Tail Behavior of Sphere-Decoding Complexity in Random Lattices

D. Seethaler*, J. Jaldén°, C. Studer*, and H. Bölcskei*

* Communication Technology Laboratory
ETH Zurich, 8092 Zurich, Switzerland, email:{seethal,studerc,boelcskei}@nari.ee.ethz.ch

° Institute of Communications and Radio-Frequency Engineering
Vienna University of Technology, 1040 Vienna, Austria, email: jjalden@nt.tuwien.ac.at

***Abstract*—We analyze the (computational) complexity distribution of sphere-decoding (SD) for random infinite lattices. In particular, we show that under fairly general assumptions on the statistics of the lattice basis matrix, the tail behavior of the SD complexity distribution is solely determined by the inverse volume of a fundamental region of the underlying lattice. Particularizing this result to $N \times M$, $N \geq M$, i.i.d. Gaussian lattice basis matrices, we find that the corresponding complexity distribution is of Pareto-type with tail exponent given by $N - M + 1$. We furthermore show that this tail exponent is not improved by lattice-reduction, which includes layer-sorting as a special case.**

I. Introduction

The problem of finding the closest lattice point in an infinite lattice is commonly referred to as the *closest lattice point* (CLP) problem (see, e.g., [1]). The sphere-decoding (SD) algorithm [1]–[6] is a promising approach for solving the CLP problem. The (computational) complexity of SD, as measured in terms of the number of searched lattice points, depends strongly on the lattice basis matrix and is, in general, very difficult to characterize analytically. However, if the basis matrix is assumed *random*, the complexity of SD is random as well and one can resort to a characterization of the *complexity distribution* of SD. Previous work along these lines focused on the mean and the variance of SD complexity for i.i.d. Gaussian basis matrices [7]–[9]. However, a characterization of the tails of the SD complexity distribution is, for example, important for using SD under run-time constraints (see, e.g., [6]). In this paper, we make a first attempt in this direction by analyzing the *tail behavior* (TB) of the SD complexity distribution. Our main contributions can be summarized as follows:

- Under fairly general assumptions on the statistics of the lattice basis matrix, we prove that the TB of the SD complexity distribution is solely determined by the TB of the inverse volume of a fundamental region of the underlying lattice.
- Specializing this result to the case of $N \times M$, $N \geq M$, i.i.d. circularly symmetric complex Gaussian basis matrices, we find that the complexity distribution of SD is of Pareto-type with tail exponent given by $N - M + 1$. We furthermore show that the tail exponent is not improved, i.e., increased, if lattice-reduction (see, e.g., [1]), which includes layer-sorting as a special case, is performed.

Notation: We write $A_{i,j}$ for the entry in the ith row and jth column of the matrix $\mathbf{A}$ and x_i for the ith entry of the vector $\mathbf{x}$. Slightly abusing common terminology, we call an $N \times M$, $N \geq M$, matrix $\mathbf{A}$ unitary if it satisfies $\mathbf{A}^H\mathbf{A} = \mathbf{I}$, where ^{H} denotes conjugate transposition and $\mathbf{I}$ is the identity matrix. The Euclidean- and the Frobenius norm are denoted by $\|\cdot\|$ and $\|\cdot\|_\mathrm{F}$, respectively, and $|\mathcal{X}|$ refers to the cardinality of the set $\mathcal{X}$. The ceil-function is denoted by $\lceil\cdot\rceil$. Furthermore, $\mathbb{CZ}$ stands for the set of Gaussian integers, i.e., $\mathbb{CZ} = \mathbb{Z} + \sqrt{-1}\,\mathbb{Z}$. The lattice generated by the full-rank $N \times M$ ($N \geq M$) basis matrix $\mathbf{A}$ is defined as $\mathcal{L}(\mathbf{A}) = \{\mathbf{Ad} : \mathbf{d} \in (\mathbb{CZ})^M\}$. For $N = M$, the corresponding covering radius is given by [10]

$$\mu(\mathbf{A}) = \max_{\mathbf{x}\in\mathbb{C}^M} \min_{\mathbf{d}\in(\mathbb{CZ})^M} \|\mathbf{x} - \mathbf{Ad}\|. \tag{1}$$

A circularly symmetric complex Gaussian random variable (RV) with variance σ_x^2 is denoted as $x \sim \mathcal{CN}(0, \sigma_x^2)$. The natural logarithm is referred to as $\log(\cdot)$. We write $g(x) \doteq f(x)$, $x \to x_0$, for $\lim_{x\to x_0} \log g(x)/\log(x) = \lim_{x\to x_0} \log f(x)/\log(x)$, assuming that the corresponding limits exist. The symbols $\dot{\leq}$ and $\dot{\geq}$ are defined analogously. Finally, non-polynomial behavior of $g(x)$ is captured by $\lim_{x\to x_0}\log g(x)/\log(x) = \pm\infty$ or $\lim_{x\to x_0}\log g(x)/\log(x) = 0$, for which we write $g(x) \doteq x^{\pm\infty}$, $x \to x_0$, and $g(x) \doteq x^0$, $x \to x_0$, respectively.

