Performance and Complexity Analysis of Infinity-Norm Sphere-Decoding

Dominik Seethaler and Helmut Bölcskei, *Fellow, IEEE*

***Abstract*—Promising approaches for efficient detection in multiple-input multiple-output (MIMO) wireless systems are based on sphere-decoding (SD). The conventional (and optimum) norm that is used to conduct the tree traversal step in SD is the l^2-norm. It was, however, recently observed that using the l^∞-norm instead reduces the hardware complexity of SD considerably at only a marginal performance loss. These savings result from a reduction in the length of the critical path in the circuit and the silicon area required for metric computation, but are also, as observed previously through simulation results, a consequence of a reduction in the computational (i.e., algorithmic) complexity. The aim of this paper is an analytical performance and computational complexity analysis of l^∞-norm SD. For independent and identically distributed (i.i.d.) Rayleigh fading MIMO channels, we show that l^∞-norm SD achieves full diversity order with an asymptotic SNR gap, compared to l^2-norm SD, that increases at most linearly in the number of receive antennas. Moreover, we provide a closed-form expression for the computational complexity of l^∞-norm SD based on which we establish that its complexity scales exponentially in the system size. Finally, we characterize the tree pruning behavior of l^∞-norm SD and show that it behaves fundamentally different from that of l^2-norm SD.**

***Index Terms*—Algorithmic complexity, data detection, hardware complexity, infinity norm, maximum-likelihood, multiple-input multiple-output (MIMO) wireless.**

I. Introduction

MULTIPLE-INPUT multiple-output (MIMO) wireless systems offer considerable gains over single-antenna systems, in terms of throughput and link reliability, see, e.g., [1]. These gains come, however, at a significant increase in receiver complexity. In particular, one of the most challenging problems in MIMO receiver design is the development of hardware-efficient data detection algorithms achieving (close-to) optimum performance [2]. Among the most promising approaches to the solution of this problem is the so-called sphere-decoding (SD) algorithm [3]–[10], which performs optimum, i.e., maximum-likelihood (ML), detection through a weighted tree search. SD exhibits (often significantly) smaller computational complexity than exhaustive search ML detection [2], [10].

A. Hardware Implementation Aspects of SD

Hardware implementations of several variants of the SD algorithm are described in [2] and [8]. It is argued in [8] that the overall hardware complexity of SD is essentially determined by (i) the computational (i.e., algorithmic) complexity in terms of the number of nodes visited in the tree search and (ii) the circuit complexity in terms of the length of the critical path in the circuit and the required silicon area for metric computation. The length of the critical path limits the clock frequency of the circuit [11]. One of the main findings of [8] is that traversing the search tree based on the l^∞-norm instead of the l^2-norm incurs only a small performance loss while significantly reducing the overall hardware complexity by virtue of a reduction of both the computational and the circuit complexity. In the context of precoding for MIMO systems, it was shown in [12] that the l^∞-norm can be used for sphere-encoding to reduce the peak-power.

To understand where the reduction in the circuit complexity in [8] comes from, we refer to Fig. 1 (cf., [8, Fig. 2]) showing tradeoff curves between circuit area and the length of the critical path corresponding to the computation of the metrics $x_1^2 + x_2^2$ (squared l^2-norm) and $\max\{|x_1|, |x_2|\}$ (l^∞-norm) for $x_1, x_2 \in \mathbb{R}$. These tradeoffs can be realized by choosing different hardware implementations of the corresponding metric computation circuit. From Fig. 1 it can be seen that the computation of $\max\{|x_1|, |x_2|\}$ can be implemented much more efficiently in hardware than the computation of $x_1^2 + x_2^2$. The main reason for this is that evaluating $\max\{|x_1|, |x_2|\}$, in contrast to $x_1^2 + x_2^2$, does not require squaring operations. Replacing the l^2- by the l^∞-norm also has an impact on the computational complexity of SD. In particular, it was observed in [8], through simulation results, that SD based on the l^∞-norm (referred to as SD-l^∞) exhibits lower computational complexity than SD based on the l^2-norm (referred to as SD-l^2). Furthermore, the results in [8] indicate that the overall complexity (determined by both the circuit and the computational complexity) of SD-l^∞ is up to a factor of 5 lower than the overall complexity of SD-l^2. SD-l^∞, therefore, appears to be a promising approach to near-optimum MIMO detection at low hardware complexity.

B. Contributions

The aim of this paper is to deepen the understanding of SD-l^∞ through an analytical performance and computational

Manuscript received November 18, 2008; revised October 06, 2009. Current version published March 10, 2010. This work was supported in part by the STREP project No. IST-026905 (MASCOT) within the Sixth Framework Programme of the European Commission. The material in this paper was presented in part at the IEEE ISIT 2008, Seoul, Korea, July 2008.

D. Seethaler was with the Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Vienna, Austria. He was also with the Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland. He is now at Himmelreich 5, 5071 Salzburg, Austria (e-mail: dominik.seethaler@gmail.com).

H. Bölcskei is with the Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland (e-mail: boelcskei@nari.ee.ethz.ch).

Communicated by E. Viterbo, Associate Editor for Coding Techniques.

Color versions of Figures 1–5 in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2009.2039034

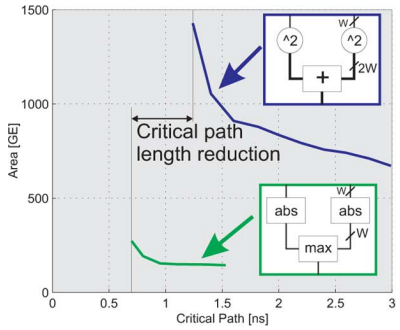

Fig. 1. Circuit area and critical path length tradeoff curves corresponding to the computation of $x_1^2+x_2^2$ (squared l^2-norm) and $\max\{|x_1|,|x_2|\}$ (l^∞-norm) for $x_1, x_2 \in \mathbb{R}$. The area is given in gate-equivalents (GE) and the length of the critical path is given in nanoseconds (ns). W denotes the word length. Figure taken from [8].

complexity analysis for i.i.d. Rayleigh fading MIMO channels. Our main contributions can be summarized as follows:

- We show that SD-l^∞ achieves the same (i.e., full) diversity order as SD-l^2.
- We show that the gap in signal-to-noise ratio (SNR) incurred by SD-l^∞, compared to SD-l^2, is bounded (as a function of SNR) and increases at most linearly in the number of receive antennas.
- We derive a closed-form expression for the complexity of SD-l^∞. Here and in the remainder of the paper, complexity is defined as the average number of nodes visited in the tree search, where averaging is performed with respect to the (random) channel, noise, and transmit signal. Corresponding results for SD-l^2 can be found in [10], [13]–[15].
- We prove that the complexity of SD-l^∞ scales exponentially in the number of transmit antennas. Our proof technique directly extends to SD-l^2 and thus yields an alternative (vis-à-vis [16]) proof of the exponential complexity scaling behavior of SD-l^2.
- Finally, we provide insights into the tree pruning behavior of SD-l^∞ relative to that of SD-l^2. In particular, based on an asymptotic (in SNR) analysis of our closed-form complexity expression, we show that SD-l^∞ prunes more aggressively than SD-l^2 at tree levels closer to the root of the search tree, whereas this behavior is reversed at tree levels closer to the leaves.

C. 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 $\mathbf{x}$. For unitary $\mathbf{A}$, we have $\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H = \mathbf{I}$, where ^{H} denotes conjugate transposition, i.e., transposition ^{T} followed by element-wise complex conjugation *, and $\mathbf{I}$ is the identity matrix. The l^2- and the l^∞-norm of a vector $\mathbf{x} = (x_1 \cdots x_M)^T \in \mathbb{C}^M$ are defined as $\|\mathbf{x}\|_2 = \sqrt{|x_1|^2+\cdots+|x_M|^2}$ and $\|\mathbf{x}\|_\infty = \max\{|x_1|,\ldots,|x_M|\}$, respectively. We will also need the $l^{\widetilde{\infty}}$-norm $\|\mathbf{x}\|_{\widetilde{\infty}} = \max\{|x_{\mathrm{R},1}|,|x_{\mathrm{I},1}|,\ldots,|x_{\mathrm{I},M}|\}$, where x_R and x_I denote the real and imaginary parts, respectively, of $x \in \mathbb{C}$. We note that the l^2-norm is invariant with respect to (w.r.t.) unitary transformations, i.e., $\|\mathbf{x}\|_2 = \|\mathbf{A}\mathbf{x}\|_2$ if $\mathbf{A}$ is unitary. $\mathbb{E}\{\cdot\}$ stands for the expectation operator and $\Phi_x(s) = \mathbb{E}\{e^{sx}\}$ refers to the moment generating function (MGF) of the random variable (RV) x. We write $x \sim \chi_a$ if the RV x is χ-distributed with $a \geq 0$ degrees of freedom and normalized such that $\mathbb{E}\{x^2\} = a$. The probability density function (pdf) of the RV $x \sim \chi_a$ is then given by [17]

$$f_x(t) = \frac{2^{1-a/2}}{\Gamma(a/2)} t^{a-1} e^{-\frac{t^2}{2}}, \qquad t \geq 0 \tag{1}$$

and $f_x(t) = 0$, $t < 0$, where $\Gamma(a) = \int_0^\infty y^{a-1}e^{-y}dy$ refers to the Gamma function. For the corresponding cumulative distribution function (cdf) we have $P[x \leq t] = \gamma_{a/2}(t^2/2)$. Here, $\gamma_a(t)$ denotes the (regularized) lower incomplete Gamma function; some important properties of $\gamma_a(t)$ are stated in Appendix C. We denote a circularly symmetric complex Gaussian RV with variance σ_x^2 as $x \sim \mathcal{CN}(0,\sigma_x^2)$; $x \sim \mathcal{N}(\mu_x,\sigma_x^2)$ refers to a real Gaussian distributed RV x with mean μ_x and variance σ_x^2. For independently and identically distributed (i.i.d.) RVs $x_i \sim \mathcal{N}(0,1)$, $i = 1,\ldots,a$, we have $z = \sqrt{x_1^2+\cdots+x_a^2} \sim \chi_a$. Furthermore, if the RV x is χ_a-distributed, x^2 is χ_a^2-distributed. We write $y \sim \chi_a^2$ if the RV y is χ_a^2-distributed with $\mathbb{E}\{y\} = a$. In particular, the MGF of the RV $y \sim \chi_a^2$ is given by

$$\Phi_y(s) = (1-2s)^{-a/2} \tag{2}$$

for any $s < 1/2$. The Q-function is defined as $Q(x) = \left(1/\sqrt{2\pi}\right)\int_x^\infty e^{-y^2/2}dy$, for $x \geq 0$. For equality in distribution we write $\stackrel{d}{=}$. Furthermore, the "Big O" notation $g(x) = \mathcal{O}(f(x))$, $x \to x_0$, denotes that $|g(x)/f(x)|$ remains bounded as $x \to x_0$ [18]. The "little o" notation $g(x) = o(f(x))$, $x \to x_0$, stands for $\lim_{x\to x_0} g(x)/f(x) = 0$, and $g(x) \stackrel{a}{\sim} f(x)$, $x \to x_0$, means that $\lim_{x\to x_0} g(x)/f(x) = 1$. By $g(x) \preceq f(x)$, $x \to x_0$, and $g(x) \succ f(x)$, $x \to x_0$, for positive functions $g(x)$ and $f(x)$, we denote $\lim_{x\to x_0} g(x)/f(x) \leq 1$ and $\lim_{x\to x_0} g(x)/f(x) > 1$, respectively. The Dirac delta function is referred to as $\delta(x)$, convolution is denoted as $*$, and $\log(\cdot)$ is the natural logarithm to the base e. The summations in $\sum_{\mathbf{x}}$ and $\sum_{\mathbf{x}\neq\mathbf{x}'}$ are over all possible values of $\mathbf{x}$ and over all possible values of $\mathbf{x}$ except for $\mathbf{x}'$, respectively. Finally, $f^{(n)}(x)$ refers to the nth derivative of the function $f(x)$ and $f'(x) = f^{(1)}(x)$.

D. System Model

We consider an $N \times M$ MIMO system with M transmit antennas and $N \geq M$ receive antennas. The corresponding complex-baseband input-output relation is given by

$$\mathbf{r} = \mathbf{H}\mathbf{d}' + \mathbf{w}$$

where $\mathbf{d}' = (d_1' \cdots d_M')^T$ denotes the transmitted data vector, $\mathbf{H}$ is the $N \times M$ channel matrix, $\mathbf{r} = (r_1 \cdots r_N)^T$ is the received vector, and $\mathbf{w} = (w_1 \cdots w_N)^T$ denotes the additive noise vector. The symbols d_m', drawn from a finite alphabet $\mathcal{A}$, have zero-mean and unit variance. Furthermore, we assume that the $H_{n,m}$ are i.i.d. $\mathcal{CN}(0,1/M)$ and the w_n are i.i.d. $\mathcal{CN}(0,\sigma^2)$. The SNR (per receive antenna) is therefore given by $\rho = 1/\sigma^2$.

E. Sphere-Decoding

We now review SD based on the l^2-norm [3]–[7], [9] and SD based on the l^∞-norm [8].

1) SD Based on the l^2-norm: SD-l^2 performs ML detection by finding

$$\widehat{\mathbf{d}}_{\mathrm{ML}} = \arg\min_{\mathbf{d}\in\mathcal{A}^M} \|\mathbf{r}-\mathbf{H}\mathbf{d}\|_2^2 \tag{3}$$

through a tree search subject to a *sphere constraint* (SC), which amounts to considering only those data vectors $\mathbf{d}$ that satisfy $\|\mathbf{r}-\mathbf{H}\mathbf{d}\|_2^2 \leq C_2^2$ (known as the Fincke-Pohst [3] strategy). Here, the radius C_2 has to be chosen sufficiently large for the corresponding search sphere to contain at least one data vector. Note, however, that if C_2 is chosen too large, too many points will satisfy the SC and the complexity of SD-l^2 will be high (for guidelines on how to choose C_2 see [10], [19] and Section III-D). The SC is then cast into a weighted tree search problem by first performing a QR-decomposition of $\mathbf{H}$ resulting in

$$\mathbf{H} = \mathbf{Q}\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}$$

where $\mathbf{Q}$ is an $N\times N$ unitary matrix, $\mathbf{R}$ is an $M\times M$ upper triangular matrix, and $\mathbf{0}$ denotes an all-zeros matrix of size $(N-M)\times M$. Then, the SC can equivalently be written as

$$\|\mathbf{z}(\mathbf{d})\|_2^2 \leq C_2^2 \tag{4}$$

where

$$\mathbf{z}(\mathbf{d}) = \mathbf{y} - \begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\mathbf{d} \quad \text{with} \quad \mathbf{y} = \mathbf{Q}^H\mathbf{r} = \begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\mathbf{d}' + \mathbf{n}. \tag{5}$$

Here, the unitarity of $\mathbf{Q}$ implies that $\mathbf{n} = \mathbf{Q}^H\mathbf{w}$ is again i.i.d. $\mathcal{CN}(0,\sigma^2)$. The length-k data subvectors $\mathbf{d}_k = (d_{M-k+1}\cdots d_M)^T$, $k=1,\ldots,M$, can be arranged in a tree with root above level $k=1$ and corresponding leaves at level $k=M$; a specific $\mathbf{d}_k$ is associated with a node in this tree at level k. Let us define

$$\mathbf{z}_k(\mathbf{d}_k) = \mathbf{y}_k - \begin{bmatrix}\mathbf{R}_k\\ \mathbf{0}\end{bmatrix}\mathbf{d}_k$$

as the vector containing the bottom $k+L$ with $L=N-M$ elements of $\mathbf{z}(\mathbf{d})$ in (5). Here, $\mathbf{R}_k$ denotes the $k\times k$ upper triangular submatrix of $\mathbf{R}$ associated with $\mathbf{d}_k$ and $\mathbf{y}_k = (y_{M-k+1}\cdots y_M\ y_{M+1}\cdots y_N)^T$. The metric $\|\mathbf{z}(\mathbf{d})\|_2^2 = \|\mathbf{z}_M(\mathbf{d}_M)\|_2^2$ can then be computed recursively according to

$$\|\mathbf{z}_k(\mathbf{d}_k)\|_2^2 = \|\mathbf{z}_{k-1}(\mathbf{d}_{k-1})\|_2^2 + |[\mathbf{z}(\mathbf{d})]_{M-k+1}|^2 \tag{6}$$

$k=1,\ldots,M$, where

$$|[\mathbf{z}(\mathbf{d})]_{M-k+1}|^2 = \left| y_{M-k+1} - \sum_{i=M-k+1}^{M} R_{M-k+1,i}d_i \right|^2. \tag{7}$$

Thus, with (6), a necessary condition for $\mathbf{d}$ to satisfy the SC (4) is that any associated $\mathbf{d}_k$ satisfies the *partial SC* (PSC)

$$\|\mathbf{z}_k(\mathbf{d}_k)\|_2^2 \leq C_2^2. \tag{8}$$

Consequently, we can find all data vectors inside the search sphere, i.e., all data vectors satisfying the SC (4), through a weighted tree search. The tree is traversed starting at level $k=1$. If the PSC is violated by a given $\mathbf{d}_k$, the node associated with that $\mathbf{d}_k$ along with all its children is pruned from the tree. The ML solution (3) is found by choosing, among all surviving leaf nodes $\mathbf{d}=\mathbf{d}_M$, the one with minimum $\|\mathbf{z}(\mathbf{d})\|_2$.

2) SD Based on the l^∞-Norm: We define SD-l^∞ as the algorithm obtained by replacing the SC (4) by the *box constraint* (BC) $\|\mathbf{z}(\mathbf{d})\|_\infty \leq C_\infty$. The metric $\|\mathbf{z}(\mathbf{d})\|_\infty$ can be computed recursively according to $\|\mathbf{z}_k(\mathbf{d}_k)\|_\infty = \max\{\|\mathbf{z}_{k-1}(\mathbf{d}_{k-1})\|_\infty, |[\mathbf{z}(\mathbf{d})]_{M-k+1}|\}$. Consequently, the PSC is replaced by the partial box constraint (PBC)

$$\|\mathbf{z}_k(\mathbf{d}_k)\|_\infty \leq C_\infty. \tag{9}$$

If the PBC is violated by a given $\mathbf{d}_k$, the node associated with that $\mathbf{d}_k$ along with all its children is pruned from the tree. The l^∞-optimal solution is obtained by choosing, among all surviving leaf nodes $\mathbf{d}=\mathbf{d}_M$, the one with minimum $\|\mathbf{z}(\mathbf{d})\|_\infty$, i.e.,

$$\widehat{\mathbf{d}}_\infty = \arg\min_{\mathbf{d}\in\mathcal{A}^M} \|\mathbf{z}(\mathbf{d})\|_\infty. \tag{10}$$

Slightly abusing terminology, we call the side length C_∞ of the search box the "*radius*" associated with SD-l^∞. Like in the SD-l^2 case with C_2, here the radius C_∞ has to be chosen large enough to ensure that at least one data vector is found by the algorithm. Again, however, choosing C_∞ too large will in general result in a high complexity of SD-l^∞ (for guidelines on how to choose C_∞ we refer to Section III-D).

