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. Characterizing 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.

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

Notation: We write $A_{i,j}$ for the entry in the $i$th row and $j$th column of the matrix $A$ and $x_i$ for the $i$th entry of the vector $x$. Slightly abusing common terminology, we call an $N \times M$, $N \geq M$, matrix $A$ unitary if it satisfies $A^H A = I$, where $^H$ denotes conjugate transposition and $I$ is the identity matrix. The Euclidean- and the Frobenius norm are denoted by $\| \cdot \|$ and $\| \cdot \|_F$, respectively, and $|X|$ refers to the cardinality of the set $X$. The ceiling 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 $A$ is defined as $L(A) = \{ A d : d \in (\mathbb{CZ})^M \}$. For $N = M$, the corresponding covering radius is given by [10]

$$\mu(A) = \max_{x \in \mathbb{C}^N} \min_{d \in (\mathbb{CZ})^M} \| x - A d \|_2.$$ (1)

A circularly symmetric complex Gaussian random variable (RV) with variance $\sigma_2^2$ is denoted as $x \sim \mathcal{CN}(0, \sigma_2^2)$. The natural logarithm is referred to as log($\cdot$). We write $g(x) = f(x)$, $x \rightarrow x_0$, for $\lim_{x \rightarrow x_0} \log g(x)/\log f(x) = \lim_{x \rightarrow x_0} \log f(x)/\log f(x)$, assuming that the corresponding limits exist. The symbols $\leq$ and $\geq$ are defined analogously. Finally, non-polynomial behavior of $g(x)$ is captured by $\lim_{x \rightarrow x_0} \log g(x)/\log f(x) = \pm \infty$ or $\lim_{x \rightarrow x_0} \log g(x)/\log f(x) = 0$, for which we write $g(x) = x^\pm \infty$, $x \rightarrow x_0$, and $g(x) = x^0$, $x \rightarrow x_0$, respectively.

A. Sphere-Decoding

The CLP problem refers to computing

$$\hat{d} = \arg \min_{d \in (\mathbb{CZ})^M} \| r - H d \|_2$$ (2)

for a given vector $r \in \mathbb{C}^N$ and a given full-rank matrix $H \in \mathbb{C}^{N \times M}$, $N \geq M$. In words, solving (2) amounts to finding the point in the lattice $L(H)$ that is closest (in Euclidean distance) to $r$. In communications, (2) is known as the maximum-likelihood (ML) detection problem for detecting $d' \in (\mathbb{CZ})^M$ based on the linear model $r = H d' + w$ with $H$ known at the receiver and $w$ 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) $H = QR$, where $R$ denotes an $M \times M$ upper triangular matrix.

This work was supported...
with positive real-valued elements on its main diagonal and \( Q \) of size \( N \times M \) is unitary. Then, (2) can equivalently be written as

\[
\hat{d} = \arg \min_{d \in \mathbb{C}^M} \| y - Rd \|^2 \tag{3}
\]

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

\[
\| y - Rd \|^2 \leq \rho^2. \tag{4}
\]

Here, the sphere radius \( \rho \) has to be chosen sufficiently large for the search sphere to contain at least one lattice point \( Rd \). Note, however, that if \( \rho \) 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 \( d_k \in \mathbb{C}^k \) of \( d \) defined as \( d_k = (d_{M-k+1} \cdots d_M)^T \), \( k = 1, \ldots, M \), where \( k \) is a layer index. The metric \( \| y - Rd \|^2 = \| y - R_M d_M \|^2 \) can be computed recursively according to

\[
\| y_k - R_k d_k \|^2 = \| y_{k-1} - R_{k-1} d_{k-1} \|^2 + \| \Delta_k (d_k) \|^2 \tag{5}
\]

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

\[
\| y_k - R_k d_k \|^2 \leq \rho^2. \tag{6}
\]

This formulation now enables finding all integer vectors \( d \) 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 \( d_k \in \mathbb{C}^k \), \( k = 1, \ldots, M \), that satisfy the partial SCs in (6) (cf. [7], [11]). Specifically, we define the \( k \)th layer complexity of SD as

\[
S_k = \{ d_k \in \mathbb{C}^k : \| y_k - R_k d_k \|^2 \leq \rho^2 \} \tag{7}
\]