A. Sphere-Decoding

The CLP problem refers to computing

$$\widehat{\mathbf{d}} = \arg\min_{\mathbf{d}\in(\mathbb{CZ})^M} \|\mathbf{r} - \mathbf{Hd}\|^2 \tag{2}$$

for a given vector $\mathbf{r} \in \mathbb{C}^N$ and a given full-rank matrix $\mathbf{H} \in \mathbb{C}^{N\times M}$, $N \geq M$. In words, solving (2) amounts to finding the point in the lattice $\mathcal{L}(\mathbf{H})$ that is closest (in Euclidean distance) to $\mathbf{r}$. In communications, (2) is known as the maximum-likelihood (ML) detection problem for detecting $\mathbf{d}' \in (\mathbb{CZ})^M$ based on the linear model $\mathbf{r} = \mathbf{Hd}' + \mathbf{w}$ with $\mathbf{H}$ known at the receiver and $\mathbf{w}$ being i.i.d. circularly symmetric complex Gaussian noise.

A prominent approach for solving (2) is the SD algorithm [1]–[6]. In the following, we consider Fincke-Pohst SD [2] without radius reduction (see, e.g., [7]). The algorithm starts by computing the (unique) QR-decomposition (QRD) $\mathbf{H} = \mathbf{QR}$, where $\mathbf{R}$ denotes an $M \times M$ upper triangular matrix

This work was supported in part by the STREP project No. IST-026905 (MASCOT) within the Sixth Framework Programme of the European Commission.

with positive real-valued elements on its main diagonal and $\mathbf{Q}$ of size $N \times M$ is unitary. Then, (2) can equivalently be written as

$$\widehat{\mathbf{d}} = \arg \min_{\mathbf{d} \in (\mathbb{CZ})^M} \|\mathbf{y} - \mathbf{Rd}\|^2 \tag{3}$$

where $\mathbf{y} = \mathbf{Q}^H \mathbf{r}$. Next, (3) is solved subject to a *sphere constraint* (SC), which amounts to considering only those $\mathbf{d} \in (\mathbb{CZ})^M$ that lie within a hypersphere of radius ρ around $\mathbf{y}$, i.e., all $\mathbf{d}$ that satisfy

$$\|\mathbf{y} - \mathbf{Rd}\|^2 \leq \rho^2. \tag{4}$$

Here, the sphere radius ρ has to be chosen sufficiently large for the search sphere to contain at least one lattice point $\mathbf{Rd}$. Note, however, that if ρ is chosen too large, too many points will satisfy the SC and the complexity of SD will be high. As detailed next, imposing a SC enables an efficient recursive solution of the triangularized CLP problem (3).

Consider the length-k subvectors $\mathbf{d}_k \in (\mathbb{CZ})^k$ of $\mathbf{d}$ defined as $\mathbf{d}_k = (d_{M-k+1} \cdots d_M)^T$, $k = 1, \ldots, M$, where k is a *layer* index. The metric $\|\mathbf{y} - \mathbf{Rd}\|^2 = \|\mathbf{y}_M - \mathbf{R}_M \mathbf{d}_M\|^2$ can be computed recursively according to

$$\|\mathbf{y}_k - \mathbf{R}_k \mathbf{d}_k\|^2 = \|\mathbf{y}_{k-1} - \mathbf{R}_{k-1} \mathbf{d}_{k-1}\|^2 + |\Delta_k(\mathbf{d}_k)|^2 \tag{5}$$

where $\Delta_k(\mathbf{d}_k) = y_{M-k+1} - \sum_{i=M-k+1}^{M} R_{M-k+1,i}\, d_i$, $\mathbf{R}_k$ denotes the $k \times k$ bottom right (upper triangular) submatrix of $\mathbf{R}$ associated with $\mathbf{d}_k$, and $\mathbf{y}_k = (y_{M-k+1} \cdots y_M)^T$. Thus, with (5), a necessary condition for $\mathbf{d}$ to satisfy the SC is that any associated $\mathbf{d}_k$ satisfies the *partial SC*

$$\|\mathbf{y}_k - \mathbf{R}_k \mathbf{d}_k\|^2 \leq \rho^2. \tag{6}$$

This formulation now enables finding all integer vectors $\mathbf{d}$ that satisfy (4) in an efficient (recursive) manner as detailed, e.g., in [5], [7].

II. Complexity Distribution of SD

We define the computational complexity of SD as the number of lattice points searched by the algorithm, i.e., the number of vectors $\mathbf{d}_k \in (\mathbb{CZ})^k$, $k = 1, \ldots, M$, that satisfy the partial SCs in (6) (cf. [7], [11]). Specifically, we define the kth *layer complexity* of SD as

$$S_k = \left|\left\{\mathbf{d}_k \in (\mathbb{CZ})^k : \|\mathbf{y}_k - \mathbf{R}_k \mathbf{d}_k\|^2 \leq \rho^2\right\}\right| \tag{7}$$

with the *total complexity* $S = \sum_{k=1}^{M} S_k$. It was shown in [5] – for the finite lattice case – that S is proportional to the run-time complexity of a corresponding VLSI implementation.