We emphasize the following aspects of SD-l^∞:

- The SD-l^∞ hardware implementation reported in [8] is actually based on the $l^{\widetilde{\infty}}$-norm $\|\mathbf{x}\|_{\widetilde{\infty}} = \max\{|x_{\mathrm{R},1}|, |x_{\mathrm{I},1}|, \ldots, |x_{\mathrm{I},M}|\}$ rather than the l^∞-norm $\|\mathbf{x}\|_\infty = \max\{|x_1|,\ldots,|x_M|\}$, $\mathbf{x}\in\mathbb{C}^M$. Here, the essential aspect is that the computation of the $l^{\widetilde{\infty}}$-norm, as opposed to the l^∞- and the l^2-norm, does not require squaring operations, which, as already noted in Section I-A, results in significantly smaller circuit complexity. Nevertheless, in the following, for the sake of simplicity of exposition, we first analyze SD-l^∞, i.e., SD based on the conventional l^∞-norm, thereby revealing the fundamental aspects (w.r.t. performance and complexity) of SD using the $l^{\widetilde{\infty}}$-norm (referred to as SD-$l^{\widetilde{\infty}}$). The modifications of the results on SD-l^∞ needed to account for the use of the $l^{\widetilde{\infty}}$-norm are described in Section V.
- The tree search strategy underlying SD-l^∞ is identical to that of *Kannan's strategy* (see, e.g., [6], [20]), which also finds all data vectors inside a hypercube. The difference between SD-l^∞ and Kannan's algorithm lies in the calculation of the final detection result. SD-l^∞ implements (10) while Kannan's approach is optimum as it implements (3). Optimality of Kannan's algorithm is achieved through (i) guaranteeing that the solution of (3) is contained inside the search hypercube (which, in general, necessitates choosing the search radius to be larger than the corresponding radius for SD-l^∞ and hence incurs a higher complexity) and (ii) in the last step comparing all found data vectors with respect to their l^2-distance $\|\mathbf{r}-\mathbf{H}\mathbf{d}\|_2$ (which, in contrast to SD-$l^{\widetilde{\infty}}$, necessitates squaring operations).
- We note that the use of the l^∞-norm instead of the l^2-norm for tree traversal is not limited to the basic Fincke-Pohst variant of the SD algorithm as considered in this paper. For example, in [8] it was proposed to replace the l^2-norm by

the l^∞-norm for SD based on Schnorr-Euchner enumeration with radius updating [6], [21]. The resulting complexity savings turned out to be significant whereas the performance loss incurred was found to be small [8]. In fact, performing tree traversal based on the l^∞-norm instead of the l^2-norm can be directly applied to more sophisticated variants of SD as well (this, for example, includes SD using statistical pruning [15] or SD with advanced layer-sorting that not only depends on the channel matrix but also on the received vector [22]).

- The use of l^∞-norm *sphere-encoding* was proposed in [12] in the context of lattice precoding for peak-power reduction. The problem considered in [12] amounts to l^∞-norm decoding on the "full" channel matrix $\mathbf{H}$ according to

$$\widehat{\mathbf{d}}_{\infty,\text{full}} = \arg\min_{\mathbf{d}\in\mathcal{A}^M} \|\mathbf{r}-\mathbf{H}\mathbf{d}\|_\infty. \tag{11}$$

We emphasize that (11) is *not* equivalent to SD-l^∞ performed on the upper triangular matrix $\mathbf{R}$ as defined in (10) since the l^∞-norm (in contrast to the l^2-norm) is not invariant with respect to unitary transformations. In [12] the solution of (11) is obtained by using SD-l^2 with a larger sphere radius than that used for solving (3). In fact, the sphere radius has to be chosen such that the corresponding hypersphere circumscribes the minimum hypercube that contains the solution of (11) (see also [12, Fig. 1]). The reduction in peak-power therefore comes at the cost of an increase in algorithmic complexity (at the same circuit complexity) as compared to that required for solving the corresponding l^2-norm problem (3) (see [12, Fig. 5]). This approach is different from SD-l^∞, which is based on approximating (3) by (10) in order to reduce the circuit complexity (with a possible reduction in the algorithmic complexity) at the cost of a performance loss (see Section II).

II. Error Probability of SD-l^∞

In this section, we show that SD-l^∞ achieves the same diversity order as ML (i.e., SD-l^2) detection and we quantify the SNR loss incurred by SD-l^∞.

A. Distance Properties

We start by investigating distance properties of the SD-l^∞ solution $\widehat{\mathbf{d}}_\infty$. Using the bounds

$$\frac{1}{N}\|\mathbf{x}\|_2^2 \le \|\mathbf{x}\|_\infty^2 \le \|\mathbf{x}\|_2^2 \tag{12}$$

valid for any vector $\mathbf{x}\in\mathbb{C}^N$, we obtain

$$\begin{aligned}\left\|\mathbf{r}-\mathbf{H}\widehat{\mathbf{d}}_\infty\right\|_2^2 &= \left\|\mathbf{y}-\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\widehat{\mathbf{d}}_\infty\right\|_2^2 \le N\left\|\mathbf{y}-\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\widehat{\mathbf{d}}_\infty\right\|_\infty^2\\ &\le N\left\|\mathbf{y}-\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\widehat{\mathbf{d}}_{\text{ML}}\right\|_\infty^2\\ &\le N\left\|\mathbf{y}-\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\widehat{\mathbf{d}}_{\text{ML}}\right\|_2^2\\ &= N\left\|\mathbf{r}-\mathbf{H}\widehat{\mathbf{d}}_{\text{ML}}\right\|_2^2.\end{aligned} \tag{13}$$

A corresponding result for the l^∞-norm optimum solution with respect to the full channel matrix has been reported in [12, Corollary 1]. From (13) we can conclude that $\|\mathbf{r}-\mathbf{H}\widehat{\mathbf{d}}_\infty\|_2$ lies within a factor of $\sqrt{N}$ of the minimum distance $\|\mathbf{r}-\mathbf{H}\widehat{\mathbf{d}}_{\text{ML}}\|_2$ realized by the ML detector (3). Trivially, SD-l^∞ is optimum for $N=1$ (simply because the l^∞-norm equals the l^2-norm in this case). For increasing N, (13) suggests an increasing performance loss incurred by SD-l^∞ when compared to the ML detector (i.e., SD-l^2). In Section III-B, we quantify this performance loss in terms of diversity order and SNR gap. We note that for suboptimum detection based on successive interference cancellation (SIC) aided by LLL lattice-reduction (SIC-LLL) [6], [23], [24], we get a result that is structurally similar to (13) when $\widehat{\mathbf{d}}_\infty$ is replaced by the SIC-LLL detection result and the factor N is replaced by $2^{(N-1)}$. Consequently, the performance loss incurred by SIC-LLL can be expected to be larger than that incurred by SD-l^∞ (for numerical results see also Section VI-A).

B. Diversity Order and SNR Gap

We denote the error probability as a function of SNR ρ as $\mathrm{P}(\rho)$. In the following, we will only encounter error probabilities of the form $\mathrm{P}(\rho)=(K\rho)^{-\delta}+o(\rho^{-\delta})$, $\rho\to\infty$, with some constant $K>0$ not depending on ρ. We can define the corresponding SNR exponent δ as [25], [26]

$$\delta = -\lim_{\rho\to\infty}\frac{\log \mathrm{P}(\rho)}{\log\rho}. \tag{14}$$

Furthermore, if $\mathrm{P}_1(\rho)$ and $\mathrm{P}_2(\rho)$ have the same SNR exponent, we can define an asymptotic SNR gap α via $\mathrm{P}_1(\rho)\overset{a}{\sim}\mathrm{P}_2(\alpha\rho)$, $\rho\to\infty$. For example, if $\mathrm{P}_1(\rho)=(K_1\rho)^{-\delta}+o(\rho^{-\delta})$ and $\mathrm{P}_2(\rho)=(K_2\rho)^{-\delta}+o(\rho^{-\delta})$, we have $\alpha=K_1/K_2$. Our analysis corresponds to multiplexing gain $r=0$ in the framework of [26], i.e., we fix the size of the symbol alphabet and let the SNR go to infinity. In the following, we first focus on the behavior of the pairwise error probability (PEP) and then analyze the total error probability.

1) Pairwise Error Probability: Assume that the data vector $\mathbf{d}'$ was transmitted. The probability of erroneously deciding in favor of another data vector $\mathbf{d}\neq\mathbf{d}'$ is denoted as $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\text{ML}}(\rho)$ in the SD-l^2 case and $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho)$ in the SD-l^∞ case. To derive (an upper bound on) $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho)$, we first present a somewhat unconventional approach for upper-bounding $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\text{ML}}(\rho)$, which lends itself naturally to an extension to the l^∞-case. We start from $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\text{ML}}(\rho)=\mathrm{P}\left[\|\mathbf{H}\mathbf{b}+\mathbf{w}\|_2\le\|\mathbf{w}\|_2\right]$ with the error (difference) vector $\mathbf{b}=\mathbf{d}'-\mathbf{d}$. Applying the inverse triangle inequality according to $\|\mathbf{H}\mathbf{b}+\mathbf{w}\|_2\ge\left|\|\mathbf{H}\mathbf{b}\|_2-\|\mathbf{w}\|_2\right|$, we further obtain

$$\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\text{ML}}(\rho)\le\mathrm{P}\left[\|\mathbf{w}\|_2\ge\frac{1}{2}\|\mathbf{H}\mathbf{b}\|_2\right] \tag{15}$$

noting that $|x|\ge x$, for all $x\in\mathbb{R}$. With $\frac{\sqrt{2}}{\sigma}\|\mathbf{w}\|_2\sim\chi_{2N}$, conditioning on $\mathbf{H}$, and applying the Chernoff upper bound yields

$$\mathrm{P}\left[\|\mathbf{w}\|_2\ge\frac{1}{2}\|\mathbf{H}\mathbf{b}\|_2\,\middle|\,\mathbf{H}\right]\le\Phi_{\chi^2_{2N}}(s)e^{-s\rho\frac{\|\mathbf{H}\mathbf{b}\|_2^2}{2}}$$

for $s \in [0, 1/2)$. Here, $\Phi_{\chi^2_{2N}}(s)$ denotes the MGF of a χ^2_{2N}-distributed RV (see (2)). Averaging over $\mathbf{H}$ then results in

$$\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\mathrm{ML}}(\rho) \le \Phi_{\chi^2_{2N}}(s)\left(1+s\rho\frac{\|\mathbf{b}\|_2^2}{2M}\right)^{-N} \tag{16}$$

because $2M\|\mathbf{Hb}\|_2^2/\|\mathbf{b}\|_2^2 \sim \chi^2_{2N}$ for a given $\mathbf{b}$. The value of s minimizing the right-hand side (RHS) of (16) is in general a function of ρ, $\|\mathbf{b}\|_2$, and M. In the remainder of this paper, we set $s = 1/4$, which minimizes the RHS of (16) for high SNR and results in

$$\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\mathrm{ML}}(\rho) \le 2^N\left(1+\rho\frac{\|\mathbf{b}\|_2^2}{8M}\right)^{-N}. \tag{17}$$

Since N is the maximum diversity order that can be achieved over an $N \times M, N \ge M$, MIMO channel with the transmission setup considered in this paper (i.e., spatial multiplexing) [26], we can immediately conclude that the SNR exponent of $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\mathrm{ML}}(\rho)$ equals N for any nonzero $\mathbf{b}$ (see also the lower bound (21) on $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\mathrm{ML}}(\rho)$ having an SNR exponent of N as well).

For SD-l^∞ we can follow a similar approach. Starting with (10), we get

$$\begin{aligned}\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho) &\le \mathrm{P}\left[\|\mathbf{z}(\mathbf{d})\|_\infty \le \|\mathbf{z}(\mathbf{d}')\|_\infty\right]\\ &= \mathrm{P}\left[\left\|\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\mathbf{b}+\mathbf{n}\right\|_\infty \le \|\mathbf{n}\|_\infty\right].\end{aligned} \tag{18}$$

Note that for SD-l^∞, unlike for SD-l^2, the event $\|\mathbf{z}(\mathbf{d})\|_\infty = \|\mathbf{z}(\mathbf{d}')\|_\infty$ can, in general, occur with nonzero probability. Declaring an error in this case certainly yields an upper bound on $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho)$. Next, we apply the upper and lower bounds in (12), exploit the invariance of the l^2-norm to unitary transformations, and apply the inverse triangle inequality to get

$$\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho) \le \mathrm{P}\left[\|\mathbf{w}\|_2 \ge \frac{1}{\sqrt{N}+1}\|\mathbf{Hb}\|_2\right]. \tag{19}$$

Note the structural similarity of (19) and (15). Employing the same arguments as in the SD-l^2 case, we obtain

$$\begin{aligned}\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho) &\le 2^N\left(1+\rho\frac{\|\mathbf{b}\|_2^2}{2\left(\sqrt{N}+1\right)^2 M}\right)^{-N}\\ &= \mathrm{UB}_\infty(\rho).\end{aligned} \tag{20}$$

As in the SD-l^2 case for $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\mathrm{ML}}(\rho)$, we can immediately conclude that the SNR exponent of $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho)$ equals N for any nonzero $\mathbf{b}$.

Let us next quantify the SNR gap between $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho)$ and $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\mathrm{ML}}(\rho)$. We start by evaluating [27, eq. (20)] for the case at hand to get

$$\begin{aligned}\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\mathrm{ML}}(\rho) &\ge \frac{1}{2}\frac{1}{4^N}\binom{2N}{N}\left(1+\rho\frac{\|\mathbf{b}\|_2^2}{4M}\right)^{-N}\\ &= \mathrm{LB}_{\mathrm{ML}}(\rho).\end{aligned} \tag{21}$$

The asymptotic SNR gap between $\mathrm{UB}_\infty(\rho)$ and $\mathrm{LB}_{\mathrm{ML}}(\rho)$, denoted as β, i.e., $\mathrm{UB}_\infty(\rho) \overset{a}{\sim} \mathrm{LB}_{\mathrm{ML}}(\rho/\beta)$, $\rho \to \infty$, is directly obtained as

$$\beta = 4\left(\sqrt{N}+1\right)^2\left[\frac{1}{2}\binom{2N}{N}\right]^{-\frac{1}{N}}. \tag{22}$$

We can thus conclude that the asymptotic SNR gap between the PEP for SD-l^∞ and the PEP for SD-l^2 is upper-bounded by β, or, equivalently, we have

$$\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho) \preceq \mathrm{P}_{\mathbf{d}'\to\mathbf{d},\mathrm{ML}}(\rho/\beta), \qquad \rho\to\infty. \tag{23}$$

2) Total Error Probability: In the following, we consider the total error probability $\mathrm{P}_{\mathcal{E}}(\rho) = \mathrm{P}\left[\mathbf{d}' \ne \widehat{\mathbf{d}}\right]$ assuming equally likely transmitted data vectors $\mathbf{d}'$. If not specified, $\mathrm{P}_{\mathcal{E}}(\rho)$ stands either for the total error probability $\mathrm{P}_{\mathcal{E}_\infty}(\rho)$ of SD-l^∞ or for the total error probability $\mathrm{P}_{\mathcal{E}_{\mathrm{ML}}}(\rho)$ of SD-l^2. We start by noting that

$$\mathrm{P}_{\mathcal{E}}(\rho) = |\mathcal{A}|^{-M}\sum_{\mathbf{d}'}\mathrm{P}_{\mathcal{E}|\mathbf{d}'}(\rho). \tag{24}$$

Here, $\mathrm{P}_{\mathcal{E}|\mathbf{d}'}(\rho)$ refers to the total error probability conditioned on $\mathbf{d}'$ being transmitted, which can be bounded as

$$\mathrm{P}_{\mathbf{d}'\to\text{any }\mathbf{d}}(\rho) \le \mathrm{P}_{\mathcal{E}|\mathbf{d}'}(\rho) \le \sum_{\mathbf{d}\ne\mathbf{d}'}\mathrm{P}_{\mathbf{d}'\to\mathbf{d}}(\rho). \tag{25}$$

It follows that

$$\mathrm{P}_{\mathcal{E}}(\rho) \le |\mathcal{A}|^{-M}\sum_{\mathbf{d}'}\sum_{\mathbf{d}\ne\mathbf{d}'}\mathrm{P}_{\mathbf{d}'\to\mathbf{d}}(\rho) \tag{26}$$

and

$$\mathrm{P}_{\mathcal{E}}(\rho) \ge |\mathcal{A}|^{-M}\sum_{\mathbf{d}'}\mathrm{P}_{\mathbf{d}'\to\text{any }\mathbf{d}}(\rho). \tag{27}$$

As the SNR exponent of $\mathrm{P}_{\mathbf{d}'\to\mathbf{d},\infty}(\rho)$ equals N for all $\mathbf{b} = \mathbf{d}-\mathbf{d}' \ne \mathbf{0}$ [cf. (20)], we can conclude that SD-l^∞ achieves full diversity order N and hence the same diversity order as ML detection. The corresponding asymptotic SNR gap is obtained as follows. From (23)–(26), we get

$$\mathrm{P}_{\mathcal{E}_\infty}(\rho) \preceq |\mathcal{A}|^M\mathrm{P}_{\mathcal{E}_{\mathrm{ML}}}(\rho/\beta), \qquad \rho\to\infty. \tag{28}$$

From (26) together with (17) and (27) together with (21), we can conclude that $\mathrm{P}_{\mathcal{E}_{\mathrm{ML}}}(\rho)$ has SNR exponent N and can be written as $\mathrm{P}_{\mathcal{E}_{\mathrm{ML}}}(\rho) = (K_{\mathrm{ML}}\rho)^{-N} + o(\rho^{-N})$, $\rho\to\infty$, with some constant $K_{\mathrm{ML}} > 0$ that does not depend on ρ. With $\mathrm{P}_{\mathcal{E}_\infty}(\rho) \preceq |\mathcal{A}|^M\mathrm{P}_{\mathcal{E}_{\mathrm{ML}}}(\rho/\beta)$ from (28) and $N \ge M$, this yields $\mathrm{P}_{\mathcal{E}_\infty}(\rho) \preceq \mathrm{P}_{\mathcal{E}_{\mathrm{ML}}}(\rho/(|\mathcal{A}|\beta))$. Furthermore, using $\binom{m}{l} \ge \left(\frac{m}{l}\right)^l$, we have $\binom{2N}{N} \ge 2^N$, which, when employed in (22), shows that $\beta \le 4\left(\sqrt{N}+1\right)^2 \le 16N$. Thus, the asymptotic SNR gap between the total error probabilities $\mathrm{P}_{\mathcal{E}_{\mathrm{ML}}}(\rho)$ and $\mathrm{P}_{\mathcal{E}_\infty}(\rho)$ is upper-bounded by $16|\mathcal{A}|N$.