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 \( H \), \( r \), and \( \rho \). In the following, we let \( H \) and \( r \) be random (potentially statistically dependent) and consider a fixed \( \rho \) that does not depend on the realizations of \( H \) and \( r \). For example, in the case of ML detection of the transmitted data vector \( d^* \) in MIMO wireless systems, \( r = H d^* + w \), where the entries of \( H \) (the channel matrix) and \( w \) (the noise vector) are typically assumed i.i.d. circularly symmetric complex Gaussian. In this setting, a reasonable fixed choice of \( \rho \) 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 \( H \) and \( r \) are random, \( S_k \) and \( S \) are random as well and can be characterized through their respective distribution functions \( P[S_k \geq L] \) and \( P[S \geq L] \). While these distributions seem hard to come by analytically, it turns out that the corresponding tail exponents \( \xi_k, k = 1, \ldots, M \), and \( \xi \), defined by

\[
P[S_k \geq L] = L^{-\xi_k}, \quad L \to \infty \quad \text{and} \quad P[S \geq L] = L^{-\xi}, \quad L \to \infty
\]

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 \( P[S \geq L] \) has tail exponent \( \xi \), one can state that \( L^{-(\xi+1)} \leq P[S \geq L] \leq L^{-(\xi-\delta)} \) for any \( \delta > 0 \) and sufficiently large \( L \). Furthermore, if \( P[S(1) \geq L] \) and \( P[S(2) \geq L] \) have tail exponents \( \xi(1) \) and \( \xi(2) \) with \( \xi(1) > \xi(2) \), we can conclude that \( P[S(1) \geq L] < P[S(2) \geq L] \) 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 \( y \) is a general function of \( r \) and \( H \) and \( R \) is a general function of \( H \). For example, the direct QRD \( H = QR \) (see Section I-A) results in the special case \( y = Q^H r \) and \( R = Q^H H \).

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

\[
P[S_k \geq L] = P\left[ \frac{1}{\det(R_k^H R_k)} \geq L \right], \quad L \to \infty \tag{9}
\]

if all of the following conditions are met:
of that the TB of $P$ provided the conditions (10) – (13) are satisfied. It is immediate where $k$ fundamental regions (each occupied by exactly one lattice dimensions. This approximation simply counts the number of $k$ is the volume of a hypersphere in $L$. It immediately follows that (9) can equivalently be written as

$$P[\mu(R) \geq L] = L^{-\infty}, \quad L \to \infty. \quad (11)$$

- **Preprocessing:** Let $\det(R_k^H R_k) = g_k(\mu)$ with $g_k : \mathbb{C}^{N \times M} \to \mathbb{R}_+$ and $\mu(R) = g_{\mu}(H)$ with $g_{\mu} : \mathbb{C}^{N \times M} \to \mathbb{R}_+$). The functions $g_k$ and $g_{\mu}$ satisfy, respectively, the scaling properties

$$g_k(bH) = b^{\alpha_k} g_k(H) \quad (12)$$

and

$$g_{\mu}(bH) = b^{\alpha_{\mu}} g_{\mu}(H) \quad (13)$$

for all $H \in \mathbb{C}^{N \times M}$ and all $b \in \mathbb{R}$, $b > 0$, with some constants $\alpha_k, \alpha_{\mu} \in \mathbb{R}$, $\alpha_k > 0$, $\alpha_{\mu} > 0$.

**Proof:** See Appendix. □

**Discussion:** Theorem 1 states that the TB of $P[S_k \geq L]$ is fully characterized by the TB of $P[1/\det(R_k^H R_k) \geq L]$ provided the conditions (10) – (13) are satisfied. It is immediate that the TB of $P[S_k \geq L]$ then depends neither on the statistics of $\mathbf{r}$ nor on the particular fixed choice of $\rho$. Consequently, $\mathbf{r}$ and $\rho$ will influence $P[S_k \geq L]$ (for example, a larger $\rho$ will certainly result in a larger value of $P[S_k \geq L]$) 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 $H$ 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 $H$ are jointly Gaussian-distributed with arbitrary non-singular covariance matrix and arbitrary finite mean, i.e., for $H$ being a Rayleigh- or Ricean-fading MIMO channel with (non-singular) covariance matrix.