A. Complexity Distribution and Tail Exponents

The quantities S_k, $k = 1, \ldots, M$, and S, defined above, are functions of $\mathbf{H}$, $\mathbf{r}$, and ρ. In the following, we let $\mathbf{H}$ and $\mathbf{r}$ be random (potentially statistically dependent) and consider a fixed ρ that does not depend on the realizations of $\mathbf{H}$ and $\mathbf{r}$. For example, in the case of ML detection of the transmitted data vector $\mathbf{d}'$ in MIMO wireless systems, $\mathbf{r} = \mathbf{Hd}' + \mathbf{w}$, where the entries of $\mathbf{H}$ (the channel matrix) and $\mathbf{w}$ (the noise vector) are typically assumed i.i.d. circularly symmetric complex Gaussian. In this setting, a reasonable fixed choice of ρ can be based on the noise statistics such that the probability of finding the transmitted data vector inside the search hypersphere is sufficiently high (see, e.g., [7]). Note, however, that this results in a nonzero probability of failing to find an integer vector inside the search hypersphere, which, in the absence of restarting the search with a larger sphere radius (see, e.g., [7]), would entail an error floor.

Since both $\mathbf{H}$ and $\mathbf{r}$ are random, S_k and S are random as well and can be characterized through their respective *distribution functions* $\mathrm{P}[S_k \geq L]$ and $\mathrm{P}[S \geq L]$. While these distributions seem hard to come by analytically, it turns out that the corresponding *tail exponents* ξ_k, $k = 1, \ldots, M$, and ξ, defined by

$$\mathrm{P}[S_k \geq L] \doteq L^{-\xi_k}, \quad L \to \infty$$

and

$$\mathrm{P}[S \geq L] \doteq L^{-\xi}, \quad L \to \infty$$

are amenable to an analytical characterization. We note that $S = \sum_{k=1}^{M} S_k$ implies

$$\xi = \min\{\xi_1, \xi_2, \ldots, \xi_M\}. \tag{8}$$

The tail exponents characterize the TB of the corresponding complexity distributions in terms of polynomial decay rates in L for $L \to \infty$. We note that for finite (non-zero) tail exponents, the corresponding complexity distributions are of Pareto-type meaning that they decay polynomially in L. In particular, if the complexity distribution $\mathrm{P}[S \geq L]$ has tail exponent ξ, one can state that $L^{-(\xi+\delta)} \leq \mathrm{P}[S \geq L] \leq L^{-(\xi-\delta)}$ for any $\delta > 0$ and sufficiently large L. Furthermore, if $\mathrm{P}[S^{(1)} \geq L]$ and $\mathrm{P}[S^{(2)} \geq L]$ have tail exponents $\xi^{(1)}$ and $\xi^{(2)}$ with $\xi^{(1)} > \xi^{(2)}$, we can conclude that $\mathrm{P}[S^{(1)} \geq L] < \mathrm{P}[S^{(2)} \geq L]$ for sufficiently large L. Hence, larger tail exponents are desirable since this implies that the probability of the complexity being atypically large is smaller. This, for example, is advantageous in the context of MIMO detection under run-time constraints (i.e., under limits on the number of lattice points that can be searched). We emphasize, however, that the complexity tail exponents, as defined above, do not capture multiplicative constants and do not characterize the small L behavior of the corresponding complexity distributions.

B. Main Result

The complexity of SD can often be reduced by employing preprocessing techniques such as lattice-reduction (LR) or layer-sorting (LS) (see, e.g., [1]). In the remainder of this paper, we account for preprocessing by assuming that $\mathbf{y}$ is a general function of $\mathbf{r}$ and $\mathbf{H}$ and $\mathbf{R}$ is a general function of $\mathbf{H}$. For example, the direct QRD $\mathbf{H} = \mathbf{QR}$ (see Section I-A) results in the special case $\mathbf{y} = \mathbf{Q}^H \mathbf{r}$ and $\mathbf{R} = \mathbf{Q}^H \mathbf{H}$.

Theorem 1: Consider SD with fixed ρ ($0 < \rho < \infty$) and let $\mathbf{H}$ and $\mathbf{r}$ be random (potentially statistically dependent). The corresponding kth layer complexity S_k, defined in (7), satisfies

$$\mathrm{P}[S_k \geq L] \doteq \mathrm{P}\left[\frac{1}{\det(\mathbf{R}_k^H \mathbf{R}_k)} \geq L\right], \quad L \to \infty \tag{9}$$

if all of the following conditions are met:

- *Statistics of* $\mathbf{H}$: The probability density function (pdf) $f(\mathbf{H})$ of $\mathbf{H}$ satisfies the scaling property
$$f(\mathbf{H}) \geq \beta f(a\mathbf{H}) \tag{10}$$
for all $\mathbf{H} \in \mathbb{C}^{N\times M}$ and all $a \in \mathbb{R}$, $a > 1$, with some constant $\beta \in \mathbb{R}$, $\beta > 0$.
- *Statistics of* $\mathbf{H}$ *and preprocessing*: The covering radius $\mu(\mathbf{R})$ of $\mathcal{L}(\mathbf{R})$ satisfies
$$\mathrm{P}[\mu(\mathbf{R}) \geq L] \doteq L^{-\infty}, \quad L \to \infty. \tag{11}$$
- *Preprocessing*: Let $\det(\mathbf{R}_k^H\mathbf{R}_k) = g_k(\mathbf{H})$ with $g_k : \mathbb{C}^{N\times M} \mapsto \mathbb{R}_+$ and $\mu(\mathbf{R}) = g_\mu(\mathbf{H})$ with $g_\mu : \mathbb{C}^{N\times M} \mapsto \mathbb{R}_+$. The functions $g_k(\mathbf{H})$ and $g_\mu(\mathbf{H})$ satisfy, respectively, the scaling properties
$$g_k(b\mathbf{H}) = b^{\alpha_k} g_k(\mathbf{H}) \tag{12}$$
and
$$g_\mu(b\mathbf{H}) = b^{\alpha} g_\mu(\mathbf{H}) \tag{13}$$
for all $\mathbf{H} \in \mathbb{C}^{N\times M}$ and all $b \in \mathbb{R}$, $b > 0$, with some constants $\alpha_k, \alpha \in \mathbb{R}$, $\alpha_k > 0$, $\alpha > 0$.