3) Discussion: From the results above we can conclude that SD-l^∞ achieves full diversity order N with an asymptotic SNR gap to SD-l^2 that is *bounded* (as a function of SNR) and scales at most *linearly* in the number of receive antennas. We emphasize that SD-l^∞ actually performs much better than the simple

upper bound $16|\mathcal{A}|N$ on the SNR gap suggests (see the numerical results in Fig. 2 of Section VI-A). We furthermore note that conventional suboptimum detection schemes like linear equalization-based or SIC detectors only achieve a diversity order of $N - M + 1$ [1], [26], [28]. It has been shown, however, that linear equalization-based and SIC detectors in combination with LLL lattice-reduction [6], [23], [24] achieve full diversity order [29] (see also [30]), i.e., the same diversity order as ML detection. For the infinite lattice case, the corresponding SNR gaps were quantified in [31] and shown to scale at most exponentially in the number of transmit antennas. Computationally more intensive lattice-reduction algorithms such as Korkine-Zolotarev lattice-reduction lead to SNR gaps that scale subexponentially or, when applied to the dual lattice, that scale even polynomially in the number of antennas [31] for the infinite lattice case but require to solve multiple shortest vector problems simultaneously. The SNR gap of LLL lattice-reduction aided detection relative to ML detection can be further reduced by performing multiple lattice-reduction steps as suggested in [33]. However, for the finite lattice case considered in this paper, the asymptotic SNR gap incurred by lattice-reduction aided detectors relative to SD-l^2 is unbounded [32] when so-called naive lattice-decoding is employed. This is in stark contrast to SD-l^∞ where the asymptotic SNR gap, as shown above, is bounded (as a function of SNR) for the finite lattice case.

For numerical results comparing the performance of SD-l^∞ to that of LLL-aided linear and SIC detectors, we refer to Fig. 2(b) in Section VI-A.

Next, we note that applying [34, Prop. 1] shows that the statement on SD-l^∞ achieving full diversity order for i.i.d. Rayleigh fading channels directly extends to more general fading statistics such as spatially correlated Rayleigh or Ricean fading. Furthermore, we remark that if SD is used for decoding general space-time codes, replacing the l^2-norm by the l^∞-norm for tree traversal achieves the same diversity order as the ML detector (in this case the diversity order of the ML detector is given by the product of N with the minimum rank of the corresponding codeword difference matrices, see, e.g., [1]). Finally, we note that l^∞-norm decoding on the full channel matrix $\mathbf{H}$ according to (11) can be shown to achieve full diversity order with an asymptotic SNR gap, vis-à-vis SD-l^2, that is bounded and increases at most linearly in the number of receive antennas.

III. Complexity of SD-l^∞

In this section, we analyze the complexity of SD-l^∞ by deriving an analytic expression for the average number of nodes visited in the tree search when pruning according to the PBC (9) is performed. A node $\mathbf{d}_k$ is visited if its corresponding PBC (9) is satisfied. We average w.r.t. channel, noise, and transmit signal and consider a fixed choice of C_∞ that is independent of the channel, noise, and transmit signal realizations and is not updated during the tree search. This setup was used in [10] to analyze the complexity of SD-l^2. We note, however, that in practice, SD is often implemented with (depth-first) Schnorr-Euchner enumeration [6], [8], [21], which includes updating of the sphere radius during the tree search. This setting, for example, is not captured by our complexity analysis. The assumption of a fixed sphere radius and no restarting—made for analytical tractability—results in an error floor and implies that the corresponding complexity expressions are a lower bound on the complexity of SD-l^∞ with restarting in the case that no leaf node is found in the search hypercube (see Section III-D for more details).

A. Basic Approach

For a given C_∞, a simple counting argument yields the number of nodes $S_{\infty,k}$ visited at tree level k, $k = 1, \ldots, M$, as

$$S_{\infty,k} = \sum_{\mathbf{d}_k} I(\mathbf{z}_k(\mathbf{d}_k)) \tag{29}$$

where

$$I(\mathbf{z}_k(\mathbf{d}_k)) = \begin{cases} 1, & \text{if } \|\mathbf{z}_k(\mathbf{d}_k)\|_\infty \leq C_\infty \\ 0, & \text{otherwise.} \end{cases}$$

We trivially have $S_{\infty,k} \leq |\mathcal{A}|^k$. First, we note that $\mathbb{E}\{I(\mathbf{z}_k(\mathbf{d}_k))\} = \mathrm{P}\left[\|\mathbf{z}_k(\mathbf{d}_k)\|_\infty \leq C_\infty\right]$, where the expectation is w.r.t. the channel $\mathbf{R}$, noise $\mathbf{n}$, and data vector $\mathbf{d}'$. Consequently, we have

$$\mathbb{E}\{S_{\infty,k}\} = \sum_{\mathbf{d}_k} \mathrm{P}\left[\|\mathbf{z}_k(\mathbf{d}_k)\|_\infty \leq C_\infty\right] \tag{30}$$

with the total complexity

$$\mathbb{E}\{S_\infty\} = \sum_{k=1}^{M} \mathbb{E}\{S_{\infty,k}\}. \tag{31}$$

Next, we condition on the data subvector $\mathbf{d}'_k \in \mathcal{A}^k$ and write $\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{d}_k)\|_\infty \leq C_\infty | \mathbf{d}'_k\right] = \mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \leq C_\infty\right]$ with

$$\mathbf{z}_k(\mathbf{b}_k) = \begin{bmatrix} \mathbf{R}_k \\ \mathbf{0} \end{bmatrix} \mathbf{b}_k + \begin{bmatrix} \mathbf{n}_k \\ \mathbf{n}_L \end{bmatrix} \tag{32}$$

where $\mathbf{b}_k = \mathbf{d}'_k - \mathbf{d}_k$ is a pairwise error subvector, $\mathbf{n}_k = (n_{M-k+1} \cdots n_M)^T$, and $\mathbf{n}_L = (n_{M+1} \cdots n_N)^T$. We set $\mathbf{z}(\mathbf{b}) = \mathbf{z}_M(\mathbf{b}_M)$ and note that $\mathbf{z}_k(\mathbf{b}_k) = ([\mathbf{z}(\mathbf{b})]_{M-k+1} \cdots [\mathbf{z}(\mathbf{b})]_N)^T$. Formally, for a given $\mathbf{d}'_k$, we will often speak of "a node" $\mathbf{b}_k$, which, in a one-to-one fashion, refers to the node $\mathbf{d}_k = \mathbf{d}'_k - \mathbf{b}_k$ in the search tree. For example, the node $\mathbf{d}_k = \mathbf{d}'_k$ corresponding to the transmitted data subvector $\mathbf{d}'_k$ is equivalent to the node $\mathbf{b}_k = \mathbf{0}$. If we speak of a node $\mathbf{b}_k$ without specifying $\mathbf{d}'_k$, the corresponding statements hold for all pairs $\mathbf{d}'_k, \mathbf{d}_k \in \mathcal{A}^k$ satisfying $\mathbf{b}_k = \mathbf{d}'_k - \mathbf{d}_k$. It follows that (30) can be written as

$$\mathbb{E}\{S_{\infty,k}\} = \frac{1}{|\mathcal{A}|^k} \sum_{\mathbf{b}_k} \mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \leq C_\infty\right] \tag{33}$$

where we assumed equally likely transmitted data subvectors $\mathbf{d}'_k$ for all tree levels $k = 1, \ldots, M$; this holds, e.g., for statistically independent (across the transmit antennas) and equally likely data symbols. The sum (33) is taken over all possible combinations of pairwise error subvectors $\mathbf{b}_k$.

Equivalently, the complexity at the kth tree level $\mathbb{E}\{S_{2,k}\}$ for SD-l^2 is given by (33) with the l^∞-norm replaced by the l^2-norm and C_∞ replaced by C_2, i.e.

$$\mathbb{E}\{S_{2,k}\} = \frac{1}{|\mathcal{A}|^k} \sum_{\mathbf{b}_k} \mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_2 \leq C_2\right]. \tag{34}$$

We finally note that $\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \leq C_\infty\right]$ in (33) and $\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_2 \leq C_2\right]$ in (34) express the probability that node $\mathbf{b}_k$ is visited by SD-l^∞ and SD-l^2.

B. Computation of the Partial Metric Cdfs

From (33) and (34) we can see that the computation of $\mathbb{E}\{S_{\infty,k}\}$ and $\mathbb{E}\{S_{2,k}\}$ requires knowledge of the cdfs of the partial metrics $\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty$ and $\|\mathbf{z}_k(\mathbf{b}_k)\|_2$, respectively. Hence, the key difference in deriving an analytic expression for the complexity of SD-l^∞ to that for SD-l^2 as carried out in (10) lies in the computation of these partial metric cdfs. An analytic expression for $\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_2 \le C_2\right]$ was provided in [10], [13]. More specifically, it is shown in [10, Lemma 1] that thanks to the invariance of the l^2-norm w.r.t. unitary transformations

$$\|\mathbf{z}_k(\mathbf{b}_k)\|_2 \stackrel{d}{=} \left\|\mathbf{H}_k\mathbf{b}_k + \begin{bmatrix}\mathbf{n}_k\\ \mathbf{n}_L\end{bmatrix}\right\|_2$$

where the $(k+L)\times k$ matrix $\mathbf{H}_k$ with $L = N - M$ has i.i.d. $\mathcal{CN}(0,1/M)$ entries. Conditioned on $\mathbf{b}_k$, the RV $\left\|\mathbf{H}_k\mathbf{b}_k + \begin{bmatrix}\mathbf{n}_k\\ \mathbf{n}_L\end{bmatrix}\right\|_2$ is then easily found to be χ_{k+L}-distributed, which leads to [10], [13]

$$\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_2 \le C_2\right] = \gamma_{k+L}\left(\frac{C_2^2}{\|\mathbf{b}_k\|_2^2/M+\sigma^2}\right). \quad (35)$$

As the l^∞-norm is not invariant w.r.t. unitary transformations, this approach does not carry over to the l^∞-case considered here. Instead, we follow a direct approach as detailed below.

1) Cdf of $\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty$: Since the nonzero entries in $\mathbf{R}$ are statistically independent [35, Lemma 2.1], the elements of $\mathbf{z}_k(\mathbf{b}_k)$ conditioned on $\mathbf{b}_k$ are statistically independent as well. We, thus, obtain

$$\begin{aligned}\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \le C_\infty\right] &= \left[\gamma_1\left(\frac{C_\infty^2}{\sigma^2}\right)\right]^L \prod_{m=1}^{k}\mathrm{P}\left[\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right| \le C_\infty\right] \quad (36)\end{aligned}$$

where we used the fact that the bottom L elements of $\mathbf{z}_k(\mathbf{b}_k)$ are given by the i.i.d. $\mathcal{CN}(0,\sigma^2)$ vector $\mathbf{n}_L$ (see (32)) so that $\mathrm{P}\left[\left|[\mathbf{z}_k(\mathbf{b}_k)]_i\right| \le C_\infty\right] = \gamma_1\left(C_\infty^2/\sigma^2\right), i = k+1,\ldots,k+L$.

2) Cdf of $\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right|$: An analytic expression for $\mathrm{P}\left[\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right| \le C_\infty\right]$ can be obtained via direct integration using the fact that the nonzero entries of $\mathbf{R}$ are statistically independent with $\sqrt{2M}R_{i,i} \sim \chi_{2(N-i+1)}$ and $R_{i,j} \sim \mathcal{CN}(0,1/M)$, for $i = 1,\ldots,M$, $j > i$ [35, Lemma 2.1]. In Appendix A it is shown that $\mathrm{P}\left[\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right| \le C_\infty\right]$ is a binomial mixture of χ-distributions with degrees of freedom reaching from 2 up to $2(m+L)$. More specifically, for $m = 1,\ldots,M$, we have

$$\begin{aligned}\mathrm{P}\left[\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right| \le C_\infty\right] &= \sum_{l=0}^{m+L-1} B_l(\mathbf{b}_m)\gamma_{m+L-l}\left(\frac{C_\infty^2}{\|\mathbf{b}_m\|_2^2/M+\sigma^2}\right) \quad (37)\end{aligned}$$

with the coefficients

$$B_l(\mathbf{b}_m) = \binom{m+L-1}{l}(p(\mathbf{b}_m))^l(1-p(\mathbf{b}_m))^{m+L-1-l} \quad (38)$$

with parameter

$$p(\mathbf{b}_m) = \frac{\|\mathbf{b}_{m-1}\|_2^2 + M\sigma^2}{\|\mathbf{b}_m\|_2^2 + M\sigma^2} \quad (39)$$

and $\|\mathbf{b}_0\|_2^2 = 0$. In [15] the pdf of the RV $\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right|^2$ associated with the distribution (37) was obtained in a different form (i.e., not in terms of a binomial mixture of χ-distributions) using an alternative derivation. More specifically, the derivation in [15] exploits the property $\|\mathbf{z}_m(\mathbf{b}_m)\|_2^2 = \|\mathbf{z}_{m-1}(\mathbf{b}_{m-1})\|_2^2 + \left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right|^2$ with $\|\mathbf{z}_{m-1}(\mathbf{b}_{m-1})\|_2^2$ and $\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right|^2$ being statistically independent and the MGFs of $\|\mathbf{z}_m(\mathbf{b}_m)\|_2^2$ and $\|\mathbf{z}_{m-1}(\mathbf{b}_{m-1})\|_2^2$ being known from [10]. This allows to compute the MGF of $\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right|^2$ and, via the inverse Fourier transform, the corresponding pdf, which can then be used to establish (37). Finally, we note that the direct integration approach used in this paper to obtain (37) can, in contrast to the approach employed in [15], be applied to derive the distributions of $\left|[\mathbf{z}(\mathbf{b})]_{\mathrm{R},M-m+1}\right|$ and $\left|[\mathbf{z}(\mathbf{b})]_{\mathrm{I},M-m+1}\right|$, which are needed to compute (bounds on) the complexity of SD-$\widetilde{l^\infty}$ (see Section V-B for more details).

3) Sum Representation and Moment Generating Function: The binomial mixture representation (37) allows for an interesting alternative representation of the RV $\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right|^2$ as the sum of independent RVs. In particular, using results from [36], it is shown in Appendix B that

$$\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right|^2 \stackrel{d}{=} t_m^2 \quad (40)$$

where

$$t_m^2 = \frac{\|\mathbf{b}_m\|_2^2/M+\sigma^2}{2}\left(\gamma^2 + \sum_{i=1}^{m+L-1}\lambda_i^2\right) \quad (41)$$

with the independent RVs γ^2 and λ_i^2, $i = 1,\ldots,m+L-1$. Here, $\gamma^2 \sim \chi_2^2$ with pdf $f_{\chi_2^2}(x)$ and the λ_i^2 have the mixture pdf

$$f_{\lambda_i^2}(x) = (1-p(\mathbf{b}_m))f_{\chi_2^2}(x) + p(\mathbf{b}_m)\delta(x) \quad (42)$$

or, equivalently, with probability $p(\mathbf{b}_m)$ the λ_i^2 come from a population having pdf $\delta(x)$ (i.e., they are equal to zero with probability $p(\mathbf{b}_m)$) and with probability $1-p(\mathbf{b}_m)$ they come from a population having a χ_2^2 distribution.

Besides being interesting in its own right, the representation (41) allows to compute the MGF of $\left|[\mathbf{z}(\mathbf{b})]_{M-m+1}\right|^2$ or, equivalently, of t_m^2 in a straightforward manner, by using (39) and (42), as [cf. (2)]

$$\begin{aligned}\Phi_{t_m^2}(s) &= \mathbb{E}\left\{e^{st_k^2}\right\} \\ &= \frac{[1-(\|\mathbf{b}_{m-1}\|_2^2/M+\sigma^2)s]^{m+L-1}}{[1-(\|\mathbf{b}_m\|_2^2/M+\sigma^2)s]^{m+L}}. \quad (43)\end{aligned}$$

We note that this result can be directly used to recover the cdf of $\|\mathbf{z}_k(\mathbf{b}_k)\|_2$ in (35) without explicitly using (as in [10]) the rotational invariance of the l^2-norm.

C. Final Complexity Expressions

We are now ready to assemble our results to get the final complexity expressions for SD-l^∞ and SD-l^2. Inserting (37) into (36) and using (33), we obtain

$$\begin{aligned}\mathbb{E}\{S_{\infty,k}\} = \frac{1}{|\mathcal{A}|^k}\left[\gamma_1\left(\frac{C_\infty^2}{\sigma^2}\right)\right]^L \sum_{\mathbf{b}_k}\prod_{m=1}^{k}\sum_{l=0}^{m+L-1} B_l(\mathbf{b}_m) \\ \cdot\gamma_{m+L-l}\left(\frac{C_\infty^2}{\|\mathbf{b}_m\|_2^2/M+\sigma^2}\right). \quad (44)\end{aligned}$$

The corresponding total complexity follows from (31). In comparison, for SD-l^2, using (34) and (35) yields [10], [13]

$$\mathbb{E}\{S_{2,k}\} = \frac{1}{|\mathcal{A}|^k} \sum_{\mathbf{b}_k} \gamma_{k+L}\left(\frac{C_2^2}{||\mathbf{b}_k||_2^2/M + \sigma^2}\right). \tag{45}$$

The total complexity for SD-l^2 is then obtained as

$$\mathbb{E}\{S_2\} = \sum_{k=1}^{M} \mathbb{E}\{S_{2,k}\}. \tag{46}$$

D. Choice of Radii

For a meaningful comparison of the complexity of SD-l^∞ and SD-l^2, the radii C_∞ and C_2 have to be chosen carefully. In our analysis below, we use the approach proposed in [5], [10], [19] for SD-l^2, where the choice of C_2 is based on the noise statistics such that the probability of finding the transmitted data vector inside the search hypersphere is sufficiently high. Recall that our complexity analysis assumes a fixed choice of the radii that does not depend on the channel, noise, and data realizations. We start by noting that $||\mathbf{z}(\mathbf{d}')||_2 = ||\mathbf{n}||_2$, which is χ_{2N}-distributed. Choosing the radius C_2 such that the transmitted data vector $\mathbf{d}'$ is found inside the search hypersphere with probability $1-\epsilon$, $\epsilon \in [0,1]$, is accomplished by setting

$$\mathrm{P}\left[||\mathbf{n}||_2 \leq C_2\right] = \gamma_N\left(\frac{C_2^2}{\sigma^2}\right) = 1-\epsilon. \tag{47}$$

Solving (47) for C_2^2 yields

$$C_2^2 = \sigma^2 \gamma_N^{-1}(1-\epsilon). \tag{48}$$

For the SD-l^∞ case, we adopt an analogous approach by choosing C_∞ such that $\mathbf{d}'$ is contained in the search hypercube with sufficiently high probability. Specifically, for the complexity comparisons in the remainder of the paper, we choose the radius C_∞ such that the probability of finding the transmitted data vector $\mathbf{d}'$ through SD-l^∞ equals that for SD-l^2. This is accomplished by setting

$$\mathrm{P}\left[||\mathbf{n}||_\infty \leq C_\infty\right] = \left[\gamma_1\left(\frac{C_\infty^2}{\sigma^2}\right)\right]^N = 1-\epsilon \tag{49}$$

which results in

$$C_\infty^2 = -\sigma^2 \log\left(1 - \sqrt[N]{1-\epsilon}\right). \tag{50}$$

For both SD-l^2 and SD-l^∞, for any $\epsilon > 0$, there is a nonzero probability that no leaf node is found by the detector, i.e., $||\mathbf{z}(\mathbf{d})||_2 > C_2$ or $||\mathbf{z}(\mathbf{d})||_\infty > C_\infty$, respectively, for all $\mathbf{d} \in \mathcal{A}^M$. If C_2^2 and C_∞^2 are given by (48) and (50), respectively, with some $\epsilon > 0$, stopping the detection procedure and declaring an error in this case can easily be shown to entail an error floor of ϵ and hence does not implement exact SD-l^2 (i.e., ML) or exact SD-l^∞ detection, respectively. However, if the system operates at a target error rate that is much higher than this error floor, a fixed radius and the absence of restarting will have negligible impact on the total error probability. To obtain exact ML or SD-l^∞ performance, the corresponding SD algorithm has to be restarted using a schedule of increasing radii (or equivalently a schedule of decreasing values for ϵ) until a leaf node is found within the search hypersphere or hypercube, respectively (see, e.g., [10], [37] for SD-l^2). In this case, the complexity expressions provided in Section III-C represent lower bounds on the complexity.