It is interesting to note that $\det(R_k^H R_k)$ is the volume of a fundamental region of $\mathcal{L}(R_k)$ [10]. A well-known approximation for $S_k$ is given by [13]

$$\hat{S}_k = \frac{V_k(\rho)}{\det(R_k^H R_k)}$$

where

$$V_k(\rho) = \frac{\pi^{k(\rho^2)k}}{k!}$$

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 $\mathcal{L}(R_k)$ [13]. Motivated by this result, $\hat{S}_k$ has been used in [1] and [11] to assess the complexity of various SD variants. For the TB, it immediately follows that (10) can equivalently be written as $P[S_k \geq L] = P[\hat{S}_k \geq L], L \to \infty$, and no averaging argument is required.

### III. Tail Exponents for i.i.d. Gaussian $H$

Specializing Theorem 1 to lattice basis matrices $H$ whose entries are i.i.d. $\mathcal{C}\mathcal{N}(0,\sigma^2_R)$ (the model typically used in MIMO detection) leads to particularly interesting results. In this case, the pdf of $H$ is given by $f(H) = c_1 e^{-c_2 ||H||_F^2}$ 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 $H$ (i.e., $R$ is obtained through $H = \mathbf{Q}\mathbf{R}$), $\mu^2(R)$ can be upper-bounded as (see, e.g., [11, Prop. 1]) extended to the complex-valued case) $\mu^2(R) \leq \frac{1}{2} \sum_{i,j=1}^{M} r_{i,j}^2 = z^2$. It follows from [14, Lemma 2.1] that $z$ is a $\chi$-distributed RV, which, upon noting that $P[\mu(R) \geq L] \leq P[z \geq L]$ 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 $bH$ results in $b\mathbf{R}$, which gives $g_k(bH) = b^{2k} g_k(H)$. Condition (13) is shown to be satisfied by noting that (11) implies $\mu(b\mathbf{R}) = b \mu(\mathbf{R})$ and hence $g_{\mu}(bH) = b^{\alpha_{\mu}} g_{\mu}(H)$. 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 $H$ can be established as [12]

$$P[S_k \geq L] \sim L^{-(N-M+1)}, \quad L \to \infty, \quad k = 1, \ldots, M \quad (15)$$

and, consequently,

$$P[S \geq L] \sim L^{-(N-M+1)}, \quad L \to \infty. \quad (16)$$

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 $H = \mathbf{HT}$, where $\mathbf{T}$ is an $M \times M$ unimodular matrix, i.e., $T_{i,j} \in \mathbb{C}, \forall i, j$, and $|\det(\mathbf{T})| = 1$. The matrix $\mathbf{T}$ is obtained, for example, through the LLL algorithm [16], which finds a basis matrix $\mathbf{H}$ of the lattice $\mathcal{L}(\mathbf{H})$ that is “closer” than $\mathbf{H}$ to an orthogonal matrix. Another important preprocessing technique is $\mathbf{LS}$ (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 (5) with $\mathbf{R}$ and $\mathbf{y}$ replaced by $\mathbf{R}$ and $\mathbf{y} = \mathbf{Q}^H \mathbf{r}$, respectively, where $\mathbf{Q}$ and $\mathbf{R}$ are the QR-factors of $\mathbf{H}$, i.e., $\mathbf{H} = \mathbf{Q}\mathbf{R}$. If we denote the corresponding solution of (6) as $\hat{\mathbf{d}}$, the final solution of (2) is $\hat{\mathbf{d}} = \mathbf{T}\hat{\mathbf{d}}$. We now consider Theorem 1 with $\mathbf{R}$ replaced by $\mathbf{R}$. Let us write $\mathbf{RT} = \mathbf{Q}' \mathbf{R}'$, where $\mathbf{Q}'$ and $\mathbf{R}'$ are the QR-factors of $\mathbf{RT}$. Noting that $\mathbf{H} = \mathbf{QRT} = \mathbf{Q}\mathbf{R}$, we obtain $\mathbf{QQ}' \mathbf{R}' = \mathbf{QR}$. Since $\mathbf{QQ}'$ is unitary and the QR-factors are unique, it follows that $\mathbf{QQ}' = \mathbf{Q}$ and $\mathbf{R}' = \mathbf{R}$, which implies

$$\mathbf{R} = \mathbf{Q}' \mathbf{R}^{-1}. \quad (17)$$
As $T^{-1}$ is unimodular (since $T$ is unimodular) and $Q'$ is unitary, it can be verified that $\mu(R) = \mu(R')$ and $\det(R'HR)$ is $\det(RHHR)$. Due to $\mu(R) = \mu(R')$, the results for direct QRD in Section II-A conditions (11) and (13) are satisfied for LR-based preprocessing.