Proof: See Appendix. □

Discussion: Theorem 1 states that the TB of $\mathrm{P}[S_k \geq L]$ is fully characterized by the TB of $\mathrm{P}[1/\det(\mathbf{R}_k^H\mathbf{R}_k) \geq L]$ provided the conditions (10) – (13) are satisfied. It is immediate that the TB of $\mathrm{P}[S_k \geq L]$ then depends neither on the statistics of $\mathbf{r}$ nor on the particular fixed choice of ρ. Consequently, $\mathbf{r}$ and ρ will influence $\mathrm{P}[S_k \geq L]$ (for example, a larger ρ will certainly result in a larger value of $\mathrm{P}[S_k \geq L]$) but do *not* affect the corresponding complexity tail exponent. The conditions (10) – (13) constitute fairly general requirements on the statistics of the lattice basis matrix $\mathbf{H}$ and on the preprocessing algorithm. For example, for direct QRD or conventional LR (see, e.g., [1]), it can be shown [12] that all the conditions above are satisfied if the entries of $\mathbf{H}$ are jointly Gaussian-distributed with arbitrary non-singular covariance matrix and arbitrary finite mean, i.e., for $\mathbf{H}$ being a Rayleigh- or Ricean-fading MIMO channel with (non-singular) covariance matrix.

It is interesting to note that $\det(\mathbf{R}_k^H\mathbf{R}_k)$ is the volume of a fundamental region of $\mathcal{L}(\mathbf{R}_k)$ [10]. A well-known approximation for S_k is given by [13]
$$\widehat{S}_k = \frac{V_k(\rho)}{\det(\mathbf{R}_k^H\mathbf{R}_k)}$$
where
$$V_k(\rho) = \frac{\pi^k(\rho^2)^k}{k!} \tag{14}$$
is the volume of a hypersphere in k complex-valued dimensions. This approximation simply counts the number of fundamental regions (each occupied by exactly one lattice point) that fit into the k-dimensional search sphere and becomes exact if an averaging of S_k is performed over $\mathbf{y}_k$ uniformly distributed over a fundamental region of $\mathcal{L}(\mathbf{R}_k)$ [13]. Motivated by this result, $\widehat{S}_k$ has been used in [1] and [11] to assess the complexity of various SD variants. For the TB, it immediately follows that (9) can equivalently be written as $\mathrm{P}[S_k \geq L] \doteq \mathrm{P}\big[\widehat{S}_k \geq L\big]$, $L \to \infty$, and no averaging argument is required.

III. Tail Exponents for i.i.d. Gaussian $\mathbf{H}$

Specializing Theorem 1 to lattice basis matrices $\mathbf{H}$ whose entries are i.i.d. $\mathcal{CN}(0, \sigma_H^2)$ (the model typically used in MIMO detection) leads to particularly interesting results. In this case, the pdf of $\mathbf{H}$ is given by $f(\mathbf{H}) = c_1\, e^{-c_2\|\mathbf{H}\|_{\mathrm{F}}^2}$ with some constants $c_1, c_2 > 0$, which directly implies that condition (10) is satisfied with $\beta = 1$.

A. Tail Exponents for Direct QRD

For direct QRD of $\mathbf{H}$ (i.e., $\mathbf{R}$ is obtained through $\mathbf{H} = \mathbf{QR}$), $\mu^2(\mathbf{R})$ can be upper-bounded as (see, e.g., [11, Prop. 1] extended to the complex-valued case) $\mu^2(\mathbf{R}) \leq \frac{1}{2}\sum_{i=1}^{M} R_{i,i}^2 = z^2$. It follows from [14, Lemma 2.1] that z is a χ-distributed RV, which, upon noting that $\mathrm{P}\,[\mu(\mathbf{R}) \geq L] \leq \mathrm{P}\big[z \geq L\big]$ implies that condition (11) is satisfied (see [12] for a detailed proof of this statement). Condition (12) is verified by observing that the QRD of $b\mathbf{H}$ results in $b\mathbf{R}$, which gives $g_k(b\mathbf{H}) = b^{2k} g_k(\mathbf{H})$. Condition (13) is shown to be satisfied by noting that (1) implies $\mu(b\mathbf{R}) = b\mu(\mathbf{R})$ and hence $g_\mu(b\mathbf{H}) = b g_\mu(\mathbf{H})$. Therefore, all the conditions of Theorem 1 are met. Finally, using results from [15], the TB of the distributions of the layer complexities of SD for direct QRD and i.i.d. Gaussian $\mathbf{H}$ can be established as [12]
$$\mathrm{P}[S_k \geq L] \doteq L^{-(N-M+1)}, \quad L \to \infty, \quad k = 1, \ldots, M \tag{15}$$
and, consequently,
$$\mathrm{P}[S \geq L] \doteq L^{-(N-M+1)}, \quad L \to \infty. \tag{16}$$
We conclude that the distributions of the layer and total complexities are of Pareto-type with tail exponents $\xi_k = \xi = N - M + 1$, $k = 1, \ldots, M$. These results show that increasing N (e.g., the number of receive antennas in a MIMO context) for given M (e.g., the number of transmit antennas) results in improved tail exponents.

