Performance-Complexity Analysis for MIMO Multiple-Access Channels with Lattice based Sphere Decoders and User Selection

Arun Singh, Petros Elia, and Hsiao-feng (Francis) Lu

***Abstract*—This paper extends the complexity analysis [1]–[3] for MIMO point-to-point systems, to the case of the multiple access channel (MAC). Specifically we derive upper bounds on the *complexity exponent* required to achieve the optimal diversity multiplexing-gain tradeoff (DMT) performance $d^*_{\text{mac}}(r)$ of the symmetric MAC with K single antenna transmitters and a receiver having n_r antennas. Finally we show how utilizing a few bits of feedback to allow for user selection, can in fact result in substantial complexity reductions, while at the same time improving the DMT performance.**

I. Introduction

This work extends recent studies on the rate-reliability-complexity limits of quasi-static point-to-point MIMO communications [1]–[3] and focuses on the multiple access channel (MAC). Specifically we are interested in establishing bounds on the computational costs required for the broad families of maximal likelihood (ML)- and lattice-based sphere decoders to achieve the optimal diversity-multiplexing tradeoff (DMT) of the symmetric MAC. We here note that while lattice reduction (LR) indeed allows for DMT optimal behavior at very manageable complexity [4], there exist scenarios for which these same LR methods cannot be readily applied. It is for this exact reason that we focus on the complexity analysis of non-LR-aided schemes which remain of strong interest for many pertinent communication scenarios.

Consider a symmetric MAC with K single-antenna users, each operating at the same multiplexing gain r, and a receiver at base station having n_r antennas. We assume that the receiver employs a joint ML (or joint lattice) decoder. The emphasis on joint decoders stems from their DMT optimality for symmetric MAC [5], [6]. These joint ML or lattice decoders are implemented as bounded search sphere decoding (SD) algorithms with a search radius $\delta := \sqrt{z \log_2 \text{SNR}}$, for a properly chosen $z > 0$, and where SNR denotes the signal to noise ratio. For SD algorithm details, readers are referred to [1], [2] however for clarity of exposition, wherever necessary, essential details are provided during the complexity analysis. The derivations focus on ML-based decoding, but given the performance-and-complexity equivalence of ML and regularized lattice based decoding [2], these same results extend automatically to the latter. We note that the validity of the presented bounds depends on the existence of actual coding schemes that meet them. These schemes will be here provided, together with the associated lattice designs, decoders and halting polices.

A. System Model

The received signal matrix Y at the base station for the K-user MAC under consideration is

$$Y = \sqrt{\text{SNR}} \sum_{i=1}^{K} \underline{h}_i \, \underline{x}_i^\top + W \;=\; \sqrt{\text{SNR}} H_{\text{eq}} X + W, \qquad (1)$$

where $\underline{h}_i \sim \mathbb{C}\mathcal{N}(\underline{0}, \mathtt{I}_{n_r})$ is the length-n_r channel vector of the ith user. It is modeled as a complex Gaussian random vector with zero mean and covariance $\mathtt{I}_{n_r}$. $H_{\text{eq}} = [\underline{h}_1 \ \cdots \ \underline{h}_K]$ represents the equivalent $(n_r \times K)$ channel fading matrix. The length-T vector $\underline{x}_i$ is the code vector sent by the ith user satisfying an average power constraint $\mathbb{E}\, \|\underline{x}_i\|^2 \le T$ for all $i = 1, \ldots, K$, and $X = [\underline{x}_1 \ \cdots \underline{x}_K]^\top$. W is the noise matrix with i.i.d. $\mathbb{C}\mathcal{N}(0,1)$ entries at the receiver. After vectorization, the real valued representation of (1) takes the form

$$\underline{y} = \sqrt{\text{SNR}} H \underline{x} + \underline{w} \qquad (2)$$

where $\underline{x} = \left[\text{Re}\{\underline{x}_1^\top\} \, \text{Im}\{\underline{x}_1^\top\} \cdots \text{Re}\{\underline{x}_K^\top\} \, \text{Im}\{\underline{x}_K^\top\}\right]^\top$,

$$H = \mathtt{I}_T \otimes \begin{bmatrix} \text{Re}\{H_{\text{eq}}\} & -\text{Im}\{H_{\text{eq}}\} \\ \text{Im}\{H_{\text{eq}}\} & \text{Re}\{H_{\text{eq}}\} \end{bmatrix}, \qquad (3)$$

and where $\underline{y}$ and $\underline{w}$ are defined similar to $\underline{x}$.

For a rate $R = r \log_2 \text{SNR}$ that scales with SNR as a function of the multiplexing gain r, we consider the case when the overall codeword $\underline{x}$ is taken from a (sequence of) full-rate linear (lattice) code(s) $\mathcal{X}_r = \mathcal{X}_{r,1} \oplus \cdots \oplus \mathcal{X}_{r,K}$, where for the ith user, $\mathcal{X}_{r,i} = \Lambda_{r,i} \cap \mathcal{R}_{r,i} \subset \mathbb{R}^{2T}$ is corresponding lattice code. The ith component code $\mathcal{X}_{r,i}$ consists of elements in a rank $2T$ lattice $\Lambda_{r,i}$ that lie inside the *shaping region* $\mathcal{R}_{r,i}$, which is properly chosen to meet the rate requirement $|\mathcal{X}_{r,i}| = 2^{RT}$ as well as the average power constraint. The region $\mathcal{R}_{r,i}$ is a compact convex subset of $\mathbb{R}^{2T}$.

Specifically, we set $\Lambda_{r,i} := \text{SNR}^{-\frac{r}{2}} \Lambda_i$, a scaled lattice of another lattice Λ_i, whose generator matrix is denoted by G_i. Set $G = \text{diag}(G_1, \ldots, G_K)$; then the overall codeword is given by $\underline{x} = \text{SNR}^{-\frac{r}{2}} G \underline{s}$ for some $\underline{s} \in \mathbb{Z}^{2KT}$. Substituting this into (2) yields the following equivalent channel input-output relatio,n which will be used for sphere decoding of $\underline{s}$

$$\underline{y} = M \underline{s} + \underline{w} \qquad (4)$$

where $M := \text{SNR}^{\frac{1-r}{2}} H G$.

B. Rate-reliability-complexity measures in outage-limited communications

In the high SNR regime, a given encoder $\mathcal{X}_r$ and decoder $\mathcal{D}_r$ are said to achieve a *multiplexing gain* r and *diversity gain* $d_{\mathcal{D}}(r)$ [7] if

$$-\lim_{\text{SNR}\to\infty} \frac{\log_2 P_e(r)}{\log_2 \text{SNR}} = d_{\mathcal{D}}(r), \tag{5}$$

where $P_e(r)$ denotes the probability of codeword error with an SD-based ML decoder $\mathcal{D}_r$ employing time-out policies.