**LR for $k = M$:** Due to $\det(R'HR) = \det(RHHR)$, condition (12) is satisfied for $k = M$ and LR-based preprocessing. Now, applying Theorem 1, we can immediately conclude that LR-based preprocessing results in the same $M$th layer complexity tail exponent as direct QRD, i.e.,

$$P[S_M \geq L] = L^{-(N-M+1)} \cdot L \to \infty$$

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)

$$P[S_k \geq L] \leq L^{-\frac{\xi_k}{2}} \cdot L \to \infty, \quad k = 1, \ldots, M$$

or, equivalently, $\xi_k \geq N/k$. Compared with $\xi_k = N - M + 1$ for the LR QRD (cf. (15)), we can conclude that LLL preprocessing improves (i.e., increases) the tail exponents up to $N$. We note that this is the case for any reasonable LR algorithm we are aware of. Prominent examples are the LLL-based processing algorithms [16] and LS according to the V-BLAST algorithm [17]. Here, Theorem 1 therefore implies that

$$P[S_k \geq L] \geq \frac{1}{\det(R^H_k R_k)} \cdot L \to \infty$$

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

**Figure 1:** Distribution of total complexity $P[S \geq L]$ of SD with direct QRD, V-BLAST LS, and LLL preprocessing for $4 \times 4$ and $5 \times 4$ 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 $P[S_k \geq L] \geq P[1/\det(R^H_k R_k) \geq L], L \to \infty$, and the exponential upper bound $P[S_k \geq L] \leq P[1/\det(R^H_k R_k) \geq L], L \to \infty$, which then combine to $P[S_k \geq L] \geq P[1/\det(R^H_k R_k) \geq L]$.

**A. Exponential Lower Bound**

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

$$S_k \geq \frac{V_k(\rho) - \mu(R_k)A_k(\rho)}{\det(R^H_k R_k)}$$

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.

Using $\mu(R_k) \leq \mu(R), k = 1, \ldots, M$, [12], we obtain

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

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

$$P[S_k \geq L] \geq P[H \in B],$$

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

Note that condition (11) implies full-rank $R_k$ with probability one. In particular, $\det(R^H_k R_k) > 0$ and $\mu(R_k) < \infty$ with probability one.
where \( B = \{ H : (c' / g_k(H)) \geq L \cap (g_\mu(H) \leq c) \} \) with
\[
\det(R^H_k R_k) = g_k(H) \quad \text{and} \quad \mu(R) = g_\mu(H).
\]
With property (10), we further obtain
\[
P[H \in B] = \int_{H \in B} f(H) dH \geq \beta \int_{H \in B} f(L^\delta H) dH
\]
for all \( \delta > 0, L > 1 \), and some \( \beta > 0 \). Performing the change of
variables \( H' = L^\delta H \) and invoking conditions (12) and (13) yields
\[
P[H \in B] \geq \beta L^{-2MN\delta} P[H \in B']
\]
where
\[
B' = \{ H : \left( \frac{c'}{g_k(H)} \right) \geq L^{1-\delta a} \cap (g_\mu(H) \leq cL^{\delta a}) \}.
\]
Next, noting that for two events \( A_1 \) and \( A_2 \), by the inclusion-exclusion principle,
\( P[A_1 \cap A_2] \geq P[A_1] - P[A_2] \), where \( A_2 \) denotes the complementary event of \( A_2 \), we get
\[
P[H \in B'] \geq P\left[ \frac{c'}{g_k(H)} \geq L^{1-\delta a} \right] - P[g_\mu(H) > cL^{\delta a}].
\]
Now (11) with \( \mu(R) = g_k(H) \) and \( \delta, a > 0 \) implies
\[
P[g_\mu(H) > cL^{\delta a}] \leq L^{-\infty}, \quad L \to \infty
\]
which, together with (20) and (21), yields
\[
P[S_k \geq L] \geq L^{-2MN\delta} P\left[ \frac{c'}{g_k(H)} \geq L^{1-\delta a} \right], \quad L \to \infty.
\]
Let us write \( P[1/g_k(H) \geq L] \leq L^{-a}, \quad L \to \infty \), for some constant \( a \geq 0 \). We then have \( P[S_k \geq L] \geq L^{-2MN\delta - (1-\delta a)a}, \quad L \to \infty \). As this result holds for arbitrarily small values of \( \delta \), we can conclude that \( P[S_k \geq L] \geq L^{-a} \Rightarrow P[1/g_k(H) \geq L], \quad L \to \infty \), which establishes the exponential lower bound.