B. Tail Exponents for LR-Based Preprocessing

We define LR-based preprocessing (see, e.g., [1]) as applying, prior to the QRD, the transformation $\widetilde{\mathbf{H}} = \mathbf{HT}$, where $\mathbf{T}$ is an $M \times M$ unimodular matrix, i.e., $T_{i,j} \in \mathbb{CZ}$, $\forall i, j$, and $|\det(\mathbf{T})| = 1$. The matrix $\mathbf{T}$ is obtained, for example, through the LLL algorithm [16], which finds a basis matrix $\widetilde{\mathbf{H}}$ of the lattice $\mathcal{L}(\mathbf{H})$ that is "closer" than $\mathbf{H}$ to an orthogonal matrix. Another important preprocessing technique is LS (e.g., with the V-BLAST algorithm [17]), which is just a special case of LR obtained by restricting $\mathbf{T}$ to be a permutation matrix.

The triangularized form of the CLP problem based on LR preprocessing is given by (3) with $\mathbf{R}$ and $\mathbf{y}$ replaced by $\widetilde{\mathbf{R}}$ and $\widetilde{\mathbf{y}} = \widetilde{\mathbf{Q}}^H\mathbf{r}$, respectively, where $\widetilde{\mathbf{Q}}$ and $\widetilde{\mathbf{R}}$ are the QR-factors of $\widetilde{\mathbf{H}}$, i.e., $\widetilde{\mathbf{H}} = \widetilde{\mathbf{Q}}\widetilde{\mathbf{R}}$. If we denote the corresponding solution of (3) as $\widetilde{\mathbf{d}}$, the final solution of (2) is $\widehat{\mathbf{d}} = \mathbf{T}\widetilde{\mathbf{d}}$. We now consider Theorem 1 with $\mathbf{R}$ replaced by $\widetilde{\mathbf{R}}$. Let us write $\mathbf{RT} = \mathbf{Q}'\mathbf{R}'$, where $\mathbf{Q}'$ and $\mathbf{R}'$ are the QR-factors of $\mathbf{RT}$. Noting that $\widetilde{\mathbf{H}} = \mathbf{QRT} = \widetilde{\mathbf{Q}}\widetilde{\mathbf{R}}$, we obtain $\mathbf{QQ}'\mathbf{R}' = \widetilde{\mathbf{Q}}\widetilde{\mathbf{R}}$. Since $\mathbf{QQ}'$ is unitary and the QR-factors are unique, it follows that $\mathbf{QQ}' = \widetilde{\mathbf{Q}}$ and $\mathbf{R}' = \widetilde{\mathbf{R}}$, which implies
$$\mathbf{R} = \mathbf{Q}'\widetilde{\mathbf{R}}\mathbf{T}^{-1}. \tag{17}$$

As $\mathbf{T}^{-1}$ is unimodular (since $\mathbf{T}$ is unimodular) and $\mathbf{Q}'$ is unitary, it can be verified that $\mu(\mathbf{R}) = \mu(\widetilde{\mathbf{R}})$ and $\det(\widetilde{\mathbf{R}}^H\widetilde{\mathbf{R}}) = \det(\mathbf{R}^H\mathbf{R})$. Due to $\mu(\mathbf{R}) = \mu(\widetilde{\mathbf{R}})$ and the results for direct QRD in Section III-A, conditions (11) and (13) are satisfied for LR-based preprocessing.

LR for $k = M$: Due to $\det(\widetilde{\mathbf{R}}^H\widetilde{\mathbf{R}}) = \det(\mathbf{R}^H\mathbf{R})$, condition (12) is satisfied for $k = M$ and LR-based preprocessing. Now, applying Theorem 1, we can immediately conclude that LR-based processing results in the *same* Mth layer complexity tail exponent as direct QRD, i.e.,

$$\mathrm{P}[S_M \geq L] \doteq L^{-(N-M+1)}, \quad L \to \infty \tag{18}$$

or, equivalently, $\xi_M = N - M + 1$. From (8), we can therefore conclude that $\xi \leq N - M + 1$ for LR-based preprocessing, which shows that LR (including LS) does *not* improve (i.e., increase) the total complexity tail exponent as compared to that obtained for direct QRD.

LR for $k < M$: It can be shown [12] that all LR algorithms delivering a unimodular transformation matrix $\mathbf{T}$, which is invariant to a positive scaling of $\mathbf{H}$, i.e., $\mathbf{H}$ and $b\mathbf{H}$ for all $b \in \mathbb{R}$, $b > 0$, result in the same $\mathbf{T}$, satisfy condition (12) for $k = 1, \ldots, M$. We note that this is the case for any reasonable LR algorithm we are aware of. Prominent examples are the LLL algorithm [16] and LS according to the V-BLAST algorithm [17]. Here, Theorem 1 therefore implies that

$$\mathrm{P}[S_k \geq L] \doteq \mathrm{P}\left[\frac{1}{\det(\widetilde{\mathbf{R}}_k^H\widetilde{\mathbf{R}}_k)} \geq L\right], \quad L \to \infty \tag{19}$$

for $k = 1, \ldots, M$.