The complexity characterization follows from [1], [2]. Given multiplexing gain r, let $N_{\max}(r)$ denote the amount of computational reserves, in floating point operations (flops) per T channel uses, that the decoder $\mathcal{D}_r$ is endowed with, in the sense that after $N_{\max}(r)$ flops, the decoder must simply terminate, potentially prematurely and before completion of its task. The complexity exponent then takes the form [2]

$$c_{\text{mac}}(r) := \lim_{\text{SNR}\to\infty} \frac{\log_2 N_{\max}(r)}{\log_2 \text{SNR}}. \tag{6}$$

C. Notation

Following [7], we use $\doteq$ to denote the *exponential equality*, i.e., a function $f(\text{SNR})$ is said to be $f(\text{SNR}) \doteq \text{SNR}^b$ if and only if $\lim_{\text{SNR}\to\infty} \frac{\log_2 f(\text{SNR})}{\log_2 \text{SNR}} = b$. Exponential inequalities such as $\dot{\leq}$, $\dot{\geq}$ are similarly defined. By $s = \lceil x \rceil$ we mean the smallest integer $s \geq x$, and by $t = \lfloor x \rfloor$ we mean the largest integer $t \leq x$. $A^\dagger$ is the Hermitian transpose of matrix A, and $(x)^+ := \max\{x, 0\}$.

II. Complexity Analysis for Multiple-Access Channel

In order to establish complexity requirements for the symmetric MAC, we briefly recall that the optimal DMT performance of the K-user MAC under consideration is given by [5]

$$d^*_{\text{mac}}(r) = \begin{cases} n_r(1-r) & \text{if } 0 < r \leq \frac{n_r}{K+1}, \\ d^*_{K,n_r}(Kr) & \text{if } \frac{n_r}{K+1} < r \leq \frac{n_r}{K}, \end{cases} \tag{7}$$

where $d^*_{m,n}(r)$ denotes the optimal DMT of an $(n \times m)$ MIMO channel, see [7] for its exact characterization. The regime $0 < r \leq \frac{n_r}{K+1}$ is termed the the lightly-loaded regime, where single user DMT performance $d^*_{1,n_r}(r) = n_r(1-r)$ can be achieved, as if there was no multiuser interference. The regime $\frac{n_r}{K+1} < r \leq \frac{n_r}{K}$ is termed the heavily-loaded regime, also known as the antenna-pooling regime [5]. Depending on the values of K and n_r, constructions of lattice coding schemes that achieve the optimal DMT performance $d^*_{\text{mac}}(r)$ can be found in [6], [8], [9].

We focus on establishing upper bounds on the complexity exponent that guarantees DMT optimal ML-based (or lattice-based) decoding. This will be achieved by considering specific codes, decoders and halting policies, as will be seen in the following theorem.

Theorem 1: For the K-suer MAC subject to i.i.d. Rayleigh fading statistics, the minimum, over all lattice designs and halting and decoding order policies, complexity exponent $c^*_{\text{mac}}(r)$ required to achieve the optimal DMT $d^*_{\text{mac}}(r)$, is upper bounded by

$$\bar{c}_{\text{mac}}(r) = \begin{cases} \sup\limits_{\underline{\mu}\in\mathcal{B}(r)} (K-n_r)r + \sum_{i=1}^{\nu} \left(r - (1-\mu_i)^+\right)^+, \\ \qquad\qquad\qquad\qquad \text{if } K \geq n_r \\ \sup\limits_{\underline{\mu}\in\mathcal{B}(r)} \sum_{i=1}^{\nu} \left[\min\{r, r+\mu_i - 1\}\right]^+, \\ \qquad\qquad\qquad\qquad \text{if } K < n_r, \end{cases} \tag{8}$$

where $\nu := \min\{K, n_r\}$ and

$$\mathcal{B}(r) := \left\{ \underline{\mu} : \begin{array}{l} \mu_1 \geq \cdots \geq \mu_\nu, 0 \leq \mu_i \in \mathbb{R} \\ \sum_{i=1}^{\nu} \left(|K-n_r| + 2i - 1\right)\mu_i \leq d^*_{\text{mac}}(r) \end{array} \right\}.$$

Moreover, the uncoded QAM signaling achieves the complexity upper bound $\bar{c}_{\text{mac}}(r)$ and delivers the optimal DMT $d^*_{\text{mac}}(r)$, using a sphere decoder with a search radius $\delta > \sqrt{d^*_{\text{mac}}(r)\log_2 \text{SNR}}$, a decoding halting policy that halts decoding if $N_{\max}(r) \doteq \text{SNR}^{c^*_{\text{mac}}(r)}$, and any decoding order policy.

Proof: See Appendix A. ■

We remark that for an underdetermined MAC, i.e., $n_r < K$, there is an intuitive explanation for the term $(K-n_r)r$ appearing in (8). Note that with $n_r < K$, the QR decomposition of matrix M defined in (4) results in an upper trapezoid matrix R, whose bottom row contains $2T(K-n_r)+1$ nonzero entries. Therefore, prior to processing the root node of a sphere-decoding tree, the sphere decoder must first search exhaustively among $N^{2T(K-n_r)}$ combinations of N-ary PAM constellation points as entries of $\underline{s}$ are in $\mathbb{Z}$ in (4). In particular, for an uncoded QAM signaling we have $N = \text{SNR}^{\frac{r}{2}}$ and $T = 1$, thereby yielding the first term $(K-n_r)r$ in (8).

To provide more meaningful insights regarding the upper bound presented in Theorem 1, we present examples for two specific cases with $n_r = 1$ and $n_r = K$.

Example 1: For the specific case of $n_r = 1$, the bound $\bar{c}_{\text{mac}}(r)$ in (8) simplifies to

$$\bar{c}_{\text{mac}}(r) = (K-1)r, \quad \text{for } 0 \leq r \leq \frac{1}{K}. \tag{9}$$

It is clear that the complexity exponent upper bound grows almost linearly in K, somehow unfavorable in practice. Fig. 1 plots the upper bounds for the $K = 4$ and $K = 5$-user cases with a single-antenna receiver, i.e. $n_r = 1$. ■

Example 2: For the specific case of $n_r = K$, the optimal DMT of (7) can be achieved by V-BLAST [5], [8]. The complexity exponent upper bound for this case simplifies after some work to

$$\bar{c}_{\text{mac}}(r) = r\left\lfloor\sqrt{K(1-r)}\right\rfloor + \left(r - 1 + \frac{K(1-r) - (\left\lfloor\sqrt{K(1-r)}\right\rfloor)^2}{2\left\lfloor\sqrt{K(1-r)}\right\rfloor + 1}\right)^+. \tag{10}$$

Fig. 2 shows the complexity exponent upper bounds for $K = 3, 4, 5$ users and $n_r = K$. ■

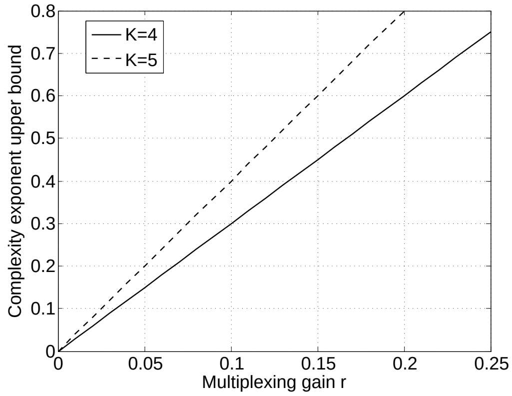

Fig. 1. Complexity exponent bounds for K-user MAC with $n_r = 1$.

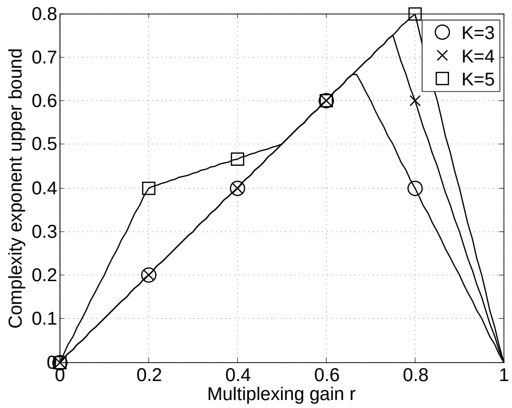

Fig. 2. Complexity exponent for K-user MAC with $n_r = K$.