E. Asymptotic Complexity Analysis

In [16] it is shown that the complexity of SD-l^2 scales exponentially in the number of transmit antennas M. Motivated by this result, we will next show that the complexity scaling behavior of SD-l^∞ is also exponential in M. For simplicity of exposition, we set $M = N$ in the following.

1) Impact of Choice of Radius: The asymptotic complexity scaling behavior of SD-l^2 is studied in detail in [16], where it is shown that $\mathbb{E}\{S_2\} \geq e^{\gamma M}$ for large M and some $\gamma > 0$. This result is derived under the assumption of C_2^2 being fixed and increasing (at least) *linearly* in M, which guarantees a nonvanishing probability of finding at least one leaf node inside the search hypersphere [38, Theorem 1]. It is furthermore shown in [38] that the exponential complexity scaling behavior of SD-l^2 extends to the case where the sphere radius is chosen optimally, i.e., the radius is set to the minimum value still guaranteeing that at least one leaf node is found (this would, of course, correspond to a genie-aided choice of the sphere radius since it essentially necessitates knowledge of the ML detection result). It can be shown that C_2^2 chosen according to (48) results in linear scaling in M for large M. Linear scaling of C_2^2 in M is also obtained, for example, by setting $C_2^2 \propto \mathbb{E}\{||\mathbf{n}||_2^2\} = \sigma^2 M$ as was done in [10], [16].

For SD-l^∞ the radius C_∞^2 according to (50) scales logarithmically in M for large M.[1] Based on [39, Eq. (2.5.5)], it can be shown that this is also the case if C_∞^2 is chosen to be proportional to $\mathbb{E}\{||\mathbf{n}||_\infty^2\}$. At first sight, the logarithmic scaling of C_∞^2 in M versus the linear scaling of C_2^2 suggests a difference in the asymptotic complexity behavior of SD-l^∞ and SD-l^2. While the complexity and pruning (see Section IV) behavior for finite M are indeed quite different in general, we will, however, next show that SD-l^∞ also exhibits exponential complexity scaling in M.

2) Lower Bound on Complexity: Computing the asymptotics of the exact SD-l^∞ complexity expression ((44) together with (31)) seems involved. We therefore tackle the problem by computing a lower bound on complexity and by showing that this lower bound scales exponentially in the problem size M. Our technique can readily be extended to the SD-l^2 case resulting in an alternative, w.r.t. [16], proof of the exponential complexity scaling behavior of SD-l^2. We note, however, that while our proof seems to be shorter and more direct, the result in [16] is more general in the sense that it applies to MIMO channels with very general fading statistics. Our approach, in contrast, explicitly hinges on the channel matrix $\mathbf{H}$ being i.i.d. Rayleigh fading. On a conceptual basis, our proof is more closely related to the

[1]For fixed SNR (i.e., fixed σ^2) and fixed ϵ, the asymptotic ($N \to \infty$) behavior of $C_\infty^2 = -\sigma^2 \log\left(1 - \sqrt[N]{1-\epsilon}\right)$ is obtained as follows. We have $\sqrt[N]{1-\epsilon} = 1 + \mathcal{O}(1)/N, N \to \infty$. Thus, $C_\infty^2 = \sigma^2 \log(N) + \mathcal{O}(1), N \to \infty$, which shows that $C_\infty^2 \overset{a}{\sim} \sigma^2 \log(N), N \to \infty$.

approach in [15], where bounds on the complexity of SD-l^2 (and variants thereof) are studied.

We start by focusing on the expression for $\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \leq C_\infty\right]$, $k = 1,\ldots,M$, obtained by inserting (37) into the RHS of (36). Considering only the summand with index $l = m-1$ in (37), we obtain (recall that $L = N - M = 0$)

$$\mathrm{P}[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \leq C_\infty] \geq \prod_{m=1}^{k} \left(\frac{\|\mathbf{b}_{m-1}\|_2^2/M+\sigma^2}{\|\mathbf{b}_m\|_2^2/M+\sigma^2}\right)^{m-1} \cdot\gamma_1\left(\frac{C_\infty^2}{\|\mathbf{b}_m\|_2^2/M+\sigma^2}\right). \quad (51)$$

Using (113) in Appendix C according to

$$\gamma_1\left(\frac{C_\infty^2}{\|\mathbf{b}_m\|_2^2/M+\sigma^2}\right) \geq \gamma_1\left(\frac{C_\infty^2}{\sigma^2}\right)\frac{\sigma^2}{\|\mathbf{b}_m\|_2^2/M+\sigma^2}$$

we get

$$\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \leq C_\infty\right] \geq \left[\gamma_1\left(\frac{C_\infty^2}{\sigma^2}\right)\right]^k \left(1+\frac{\|\mathbf{b}_k\|_2^2}{M\sigma^2}\right)^{-k}. \quad (52)$$

Furthermore, it is easy to show that $\left[\gamma_1\left(C_\infty^2/\sigma^2\right)\right]^k \geq 1-\epsilon$ for the specific choice of C_∞^2 according to (50). Hence, (52) becomes

$$\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \leq C_\infty\right] \geq (1-\epsilon)\left(1+\frac{\|\mathbf{b}_k\|_2^2}{M\sigma^2}\right)^{-k}. \quad (53)$$

With (33) and (31) we then obtain

$$\mathbb{E}\{S_\infty\} \geq (1-\epsilon)\sum_{k=1}^{M}\frac{1}{|\mathcal{A}|^k}\sum_{\mathbf{b}_k}\left(1+\frac{\|\mathbf{b}_k\|_2^2}{M\sigma^2}\right)^{-k} \quad (54)$$

which can be further simplified using $\|\mathbf{b}_k\|_2^2 \leq B^2\xi(\mathbf{b}_k)$, where $B^2 = \max_{d,d'\in\mathcal{A}}|d'-d|^2$ is the maximum Euclidean distance in the scalar symbol constellation and $\xi(\mathbf{b}_k) = \xi(\mathbf{d}_k', \mathbf{d}_k)$ denotes the Hamming distance between $\mathbf{d}_k$ and $\mathbf{d}_k'$, i.e., the number of nonzero entries (symbol errors) in $\mathbf{b}_k = \mathbf{d}_k' - \mathbf{d}_k$. Note that every data vector $\mathbf{d}_k'$ induces the same set of Hamming distances $\{\xi(\mathbf{d}_k', \mathbf{d}_k), \mathbf{d}_k \in \mathcal{A}^k\}$. From (54), we therefore get

$$\mathbb{E}\{S_\infty\} \geq (1-\epsilon)\sum_{k=1}^{M}\sum_{i=0}^{k} W_i\left(1+\frac{B^2 i}{M\sigma^2}\right)^{-k}$$

where all terms having the same Hamming distance $\xi(\mathbf{d}_k', \mathbf{d}_k) = i$ have been merged and $W_i = \binom{k}{i}(|\mathcal{A}|-1)^i \geq \binom{k}{i}$ denotes the number of data vectors $\mathbf{d}_k \in \mathcal{A}^k$ that have Hamming distance i from $\mathbf{d}_k'$. Furthermore, with $\binom{k}{i} \geq \left(\frac{k}{i}\right)^i$, $i = 1,\ldots,k$, we get $W_i \geq \left(\frac{k}{i}\right)^i$, so that

$$\mathbb{E}\{S_\infty\} \geq (1-\epsilon)\sum_{k=1}^{M}\sum_{i=1}^{k}\left(\frac{k}{i}\right)^i\left(1+\frac{B^2 i}{M\sigma^2}\right)^{-k} \quad (55)$$

where the $i = 0$ ($W_0 = 1$) term is omitted for all k. We note that the RHS of (55) can be shown to also be a lower bound on $\mathbb{E}\{S_2\}$.

3) Asymptotic Analysis of Lower Bound: In the following, we show that the lower bound (55) exhibits exponential scaling in the system size $M = N$, which, together with the trivial upper bound $\mathbb{E}\{S_\infty\} \leq |\mathcal{A}|^{M+1}$, establishes exponential complexity scaling of SD-l^∞ (and of SD-l^2 together with $\mathbb{E}\{S_2\} \leq |\mathcal{A}|^{M+1}$).

We start by noting that a trivial lower bound on $\mathbb{E}\{S_\infty\}$ is obtained by considering only one term in the RHS of (55), resulting in

$$\mathbb{E}\{S_\infty\} \geq (1-\epsilon)\left(\frac{k}{i}\right)^i\left(1+\frac{B^2 i}{M\sigma^2}\right)^{-k} = f(M). \quad (56)$$

Evidently, establishing that

$$\lim_{M\to\infty}\frac{\log f(M)}{M} > 0 \quad (57)$$

is sufficient to prove that SD-l^∞ (and SD-l^2) exhibits exponential complexity scaling. To this end, we set $k = \lceil \alpha M\rceil$ and $i = \lceil \beta M\rceil$ with $\alpha \in]0,1]$ and $\beta \in]0,\alpha]$. We then have

$$\frac{\log f(M)}{M} = \frac{\log(1-\epsilon)}{M} + \frac{\lceil\beta M\rceil}{M}\log\left(\frac{\lceil\alpha M\rceil}{\lceil\beta M\rceil}\right) - \frac{\lceil\alpha M\rceil}{M}\log\left(1+\frac{B^2\lceil\beta M\rceil}{M\sigma^2}\right).$$

Furthermore, writing $\lceil\alpha M\rceil = \alpha M + \Delta_\alpha$ and $\lceil\beta M\rceil = \beta M + \Delta_\beta$ for some values Δ_α and Δ_β satisfying $0 \leq \Delta_\alpha, \Delta_\beta < 1$ gives

$$\frac{\log f(M)}{M} = \frac{\log(1-\epsilon)}{M} + (\beta+\Delta_\beta/M)\log\left(\frac{\alpha+\Delta_\alpha/M}{\beta+\Delta_\beta/M}\right) - (\alpha+\Delta_\alpha/M)\log\left(1+\frac{B^2\beta}{\sigma^2}+\frac{B^2}{\sigma^2}\Delta_\beta/M\right)$$

which results in

$$\lim_{M\to\infty}\frac{\log f(M)}{M} = \beta\log(\alpha/\beta) - \alpha\log(1+B^2\beta/\sigma^2) = \gamma(\alpha,\beta).$$

Indeed, for any SNR (i.e., for any σ^2), there exist values of α and β for which $\gamma(\alpha,\beta) > 0$. For example, with $\beta = \alpha/2$, any α satisfying

$$0 < \alpha < \min\left\{\frac{1}{B^2}2\sigma^2(\sqrt{2}-1), 1\right\} \quad (58)$$

results in $\gamma(\alpha,\beta) > 0$, which establishes the desired result.

IV. Tree Pruning Behavior

In the previous section, we showed that both SD-l^∞ and SD-l^2 exhibit exponential complexity scaling in M. The analytic results for $\mathbb{E}\{S_\infty\}$ and $\mathbb{E}\{S_2\}$ in Section III-C indicate, however, that the finite-M complexity can be very different for SD-l^∞ and SD-l^2. While it seems difficult, based on the analytic expressions for $\mathbb{E}\{S_\infty\}$ and $\mathbb{E}\{S_2\}$, to draw general conclusions on the finite-M complexity, we can obtain interesting insights on the differences in the tree pruning behavior (TPB). In the following, we study the TPB of SD-l^∞ and SD-l^2 through a high-SNR analysis of the probabilities $\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_\infty \leq C_\infty\right]$ [see (36) with (37)] and $\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_2 \leq C_2\right]$ (35) of a certain node $\mathbf{b}_k$ being visited by SD-l^∞ and SD-l^2, respectively. While (35) shows that the probability of a node $\mathbf{b}_k$ being visited by

SD-l^2 depends only on $||\mathbf{b}_k||_2$, i.e., on the Euclidean distance between $\mathbf{d}_k$ and $\mathbf{d}'_k$, in the SD-l^∞ case this dependence on $\mathbf{b}_k$ seems in general rather involved. However, the high-SNR analysis of $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_\infty \leq C_\infty\right]$ reveals simple characteristics of $\mathbf{b}_k$ that determine the probability of node $\mathbf{b}_k$ being visited by SD-l^∞, which then enables us to characterize the fundamental differences in the TPB of SD-l^∞ and SD-l^2. The corresponding results will be supported by simple geometric insights. Throughout this section, the radii C_∞ and C_2 are chosen according to (50) and (48), respectively, and we define

$$\kappa_\infty = \frac{C_\infty^2}{\sigma^2} = -\log\left(1 - \sqrt[N]{1-\epsilon}\right)$$
$$\kappa_2 = \frac{C_2^2}{\sigma^2} = \gamma_N^{-1}(1-\epsilon). \tag{59}$$

A. High-SNR Analysis

Consider a node[2] $\mathbf{b}_k \neq \mathbf{0}$ at tree level k and denote the index of the corresponding first tree level exhibiting a symbol error by $\widehat{m}(\mathbf{b}_k)$ with $\widehat{m}(\mathbf{b}_k) \in [1,\ldots,k]$. More precisely, we have $[\mathbf{b}_k]_{k-i+1} = 0$, for $i = 1,\ldots,\widehat{m}(\mathbf{b}_k)-1$ and $[\mathbf{b}_k]_{k-\widehat{m}(\mathbf{b}_k)+1} \neq 0$. In Appendix D, it is shown that

$$\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_\infty \leq C_\infty\right] \overset{a}{\sim} A\left(\widehat{m}(\mathbf{b}_k)\right)\kappa_\infty^{k+L}\left(\rho||\mathbf{b}_k||_2^2/M\right)^{-(k+L)} \tag{60}$$

$\rho \to \infty$, where

$$A\left(\widehat{m}(\mathbf{b}_k)\right) = \left[\gamma_1(\kappa_\infty)\right]^{\widehat{m}(\mathbf{b}_k)+L-1} \cdot \sum_{l=0}^{\widehat{m}(\mathbf{b}_k)+L-1}\binom{\widehat{m}(\mathbf{b}_k)+L-1}{l} \cdot \frac{1}{(\widehat{m}(\mathbf{b}_k)+L-l)!}\kappa_\infty^{-l}. \tag{61}$$

Note that $A\left(\widehat{m}(\mathbf{b}_k)\right)$ does not depend on SNR ρ. Furthermore, using (104) in (35), we directly obtain a corresponding result for SD-l^2 as

$$\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_2 \leq C_2\right] \overset{a}{\sim} \frac{1}{(k+L)!}\kappa_2^{k+L}\left(\rho||\mathbf{b}_k||_2^2/M\right)^{-(k+L)} \tag{62}$$

$\rho \to \infty$. From (60) and (62) we can infer that the only characteristics of $\mathbf{b}_k$, which determine the high-SNR probability of node $\mathbf{b}_k$ being visited, are $||\mathbf{b}_k||_2^2$ in the case of SD-l^2 and $||\mathbf{b}_k||_2^2$ and $\widehat{m}(\mathbf{b}_k)$ in the case of SD-l^∞. Moreover, for SD-l^∞ the dependence on $\widehat{m}(\mathbf{b}_k)$ is through the function $A\left(\widehat{m}(\mathbf{b}_k)\right)$ (61), which has the following properties. By inspection we get

$$\lim_{\kappa_\infty \to \infty} A\left(\widehat{m}(\mathbf{b}_k)\right) = \frac{1}{(\widehat{m}(\mathbf{b}_k)+L)!} \tag{63}$$

and it is easy to show that

$$\lim_{\kappa_\infty \to 0} A\left(\widehat{m}(\mathbf{b}_k)\right) = 1. \tag{64}$$

Note that $\kappa_\infty \to \infty$ for $\epsilon \to 0$ and $\kappa_\infty \to 0$ for $\epsilon \to 1$. In Appendix E, it is furthermore shown that $A\left(\widehat{m}(\mathbf{b}_k)\right)$ is a nonincreasing function of κ_∞, which, together with (63) and (64), yields

$$\frac{1}{(\widehat{m}(\mathbf{b}_k)+L)!} \leq A\left(\widehat{m}(\mathbf{b}_k)\right) \leq 1. \tag{65}$$

The lower bound in (65) allows us to conclude that the high-SNR probability of SD-l^∞ visiting node $\mathbf{b}_k$ decreases at most as $1/((\widehat{m}(\mathbf{b}_k)+L)!)$ for increasing $\widehat{m}(\mathbf{b}_k)$. This suggests that nodes corresponding to a first symbol error at high tree levels, i.e, nodes with large $\widehat{m}(\mathbf{b}_k)$, are in general pruned with higher probability than nodes corresponding to a first symbol error at low tree levels, i.e., nodes with small $\widehat{m}(\mathbf{b}_k)$ (provided, of course, that $||\mathbf{b}_k||_2$ is constant in this comparison).