LR Based on LLL: For LR carried out through the LLL algorithm [16] (see [18] for its complex-valued extension), based on (19), one can show the more specific result [12]

$$\mathrm{P}[S_k \geq L] \;\dot{\leq}\; L^{-\frac{N}{k}}, \quad L \to \infty, \quad k = 1, \ldots, M$$

or, equivalently, $\xi_k \geq N/k$. Compared with $\xi_k = N - M + 1$ for direct QRD (cf. (15)), we can conclude that LLL preprocessing improves (i.e., increases) the tail exponents up to layer $k \leq \lceil N/(N - M + 1) \rceil - 1$. In the following, consider $N = M$. We have $\xi_k = M/k > 1$, $k = 1, \ldots, M - 1$, and $\xi_M = 1$ (see (18)), which, in this case, establishes that the TB of the distribution of the total complexity of SD with LLL preprocessing is dominated by the TB of the distribution of the Mth layer complexity; in particular, we have $\xi = \xi_M = 1$, as in the case of direct QRD (cf. (16)).

IV. Numerical Results

We consider SD for data detection in $N \times M$ MIMO wireless systems, where $\mathbf{r} = \mathbf{H}\mathbf{d}' + \mathbf{w}$ with the entries of $\mathbf{H}$ and $\mathbf{w}$ assumed i.i.d. $\mathcal{CN}(0, 1/M)$ and i.i.d. $\mathcal{CN}(0, \sigma^2)$, respectively, and with the transmitted vector $\mathbf{d}' \in (\mathbb{CZ})^M$. We choose the radius ρ in (4) such that $\mathbf{d}'$ is found by the SD algorithm with probability 0.99 and for $1/\sigma^2$ we assume a value of 15 dB. We note that the complexity of SD is random in $\mathbf{H}$ and $\mathbf{w}$ and does not depend on $\mathbf{d}'$.

Fig. 1 shows the distribution of the total complexity $\mathrm{P}[S \geq L]$ in double log-scale for SD with direct QRD, V-BLAST LS [17], and with (complex-valued) LLL preprocessing [18, with parameter $\delta = 3/4$] for 4×4 and 5×4 MIMO systems. We can see that the results reflect our analytic findings. For example, for direct QRD in the 4×4 case, the distribution of the total complexity in Fig. 1 exhibits a large-L behavior of L^{-1} as predicted by (16) for $N = M$. Furthermore, it can be seen that adding one receive antenna, indeed, improves the TB and leads to a large-L behavior of L^{-2} (cf. (16)). Finally, the numerical results indicate that LLL preprocessing and V-BLAST LS do reduce the complexity, as compared to direct QRD, but do not improve the tail exponents.

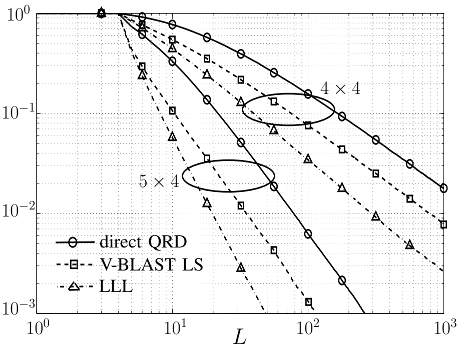

Fig. 1. Distribution of total complexity $\mathrm{P}[S \geq L]$ of SD with direct QRD, V-BLAST LS, and LLL preprocessing for 4×4 and 5×4 MIMO systems.

Acknowledgment

The authors would like to thank C. Akçaba and P. Coronel for helpful discussions.

Appendix
Proof of Theorem 1

The proof of Theorem 1 is based on separately establishing the exponential lower bound $\mathrm{P}[S_k \geq L] \;\dot{\geq}\; \mathrm{P}\big[1/\det(\mathbf{R}_k^H\mathbf{R}_k) \geq L\big]$, $L \to \infty$, and the exponential upper bound $\mathrm{P}[S_k \geq L] \;\dot{\leq}\; \mathrm{P}\big[1/\det(\mathbf{R}_k^H\mathbf{R}_k) \geq L\big]$, $L \to \infty$, which then combine to $\mathrm{P}[S_k \geq L] \doteq \mathrm{P}\big[1/\det(\mathbf{R}_k^H\mathbf{R}_k) \geq L\big]$.

A. Exponential Lower Bound

We start by noting that [13, Ch. 3.2, Eq. (3.3)]

$$S_k \geq \frac{V_k(\rho) - \mu(\mathbf{R}_k)A_k(\rho)}{\det(\mathbf{R}_k^H\mathbf{R}_k)}$$

where $V_k(\rho)$ and $A_k(\rho)$ denote the volume (cf. (14)) and the surface area of the search sphere at layer k, respectively.[1] Using $\mu(\mathbf{R}_k) \leq \mu(\mathbf{R})$, $k = 1, \ldots, M$, [12], we obtain

$$\mathrm{P}[S_k \geq L] \geq \mathrm{P}\left[\frac{V_k(\rho) - \mu(\mathbf{R})A_k(\rho)}{\det(\mathbf{R}_k^H\mathbf{R}_k)} \geq L\right].$$