We shall mention that the proof of Theorem 1 in Appendix A shows the uncoded QAM is indeed DMT optimal for the present MAC, for all possible values of K and n_r, as long as the equivalent channel matrix H_{eq} has an isotropic probability distribution, i.e., invariant under any unitary transformation. It improves the result reported by Tse and Viswanath [10, p. 3246] who showed that plain uncoded QAM using in V-BLAST is DMT optimal when the multiplexing gain r lies in the lightly-loaded regime, or equivalently when $n_r \geq K$. The proof in Appendix A shows further that for $n_r < K$ the resulting DMT becomes dominated by $d^*_{K,n_r}(Kr)$, as if the system operates in the antenna-pooling regime. Similarly, the proof in Appendix A improves one of our previous results [8]. We summarize the above discussion in the theorem below.

Theorem 2: Consider a K-user symmetric MAC with single antenna at each user and n_r antennas at base station. The uncoded QAM for each user is approximately universal in terms of DMT for all isotropic channel probability distributions. ■

III. User Selection for the Multiple-Access Channel

In this section we present a user selection scheme for the present MAC system that can achieve a substantial improvement over $d^*_{\text{mac}}(r)$ for many values of multiplexing gain r, together with an exponential reduction in the complexity costs compared to ML/lattice based sphere decoding complexity costs $c^*_{\text{mac}}(r)$, as these were presented in Theorem 1. The user selection scheme is based on the antenna selection algorithm of [11] by Jiang and Varanasi to select L out of K users for transmission throughout a block fading channel of length T_c channel uses. Hence such selection scheme requires a channel feedback at rate $\frac{1}{T_c}\log_2\binom{K}{L}$ in bits per channel use. As T_c is very large in practice, the required feedback rate is extremely low. Assuming L out of K users are chosen for transmission, as each user is selected with probability $\frac{L}{K}$, it means each selected user has to transmit at multiplexing gain $\frac{K}{L}r$.

It should be noted that there is a potential limitation of the Jiang-Varanasi antenna selection algorithm when it is used for user selection. More precisely, the Jiang-Varanasi algorithm is based on a specific kind of QR decomposition of the overall channel matrix H_{eq} and then makes a selection according to the order of the diagonal entries of the R matrix, which can be either a tall or a flat matrix, depending on the values of K and n_r. This means that the algorithm can make a selection of L users if and only if $L \leq \nu = \min\{K, n_r\}$.

A. DMT and Complexity of User Selection

The performance of extent of system reduction is naturally limited by the rate-reliability requirements. The theorem below provides an upper bound on the DMT performance achieved by the user selection scheme.

Theorem 3: Given L with $1 \leq L \leq \nu = \min\{K, n_r\}$, the DMT achieved by selecting L users based on Jiang-Varanasi algorithm and transmitting over the present MAC is upper bounded by

$$d^*_{\text{us},L}(r) \;\leq\; \min_{\substack{k\geq 0,,\ell\geq 1\\ k+\ell\leq L}} d_{k,\ell}\left(\frac{\ell K r}{L}\right) \;:=\; \bar{d}_{\text{us},L}(r), \qquad (11)$$

where

$$d_{k,\ell}(r) := \inf_{\mathcal{A}_\ell(r)} D_{k,\ell}(\underline{\alpha}),$$

$$\mathcal{A}_\ell(r) := \left\{0 \leq \alpha_1 \leq \cdots \leq \alpha_\ell : \sum_{i=1}^{\ell}(1-\alpha_i)^+ \leq r\right\},$$

and

$$D_{k,\ell}(\underline{\alpha}) := \sum_{i=1}^{\ell}(n_r + \ell - 2i + 1)\alpha_i + \sum_{i=1}^{\ell-1}(K-k-\ell)\alpha_i + \alpha_\ell(K-k-\ell)(n_r-k-\ell+1) \quad (12)$$

Proof: See Appendix B. ■

Having obtained an upper bound on the DMT performance achieved by the L-user selection algorithm, we next explore the complexity ramifications of user selection. The following result is a direct consequence of Theorem 1.

Theorem 4: For the K-suer MAC subject to i.i.d. Rayleigh fading statistics, the minimum, over all lattice designs and halting and decoding order policies, complexity exponent $c^*_{\text{us},L}(r)$ based on the proposed L-user selection algorithm,

$1 \leq L \leq \nu = \min\{K, n_r\}$, required to achieve the optimal DMT $d^*_{\text{us},L}(r)$, is upper bounded by

$$\bar{c}_{\text{us},L}(r) = \sup_{\underline{\alpha} \in \mathcal{F}(r)} \sum_{i=1}^{L} \left[\min\left\{ \frac{K}{L} r, \frac{K}{L} r + \alpha_i - 1 \right\} \right]^+, \quad (13)$$

where

$$\mathcal{F}(r) := \left\{ \underline{\alpha} : \begin{array}{l} \alpha_1 \leq \cdots \leq \alpha_L, 0 \leq \alpha_i \in \mathbb{R} \\ D_{0,L}(\underline{\alpha}) \leq \bar{d}_{\text{us},L}(r) \end{array} \right\}.$$

where $D_{k,\ell}(\underline{\alpha})$ is defined in (12). Moreover, the uncoded QAM signaling achieves the complexity upper bound $c^*_{\text{us},L}(r)$ and delivers the optimal DMT $d^*_{\text{us},L}(r)$, given given a sphere decoder with a search radius $\delta > \sqrt{d^*_{\text{us},L}(r) \log_2 \text{SNR}}$, a decoding halting policy that halts decoding if $N_{\max}(r) \doteq \text{SNR}^{c^*_{\text{us},L}(r)}$, and any decoding order policy.

Proof: The claim on the upper bound $\bar{c}_{\text{us},L}(r)$ follows from Theorem 1 and the fact that $\bar{d}_{\text{us},L}(r) \geq d^*_{\text{us},L}(r)$, which shows to achieve the smaller diversity value $d^*_{\text{us},L}(r)$, only less computational complexity is needed. The claim on the optimality of uncoded QAM follows from Theorem 2 of being approximately universal. ■

In Fig. 3 we consider the underdetermined MAC for $K = 4$ and $n_r = 3$ and plot the DMT upper bounds $\bar{d}_{\text{us},L}(r)$ giving in Theorem 3 for $L = 1, 2, 3$. We also compare these DMT upper bounds to the optimal MAC DMT $d^*_{\text{mac}}(r)$ without selection. It can be seen that the user selection algorithm can provide a significant increase in diversity whenever multiplexing gain $r \leq 0.42$.

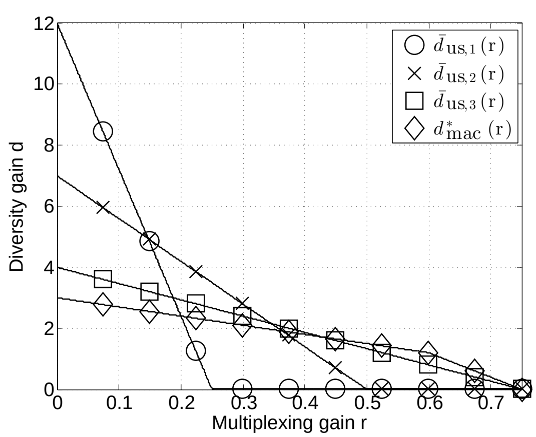

Fig. 3. The DMT upper bounds $\bar{d}_{\text{us},L}(r)$ for user selection and the optimal MAC DMT $d^*_{\text{mac}}(r)$ without selection for $K = 4$, $n_r = 3$, and $L = 1, 2, 3$.