B. TPB Comparison

Let us next compare the high-SNR TPB of SD-l^∞ to that of SD-l^2. We start by defining

$$\rho_C = \frac{C_2^2}{C_\infty^2}. \tag{66}$$

For $\mathbf{b}_k \neq \mathbf{0}$, the results in (60) and (62) imply

$$\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_\infty \leq C_\infty\right] \preceq \mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_2 \leq C_2\right], \qquad \rho \to \infty \tag{67}$$

if

$$A(\widehat{m}(\mathbf{b}_k)) \leq \frac{1}{(k+L)!}\rho_C^{k+L} \tag{68}$$

and

$$\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_\infty \leq C_\infty\right] \succ \mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_2 \leq C_2\right], \qquad \rho \to \infty \tag{69}$$

if

$$A(\widehat{m}(\mathbf{b}_k)) > \frac{1}{(k+L)!}\rho_C^{k+L}. \tag{70}$$

Hence, the high-SNR average pruning probability of a node $\mathbf{b}_k \neq \mathbf{0}$ for SD-l^∞ as compared to SD-l^2 is entirely described by the two functions $A(\widehat{m}(\mathbf{b}_k))$ and $1/(k+L)!\rho_C^{k+L}$, $k = 1,\ldots,M$. Since $A(\widehat{m}(\mathbf{b}_k)) \leq 1$, the condition in (68) is certainly satisfied for *all* nodes $\mathbf{b}_k \neq \mathbf{0}$ and tree levels $k = 1,\ldots,\bar{k}$, with $\bar{k}$ being the largest integer satisfying

$$\sqrt[k+L]{(k+L)!} \leq \rho_C. \tag{71}$$

We set $\bar{k} = 0$ if no integer $k \geq 1$ satisfies (71). Next, using (67) for $\mathbf{b}_k \neq \mathbf{0}$ and[3] $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_\infty \leq C_\infty\right] \leq \mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_2 \leq C_2\right]$ for $\mathbf{b}_k = \mathbf{0}$ in the expressions for $\mathbb{E}\{S_{\infty,k}\}$ in (33) and $\mathbb{E}\{S_{2,k}\}$ in (34) yields

$$\mathbb{E}\{S_{\infty,k}\} \preceq \mathbb{E}\{S_{2,k}\}, \qquad \rho \to \infty \tag{72}$$

[2]For $\mathbf{b}_k = \mathbf{0}$ the high-SNR behavior of $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_\infty \leq C_\infty\right]$ and $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_2 \leq C_2\right]$ is trivial since the expressions $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{0})||_\infty \leq C_\infty\right] = \left[\gamma_1(\kappa_\infty)\right]^{k+L}$ and $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{0})||_2 \leq C_2\right] = \gamma_{k+L}(\kappa_2)$ do not depend on the SNR ρ.

[3]As already noted in footnote 2, we have $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{0})||_\infty \leq C_\infty\right] = \left[\gamma_1(\kappa_\infty)\right]^{k+L}$ and $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{0})||_2 \leq C_2\right] = \gamma_{k+L}(\kappa_2)$, where $\left[\gamma_1(\kappa_\infty)\right]^N = \gamma_N(\kappa_2) = 1-\epsilon$ due to (48) and (47). It follows that the condition $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{0})||_\infty \leq C_\infty\right] \leq \mathrm{P}\left[||\mathbf{z}_k(\mathbf{0})||_2 \leq C_2\right]$ is equivalent to $\left[\gamma_1(\kappa_\infty)\right]^{k+L} \leq \gamma_{k+L}(\kappa_2)$. Furthermore, using $\gamma_1(\kappa_\infty) = (1-\epsilon)^{1/N} = \left[\gamma_N(\kappa_2)\right]^{1/N}$, this condition can be written as $\left[\gamma_N(\kappa_2)\right]^{1/N} \leq \left[\gamma_{k+L}(\kappa_2)\right]^{1/(k+L)}$, which according to (108) holds for all $k = 1,\ldots,M$.

for $k = 1, \ldots, \bar{k}$. Equivalently, in the high-SNR regime, the average number of nodes visited by SD-l^∞ up to tree level $\bar{k}$ (corresponding to tree levels closer to the root) is smaller than that for SD-l^2. Furthermore, if $\bar{k} = M$, we have

$$\mathbb{E}\{S_\infty\} \preceq \mathbb{E}\{S_2\}, \qquad \rho \to \infty \tag{73}$$

since (72) then holds for all tree levels $k = 1, \ldots, M$. We will next show that for the radii chosen according to (50) and (48), we can, indeed, have $\bar{k} = M$. In the following, we write $\rho_C(\epsilon)$ to emphasize the dependence of the radii ratio on the parameter $\epsilon \in [0, 1]$. In Appendix F it is shown that $\rho_C(\epsilon)$ is a nondecreasing function of ϵ and furthermore

$$\lim_{\epsilon \to 0} \rho_C(\epsilon) = 1 \quad \text{and} \quad \lim_{\epsilon \to 1} \rho_C(\epsilon) = \sqrt[N]{N!} \tag{74}$$

which implies

$$1 \leq \rho_C(\epsilon) \leq \sqrt[N]{N!}. \tag{75}$$

We can therefore conclude that $\bar{k}$ is a nondecreasing function of ϵ taking on any value in $[0, M]$ (achieved by varying the parameter ϵ) with the following two extreme cases.

- For $\epsilon \to 1$, we get $\bar{k} \to M$ so that (73) holds. This indicates that in the high-SNR regime SD-l^∞ will have a smaller total complexity than SD-l^2 if ϵ is sufficiently close to 1.
- For $\epsilon \to 0$, we have $\bar{k} \to 1$ for $L = 0$ and $\bar{k} \to 0$ for $L > 0$. Equivalently, if ϵ is sufficiently close to 0, (72) holds for the first tree level if $L = 0$ and holds for none of the tree levels if $L > 0$. In particular, for $\epsilon \to 0$, we have $\rho_C(\epsilon) \to 1$ and hence $C_\infty \overset{a}{\sim} C_2$, $\epsilon \to 0$. This implies that the hypercube of radius C_∞ contains the hypersphere of radius C_2 and the total complexity of SD-l^∞ will trivially be higher than that of SD-l^2. In general, SD-l^∞ will have a higher total complexity than SD-l^2 if ϵ is small.

In summary, varying the parameter ϵ has a significant impact on the total complexity of SD-l^∞ relative to that of SD-l^2. In particular, the total complexity of SD-l^∞ can be higher or lower than that of SD-l^2.

Finally, we want to emphasize that the results on the average high-SNR TPB are nicely supported by simple geometric considerations if we assume that $L = N - M = 0$ and argue that the average number of visited nodes at tree level k is roughly determined by the volume of the involved search space of dimension k (see, e.g., [10]). Indeed, for SD-l^2, the search space at tree level k is a hypersphere of radius C_2 with volume $V_{2,k} = \pi^k (C_2^2)^k / k!$ [40] whereas the corresponding search space for SD-l^∞ has volume $V_{\infty,k} = \pi^k (C_\infty^2)^k$. It follows that $V_{\infty,k} \leq V_{2,k}$ for all tree levels $k = 1, \ldots, \bar{k}$ with $\bar{k}$ being the largest integer satisfying $\sqrt[k]{k!} \leq \rho_C$ and $V_{\infty,k} > V_{2,k}$ for $k = \bar{k} + 1, \ldots, M$. This indicates that SD-l^∞ prunes more nodes than SD-l^2 at tree levels closer to the root, whereas this behavior is reversed at tree levels closer to the leaves. Even more, the threshold tree level $\bar{k}$ (found through analyzing the volume behavior of the search spaces) equals the threshold tree level (71) found through a high-SNR analysis of the pruning probabilities.

V. The Truth and the Beautiful: $l^{\widetilde{\infty}}$-Norm SD

As already mentioned, the SD-l^∞ VLSI implementation in [8] is actually based on the $l^{\widetilde{\infty}}$-norm rather than the l^∞-norm; the corresponding tree search is conducted using the recursive metric computation rule $\|\mathbf{z}_k(\mathbf{d}_k)\|_{\widetilde{\infty}} = \max\left\{\|\mathbf{z}_{k-1}(\mathbf{d}_{k-1})\|_{\widetilde{\infty}}, \|[\mathbf{z}(\mathbf{d})]_{M-k+1}\|_{\widetilde{\infty}}\right\}$ together with the partial BC

$$\|\mathbf{z}_k(\mathbf{d}_k)\|_{\widetilde{\infty}} \leq C_{\widetilde{\infty}} \tag{76}$$

where $C_{\widetilde{\infty}}$ denotes the "radius" associated with SD-$l^{\widetilde{\infty}}$. Consequently, SD-$l^{\widetilde{\infty}}$ finds all data vectors $\mathbf{d}$ satisfying $\|\mathbf{z}(\mathbf{d})\|_{\widetilde{\infty}} \leq C_{\widetilde{\infty}}$ and chooses, within this set, the vector

$$\widehat{\mathbf{d}}_{\widetilde{\infty}} = \arg\min_{\mathbf{d} \in \mathcal{A}^M} \|\mathbf{z}(\mathbf{d})\|_{\widetilde{\infty}}. \tag{77}$$

We next show how the error probability (see Section II) and complexity (see Section III) results obtained for SD-l^∞ carry over to SD-$l^{\widetilde{\infty}}$. Most results in this section are based on the simple inequalities

$$\frac{1}{2}\|\mathbf{x}\|_\infty^2 \leq \|\mathbf{x}\|_{\widetilde{\infty}}^2 \leq \|\mathbf{x}\|_\infty^2, \qquad \mathbf{x} \in \mathbb{C}^N. \tag{78}$$

A. Error Probability of SD-$l^{\widetilde{\infty}}$

With (77), an upper bound on the PEP of SD-$l^{\widetilde{\infty}}$ is given by

$$\mathrm{P}_{\mathbf{d}' \to \mathbf{d}, \widetilde{\infty}}(\rho) \leq \mathrm{P}\left[\|\mathbf{z}(\mathbf{d})\|_{\widetilde{\infty}} \leq \|\mathbf{z}(\mathbf{d}')\|_{\widetilde{\infty}}\right].$$

Next, following the steps in (18)–(19) for SD-l^∞, using the bounds (78) and (12), yields

$$\mathrm{P}_{\mathbf{d}' \to \mathbf{d}, \widetilde{\infty}}(\rho) \leq \mathrm{P}\left[\|\mathbf{w}\|_2 \geq \frac{1}{\sqrt{2N+1}}\|\mathbf{H}\mathbf{b}\|_2\right]. \tag{79}$$

Employing the same arguments as in the SD-l^∞ or in the SD-l^2 case in Section II-B-I, we can conclude that $\mathrm{P}_{\mathbf{d}' \to \mathbf{d}, \widetilde{\infty}}(\rho)$ has the same SNR exponent as $\mathrm{P}_{\mathbf{d}' \to \mathbf{d}, \infty}(\rho)$ and $\mathrm{P}_{\mathbf{d}' \to \mathbf{d}, \mathrm{ML}}(\rho)$. Furthermore, from (79) we obtain $\mathrm{P}_{\mathbf{d}' \to \mathbf{d}, \widetilde{\infty}}(\rho) \leq \mathrm{UB}_{\widetilde{\infty}}(\rho)$, where $\mathrm{UB}_{\widetilde{\infty}}(\rho)$ is given by $\mathrm{UB}_\infty(\rho)$ in (20) with the factor $\sqrt{N}$ replaced by $\sqrt{2N}$. Accordingly, the asymptotic SNR gap $\widetilde{\beta}$ between $\mathrm{UB}_{\widetilde{\infty}}(\rho)$ and $\mathrm{LB}_{\mathrm{ML}}(\rho)$, as defined in (21), i.e., $\mathrm{UB}_{\widetilde{\infty}}(\rho) \overset{a}{\sim} \mathrm{LB}_{\mathrm{ML}}(\rho/\widetilde{\beta})$, $\rho \to \infty$, is given by β in (22) with the factor $\sqrt{N}$ replaced by $\sqrt{2N}$. This corresponds to an increase of a factor of roughly two in the corresponding upper bound on the SNR gap as compared to that achieved by SD-l^∞(22). Finally, employing the arguments used in Section II-B2, these statements carry over to the total error probability in a straightforward fashion showing that SD-$l^{\widetilde{\infty}}$ (like SD-l^∞ and SD-l^2) achieves full diversity order N with an asymptotic SNR gap to ML detection that is bounded and that increases at most linearly in N.

B. Complexity of SD-$l^{\widetilde{\infty}}$

With (76) and following the steps (29)–(33), we obtain the complexity $\mathbb{E}\{S_{\widetilde{\infty},k}\}$ of SD-$l^{\widetilde{\infty}}$ at tree level k as

$$\mathbb{E}\{S_{\widetilde{\infty},k}\} = \frac{1}{|\mathcal{A}|^k} \sum_{\mathbf{b}_k} \mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_{\widetilde{\infty}} \leq C_{\widetilde{\infty}}\right] \tag{80}$$

with the total complexity given by $\mathbb{E}\{S_{\widetilde{\infty}}\} = \sum_{k=1}^M \mathbb{E}\{S_{\widetilde{\infty},k}\}$. As in the case of SD-$l^\infty$, invoking the fact that the elements

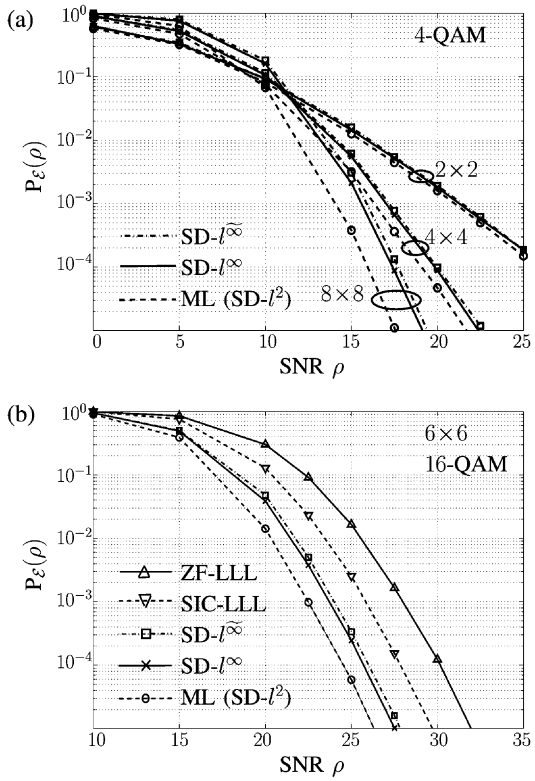

Fig. 2. Uncoded total error probability $\mathrm{P}_\mathcal{E}(\rho)$ as a function of SNR ρ for SD-$l^{\widetilde{\infty}}$, SD-l^{∞}, and SD-l^2 (ML) detection for (a) a 2×2, 4×4, and 8×8 MIMO system, respectively, and a 4-QAM symbol alphabet, and (b) for a 6×6 MIMO system with a 16-QAM symbol alphabet as well as for ZF-LLL and SIC-LLL.

of $\mathbf{z}_k(\mathbf{b}_k)$ (conditioned on $\mathbf{b}_k$) are statistically independent (cf. (36)), we get

$$\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right] = \prod_{i=1}^{k+L} \mathrm{P}\left[\|[\mathbf{z}_k(\mathbf{b}_k)]_i\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right]. \tag{81}$$

The real and imaginary parts of the bottom L elements of $\mathbf{z}_k(\mathbf{b}_k)$ are i.i.d. $\mathcal{N}(0,\sigma^2/2)$ so that

$$\mathrm{P}\left[\|[\mathbf{z}_k(\mathbf{b}_k)]_i\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right] = \left[\gamma_{\frac{1}{2}}\left(\frac{C_{\widetilde{\infty}}}{\sigma^2}\right)\right]^2$$

$i = k+1,\ldots,k+L$, which, upon insertion into (81), yields

$$\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right] = \left[\gamma_{\frac{1}{2}}\left(\frac{C_{\widetilde{\infty}}}{\sigma^2}\right)\right]^{2L} \cdot \prod_{m=1}^{k} \mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right] \tag{82}$$

analogously to (36). An analytic expression for $\mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right]$ can be obtained if b_{M-m+1} is purely real, purely imaginary, or equal to zero. For these cases the real and imaginary parts of $[\mathbf{z}(\mathbf{b})]_{M-m+1}$ are statistically independent, which gives (see Appendix G)

$$\mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right] = \gamma_{\frac{1}{2}}\left(\frac{C_{\widetilde{\infty}}^2}{\sigma_m^2}\right)\sum_{s=0}^{\infty} D_s(\mathbf{b}_m)\gamma_{s+\frac{1}{2}}\left(\frac{C_{\widetilde{\infty}}^2}{\sigma_m^2}\right) \tag{83}$$

where $D_s(\mathbf{b}_m)$ is defined in (94) and σ_m^2 is specified in (90). For the general case of b_{M-m+1} having a nonzero real and a nonzero imaginary part, i.e., $b_{\mathrm{R},M-m+1} \neq 0$ and $b_{\mathrm{I},M-m+1} \neq 0$, the real and imaginary parts of $[\mathbf{z}(\mathbf{b})]_{M-m+1}$ are statistically dependent, which seems to make it difficult to find a closed-form expression for $\mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right]$. One can, however, resort to upper and lower bounds. In particular, it follows from (78) that

$$\mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right] \ge \mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \le C_{\widetilde{\infty}}\right] \tag{84}$$

and

$$\mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right] \le \mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \le \sqrt{2}C_{\widetilde{\infty}}\right]. \tag{85}$$

The RHS expressions of (84) and (85) can now be expressed analytically using (37). Together with (82) and (83) this provides upper and lower bounds on $\mathrm{P}\left[\|\mathbf{z}_k(\mathbf{b}_k)\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right]$ and thus on $\mathbb{E}\{S_{\widetilde{\infty},k}\}$, $k = 1,\ldots,M$, and $\mathbb{E}\{S_{\widetilde{\infty}}\}$. We do not display the resulting final expressions as they are rather involved and do not contribute to deepening the understanding of the computational complexity of SD-$l^{\widetilde{\infty}}$. Corresponding numerical results are provided in Section VI-C.

Following the choice of the radii for SD-l^{∞} and SD-l^2 in (49) and (47), respectively, $C_{\widetilde{\infty}}$ is obtained by setting

$$\mathrm{P}\left[\|\mathbf{n}\|_{\widetilde{\infty}} \le C_{\widetilde{\infty}}\right] = \left[\gamma_{\frac{1}{2}}\left(\frac{C_{\widetilde{\infty}}^2}{\sigma^2}\right)\right]^{2N} = 1-\epsilon \tag{86}$$

which results in (cf. (50) and (48))

$$C_{\widetilde{\infty}}^2 = \sigma^2\gamma_{\frac{1}{2}}^{-1}\left(\sqrt[2N]{1-\epsilon}\right). \tag{87}$$

Finally, we note that SD-$l^{\widetilde{\infty}}$ with $C_{\widetilde{\infty}}$ chosen according to (87) also exhibits exponential complexity scaling in the problem size M. This can be established by following the same approach as for SD-l^{∞} and SD-l^2 (see Sections III-E2 and III-E3, respectively), i.e., by developing an analytically tractable lower bound on $\mathbb{E}\{S_{\widetilde{\infty}}\}$ and then establishing that this bound scales exponentially in M.

VI. Numerical Results

In this section, we provide numerical results quantifying some of our analytical findings. All the results in the remainder of this section are based on independently and equally likely transmitted data symbols.