Consider a constant $c \in \mathbb{R}$, $c > 0$, such that $V_k(\rho) - cA_k(\rho) > 0$ and define $c' = V_k(\rho) - cA_k(\rho) > 0$. We then have

$$\mathrm{P}[S_k \geq L] \geq \mathrm{P}\left[\mathbf{H} \in \mathcal{B}\right] \tag{20}$$

[1] Note that condition (11) implies full-rank $\mathbf{R}_k$ with probability one. In particular, $\det(\mathbf{R}_k^H\mathbf{R}_k) > 0$ and $\mu(\mathbf{R}_k) < \infty$ with probability one. However, it is straightforward to show that Theorem 1 also holds in the case where $\mathbf{R}_k$ is rank-deficient with non-zero probability, which leads to $\mathrm{P}[S_k \geq L] \doteq \mathrm{P}\big[1/\det(\mathbf{R}_k^H\mathbf{R}_k) \geq L\big] \doteq L^0$, $L \to \infty$.

where $\mathcal{B} = \left\{\mathbf{H}\colon \left(\frac{c'}{g_k(\mathbf{H})} \geq L\right) \cap (g_\mu(\mathbf{H}) \leq c)\right\}$ with $\det(\mathbf{R}_k^H\mathbf{R}_k) = g_k(\mathbf{H})$ and $\mu(\mathbf{R}) = g_\mu(\mathbf{H})$. With property (10), we further obtain

$$\mathrm{P}[\mathbf{H}\in\mathcal{B}] = \int_{\mathbf{H}\in\mathcal{B}} f(\mathbf{H})d\mathbf{H} \geq \beta \int_{\mathbf{H}\in\mathcal{B}} f(L^{\delta}\mathbf{H})d\mathbf{H}$$

for all $\delta > 0$, $L > 1$, and some $\beta > 0$. Performing the change of variables $\mathbf{H}' = L^{\delta}\mathbf{H}$ and invoking conditions (12) and (13) yields

$$\mathrm{P}[\mathbf{H}\in\mathcal{B}] \geq \beta\, L^{-2MN\delta}\,\mathrm{P}[\mathbf{H}\in\mathcal{B}'] \tag{21}$$

where

$$\mathcal{B}' = \left\{\mathbf{H}\colon \left(\frac{c'}{g_k(\mathbf{H})} \geq L^{1-\delta\alpha_k}\right) \cap \left(g_\mu(\mathbf{H}) \leq cL^{\delta\alpha}\right)\right\}.$$

Next, noting that for two events A_1 and A_2, by the inclusion-exclusion principle, $\mathrm{P}[A_1 \cap A_2] \geq \mathrm{P}[A_1] - \mathrm{P}[\bar{A}_2]$, where $\bar{A}_2$ denotes the complementary event of A_2, we get

$$\mathrm{P}[\mathbf{H}\in\mathcal{B}'] \geq \mathrm{P}\left[\frac{c'}{g_k(\mathbf{H})} \geq L^{1-\delta\alpha_k}\right] - \mathrm{P}\big[g_\mu(\mathbf{H}) > cL^{\delta\alpha}\big]\,.$$

Now (11) with $\mu(\mathbf{R}) = g_\mu(\mathbf{H})$ and $\delta, \alpha > 0$ implies $\mathrm{P}\big[g_\mu(\mathbf{H}) > cL^{\delta\alpha}\big] \doteq L^{-\infty}$, $L \to \infty$, which, together with (20) and (21), yields

$$\mathrm{P}[S_k \geq L] \mathrel{\dot{\geq}} L^{-2MN\delta}\,\mathrm{P}\left[\frac{c'}{g_k(\mathbf{H})} \geq L^{1-\delta\alpha_k}\right], \quad L \to \infty.$$

Let us write $\mathrm{P}[1/g_k(\mathbf{H}) \geq L] \doteq L^{-a}$, $L \to \infty$, for some constant $a \geq 0$. We then have $\mathrm{P}[S_k \geq L] \mathrel{\dot{\geq}} L^{-2MN\delta-(1-\delta\alpha_k)a}$, $L \to \infty$. As this result holds for arbitrarily small values of δ, we can conclude that $\mathrm{P}[S_k \geq L] \mathrel{\dot{\geq}} L^{-a} \doteq \mathrm{P}[1/g_k(\mathbf{H}) \geq L]$, $L \to \infty$, which establishes the exponential lower bound.

B. Exponential Upper Bound

From [13, Ch. 3.2, Eq. (3.3)]

$$S_k \leq \frac{V_k(\rho + \mu(\mathbf{R}_k))}{\det(\mathbf{R}_k^H\mathbf{R}_k)}$$

which, again using $\mu(\mathbf{R}_k) \leq \mu(\mathbf{R})$, $k = 1, \ldots, M$, results in

$$\mathrm{P}[S_k \geq L] \leq \mathrm{P}\left[\frac{V_k(\rho + \mu(\mathbf{R}))}{\det(\mathbf{R}_k^H\mathbf{R}_k)} \geq L\right]. \tag{22}$$