Another example is given in Fig. 4 for the case of $K = 3$, $n_r = 4$, and $L = 1, 2, 3$, which represents the overdetermined MAC case. Besides a significant improvement on the diversity gain by using the user selection scheme, it is seen that $\bar{d}_{\text{us},3}(r) = d^*_{\text{mac}}(r)$ for $L = 3$, i.e. all users are selected for transmission. This is an indication that the DMT upper bounds given in Theorem 3 might be very tight.

Moreover, we shall remark that for the case of $L < \nu$ in the user selection, the maximal multiplexing gain for a nonzero diversity gain is $\frac{L}{K}$. This means that for instance, if the desired multiplexing gain equals $r = 0.3$ in the case of $K = 3$ and $n_r = 4$, then one shall not consider the user selection with $L = 1$.

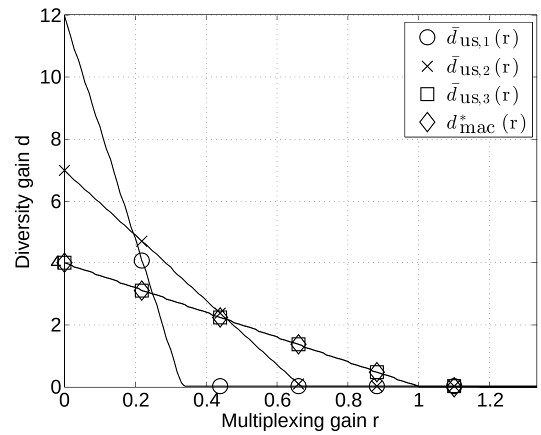

Fig. 4. The DMT upper bounds $\bar{d}_{\text{us},L}(r)$ for user selection and the optimal MAC DMT $d^*_{\text{mac}}(r)$ without selection for $K = 3$, $n_r = 4$, and $L = 1, 2, 3$.

B. Complexity Ramification of User Selection

In the previous section we have seen that by allowing a very low-rate feedback to the users, the resulting DMT performance can be significantly improved in the regime of low and moderate multiplexing gain values. Thus, a natural way for designing a communication scheme for MAC would be to take advantage of user selection scheme and the conventional MAC. Specifically, let $d^*(r)$ be the maximal diversity gain that can be provided by using either user selection or conventional MAC, i.e.

$$d^*_{\text{mac-us}}(r) = \max\left\{ \max_{1 \leq L \leq \nu} d^*_{\text{us},L}(r), d^*_{\text{mac}}(r) \right\}. \quad (14)$$

Clearly, $d^*_{\text{mac-us}}(r)$ is upper bounded by

$$\bar{d}_{\text{mac-us}}(r) = \max\left\{ \max_{1 \leq L \leq \nu} \bar{d}_{\text{us},L}(r), d^*_{\text{mac}}(r) \right\}. \quad (15)$$

Also given the desired multiplexing gain, let $L^*_{\text{mac-us}}(r)$ denote the optimal number of users selected for transmission such that the upper bound DMT $\bar{d}(r)$ can be achieved, i.e.,

$$L^*_{\text{mac-us}}(r) := \begin{cases} \arg\max_{1 \leq L \leq \nu} \bar{d}_{\text{us},L}(r), \\ \quad \text{if user selection is used for } \bar{d}_{\text{mac-us}}(r), \\ \\ K, \\ \quad \text{if conventional MAC is used for } \bar{d}_{\text{mac-us}}(r). \end{cases} \quad (16)$$

For instance, the optimal number of selected users $L^*_{\text{mac-us}}(r)$ for the underdetermined MAC case of $K = 4$ and $n_r = 3$ is given in Fig. 5.

Results in Theorems 1 and 4 then provides us with an explicit characterization of the computational complexity required by the joint consideration of conventional MAC and

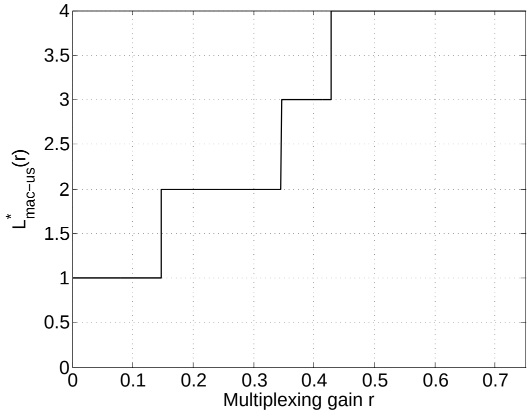

Fig. 5. Optimal number of selected users $L^*_{\text{mac-us}}(r)$ for $K = 4$-user MAC with $n_r = 3$

user selection schemes. More precisely, the upper bound on the computational complexity exponent of the corresponding SD algorithm is upper bounded by

$$\bar{c}_{\text{mac-us}}(r) = \begin{cases} \bar{c}_{\text{us},L}(r), & \text{if } L^*_{\text{mac-us}}(r) = L < \nu, \\ \bar{c}_{\text{mac}}(r), & \text{if } L^*_{\text{mac-us}}(r) = K. \end{cases} \quad (17)$$

In Figs. 6 and 7 we plot the DMT upper bound $\bar{d}_{\text{mac-us}}(r)$ as well as the complexity exponent upper bound $\bar{c}_{\text{mac-us}}(r)$ for $K = 4$-user MAC with $n_r = 3$, respectively. The results are also compared to those of conventional MAC without selection. It can be seen that the user selection scheme provides a significant improvement on the diversity gain and at the same time offers an exponentially large reduction on SD complexity. When only user selection is considered, we also provide in Figs. 6 and 7 the optimal DMT

$$\bar{d}_{\text{us}}(r) := \max_{1 \le L \le \nu} \bar{d}_{\text{us},L}(r)$$

and the corresponding complexity exponent $\bar{c}_{\text{us}}(r)$. It can be seen that if a very small cutback of diversity gain is allowed at high multiplexing gain regime, then the user selection scheme can provide a substantial and exponentially large reduction on complexity cost.

IV. Conclusions

In this paper we computed bounds on the complexity costs that are sufficient to achieve the optimal DMT of the symmetric MAC, as well as provided analysis for the DMT and complexity resulting from user selection.

Appendix A
Proof of Theorem 1

In the following we establish an upper bound on the minimum complexity exponent $c^*_{\text{mac}}(r)$ required by ML-based decoding to achieve the optimal DMT $d^*_{\text{mac}}(r)$ of the MAC. The joint ML decoder sees a $K \times n_r$ equivalent MIMO-MAC and a sum of multiplexing gain of Kr. Yet, as the users do not cooperator with each other, for any distinct pair of codeword matrices X and X' (cf. (1)), the different matrix $X - X'$ might not be nonsingular as in the point-to-point case.