A. Error Probability

We compare the uncoded error-rate performance of SD-l^{∞} and SD-$l^{\widetilde{\infty}}$ to that of SD-l^2 (ML) detection by means of

Monte-Carlo simulations. Fig. 2(a) shows total error probabilities $\mathrm{P}_{\mathcal{E}}(\rho)$ as functions of SNR ρ for a 2×2, 4×4, and 8×8 MIMO system, respectively, using 4-QAM symbols in all three cases. We can observe that both SD-l^∞ and SD-$l^{\widetilde{\infty}}$ achieve full diversity order and show near-ML performance. Indeed, SD-l^∞ and SD-$l^{\widetilde{\infty}}$ perform much better than suggested by the corresponding upper bounds on the SNR gap (i.e., $|\mathcal{A}|\beta$ with β given by (22) for SD-l^∞ and $|\mathcal{A}|\tilde{\beta}$ with $\tilde{\beta}$ given by (22) with the factor $\sqrt{N}$ replaced by $\sqrt{2N}$ for SD-$l^{\widetilde{\infty}}$). Consistent with the $\sqrt{2}$-difference in the upper bounds on the corresponding SNR gaps, we can observe that SD-$l^{\widetilde{\infty}}$ performs slightly worse than SD-l^∞. The results in Fig. 2(a) also show that the performance loss incurred by SD-l^∞ and SD-$l^{\widetilde{\infty}}$ increases for increasing $M=N$. Fig. 2(b) shows $\mathrm{P}_{\mathcal{E}}(\rho)$ for a 6×6 MIMO system and a 16-QAM symbol alphabet as well as $\mathrm{P}_{\mathcal{E}}(\rho)$ for zero-forcing (ZF) and SIC detection in combination with LLL lattice-reduction [24] (referred to as ZF-LLL and SIC-LLL, respectively). For LLL lattice-reduction the complex-valued variant according to [41, with parameter $\delta=3/4$] was used. It can be observed that all detectors achieve full diversity order and that SD-l^∞ and SD-$l^{\widetilde{\infty}}$ outperform ZF-LLL and SIC-LLL.

B. Complexity

Next we consider the complexity of SD-l^∞ and SD-l^2 for the case of fixed radii C_∞ and C_2 chosen according to (50) and (48), respectively, with the same value of ϵ in both cases (for numerical results on the complexity of SD-$l^{\widetilde{\infty}}$, we refer to Section VI-C). The total complexity $\mathbb{E}\{S\}$ as a function of ϵ for SD-l^∞ [see (31) with (44)] and SD-l^2 [see (46) with (45)] is shown in Fig. 3 for a 4×4, 6×6, and 8×8 MIMO system, respectively, using a 4-QAM alphabet and an SNR of $\rho=15$ dB. The following conclusions can be drawn from these results:

- For a given ϵ, the complexity of SD-l^∞ can be higher or lower than that of SD-l^2.
- SD-l^∞ exhibits a lower complexity than SD-l^2 for larger values of ϵ, while for smaller values of ϵ, SD-l^∞ has a higher complexity than SD-l^2. This behavior was indicated by the high SNR-analysis of the TPB of SD-l^∞ and SD-l^2 (in particular, see the discussion on the two extreme cases $\epsilon\to 0$ and $\epsilon\to 1$ in Section IV-B).
- The complexity savings of SD-l^∞ over SD-l^2 for values of ϵ close to 1 are more pronounced for increasing $M=N$.
- In practice, ϵ is set equal to the target error rate of the system (see the discussion in Section III-D). In the present example, we operate at 15 dB SNR and the corresponding target error rate can be inferred from Fig. 2(a), which results in ϵ values for which SD-l^∞ has a lower complexity than SD-l^2 (cf. Fig. 3). For the 8×8 system, for example, our target error rate at 15 dB SNR, according to Fig. 2(a), is around 10^{-3}. For this case, the complexity savings of SD-l^∞ as compared to SD-l^2 are around 25% according to Fig. 3.

C. Tree Pruning Behavior and Complexity Bounds for SD-$l^{\widetilde{\infty}}$

The goal of this section is to quantify the level-wise complexities $\mathbb{E}\{S_k\}$ for SD-l^∞, SD-$l^{\widetilde{\infty}}$, and SD-l^2, as well as to illustrate the quality of the upper and lower bounds on the complexity of SD-$l^{\widetilde{\infty}}$ reported in Section V-B. We consider a 6×6 MIMO system and a 4-QAM symbol alphabet with the radii C_2, C_∞, and $C_{\widetilde{\infty}}$ chosen according to (48), (50), and (87), respectively, for $\epsilon=10^{-2}$ and $\epsilon=10^{-5}$. Note that for the cases of SD-l^∞ and SD-l^2 exact complexity expressions according to (44) and (45), respectively, are available. Fig. 4 shows $\mathbb{E}\{S_k\}$ as a function of the tree level k for SD-l^2 and for SD-l^∞ including the corresponding upper and lower bounds on $\mathbb{E}\{S_k\}$ for SD-$l^{\widetilde{\infty}}$ at an SNR of $\rho=15$ dB (Fig. 4(a) for $\epsilon=10^{-2}$ and Fig. 4(b) for $\epsilon=10^{-5}$). The following conclusions can be drawn from these results:

- At tree levels close to the root (i.e., for small k), SD-l^∞ (SD-$l^{\widetilde{\infty}}$) visits fewer nodes than SD-l^2 on average; at tree levels close to the leaves this behavior is reversed. This observation is supported by the results on the average TPB reported in Section IV [in particular, see (72)].
- The complexity savings of SD-l^∞ (SD-$l^{\widetilde{\infty}}$) over SD-l^2 close to the root extend to higher tree levels for the larger ϵ value of 10^{-2}. This behavior is consistent with the average TPB analysis in Section IV stating that $\mathbb{E}\{S_{\infty,k}\}\preceq\mathbb{E}\{S_{2,k}\}$, $\rho\to\infty$, up to tree level $\bar{k}$, where $\bar{k}$ was shown to be a nondecreasing function of ϵ (see Section IV-B). For

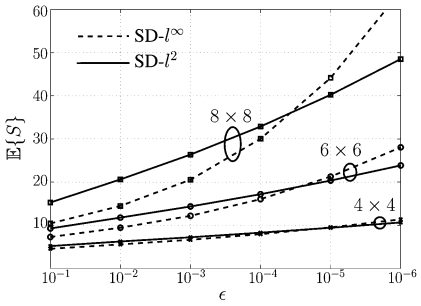

Fig. 3. Total complexity $\mathbb{E}\{S\}$ as a function of ϵ for SD-l^∞ and SD-l^2 for a 4×4, 6×6, and 8×8 MIMO system, respectively, and a 4-QAM symbol alphabet at an SNR of $\rho=15$ dB.

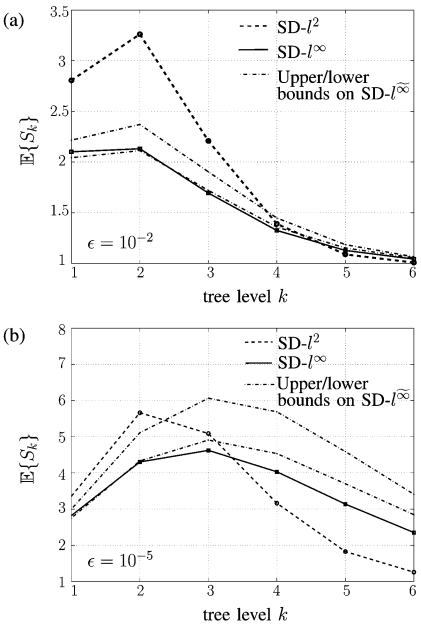

Fig. 4. Complexity $\mathbb{E}\{S_k\}$ as a function of the tree level k for SD-$l^{\widetilde{\infty}}$, SD-l^∞, and SD-l^2 with (a) $\epsilon=10^{-2}$ and (b) $\epsilon=10^{-5}$ for a 6×6 MIMO system at an SNR of $\rho=15$ dB using 4-QAM modulation. For SD-$l^{\widetilde{\infty}}$ upper and lower bounds are shown (see Section V-B). Fig. 3 shows the corresponding complexity results for SD-l^∞ and SD-l^2.

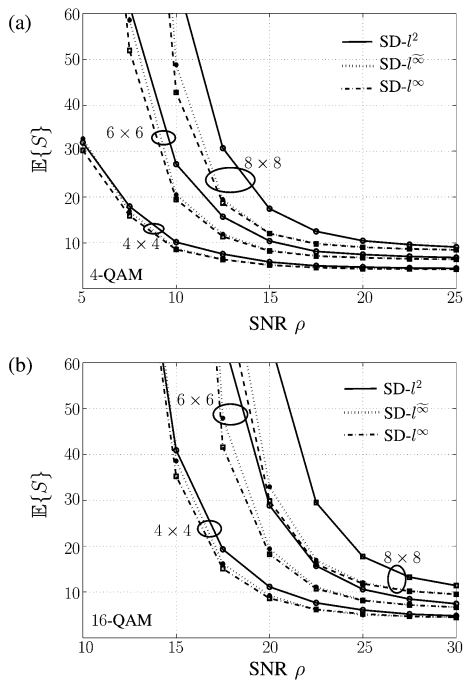

Fig. 5. Total complexity versus SNR ρ for SD-$l^{\widetilde{\infty}}$, SD-l^{∞}, and SD-l^2, all with restarting for a 4×4, 6×6, and 8×8 MIMO system (for the ϵ-schedule see text), using (a) 4-QAM modulation and (b) 16-QAM modulation.

example, we have $\bar{k}=3$ for $\epsilon=10^{-2}$, while $\bar{k}=2$ for $\epsilon=10^{-5}$.

- For $\epsilon=10^{-2}$, the complexity savings of SD-l^{∞} at tree levels close to the root are dominant enough to result in a smaller total complexity of SD-l^{∞} as compared to the complexity of SD-l^2 (cf. Fig. 3). For $\epsilon=10^{-5}$, however, the increased complexity of SD-l^{∞} at tree levels close to the leaves outweighs the savings close to the root resulting in higher total complexity of SD-l^{∞} when compared to the complexity of SD-l^2 (cf. Fig. 3).
- The upper and lower bounds on the complexity of SD-$l^{\widetilde{\infty}}$ are sufficiently tight to capture the essential aspects of the level-wise complexity of SD-$l^{\widetilde{\infty}}$. As for SD-l^{∞}, we can again observe complexity savings of SD-$l^{\widetilde{\infty}}$ over SD-l^2 close to the root, whereas this behavior is reversed at tree levels close to the leaves. Furthermore, for the examples considered, the lower bounds on the complexity of SD-$l^{\widetilde{\infty}}$ show that SD-$l^{\widetilde{\infty}}$ has a higher total complexity than SD-l^{∞} (see also Section VI-D).

D. Complexity of Sphere-Decoding With Restarting

As already mentioned in Section III-D, to guarantee ML or exact SD-l^{∞} performance the corresponding SD algorithm has to be restarted with an increased radius in cases where the initial radius was chosen too small for the search sphere (or box) to contain a valid leaf node. The same is, of course, true for SD-$l^{\widetilde{\infty}}$. To evaluate the overall (across potential multiple SD runs) complexity of SD-l^2, SD-l^{∞}, and SD-$l^{\widetilde{\infty}}$ we choose an increasing radii schedule obtained by setting $\epsilon=(0.1)^i$, $i=1,2,\ldots$, in the ith run of the SD. Corresponding average (w.r.t. channel, noise, and data) complexity results for 4×4, 6×6, and 8×8 MIMO systems using 4-QAM and 16-QAM modulation obtained through Monte-Carlo simulations can be found in Fig. 5. We note that analytical expressions for the overall complexity of SD (be it for SD-l^2 or for SD-l^{∞}) with restarting are not available since the statistics of the corresponding required number of SD runs seem to be difficult to obtain. From Fig. 5, we can observe that in the relevant SNR regime (e.g., about 10 dB to 15 dB for the 4-QAM case and 20 dB to 25 dB for the 6×6 16-QAM case corresponding to error probabilities of about 10^{-1} to 10^{-3}, cf. Fig. 2) SD-l^{∞} and SD-$l^{\widetilde{\infty}}$ exhibit lower complexity than SD-l^2. For example, at 12.5 dB, we can infer from Fig. 5(a) that the corresponding complexity savings of SD-l^{∞} and SD-$l^{\widetilde{\infty}}$ over SD-l^2 are about 30%. Furthermore, it can be observed that the complexity savings of SD-l^{∞} and SD-$l^{\widetilde{\infty}}$ over SD-l^2 are more pronounced for increasing $M=N$. We finally emphasize that these computational (algorithmic) complexity savings of SD-$l^{\widetilde{\infty}}$ over SD-l^2 go along with a significant reduction in the circuit complexity for metric computation [8] (see the discussion in Section I-A). Indeed, the results in [8] indicate that the overall (circuit and algorithmic) complexity of SD-$l^{\widetilde{\infty}}$ is up to a factor of 5 lower than the overall complexity of SD-l^2.

VII. Conclusion

We analyzed sphere-decoding (SD) based on the l^{∞}-norm and provided theoretical underpinning for the observations reported in [8]. The significance of l^{∞}-norm SD is supported by the fact that its overall implementation complexity in hardware is up to a factor of 5 lower than that for SD based on the l^2-norm (corresponding to optimum detection). In particular, we found that using the l^{∞}-norm instead of the l^2-norm does not result in a reduction of diversity order while leading to an SNR gap, compared to optimum performance, that increases at most linearly in the number of receive antennas. We furthermore showed that for many cases of practical interest l^{∞}-norm SD, besides having a smaller circuit complexity for metric computation (thanks to the fact that it avoids squaring operations) also exhibits smaller computational (algorithmic) complexity (in terms of the number of nodes visited in the search tree) than l^2-norm SD. The computational complexity of l^{∞}-norm SD was found to scale exponentially in the number of transmit antennas as is also the case for l^2-norm SD.

Appendix A
Calculation of $\mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}|\leq C_{\infty}\right]$

In the following, we derive (37). We start by introducing the RVs

$$v_m = R_{M-m+1,M-m+1}|b_{M-m+1}|$$
$$u_m = \sum_{i=M-m+2}^{M} R_{M-m+1,i}\, b_i + n_{M-m+1}. \qquad (88)$$

Since the nonzero entries in $\mathbf{R}$ and the entries in $\mathbf{n}$ are all statistically independent, v_m and u_m are statistically independent as well. Here, v_m is a $\chi_{2(m+L)}$-distributed RV with pdf [cf. (1)]

$$g_m(v) = \frac{2M^{m+L}}{\Gamma(m+L)|b_{M-m+1}|^{2(m+L)}} \cdot v^{2(m+L)-1} e^{-\frac{v^2}{|b_{M-m+1}|^2/M}}. \quad (89)$$

The RV u_m is $\mathcal{CN}(0,\sigma_m^2)$ distributed, where

$$\sigma_m^2 = ||\mathbf{b}_{m-1}||^2/M + \sigma^2. \quad (90)$$

Exploiting the circular symmetry of u_m, we have $|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \stackrel{d}{=} |v_m + u_m|$ and, thus

$$\mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \le C_\infty\right] = \int_0^\infty \mathrm{P}\left[|v_m + u_m| \le C_\infty | \, v_m = v\right] g_m(v). \quad (91)$$

For given $v_m = v$, the RV $\frac{2}{\sigma_m^2}|v + u_m|^2$ is noncentral χ_2^2-distributed with noncentrality parameter $\frac{2v^2}{\sigma_m^2}$. Thus, (see [42, Corollary 1.3.5])

$$\mathrm{P}\left[|v_m + u_m| \le C_\infty | \, v_m = v\right] = \sum_{s=0}^{\infty} e^{-\frac{v^2}{\sigma_m^2}} \left(\frac{v}{\sigma_m}\right)^{2s} \frac{1}{s!} \gamma_{s+1}\left(\frac{C_\infty^2}{\sigma_m^2}\right). \quad (92)$$

Inserting (89) and (92) into (91) yields

$$\mathrm{P}\left[|v_m + u_m| \le C_\infty\right] = \sum_{s=0}^{\infty} \frac{2\gamma_{s+1}(C_\infty^2/\sigma_m^2)M^{m+L}}{s!\sigma_m^{2s}\Gamma(m+L)|b_{M-m+1}|^{2(m+L)}} \cdot \int_0^\infty v^{2(s+m+L)-1} e^{-v^2\left(\frac{M}{|b_{M-m+1}|^2} + \frac{1}{\sigma_m^2}\right)} dv.$$

Here, the integral can easily be rewritten such that the integrand is the pdf of a $\chi_{2(s+m+L)}$-distributed RV (cf. (1)), which then yields

$$\int_0^\infty v^{2(s+m+L)-1} e^{-v^2\left(\frac{M}{|b_{M-m+1}|^2} + \frac{1}{\sigma_m^2}\right)} dv = \frac{1}{2}\Gamma(s+m+L)\left(\frac{M}{|b_{M-m+1}|^2} + \frac{1}{\sigma_m^2}\right)^{-(s+m+L)}.$$

Finally, using $\Gamma(a) = (a-1)!$ for positive integers a, we get

$$\mathrm{P}\left[|v_m + u_m| \le C_\infty\right] = \sum_{s=0}^{\infty} D_s(\mathbf{b}_m)\gamma_{s+1}\left(\frac{C_\infty^2}{\sigma_m^2}\right) \quad (93)$$

where

$$D_s(\mathbf{b}_m) = \binom{s+m+L-1}{m+L-1} p(\mathbf{b}_m)^{m+L}(1-p(\mathbf{b}_m))^s \quad (94)$$

and, as defined in (39)

$$p(\mathbf{b}_m) = \frac{\sigma_m^2}{\sigma_m^2 + |b_{M-m+1}|^2/M} = \frac{||\mathbf{b}_{m-1}||_2^2 + M\sigma^2}{||\mathbf{b}_m||_2^2 + M\sigma^2}.$$

In the remainder of this section, we show that the infinite summation in (93) can be avoided. We use $p_m = p(\mathbf{b}_m)$ to simplify notation and we start by noting that (93) can be written as

$$\mathrm{P}\left[|v_m + u_m| \le C_\infty\right] = \frac{p_m^{m+L}}{(m+L-1)!} \cdot \int_0^{\frac{C_\infty^2}{\sigma_m^2}} \left[\sum_{s=0}^{\infty}\left(\prod_{i=1}^{m+L-1}(s+i)\right)\frac{[(1-p_m)t]^s}{s!}\right] e^{-t} dt \quad (95)$$

where the identity (103) for the lower incomplete Gamma function was used. With $g(x) = e^x x^{m+L-1}$ and the series expansion $e^x = \sum_{s=0}^{\infty} \frac{x^s}{s!}$, we have that

$$g^{(m+L-1)}(x) = \sum_{s=0}^{\infty}\left(\prod_{i=1}^{m+L-1}(s+i)\right)\frac{x^s}{s!}.$$

On the other hand, by Leibniz's rule for the differentiation of products of functions, we also have

$$g^{(m+L-1)}(x) = \sum_{l=0}^{m+L-1}\binom{m+L-1}{l}\frac{(m+L-1)!}{(m+L-1-l)!}e^x x^{m+L-1-l}.$$

Thus, (95) can equivalently be written as

$$\mathrm{P}\left[|v_m + u_m| \le C_\infty\right] = \sum_{l=0}^{m+L-1}\binom{m+L-1}{l} \cdot \frac{p_m^{m+L}(1-p_m)^{m+L-1-l}}{\Gamma(m+L-l)} \int_0^{\frac{C_\infty^2}{\sigma_m^2}} t^{m+L-l-1} e^{-p_m t} dt.$$

By substituting $t' = p_m t$ and again using identity (103), we finally get (37) upon noting that $p_m/\sigma_m^2 = 1/(\sigma^2 + ||\mathbf{b}_m||^2/M)$.