Note that $\mathrm{P}[xy \geq L] = \mathrm{P}\big[(xy \geq L) \cap (y < L^{\delta})\big] + \mathrm{P}\big[(xy \geq L) \cap (y \geq L^{\delta})\big] \leq \mathrm{P}\big[x \geq L^{1-\delta}\big] + \mathrm{P}\big[y \geq L^{\delta}\big]$ for any two RVs $x, y \in \mathbb{R}$ and any constant $\delta \in \mathbb{R}$, $0 < \delta < 1$. Applying this to (22) with $x = 1/\det(\mathbf{R}_k^H\mathbf{R}_k)$ and $y = V_k(\rho + \mu(\mathbf{R}))$, we get

$$\mathrm{P}[S_k \geq L] \leq \mathrm{P}\left[\frac{1}{\det(\mathbf{R}_k^H\mathbf{R}_k)} \geq L^{1-\delta}\right] + \mathrm{P}\big[V_k(\rho + \mu(\mathbf{R})) \geq L^{\delta}\,\big]\,.$$

With (14) and the binomial theorem, we can write

$$V_k\big(\rho + \mu(\mathbf{R})\big) = \frac{\pi^k}{k!}\sum_{i=0}^{2k}\binom{2k}{i}\rho^{2k-i}\big(\mu(\mathbf{R})\big)^i$$

which, using $\mathrm{P}\Big[\sum_{i=1}^{M} x_i \geq L\Big] \leq \sum_{i=1}^{M}\mathrm{P}[x_i \geq L/M]$ for any set of RVs $\{x_i\}_{i=1}^{M}$ yields

$$\mathrm{P}\big[V_k(\rho + \mu(\mathbf{R})) \geq L^{\delta}\big] \mathrel{\dot{\leq}} \sum_{i=0}^{2k}\mathrm{P}\Big[\big(\mu(\mathbf{R})\big)^i \geq L^{\delta}\Big], \quad L \to \infty.$$

Property (11) (for the terms corresponding to $i > 0$) and $\mathrm{P}[c'' \geq L] \leq e^{-(L-c'')} \doteq L^{-\infty}$, $L \to \infty$, for any constant $c'' \geq 0$ (for the term corresponding to $i = 0$) now directly imply $\mathrm{P}\big[V_k(\rho + \mu(\mathbf{R})) \geq L^{\delta}\big] \doteq L^{-\infty}$, $L \to \infty$, and, hence,

$$\mathrm{P}[S_k \geq L] \mathrel{\dot{\leq}} \mathrm{P}\left[\frac{1}{\det(\mathbf{R}_k^H\mathbf{R}_k)} \geq L^{1-\delta}\right], \quad L \to \infty.$$

As before, writing $\mathrm{P}\big[1/\det(\mathbf{R}_k^H\mathbf{R}_k) \geq L\big] \doteq L^{-a}$, $L \to \infty$, for some constant $a \geq 0$, we get $\mathrm{P}[S_k \geq L] \mathrel{\dot{\leq}} L^{-(1-\delta)a}$, $L \to \infty$. As this result holds for arbitrarily small values of δ, we can conclude that $\mathrm{P}[S_k \geq L] \mathrel{\dot{\leq}} L^{-a} \doteq \mathrm{P}\big[1/\det(\mathbf{R}_k^H\mathbf{R}_k) \geq L\big]$, $L \to \infty$, which establishes the exponential upper bound.

References

[1] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[2] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Comp.*, vol. 44, pp. 463–471, April 1985.

[3] W. H. Mow, "Maximum likelihood sequence estimation from the lattice viewpoint," in *Proc. ICCS/ISITA 1992*, vol. 1, Singapore, Nov. 1992, pp. 127–131.

[4] E. Viterbo and E. Biglieri, "A universal decoding algorithm for lattice codes," in *GRETSI 14-ème Colloq.*, Juan-les-Pins, France, Sept. 1993, pp. 611–614.

[5] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bölcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. of Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, July 2005.

[6] C. Studer, A. Burg, and H. Bölcskei, "Soft-output sphere decoding: Algorithms and VLSI implementation," *IEEE J. on Select. Areas in Comm.*, vol. 26, no. 2, pp. 290–300, Feb. 2008.

[7] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm I. Expected complexity," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[8] H. Vikalo and B. Hassibi, "On the sphere decoding algorithm II. Generalizations, second-order statistics, and applications to communications," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2819–2834, Aug. 2005.

[9] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Processing*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[10] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. Berlin, Heidelberg, New York: Springer, 1988.

[11] A. H. Banihashemi and A. K. Khandani, "On the complexity of decoding lattices using the Korkin-Zolotarev reduced basis," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 162–171, Jan. 1998.

[12] D. Seethaler, J. Jaldén, C. Studer, and H. Bölcskei, "On the complexity distribution of sphere-decoding," *in preparation*.

[13] P. M. Gruber and J. M. Wills, Eds., *Handbook of Convex Geometry*. vol. B, North Holland, Amsterdam: Elsevier, 1993.

[14] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Hanover, MA: Now Publishers Inc. 2004.

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

[16] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, "Factoring polynomials with rational coefficients," *Math. Ann.*, vol. 261, pp. 515–534, 1982.

[17] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Elect. Lett.*, vol. 35, pp. 14–16, Jan. 1999.

[18] Y. H. Gan and W. H. Mow, "Complex lattice reduction algorithms for low-complexity MIMO detection," in *Proc. IEEE GLOBECOM 2005*, vol. 5, St. Louis, MI, USA, Nov. 2005, pp. 2953–2957.