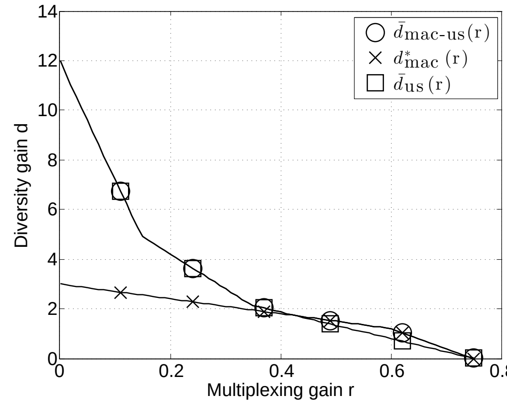

Fig. 6. DMT upper bounds $\bar{d}_{\text{us}}(r)$ for user selection, $d^*_{\text{mac}}(r)$ without selection, and $\bar{d}_{\text{mac-us}}(r)$ when both user-selection and conventional MAC are allowed, for $K = 4$-user MAC with $n_r = 3$.

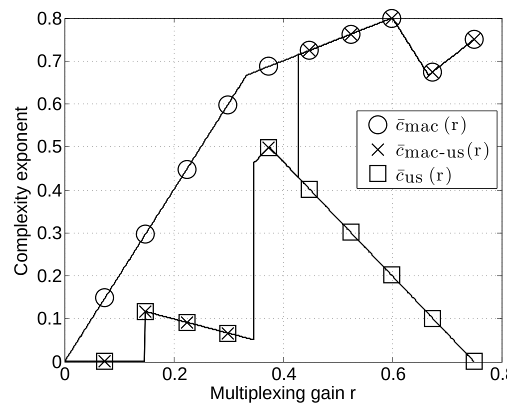

Fig. 7. Complexity exponent upper bounds $\bar{c}_{\text{us}}(r)$ for user selection, $\bar{c}_{\text{mac}}(r)$ without selection, and $\bar{c}_{\text{mac-us}}(r)$ when both user-selection and conventional MAC are allowed, for $K = 4$-user MAC with $n_r = 3$.

A. SD Complexity

For complexity analysis we note that for the K-user MAC with full-rate codes described in Section I-A, the overall lattice $\Lambda := \Lambda_1 \oplus \cdots \oplus \Lambda_K$ has rank $2KT$. For the underdetermined case of $K > n_t$, we might require an MMSE-preprocessing in order to allow for a SD implementation of ML decoding [12]. Towards this, let $\tilde{M}$ be the MMSE-preprocessed code-channel matrix given by

$$\tilde{M} = \begin{bmatrix} M \\ \alpha \mathtt{I}_u \end{bmatrix} = QR \quad (18)$$

where $u = 2(K - n_r)T$ and $\alpha = \text{SNR}^{-\frac{r}{2}}$ if $K > n_r$, and $\alpha = u = 0$ if otherwise, and where the second equality means the QR factorization of $\tilde{M}$. In which case (4) yields

$$\underline{r} \ := \ \left(R^\dagger\right)^{-1} M^\dagger \underline{y} = R\underline{s} + \underline{w}'$$

where $\underline{w}' = -\alpha^2 \left(R^\dagger\right)^{-1} \underline{s} + \left(R^\dagger\right)^{-1} M^\dagger \underline{w}$. Let $\mathbb{S}_r := \left\{ \underline{s} \in \mathbb{Z}^{2KT} : \text{SNR}^{-\frac{r}{2}} G\underline{s} \in \mathcal{X}_r \right\}$ denotes the set of coordinates of points in Λ that lie in $\mathcal{X}_r$ after certain scaling. Then the

MMSE-preprocessed SD decoder for this system then takes the form

$$\hat{\underline{s}}_{\text{MMSE-SD}} \;=\; \arg\min_{\underline{s}\in\mathbb{S}_r} \|\underline{r} - R\underline{s}\|^2 \tag{19}$$

and can be implemented by the sphere decoder which recursively enumerates all candidate vectors $\underline{s} \in \mathbb{S}_r$ within a given search sphere of radius $\delta = \sqrt{z \log_2 \text{SNR}}$ for some $z > d^*_{\text{mac}}(r)$. Moreover, it has been shown [2] that the MMSE-preprocessed SD decoder allows for a vanishing gap to the exact solution found by the ML decoder.

To compute an upper bound on the complexity exponent, we follow the approach similar to [2], i.e., let $\lambda_i = \sigma_i(H^\dagger_{\text{eq}} H_{\text{eq}})$, $i = 1, \ldots, \nu = \min\{K, n_r\}$, be the nonzero singular values, arranged in ascending order, of matrix $H^\dagger_{\text{eq}} H_{\text{eq}}$, and define $\mu_i = -\frac{\log_2 \lambda_i}{\log_2 \text{SNR}}$. It then follows that

$$\sigma_i(R) \;=\; \sigma_i(\tilde{M}) = \sqrt{\alpha^2 + \sigma_i(M^\dagger M)} \tag{20}$$

with $i = 1, \ldots, 2KT$. Moreover, we have $\sigma_i(R) \doteq \text{SNR}^{\frac{1}{2}(1-r-\mu_{\lceil \frac{i}{2T} \rceil})}$ for $K < n_r$, and

$$\sigma_i(R) \doteq \begin{cases} \text{SNR}^{-\frac{r}{2}}, & \text{if } 1 \le i \le 2T(K - n_r) \\ \text{SNR}^{-\frac{r}{2}+\frac{1}{2}(1-\mu_j)^+}, & \text{otherwise} \end{cases} \tag{21}$$

for $K \ge n_r$, where $j = \lceil \frac{i-2T(K-n_r)}{2T} \rceil$, and we have used the fact that $\sigma_i(G) \doteq \text{SNR}^0$.

The total number of visited nodes is commonly taken as a measure of the SD complexity [1]. For any given channel realization $\underline{\mu} := [\mu_1 \cdots \mu_\nu]^\top$, the total number of visited nodes is given by

$$N_{\text{SD}}(\underline{\mu}) := \sum_{k=1}^{2KT} N_k(\underline{\mu})$$

and it has been shown in [1, Lemma 1] that

$$N_k(\underline{\mu}) \;\le\; \prod_{i=1}^{k} \left[\sqrt{k} + 2\min\left\{ \frac{\delta}{\sigma_i(R)}, \sqrt{k}\text{SNR}^{\frac{r}{2}} \right\} \right]$$

It then follows that

$$N_{\text{SD}}(\underline{\mu}) \le \sum_{k=1}^{2KT} \prod_{i=1}^{k} \left[\sqrt{k} + 2\min\left\{ \frac{\delta}{\sigma_i(R)}, \sqrt{k}\text{SNR}^{\frac{r}{2}} \right\} \right],$$

and in particular, by (20) for the overdetermined case, i.e, $K \le n_r$, we have

$$N_{\text{SD}}(\underline{\mu}) \;\dot{\le}\; \text{SNR}^{T\sum_{i=1}^{K}[\min\{r, r+\mu_i-1\}]^+} \tag{22}$$

and by (21) for the underdetermined case it yields

$$N_{\text{SD}}(\underline{\mu}) \;\dot{\le}\; \text{SNR}^{(K-n_r)rT + T\sum_{i=1}^{n_r}\left(r-(1-\mu_i)^+\right)^+}, \tag{23}$$

where we have used the fact of $2T$ multiplicity of the singular values of H defined in (3).