APPENDIX B
SUM REPRESENTATION OF $|[\mathbf{z}(\mathbf{b})]_{M-m+1}|^2$

In the following, we prove (40) based on the following theorem.

Theorem [36]: Consider the RVs

$$z^{(l)} = g\left(y_1^{(l)}, \ldots, y_a^{(l)}\right), \qquad l = 0, \ldots, a \quad (96)$$

where $y_i^{(l)}$, $i = 1, \ldots, a$, for every l, are statistically independent RVs with pdfs equal to $f_1(x)$ if $i \le l$ and $f_2(x)$ otherwise. If $g(\cdot)$ is a symmetric function (i.e., $g(\cdot)$ is unchanged by any permutation of its arguments), then the pdf of

$$z = g(y_1, \ldots, y_a) \quad (97)$$

where the y_i, $i = 1, \ldots, a$, are i.i.d. with mixture pdf

$$f_{y_i}(x) = pf_1(x) + (1-p)f_2(x), \qquad 0 \le p \le 1 \quad (98)$$

is given by

$$f_z(x) = \sum_{l=0}^{a} B_l f_{z^{(l)}}(x) \tag{99}$$

with

$$B_l = \binom{a}{l} p^l (1-p)^{M-l}.$$

Here, $f_{z^{(l)}}(x), l = 0, \ldots, a$, denotes the pdf of $z^{(l)}$ specified in (96).

We apply this theorem to the case at hand by defining $f_1(x) = \delta(x)$ and $f_2(x) = f_{\chi_2^2}(x)$ and setting $a = m + L - 1$. Furthermore, we take $g(\cdot)$ as

$$g(x_1, \ldots, x_{m+L-1}) = \frac{\|\mathbf{b}_m\|_2^2/M + \sigma^2}{2} \left(\sum_{i=1}^{m+L-1} x_i \right) \tag{100}$$

which implies

$$f_{z^{(l)}}(x) = \frac{2}{\|\mathbf{b}_m\|_2^2/M + \sigma^2} \cdot f_{\chi^2_{2(m+L-1-l)}} \left(\frac{2x}{\|\mathbf{b}_m\|_2^2/M + \sigma^2} \right) \tag{101}$$

if $l < m + L - 1$ and $f_{z^{(l)}}(x) = \delta(x)$ if $l = m + L - 1$ for the pdfs of the RVs $z^{(l)}$ defined in (96). Using (99), we thus get

$$f_z(x) = \frac{2}{\|\mathbf{b}_m\|_2^2/M + \sigma^2} \cdot \sum_{l=0}^{m+L-1} B_l f_{\chi^2_{2(m+L-1-l)}} \left(\frac{2x}{\|\mathbf{b}_m\|_2^2/M + \sigma^2} \right)$$

with the corresponding cdf essentially given by the RHS of (37) but with two missing degrees of freedom in the χ^2-distributed RVs underlying the individual terms in the sum. To compensate for these two missing degrees of freedom, we construct the RV

$$t_m^2 = z + \frac{\|\mathbf{b}_m\|_2^2/M + \sigma^2}{2} \gamma^2 \tag{102}$$

with $\gamma^2 \sim \chi_2^2$ statistically independent of z. Noting that $(f_{\chi_a^2} * f_{\chi_b^2})(x) = f_{\chi^2_{a+b}}(x)$, we obtain

$$f_{t_m^2}(x) = \frac{2}{\|\mathbf{b}_m\|_2^2/M + \sigma^2} \cdot \sum_{l=0}^{m+L-1} B_l f_{\chi^2_{2(m+L-l)}} \left(\frac{2x}{\|\mathbf{b}_m\|_2^2/M + \sigma^2} \right)$$

or, equivalently,

$$\mathrm{P}\left[t_m^2 \le x\right] = \sum_{l=0}^{m+L-1} B_l \gamma_{m+L-l} \left(\frac{x}{\|\mathbf{b}_m\|_2^2/M + \sigma^2} \right)$$

thus, by comparison with (37), establishing that $t_m^2 \stackrel{d}{=} |[\mathbf{z}(\mathbf{b})]_{M-m+1}|^2$. Finally, (102) together with (97) and (100) shows (41).

Appendix C
Bounds on Lower Incomplete Gamma Function

In this section, we summarize properties of the lower (regularized) incomplete Gamma function

$$\gamma_a(x) = \frac{1}{\Gamma(a)} \int_0^x y^{a-1} e^{-y} dy, \quad x, a \in \mathbb{R},\ x, a \ge 0 \tag{103}$$

needed in this paper. In the remainder of this section, we will furthermore assume that $a \in \mathbb{N}$, which is the most relevant case for our results. We start by noting that $\gamma_a(x)$ can equivalently be written as [43, Sec. 6.5]

$$\gamma_a(x) = e^{-x} \sum_{i=a}^{\infty} \frac{x^i}{i!} \tag{104}$$

$$= 1 - e^{-x} \sum_{i=0}^{a-1} \frac{x^i}{i!}. \tag{105}$$

An immediate consequence of (105) is $\gamma_1(x) = 1 - e^{-x}$. From (104), we can directly infer that

$$\gamma_{a_1}(x) \ge \gamma_{a_2}(x), \qquad a_1 \le a_2. \tag{106}$$

Furthermore, we have [44, Eq. (5.4)]

$$\left(1 - e^{-\frac{1}{\sqrt[a]{a!}} x}\right)^a \le \gamma_a(x) \le \left(1 - e^{-x}\right)^a. \tag{107}$$

We will also need the relation

$$[\gamma_{a_1}(x)]^{\frac{1}{a_1}} \ge [\gamma_{a_2}(x)]^{\frac{1}{a_2}}, \qquad a_1 \le a_2 \tag{108}$$

which will be proved by showing that $[\gamma_a(x)]^{\frac{1}{a}}$ is a nonincreasing function of $a \in \mathbb{N}$, i.e.,

$$[\gamma_a(x)]^{\frac{1}{a}} \ge [\gamma_{a+1}(x)]^{\frac{1}{a+1}}. \tag{109}$$

The proof is by induction. For $a = 1$, we have $\gamma_1(x) \ge [\gamma_2(x)]^{\frac{1}{2}}$, which follows from (107). It remains to show that

$$[\gamma_n(x)]^{\frac{1}{n}} \ge [\gamma_{n+1}(x)]^{\frac{1}{n+1}}, \qquad n \in \mathbb{N} \tag{110}$$

implies

$$[\gamma_{n+1}(x)]^{\frac{1}{n+1}} \ge [\gamma_{n+2}(x)]^{\frac{1}{n+2}}. \tag{111}$$

To this end, we use [45, Lemma 3] which states that

$$\gamma_{n+1}(x) \ge [\gamma_n(x)]^{\frac{1}{2}} [\gamma_{n+2}(x)]^{\frac{1}{2}}. \tag{112}$$

Inserting (110) into (112), we get

$$[\gamma_{n+1}(x)]^{\frac{1}{2}\left(\frac{n+2}{n+1}\right)} \ge [\gamma_{n+2}(x)]^{\frac{1}{2}}$$

which establishes (111) and thereby concludes the proof.

Finally, using the definition (103), it can be shown that

$$\gamma_a\left(\frac{x_1}{1+x_2}\right) \ge \gamma_a(x_1)(1+x_2)^{-a} \tag{113}$$

for any $x_1, x_2 \geq 0$.

APPENDIX D
ASYMPTOTIC BEHAVIOR OF $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_\infty \leq C_\infty\right]$

In the following, we characterize the asymptotic (in SNR) behavior of $\mathrm{P}\left[||\mathbf{z}_k(\mathbf{b}_k)||_\infty \leq C_\infty\right]$. This is done by splitting the product on the RHS in (36) into three parts, which are treated separately (recall the definition of $\widehat{m}(\mathbf{b}_k)$ in Section IV-A as the index of the first erroneous tree level and the definition of κ_∞ in (59)).

- $m = 1, \ldots, \widehat{m}(\mathbf{b}_k) - 1$: We have $[\mathbf{z}(\mathbf{b})]_{M-m+1} = [\mathbf{n}]_{M-m+1}$, which is distributed as $\mathcal{CN}(0, \sigma^2)$ so that $\mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \leq C_\infty\right] = \gamma_1(\kappa_\infty)$. Hence, the first part of the RHS in (36) is given by

$$\left[\gamma_1(\kappa_\infty)\right]^L \prod_{m=1}^{\widehat{m}(\mathbf{b}_k)-1} \mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \leq C_\infty\right] = \left[\gamma_1(\kappa_\infty)\right]^{\widehat{m}(\mathbf{b}_k)-1+L}. \quad (114)$$

- $m = \widehat{m}(\mathbf{b}_k)$: The second part of the RHS in (36) corresponds to the first erroneous tree level associated with $\mathbf{b}_k$. Here, we start by noting that (39) yields $p(\mathbf{b}_m) = M\sigma^2/(||\mathbf{b}_m||_2^2 + M\sigma^2)$, where we used $||\mathbf{b}_{m-1}||_2^2 = 0$. We thus have

$$p(\mathbf{b}_m) \overset{a}{\sim} \left(\rho||\mathbf{b}_m||_2^2/M\right)^{-1}, \qquad \rho \to \infty \quad (115)$$

and

$$1 - p(\mathbf{b}_m) \overset{a}{\sim} 1, \qquad \rho \to \infty. \quad (116)$$

Furthermore,

$$\gamma_{m+L-l}\left(\frac{C_\infty^2}{||\mathbf{b}_m||_2^2/M + \sigma^2}\right) = \gamma_{m+L-l}\left(\frac{\kappa_\infty}{1 + \rho||\mathbf{b}_m||_2^2/M}\right)$$

and (104) implies that

$$\gamma_{m+L-l}\left(\frac{\kappa_\infty}{1 + \rho||\mathbf{b}_m||_2^2/M}\right) \overset{a}{\sim} \frac{1}{(m+L-l)!} \cdot \kappa_\infty^{m+L-l}\left(\rho||\mathbf{b}_m||_2^2/M\right)^{-(m+L-l)}, \qquad \rho \to \infty. \quad (117)$$

With (37) and (115)–(117), we finally arrive at

$$\mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \leq C_\infty\right] \overset{a}{\sim} D(m)\left(\rho||\mathbf{b}_m||_2^2/M\right)^{-(m+L)}, \quad \rho \to \infty \quad (118)$$

where

$$D(m) = \sum_{l=0}^{m+L-1} \binom{m+L-1}{l} \frac{1}{(m+L-l)!} \kappa_\infty^{m+L-l}.$$

- $m = \widehat{m}(\mathbf{b}_k) + 1, \ldots, k$: For these tree levels, we have $||\mathbf{b}_{m-1}||_2^2 \neq 0$, which yields

$$p(\mathbf{b}_m) \overset{a}{\sim} \frac{||\mathbf{b}_{m-1}||_2^2}{||\mathbf{b}_m||_2^2}, \qquad \rho \to \infty. \quad (119)$$

Combining this result with (117) and (37), we thus obtain

$$\mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \leq C_\infty\right] \overset{a}{\sim} \kappa_\infty \left(\frac{||\mathbf{b}_{m-1}||_2^2}{||\mathbf{b}_m||_2^2}\right)^{m+L-1} \cdot \left(\rho||\mathbf{b}_m||_2^2/M\right)^{-1}, \quad \rho \to \infty$$

so that

$$\prod_{m=\widehat{m}(\mathbf{b}_k)+1}^{k} \mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{M-m+1}| \leq C_\infty\right] \overset{a}{\sim} \kappa_\infty^{k-\widehat{m}(\mathbf{b}_k)} \cdot \rho^{-(k-\widehat{m}(\mathbf{b}_k))} \prod_{m=\widehat{m}(\mathbf{b}_k)+1}^{k} \frac{(||\mathbf{b}_{m-1}||_2^2/M)^{m+L-1}}{(||\mathbf{b}_m||_2^2/M)^{m+L}}, \qquad \rho \to \infty. \quad (120)$$

Next, note that

$$\prod_{m=\widehat{m}(\mathbf{b}_k)+1}^{k} \frac{\left(||\mathbf{b}_{m-1}||_2^2/M\right)^{m+L-1}}{\left(||\mathbf{b}_m||_2^2/M\right)^{m+L}} = \frac{\left(||\mathbf{b}_{\widehat{m}(\mathbf{b}_k)}||_2^2/M\right)^{\widehat{m}(\mathbf{b}_k)+L}}{\left(||\mathbf{b}_k||_2^2/M\right)^{k+L}}. \quad (121)$$

Combining (114), (118), (120), and (121) finally yields (60).

APPENDIX E
MONOTONICITY OF $A(\widehat{m}(\mathbf{b}_k))$

In the following, we show that $A(\widehat{m}(\mathbf{b}_k))$ in (61) is a nonincreasing function of κ_∞ (or, equivalently, noting that $\kappa_\infty = -\log\left(1 - \sqrt[N]{1-\epsilon}\right)$, $A(\widehat{m}(\mathbf{b}_k))$ is a nondecreasing function of ϵ). This will be done by setting $x = \kappa_\infty$, $\widehat{m} = \widehat{m}(\mathbf{b}_k) + L$, and by showing that

$$f(x) = \left[\gamma_1(x)\right]^{\widehat{m}-1} \sum_{l=0}^{\widehat{m}-1} \binom{\widehat{m}-1}{l} \frac{1}{(\widehat{m}-l)!} x^{-l}$$

is a nonincreasing function of $x \geq 0$, or equivalently, $f'(x) \leq 0$, for $x \geq 0$. For $\widehat{m} = 1$ this holds trivially as $f(x) = 1$. We therefore consider the case $\widehat{m} \geq 2$ in what follows. The condition $f'(x) \leq 0$, for $x \geq 0$, is equivalent to

$$\frac{e^x-1}{(\widehat{m}-1)}\sum_{l=0}^{\widehat{m}-1}\binom{\widehat{m}-1}{l}\frac{l}{(\widehat{m}-l)!}x^{-l-1}\geq \tag{122}$$

$$\sum_{l=0}^{\widehat{m}-1}\binom{\widehat{m}-1}{l}\frac{1}{(\widehat{m}-l)!}x^{-l}, \qquad x\geq 0. \tag{123}$$

Multiplying both sides of (123) by $x^{\widehat{m}}\geq 0$, and substituting $i=\widehat{m}-l$, it remains to show that

$$p(x)\geq q(x), \qquad \text{for } x\geq 0 \tag{124}$$

where

$$p(x)=(e^x-1)\sum_{i=1}^{\widehat{m}}\frac{\widehat{m}-i}{\widehat{m}-1}a_i x^{i-1} \tag{125}$$

and

$$q(x)=\sum_{i=1}^{\widehat{m}}a_i x^i \tag{126}$$

with

$$a_i=\binom{\widehat{m}-1}{i-1}\frac{1}{i!}. \tag{127}$$

Here, we used $\binom{\widehat{m}-1}{\widehat{m}-i}=\binom{\widehat{m}-1}{i-1}$. Evidently, a sufficient condition for (124) to hold is that $p(0)\geq q(0)$ and $p'(x)\geq q'(x)$, for $x\geq 0$. Successively applying this argument, (124) can be shown by proving that

$$p^{(n)}(x)\Big|_{x=0}\geq q^{(n)}(x)\Big|_{x=0}, \qquad \text{for } n=0,\ldots,\widehat{m} \tag{128}$$

and

$$p^{(\widehat{m}+1)}(x)\geq q^{(\widehat{m}+1)}(x), \qquad x\geq 0. \tag{129}$$

Condition (129) can be verified by noting that $p^{(\widehat{m}+1)}(x)\geq 0$ for $x\geq 0$ (cf. (125)) and $q^{(\widehat{m}+1)}(x)=0$ since $q(x)$ in (126) is a polynomial of degree $\widehat{m}$. It thus remains to establish (128). Since we have $p(0)=0$ and $q(0)=0$, it follows that (128) is trivially satisfied for $n=0$. It therefore remains to show (128) for $n=1,\ldots,\widehat{m}$. By Leibniz's rule for the differentiation of products of functions, we obtain

$$g^{(n)}(x)\Big|_{x=0}=\begin{cases}\binom{n}{i-1}(i-1)!, & i\leq n\\ 0, & i=n+1,\ldots,\widehat{m}\end{cases}$$

for $g(x)=(e^x-1)x^{i-1}$, which yields

$$p^{(n)}(x)\Big|_{x=0}=\sum_{i=1}^{n}\binom{n}{i-1}\frac{\widehat{m}-i}{\widehat{m}-1}a_i(i-1)!.$$

For the RHS of (128) we get $q^{(n)}(x)\big|_{x=0}=a_n n!$. Using (127), the condition (128) can, thus, be rewritten as

$$\sum_{i=1}^{n}\binom{n}{i-1}\binom{\widehat{m}-2}{i-1}\frac{1}{i}\geq\binom{\widehat{m}-1}{n-1}, \qquad n=1,\ldots,\widehat{m}. \tag{130}$$

Note that (130) is trivially satisfied for $n=1$. It, thus, remains to consider $n=2,\ldots,\widehat{m}$. The RHS of (130) can be written as

$$\binom{\widehat{m}-1}{n-1}=\binom{\widehat{m}-2}{n-2}+\binom{\widehat{m}-2}{n-1}. \tag{131}$$

The proof is concluded by showing that the sum of the two terms on the left-hand side (LHS) of (130) corresponding to $i=n$ and $i=n-1$ is greater than or equal to the RHS in (131). A direct comparison shows that this is the case if

$$\binom{n}{n-1}\frac{1}{n}\geq 1 \quad \text{and} \quad \binom{n}{n-2}\frac{1}{n-1}\geq 1$$

for $n=2,\ldots,\widehat{m}$. This is now verified by noting that $\binom{n}{n-1}/n=1$ and $\binom{n}{n-2}/(n-1)=n/2$.

Appendix F
Properties of $\rho_C(\epsilon)$

Using definition (66) with (50) and (48), we have

$$\rho_C(\epsilon)=\frac{\gamma_N^{-1}(1-\epsilon)}{\gamma_1^{-1}\left((1-\epsilon)^{1/N}\right)} \tag{132}$$

by noting that $\gamma_1(x)=1-e^{-x}$ (see Appendix C).