### B. Exponential Upper Bound

From (13, Ch. 3.2, Eq. (3.3))
\[
S_k \leq \frac{V_k(\rho + \mu(R_k))}{\det(R^H_k R_k)}
\]
which, again using \( \mu(R_k) \leq \mu(R), \quad k = 1, \ldots, M \), results in
\[
P[S_k \geq L] \leq P\left[ \frac{V_k(\rho + \mu(R))}{\det(R^H_k R_k)} \geq L \right].
\]
Note that \( P[x y \geq L] = P[(x y \geq L) \cap (y \leq L^\delta)] + P[(xy \geq L) \cap (y \geq L^\delta)] \)
\[
\leq P[x \geq L^{1-\delta}] + P[y \geq L^\delta]
\]
for any two RVs \( x, y \in R \) and any constant \( \delta \in R, \quad 0 < \delta < 1 \). Applying this to (22) with \( x = 1/\det(R^H_k R_k) \) and \( y = V_k(\rho + \mu(R)) \), we get
\[
P[S_k \geq L] \leq P\left[ \frac{1}{\det(R^H_k R_k)} \geq L^{1-\delta} \right] + P[V_k(\rho + \mu(R)) \geq L^\delta].
\]
With (14) and the binomial theorem, we can write
\[
V_k(\rho + \mu(R)) = \pi^k / k! \sum_{i=0}^{2k} \binom{2k}{i} \rho^{2k-i}(\mu(R))^i
\]
which, using \( P[\sum_{i=1}^M x_i \geq L] \leq \sum_{i=1}^M P[x_i \geq L/M] \) for any set of RVs \( \{x_i\}_{i=1}^M \) yields
\[
P[V_k(\rho + \mu(R)) \geq L^\delta] \leq \sum_{i=0}^{2k} P[(\mu(R))^i \geq L^\delta], \quad L \to \infty.
\]
Property (11) (for the terms corresponding to \( i > 0 \)) and \( P[c' \geq L] \leq e^{(L-c')/L} \rightarrow L^{-\infty}, \quad L \to \infty \), for any constant \( \delta' \geq 0 \) (for the term corresponding to \( i = 0 \)) now directly imply \( P[V_k(\rho + \mu(R)) \geq L^\delta] \rightarrow L^{-\infty}, \quad L \to \infty \), and, hence,
\[
P[S_k \geq L] \leq P\left[ \frac{1}{\det(R^H_k R_k)} \geq L^{1-\delta} \right], \quad L \to \infty.
\]
As before, writing \( P[1/\det(R^H_k R_k) \geq L] \leq L^{-a}, \quad L \to \infty \), for some constant \( a \geq 0 \), we get \( P[S_k \geq L] \leq L^{-a} \Rightarrow P[1/\det(R^H_k R_k) \geq L], \quad L \to \infty \). As this result holds for arbitrarily small values of \( \delta \), we can conclude that \( P[S_k \geq L] \leq L^{-a} \Rightarrow P[1/\det(R^H_k R_k) \geq L], \quad L \to \infty \), 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. ICCCISISITA 1992, vol. 1, Singapore, Nov. 1992, pp. 127–131.

[4] E. Viterbo and E. Biglieri, “A universal decoding algorithm for lattice codes,” in GRETSI 14ème 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.