Following the footsteps of the complexity analysis in [2] for the MMSE-preprocessed channel matrix R, the upper bound on the complexity exponent can be obtained as the solution to a constrained minimization problem of finding a value $c^*_{\text{mac}}(r)$ such that the probability of a prematurely termination of SD algorithm is no larger than the channel outage probability, i.e.,

$$\Pr\left\{ N_{\text{SD}}(\underline{\mu}) \ge N_{\max}(r) = \text{SNR}^{c^*_{\text{mac}}(r)} \right\} \le \text{SNR}^{-d^*_{\text{mac}}(r)}. \tag{24}$$

B. Optimality of Uncoded QAM

It remains to show that the upper bounds (22) and (23) can be significantly tightened by setting $T = 1$ while the probabilitic complexity constraint (24) still holds. In particular, setting $T = 1$ means that the overall code lattice Λ has rank $2K$, which is isomorphic to the rectangular lattice $\mathbb{Z}^{2K}$. Equivalently, it means that the overall code $\mathcal{X}_r$ is given by

$$\mathcal{X}_r \;=\; \mathcal{L}_r = \left\{ \text{SNR}^{-\frac{r}{2}}\underline{x} : \underline{x} \in (\mathbb{Z}[\imath])^K, |x_i|^2 \le \text{SNR}^r \right\},$$

i.e., the uncoded QAM signaling after scaling. To show that the above code $\mathcal{L}_r$ satisfies the probabilistic complexity constraint (24), following the footsteps in [6], [8], we consider a K-fold extension of the $\mathcal{L}_r$,

$$\mathcal{L}_{r,\text{ext}} = \bigoplus_{i=1}^{K} \mathcal{L}_r \subset \text{SNR}^{-\frac{r}{2}} M_K(\mathbb{Z}[\imath])$$

that is, elements of $\mathcal{L}_{r,\text{ext}}$ are square $K \times K$ matrices whose entries are independent QAM constellation points after a scaling of $\text{SNR}^{-\frac{r}{2}}$. It is obvious that the error probability of code $\mathcal{L}_{r,\text{ext}}$ is at most K times that of code $\mathcal{L}_r$, provided that the same decoder is used for decoding both codes; hence $\mathcal{L}_{r,\text{ext}}$ and $\mathcal{L}_r$ achieve the same DMT performance. Assuming a quasi-static fading channel with a channel coherence time $T_c \ge K$ channel uses, the error probability of code $\mathcal{L}_{r,\text{ext}}$ achieved by the ML decoder is upper bounded by

$$\begin{aligned} &P_{e,\text{ext}}(r) \\ &= \mathbb{E}\,\Pr\{L' \in \mathcal{E}(L) \text{ decoded}\} \\ &= \mathbb{E}\,\Pr\left\{ \bigcup_{k=1}^{K} \{L' \in \mathcal{E}_k(L) \text{ decoded}\} \right\} \\ &\overset{\text{(i)}}{\le} \sum_{k=1}^{K} \mathbb{E}\,\Pr\{L' \in \mathcal{E}_k(L) \text{ decoded}\} \\ &\overset{\text{(ii)}}{\le} \sum_{k=1}^{K} \mathbb{E}\,\Pr\left\{ \bigcup_{L' \in \mathcal{E}_k(L)} \left\{ H_{\text{eq}} : \left\| \text{SNR}^{\frac{1}{2}} H_{\text{eq}}(L - L') \right\|^2 \dot{\le} 1 \right\} \right\} \end{aligned} \tag{25}$$

where the expectation is taken over all codeword matrices $L \in \mathcal{L}_{r,\text{ext}}$, and where $\mathcal{E}(L) := \mathcal{L}_{r,\text{ext}} \setminus \{L\}$ is the set of all possible erroneous codewords given L transmitted, $\mathcal{E}_k(L) := \{L' \in \mathcal{E}(L) : \text{rank}(L - L') = k\}$ with $k = 1, \ldots, K$ is a partition of $\mathcal{E}(L)$. Step (i) follows from the union bound and step (ii) is due to the use of a suboptimal bounded distance decoder (cf. [6]). For any $L' \in \mathcal{E}_k(L)$, set $\Delta_{L'} = L - L'$ and let $\Delta_{L'}\Delta^\dagger_{L'} = U_{L'}\Sigma_{L'}U^\dagger_{L'}$ be the corresponding eigen-decomposition. Note that as $\text{rank}(\Delta_{L'}) = k$, the eigenvalue matrix $\Sigma_{L'}$ has form $\Sigma_{L'} = \text{diag}(\Omega_{L'}, 0_{K-k})$, where $\Omega_{L'}$

consists of all nonzero eigenvalues of $\Delta_{L'}\Delta_{L'}^{\dagger}$. Substituting the above into (25) we obtain

$$\Pr\left\{\bigcup_{L'\in\mathcal{E}_k(L)}\left\{H_{\text{eq}}:\left\|\text{SNR}^{\frac{1}{2}}H_{\text{eq}}(L-L')\right\|^2\dot{\leq}1\right\}\right\} = \Pr\left\{\bigcup_{L'\in\mathcal{E}_k(L)}\left\{G:\text{tr}\left(G^{\dagger}\Omega_{L'}G\right)\dot{\leq}1\right\}\right\}, \quad (26)$$

where G is an $(n_r \times k)$ random matrix with i.i.d. $\mathbb{C}\mathcal{N}(0,1)$ entries. Let $\lambda_1 \leq \cdots \leq \lambda_m$ be the nonzero eigenvalues of $G^{\dagger}G$, where $m = \min\{k, n_r\}$, and set $\mu_i = -\frac{\log_2 \lambda_i}{\log_2 \text{SNR}}$. Noting that $\det(\Omega_{L'})\dot{\geq}\text{SNR}^{-kr}$ and $\text{tr}(\Omega_{L'})\dot{\leq}1$, it can be shown using arguments similar to [6] that the condition of $\text{tr}\left(G^{\dagger}\Omega_{L'}G\right)\dot{\leq}1$ implies $\sum_{i=1}^{m}(1-\mu_i)^+ \leq kr$, which is independent of the choice of L'. Hence we have

$$\begin{aligned}
&\Pr\left\{\bigcup_{L'\in\mathcal{E}_k(L)}\left\{G:\text{tr}\left(G^{\dagger}\Omega_{L'}G\right)\dot{\leq}1\right\}\right\}\\
&\dot{\leq}\Pr\left\{\bigcup_{L'\in\mathcal{E}_k(L)}\left\{\underline{\mu}=[\mu_1\cdots\mu_k]^{\top}:\sum_{i=1}^{m}(1-\mu_i)^+\leq kr\right\}\right\}\\
&=\Pr\left\{\underline{\mu}=[\mu_1\cdots\mu_k]^{\top}:\sum_{i=1}^{m}(1-\mu_i)^+\leq kr\right\}\\
&\doteq\text{SNR}^{-d^*_{k,n_r}(kr)},
\end{aligned}$$

where the last dotted equality follows from [7]. Finally, note that the error probability of $\mathcal{L}_r$ subject to ML decoding is upper bounded by

$$P_e(r)\leq\frac{1}{K}P_{e,\text{ext}}(r)\dot{\leq}\frac{1}{K}\sum_{i=1}^{K}\text{SNR}^{-d^*_{k,n_r}(kr)}\doteq\text{SNR}^{-d^*_{\text{mac}}(r)},$$

and the proof is complete.

