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