Appendix B
Proof of Theorem 3

Let H_{eq} be the $(n_r \times K)$ equivalent channel matrix defined in (1). To select the L users, Jiang-Varanasi algorithm [11] takes L iterations of column permutation Π_i and Householder transformation T_i to obtain the following matrix R

$$R = T_L\cdots T_1 H_{\text{eq}}\Pi_1\cdots\Pi_L = \begin{bmatrix} r_{1,1} & * & \cdots & * & * & \cdots & * \\ & r_{2,2} & \cdots & * & * & \cdots & * \\ & & \ddots & \vdots & \vdots & \vdots & \vdots \\ & & & r_{L,L} & * & \cdots & * \\ & & & & \vdots & \vdots & \vdots \\ & & & & * & \cdots & * \end{bmatrix}.$$

Let u_i be the user associated with the ith column of R and denote the set of selected users by $\{u_1, u_2, \ldots, u_L\}$. Since entries of H_{eq} are i.i.d. $\mathbb{C}\mathcal{N}(0,1)$, to meet an average multiplexing gain r for each user, the selected user has to transmit at a larger multiplexing gain of $\frac{K}{L}r$. We consider a series of specific outage events,

$$\mathcal{O}_{k,\ell} := \{\text{the sum-rate of users } u_{k+1},\cdots,u_{k+\ell} \text{ is in outage}\} \quad (27)$$

for all $k \geq 0$, $\ell \geq 1$, and $k+\ell \leq L$. It will be seen that the outage event considered by Jiang and Varanasi [11, Theorem 4.1] is a special case of the above when setting $k=0$ and $\ell = L$, i.e., the outage event $\mathcal{O}_{0,L}$.

A. DMT Analysis for error event $\mathcal{O}_{k,\ell}$

To analyze the DMT for the error event $\mathcal{O}_{k,\ell}$ for any $k \geq 0$, $\ell \geq 1$, and $k+\ell \leq L$, let $R_{k,\ell}$ be the matrix resulting from applying $(k+\ell)$ iterations of Jiang-Varanasi algorithm to overall matrix H_{eq}. We partition matrix $R_{k,\ell}$ as follows

$$R_{k,\ell} = T_{k+\ell}\cdots T_1 H_{\text{eq}}\Pi_1\cdots\Pi_{k+\ell} = \begin{bmatrix} R_L & R_{C,U} & R_{R,U} \\ & R_{C,B} & R_{R,M} \\ & & R_{R,B}\end{bmatrix}, \quad (28)$$

where

$$\begin{aligned}
R_L &= \begin{bmatrix} r_{1,1} & \cdots & r_{1,k} \\ & \ddots & \vdots \\ & & r_{k,k}\end{bmatrix},\\
R_{C,U} &= \begin{bmatrix} r_{1,k+1} & \cdots & r_{1,k+\ell} \\ \vdots & \vdots & \vdots \\ r_{k,k+1} & & r_{k,k+\ell}\end{bmatrix},\\
R_{C,B} &= \begin{bmatrix} r_{k+1,k+1} & \cdots & r_{k+1,k+\ell} \\ & \ddots & \vdots \\ & & r_{k+\ell,k+\ell}\end{bmatrix},\\
R_{R,U} &= \begin{bmatrix} r_{1,k+\ell+1} & \cdots & r_{1,K} \\ \vdots & \vdots & \vdots \\ r_{k,k+\ell+1} & \cdots & r_{k,K}\end{bmatrix},\\
R_{R,M} &= \begin{bmatrix} r_{k+1,k+\ell+1} & \cdots & r_{k+1,K} \\ \vdots & \vdots & \vdots \\ r_{k+\ell,k+\ell+1} & \cdots & r_{k+\ell,K}\end{bmatrix},\\
R_{R,B} &= \begin{bmatrix} r_{k+\ell+1,k+\ell+1} & \cdots & r_{k+\ell+1,K} \\ \vdots & \vdots & \vdots \\ r_{n_r,k+\ell+1} & \cdots & r_{n_r,K}\end{bmatrix}
\end{aligned}$$

and the entries satisfy

$$|r_{i,i}|^2 \geq \sum_{m=i}^{n_r}|r_{m,j}|^2, \quad (29)$$

for $i = 1, 2, \ldots, k+\ell$ and for $j = i+1, \ldots, K$.

Set

$$R_C := \begin{bmatrix} R_{C,U} \\ R_{C,B}\end{bmatrix};$$

then the probability that the sum-rate of users $u_{k+1},\cdots,u_{k+\ell}$ is in outage is

$$\Pr\{\mathcal{O}_{k,\ell}\}$$

$$= \Pr\left\{\log_2 \det\left(\mathtt{I}_{n_r} + \text{SNR} R_C R_C^\dagger\right) \;<\; \frac{K}{L}\ell r \log_2 \text{SNR}\right\}$$

$$= \int_{\mathcal{I}(r)} c \cdot \left[\prod_{i=1}^{k+\ell} f_{\chi^2_{2(n_r-i+1)}}\left(|r_{i,i}|^2\right)\left(\prod_{j=i+1}^{K}\frac{1}{\pi}e^{-|r_{i,j}|^2}\right)\right] \times \frac{1}{\pi^{(K-k-\ell)(nr-k-\ell)}} e^{-\|R_{R,B}\|^2}\, dR \quad (30)$$

$$:\doteq \text{SNR}^{-d_{k,\ell}\left(\frac{K}{L}\ell r\right)} \quad (31)$$

for some constant c due to ordered statistics (cf. the first constraint in (32)), where $f_{\chi^2_\kappa}(\cdot)$ is the probability density function for χ^2 random variable with degree of freedom κ and mean $\frac{\kappa}{2}$, and where the region for integration is

$$\mathcal{I}(r) := \left\{ R : \begin{array}{l} |r_{i,i}|^2 \;\geq\; \sum_{m=i}^{n_r} |r_{m,j}|^2, \\ i = 1, \ldots, k+\ell, \quad j = i+1, \ldots, K \\ \det\left(\mathtt{I}_{n_r} + \text{SNR} R_C R_C^\dagger\right) < \text{SNR}^{\frac{K}{L}\ell r} \end{array} \right\} \quad (32)$$

We make the following observations.

1) In seeking the DMT $d_{k,\ell}(r)$ by applying the Laplace principle, entries in R_L and $R_{R,U}$ are not involved in the second constraint in $\mathcal{D}(r)$. This implies that the dominant sub-event in $\mathcal{I}(r)$ has $R_L, R_{R,U} \doteq \mathbf{1}$, the all-one matrix with proper size.

2) Let $\lambda_1 \geq \cdots \geq \lambda_\ell$ be the ordered singular values for R_C and let $\mu_1 \geq \cdots \geq \mu_\ell$ be the ordered singular values of $R_{C,B}$. Clearly, as $R_{C,B} R_{C,B}^\dagger \preceq R_C R_C^\dagger$, we have

$$\mu_i^2 \leq \lambda_i^2, \quad i = 1, 2, \ldots, \ell. \quad (33)$$

By [11, Lemma 3.3] we have

$$(R_{C,B})_{i,i} \;=\; r_{k+i,k+i}^2 \;\geq\; \frac{\sum_{j=i}^{\ell} \mu_j^2}{\ell - i + 1} \;\doteq\; \mu_i^2 \quad (34)$$

for $i = 1, 2, \ldots, \ell$. Moreover, by [11, Eq. (27)] that the squared diagonal elements in $R_{C,B}$ are multiplicatively majorized by its squared singular values, i.e.

$$\prod_{i=1}^{m} \mu_i^2 \;\geq\; \prod_{i=1}^{m} r_{k+i,k+i}^2 \;\dot{\geq}\; \prod_{i=1}^{m} \mu_i^2, \quad m = 1, 2, \ldots, \ell \quad (35)$$

where the second dotted inequality is due to (34). It then follows that

$$r_{k+i,k+i}^2 \doteq \mu_i^2, \quad (36)$$

for $i = 1, 2, \ldots, \ell$.

3) Finally, at the end of first k iterations of Jiang-Varanasi algorithm, we get

$$T_k \cdots T_1 H \Pi_1 \cdots \Pi_k = \begin{bmatrix} r_{1,1} & * & \cdots & * & * & \cdots & * \\ & r_{2,2} & \cdots & * & * & \cdots & * \\ & & \ddots & \vdots & \vdots & \vdots & \vdots \\ & & & r_{k,k} & * & \cdots & * \\ & & & & \vdots & \vdots & \vdots \\ & & & & * & \cdots & * \end{bmatrix} = \left[\; R_L \quad R_R \;\right],$$

where R_R is the $(n_r \times (K-k))$ matrix consisting of the rightmost $(K-k)$ columns of the above matrix. Entries of R_R can still be regarded as i.i.d. $\mathbb{CN}(0,1)$ random variables as $R_L \doteq \mathbf{1}$ due to the first remark above. It means that the singular values $\lambda_i, i = 1, \ldots, \ell$ are still of the same probability distribution for ordered singular values of an $(n_r \times \ell)$ random matrix with i.i.d. $\mathbb{CN}(0,1)$ entries.

With the above, we now proceed to analyze the integral (30) to obtain a formula for the DMT function $d_{k,\ell}(r)$. Specifically, we will show that

$$d_{k,\ell}(r) = \inf_{\mathcal{A}_\ell(r)} \left\{ \sum_{i=1}^{\ell}(n_r + \ell - 2i + 1)\alpha_i + \sum_{i=1}^{\ell-1}(K-k-\ell)\alpha_i + \alpha_\ell(K-k-\ell)(n_r-k-\ell+1) \right\} \quad (37)$$

where

$$\mathcal{A}_\ell(r) = \left\{ 0 \leq \alpha_1 \leq \cdots \leq \alpha_\ell : \sum_{i=1}^{\ell}(1-\alpha_i)^+ \leq r \right\}.$$

To see the above, set $\lambda_i^2 \doteq \text{SNR}^{-\alpha_i}$ with $\alpha_1 \leq \alpha_2 \cdots \leq \alpha_\ell$. The first summand appearing in (37) follows from the joint probability density function of ordered singular values $\lambda_1 \geq \cdots \geq \lambda_\ell$ for an $(n_r \times \ell)$ matrix with i.i.d. $\mathbb{CN}(0,1)$ entries is [7], [11]

$$f(\alpha_1, \cdots, \alpha_\ell) \;\doteq\; \text{SNR}^{-\sum_{i=1}^{\ell}(n_r+\ell-2i+1)\alpha_i}.$$

Also by (33), (36) and (29), we have the following constraints for entries in matrix $R_{k,\ell}$:

1) For $i = 1, \ldots, \ell-1$, entries $r_{k+i,j}$, $j = k+\ell+1, \ldots, K$, must satisfy

$$|r_{k+i,j}|^2 \leq |r_{k+i,k+i}|^2 \leq \lambda_i^2.$$

The constraints on $r_{k+i,j}$ contribute to (37) the term $\sum_{i=1}^{\ell-1}(K-k-\ell)\alpha_i$.

2) Entries $r_{k+i,j}$ with $i = \ell, \ldots, n_r - k$ and $j = k+\ell+1, \ldots, K$, must satisfy

$$\sum_{i=\ell}^{n_r-k} |r_{k+i,j}|^2 \leq |r_{k+\ell,k+\ell}|^2 \doteq \mu_\ell^2 \leq \lambda_\ell^2$$

Such constraints contribute to (37) the term $\alpha_\ell(K-k-\ell)(n_r-k-\ell+1)$.

Thus we have completed the proof of (37)

Remark 1: It can be shown that with the setting of $k = 0$ and $\ell = L \leq \nu$, the DMT $d_{0,L}(r)$ is exactly the antenna-selection DMT given by Jiang and Varanasi in [11, Theorem 4.1], i.e., $d_{0,L}(r)$ is a piecewise linear function connecting the following $(P+2)$ points

$$(r, (K-r)(n_r-r)), r = 0, 1, \ldots, P, \text{ and } (L, 0), \quad (38)$$

where

$$P = \arg \min_{p=0,1,\ldots,L-1} \frac{(K-p)(n_r-p)}{L-p}. \quad (39)$$

Hence the DMT result of Jiang and Varanasi in [11, Theorem 4.1] can be seen as a special case of (37). ■

Finally, the proof of Theorem 3 is complete after noting that the union of outage events $\mathcal{O}_{k,\ell}$ is a subset of the overall outage event, i.e.

$$\bigcup_{\substack{k,\ell\geq 0\\ k+\ell\leq L}} \mathcal{O}_{k,\ell} \subseteq \bigcup_{\mathcal{U}\subset\{u_1,\ldots,u_L\}} \{\text{ users in } \mathcal{U} \text{ are in outage}\}.$$

Acknowledgments

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 257616 (CONECT), from the FP7 CELTIC SPECTRA project, and from Agence Nationale de la Recherche project ANR-IMAGENET

References

[1] J. Jalden and P. Elia, "Sphere decoding complexity exponent for decoding full-rate codes over the quasi-static MIMO channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5785–5803, Sep. 2012.

[2] A. K. Singh, P. Elia, and J. Jaldén, "Achieving a vanishing SNR gap to exact lattice decoding at a subexponential complexity," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3692–3707, Jun. 2012.

[3] A. Singh, P. Elia, , and J. Jaldén, "Complexity analysis for ML-based sphere decoder achieving a vanishing performance-gap to brute force ML decoding," in *Proc. Int. Zurich Seminar on Communications (IZS)*, Mar. 2012, pp. 127–130.

[4] P. Elia and J. Jaldén, "General DMT optimality of LR-aided linear MIMO-MAC transceivers with worst-case complexity at most linear in sum-rate," in *Proc. IEEE Information Theory Workshop (ITW)*, Cairo, Egypt, Jan. 2010.

[5] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1859–1874, Sept. 2004.

[6] H. F. Lu, C. Hollanti, R. Vehkalahti, and J. Lahtonen, "DMT optimal codes constructions for multiple-access MIMO channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3594–3617, June 2011.

[7] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[8] H. F. Lu, "Remarks on diversity-multiplexing tradeoffs for multiple-access and point-to-point MIMO channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 858–863, Feb. 2012.

[9] H.-F. Lu, R. Vehkalahti, C. Hollanti, J. Lahtonen, Y. Hong, and E. Viterbo, "New space-time code construcions for two user multiple access channels," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 6, pp. 939–957, Dec. 2009.

[10] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge, University Press, 2005.

[11] Y. Jiang and M. K. Varanasi, "The RF-chain limited MIMO system-part I: optimum diversity-multiplexing tradeoff," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5238–5247, Oct. 2009.

[12] M. O. Damen, H. El Gamal, and G. Caire, "MMSE-GDFE lattice decoding for solving under-determined linear systems with integer unknowns," in *Proc. 2004 IEEE Int. Symp. Inform. Theory*, Chicago, IL, June 2004, p. 538.