A. Limits of $\rho_C(\epsilon)$

1) Limit of $\rho_C(\epsilon)$ for $\epsilon\to 1$: We want to prove that $\lim_{\epsilon\to 1}\rho_C(\epsilon)=\sqrt[N]{N!}$. This is accomplished by using

$$\gamma_N^{-1}(x)=\sqrt[N]{N!}x^{1/N}(1+o(1))^{-1}, \quad x\to 0$$

and

$$\gamma_1^{-1}\left(x^{1/N}\right)=x^{1/N}(1+o(1))^{-1}, \qquad x\to 0$$

in (132).

2) Limit of $\rho_C(\epsilon)$ for $\epsilon\to 0$: We want to prove that $\lim_{\epsilon\to 0}\rho_C(\epsilon)=1$. Again, setting $x=1-\epsilon$, this amounts to showing that

$$\gamma_N^{-1}(x)\overset{a}{\sim}\gamma_1^{-1}\left(x^{1/N}\right), \qquad x\to 1. \tag{133}$$

Starting with the RHS in (133), we first note that

$$\gamma_1^{-1}\left(x^{1/N}\right)=\log\left(\frac{1}{1-x^{1/N}}\right).$$

Next, we have

$$\begin{aligned}x^{1/N}&=(1-(1-x))^{1/N}\\&=1-\frac{1}{N}(1-x)+\mathcal{O}((1-x)^2), \qquad x\to 1\end{aligned}$$

and, hence

$$\begin{aligned}\log\left(\frac{1}{1-x^{1/N}}\right)&=\log\left(\frac{1}{1-x}\right)+\log(N)+\mathcal{O}(1-x)\\&\overset{a}{\sim}\log\left(\frac{1}{1-x}\right), \qquad x\to 1\end{aligned}$$

establishing that

$$\gamma_1^{-1}\left(x^{1/N}\right)\overset{a}{\sim}\log\left(\frac{1}{1-x}\right), \qquad x\to 1. \tag{134}$$

For the LHS in (133), we first note that $\lim_{x\to\infty}\gamma_N(x)=1$, which implies that the $x\to 1$ asymptote of the inverse function $\gamma_N^{-1}(x)$ can be obtained by characterizing the $x\to\infty$ asymptote of $\gamma_N(x)$. It follows from (105) that

$$\gamma_N(x)=1-\frac{1}{(N-1)!}e^{-x}x^{N-1}(1+o(1)),\qquad x\to\infty$$

which yields

$$\begin{aligned}&\log\left((N-1)!(1-\gamma_N(x))\right)\\&\quad=-x+(N-1)\log(x)+o(1),\qquad x\to\infty\end{aligned}$$

and hence $\log\left((N-1)!(1-\gamma_N(x))\right)\overset{a}{\sim}-x$, $x\to\infty$. Now setting $x=\gamma_N^{-1}(y)$, we finally get

$$\gamma_N^{-1}(y)\overset{a}{\sim}-\log\left((N-1)!(1-y)\right)\overset{a}{\sim}\log\left(\frac{1}{1-y}\right),\qquad y\to 1.$$

Together with (134), this implies (133).

B. Monotonicity of $\rho_C(\epsilon)$

In the following, we show that $\rho_C(\epsilon)$ in (132) is a nondecreasing function of ϵ on the interval $[0,1]$. This will be accomplished by setting $1-\epsilon=\gamma_N(x)$, $x\in\mathbb{R}$, $x\geq 0$, and showing that the function $f(x)=x/g(x)$ with $g(x)=-\log\left(1-[\gamma_N(x)]^{\frac{1}{N}}\right)$ is nonincreasing in $x\geq 0$, or equivalently

$$f'(x)=\frac{g(x)-g'(x)x}{g^2(x)}\leq 0,\qquad \text{for } x\geq 0.$$

It thus remains to show that

$$g(x)-g'(x)x\leq 0,\qquad \text{for } x\geq 0. \tag{135}$$

Next, we note that $g(x)$ is convex for $x\geq 0$ if and only if the first-order convexity condition $g(x)+g'(x)(y-x)\leq g(y)$ holds for all $x,y\geq 0$ [46, Eq. (3.2)]. This first-order convexity condition evaluated at $y=0$ becomes (135) by noting that $g(0)=0$. Consequently, it is sufficient to show that $g(x)$ is a convex function for $x\geq 0$ or, equivalently, that $1-[\gamma_N(x)]^{\frac{1}{N}}$ is *log-concave* for $x\geq 0$. The function $1-[\gamma_N(x)]^{\frac{1}{N}}$ is a complementary cdf, which can be written as $1-[\gamma_N(x)]^{\frac{1}{N}}=\int_x^\infty\left([\gamma_N(t)]^{\frac{1}{N}}\right)'dt$, where $\left([\gamma_N(x)]^{\frac{1}{N}}\right)'$ denotes the corresponding pdf. Using the fact that log-concavity of a pdf implies that the corresponding complementary cdf is also log-concave [47, Theorem 3], it is sufficient to show that

$$\begin{aligned}\left([\gamma_N(x)]^{\frac{1}{N}}\right)'&=\frac{1}{N}[\gamma_N(x)]^{\frac{1}{N}-1}\gamma_N'(x)\\&=\frac{e^{-x}}{N\Gamma(N)}\left(\frac{x}{[\gamma_N(x)]^{\frac{1}{N}}}\right)^{N-1}\end{aligned} \tag{136}$$

is log-concave for $x\geq 0$. Here, we used $\gamma_N'(x)=e^{-x}x^{N-1}/\Gamma(N)$ (cf. (103)). The log-concavity (or log-convexity) of functions is preserved by the multiplication with exponentials, by the multiplication with positive constants, and by taking positive powers [46]. Therefore, (136) is log-concave if $x^Ne^{-x}/\gamma_N(x)$ is log-concave. Equivalently, (136) is log-concave for $x\geq 0$ if $h(x)=\gamma_N(x)x^{-N}e^x$ is log-convex for $x\geq 0$. Next, with the series expansion (104) for $\gamma_N(x)$, we obtain

$$h(x)=\sum_{i=0}^{\infty}\frac{x^i}{(i+N)!}.$$

Using the series representation of the confluent hypergeometric function

$$F(a,b,x)=\sum_{i=0}^{\infty}\frac{(a)_i}{(b)_i}\frac{x^i}{i!}$$

where $(\cdot)_i$ denotes the Pochhammer symbol, i.e., $(a)_i=a(a+1)\cdots(a+i-1)$, $(b)_i=b(b+1)\cdots(b+i-1)$ with $(a)_0=(b)_0=1$, we can write $h(x)=F(1,N+1,x)/(N!)$. With the integral representation of $F(a,b,x)$ [43], we finally get

$$h(x)=\frac{1}{\Gamma(N)}\int_0^1 e^{xt}(1-t)^{N-1}dt. \tag{137}$$

Applying the integration property of log-convex functions [46, p. 106], we can conclude that $h(x)$ is log-convex for $x\geq 0$ if the integrand in (137) is log-convex in x for each $t\in[0,1]$. The proof is concluded by noting that this is the case as the integrand, for each $t\in[0,1]$, is proportional to an exponential function (which is log-convex) for all t.

Appendix G
Calculation of $\mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}}\leq C_{\widetilde{\infty}}\right]$

In the following, we derive an analytic expression for $\mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}}\leq C_{\widetilde{\infty}}\right]$ under the assumption that b_{M-m+1} is purely real, purely imaginary, or equal to zero. The real and imaginary parts of $[\mathbf{z}(\mathbf{b})]_{M-m+1}$ are given by

$$\begin{aligned}[\mathbf{z}(\mathbf{b})]_{\mathrm{R},M-m+1}&=R_{M-m+1,M-m+1}b_{\mathrm{R},M-m+1}+u_{\mathrm{R},m}\\[\mathbf{z}(\mathbf{b})]_{\mathrm{I},M-m+1}&=R_{M-m+1,M-m+1}b_{\mathrm{I},M-m+1}+u_{\mathrm{I},m}.\end{aligned}$$

Here, $u_m\sim\mathcal{CN}(0,\sigma_m^2)$ is specified in (88) (σ_m^2 is specified in (90)) and $R_{M-m+1,M-m+1}\in\mathbb{R}$. Under the assumption that b_{M-m+1} is purely real, purely imaginary, or equal to zero, $[\mathbf{z}(\mathbf{b})]_{\mathrm{R},M-m+1}$ and $[\mathbf{z}(\mathbf{b})]_{\mathrm{I},M-m+1}$ are statistically independent, which yields

$$\begin{aligned}&\mathrm{P}\left[\|[\mathbf{z}(\mathbf{b})]_{M-m+1}\|_{\widetilde{\infty}}\leq C_{\widetilde{\infty}}\right]\\&\quad=\mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{\mathrm{R},M-m+1}|\leq C_{\widetilde{\infty}}\right]\\&\qquad\times\mathrm{P}\left[|[\mathbf{z}(\mathbf{b})]_{\mathrm{I},M-m+1}|\leq C_{\widetilde{\infty}}\right].\end{aligned} \tag{138}$$

Let us first assume that b_{M-m+1} is purely real, i.e., $b_{M-m+1}=b_{\mathrm{R},M-m+1}\neq 0$. We can write $|[\mathbf{z}(\mathbf{b})]_{\mathrm{R},M-m+1}|\overset{d}{=}|v_m+u_{\mathrm{R},m}|$ and $|[\mathbf{z}(\mathbf{b})]_{\mathrm{I},M-m+1}|=|u_{\mathrm{I},m}|$, where $v_m=R_{M-m+1,M-m+1}|b_{M-m+1}|$ is a scaled $\chi_{2(m+L)}$-distributed RV with pdf (89) and $u_{\mathrm{R},m}$ and $u_{\mathrm{I},m}$ are i.i.d.

$\mathcal{N}(0, \sigma_m^2/2)$. The RV $\frac{\sqrt{2}}{\sigma_m}|u_{\mathrm{I},m}|$ is thus χ_1-distributed, which gives

$$\mathrm{P}\left[|u_{\mathrm{I},m}| \leq C_{\widetilde{\infty}}\right] = \gamma_{\frac{1}{2}}\left(\frac{C_{\widetilde{\infty}}^2}{\sigma_m^2}\right). \tag{139}$$

For given $v_m = v$, the RV $\frac{2}{\sigma_m^2}|v + u_{\mathrm{R},m}|^2$ is noncentral χ_1^2-distributed with noncentrality parameter $\frac{2v^2}{\sigma_m^2}$. Thus, following the steps (92)–(93), we obtain

$$\mathrm{P}\left[|v_m + u_{\mathrm{R},m}| \leq C_{\widetilde{\infty}}\right] = \sum_{s=0}^{\infty} D_s(\mathbf{b}_m)\gamma_{s+\frac{1}{2}}\left(\frac{C_{\widetilde{\infty}}^2}{\sigma_m^2}\right) \tag{140}$$

where $D_s(\mathbf{b}_m)$ was defined in (94). Note that the only difference between (140) and (93) is the occurrence of the factor $1/2$ instead of the factor 1 in the index of the incomplete Gamma function. As a result, however, it seems that (140) cannot be expressed as a finite sum as was done for (93) to arrive at (37). The final expression (83) now follows by combining (138)–(140). The cases $b_{M-m+1} = b_{\mathrm{I},M-m+1} \neq 0$ and $b_{M-m+1} = 0$ can be seen to result in (83) by following the steps (138)–(140) suitably modified.

ACKNOWLEDGMENT

The authors would like to thank G. Matz for suggesting the direct integration approach to derive (44) and for pointing out [36], M. Borgmann for valuable discussions on the diversity order of SD-l^∞, A. Burg for helpful discussions on VLSI implementation aspects of SD-l^∞, and S. Gerhold for pointing out the proof in Appendix F-A.

REFERENCES

[1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[2] A. Burg, *VLSI Circuits for MIMO Communication Systems*. Konstanz: Hartung-Gorre Verlag, 2006.

[3] 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, Apr. 1985.

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

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

[6] 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.

[7] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[8] 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. Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, Jul. 2005.

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

[10] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm I. Expected complexity," *IEEE Trans. Signal Process.*, vol. 53, pp. 2806–2818, Aug. 2005.

[11] H. Kaeslin, *Digital Integrated Circuit Design: From VLSI Architectures to CMOS Fabrication*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[12] F. Boccardi and G. Caire, "The p-sphere encoder: Peak-power reduction by lattice precoding for the MIMO Gaussian broadcast channel," *IEEE Trans. Commun.*, vol. 54, no. 11, pp. 2085–2091, Nov. 2006.

[13] H. Vikalo and B. Hassibi, "On the sphere decoding algorithm II. Generalizations, second-order statistics, and applications to communications," *IEEE Trans. Signal Process.*, vol. 53, pp. 2819–2834, Aug. 2005.

[14] A. D. Murugan, H. El Gamal, M. O. Damen, and G. Caire, "A unified framework for tree search decoding: Rediscovering the sequential decoder," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 933–953, Mar. 2006.

[15] R. Gowaikar and B. Hassibi, "Statistical pruning for near-maximum likelihood decoding," *IEEE Trans. Signal Process.*, vol. 55, pp. 2661–2675, Jun. 2007.

[16] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, pp. 1474–1484, Apr. 2005.

[17] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[18] D. E. Knuth, "Big omicron and big omega and big theta," *Association for Computing Machinery SIGACT News*, vol. 8, no. 2, pp. 18–24, 1976.

[19] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[20] 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.

[21] C. P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Math. Programming*, vol. 66, no. 2, pp. 181–191, Sep. 1994.

[22] K. Su and I. J. Wassell, "A new ordering for efficient sphere decoding," in *Proc. IEEE ICC 2005*, Seoul, Korea, May 2005, pp. 1906–1910.

[23] L. Babai, "On Lovász' lattice reduction and the nearest lattice point problem," *Combinatorica*, vol. 6, pp. 1–13, 1986.

[24] C. Windpassinger and R. F. H. Fischer, "Low-complexity near-maximum-likelihood detection and precoding for MIMO systems using lattice reduction," in *Proc. IEEE Inf. Theory Workshop*, Paris, France, Mar./Apr. 2003, pp. 345–348.

[25] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[26] 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.

[27] H. Lu, Y. Wang, P. Kumar, and K. Chugg, "Remarks on space-time codes including a new lower bound and an improved code," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2752–2757, Oct. 2003.

[28] S. Loyka and F. Gagnon, "Performance analysis of the V-BLAST algorithm: An analytical approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1326–1337, Jul. 2004.

[29] M. Taherzadeh, A. Mobasher, and A. Khandani, "LLL lattice-basis reduction achieves the maximum diversity in MIMO systems," in *Proc. IEEE ISIT 2005*, Adelaide, Australia, Sep. 2005, pp. 1300–1304.

[30] C. Ling, "Towards characterizing the performance of approximate lattice decoding in MIMO communications," in *Proc. ITG Conf. Source and Channel Coding*, Munich, Germany, Apr. 2006.

[31] C. Ling, "Approximate lattice decoding: Primal versus dual basis reduction," in *Proc. IEEE ISIT 2006*, Seattle, WA, Jul. 2006, pp. 1–5.

[32] A. Mobasher and A. K. Khandani, "On the limitations of the naive lattice decoding," in *Proc. IEEE ISIT 2007*, Nice, France, Jun. 2007, pp. 201–204.

[33] C. Windpassinger, L. H.-J. Lampe, and R. F. H. Fischer, "From lattice-reduction-aided detection towards maximum-likelihood detection in MIMO systems," in *Proc. IASTED Int. Conf. Wireless and Opt. Commun.*, Banff, Canada, Jul. 2003, pp. 144–148.

[34] J. Jaldén and G. Matz, "MIMO receiver diversity in general fading," in *Proc. IEEE ICASSP 2008*, Las Vegas, NV, Mar./Apr. 2008, pp. 2837–2840.

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

[36] J. Behboodian, "On the distribution of a symmetric statistic from a mixed population," *Technometrics*, vol. 14, no. 4, pp. 919–923, Nov. 1972.

[37] M. O. Damen, A. Chkeif, and J. C. Belfiore, "Lattice code decoder for space-time codes," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 161–163, May 2000.
[38] J. Jaldén and B. Ottersten, "On the limits of sphere decoding," in *Proc. IEEE ISIT 2005*, Adelaide, Australia, Sep. 2005, pp. 1691–1695.
[39] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed. Hoboken, NJ: Wiley, 2003.
[40] K. Ball, "An elementary introduction to modern convex geometry," in *MSRI Book Series, Flavors of Geometry*. Cambridge, U.K.: Cambridge Univ. Press, 1997, vol. 31, pp. 1–58.
[41] Y. H. Gan and W. H. Mow, "Complex lattice reduction algorithms for low-complexity MIMO detection," in *Proc. IEEE Globecom 2005*, St. Louis, MO, Nov. 2005, vol. 5, pp. 2953–2957.
[42] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. Hoboken, NJ: Wiley, 2005.
[43] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1965.
[44] W. Gautschi, "The incomplete gamma functions since Tricomi," in *Tricomi's Ideas and Contemporary Applied Mathematics*. Rome, Italy: Atti dei Convegni Lincei, Accademia Nationale dei Lincei, 1998, pp. 203–237.
[45] M. Merkle, "Some inequalities for the chi square distribution function and the exponential function," *Arch. Math.*, vol. 60, no. 5, pp. 451–458, 1993.
[46] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[47] M. Bagnoli and T. Bergstrom, "Log-concave probability and its applications," *Econom. Theory*, vol. 26, no. 2, pp. 445–469, Aug. 2005.

Dominik Seethaler received the Dipl.-Ing. and Dr. techn. degrees in electrical/communication engineering from Vienna University of Technology, Vienna, Austria, in 2002 and 2006, respectively.

From 2002 to 2007, he was a Research and Teaching Assistant with the Institute of Communications and Radio Frequency Engineering, Vienna University of Technology. From September 2007 to August 2009, he was a Postdoctoral Researcher with the Communication Technology Laboratory, ETH Zurich, Switzerland. Since September 2009, he has been freelancing in the area of home robotics.

Helmut Bölcskei (M'98–SM'02–F'09) was born in Mödling, Austria, on May 29, 1970. He received the Ph.D. degree in electrical engineering from Vienna University of Technology, Vienna, Austria, in 1997.

From 1999 to 2001, he was a Postdoctoral Researcher with the Information Systems Laboratory, Department of Electrical Engineering, and with the Department of Statistics, Stanford University, Stanford, CA. He was on the founding team of Iospan Wireless Inc., a Silicon Valley-based startup company (acquired by Intel Corporation in 2002) specialized in multiple-input multiple-output (MIMO) wireless systems for high-speed Internet access. From 2001 to 2002, he was an Assistant Professor of Electrical Engineering with the University of Illinois at Urbana-Champaign. He has been with ETH Zurich, Switzerland, since 2002, where he is a Professor of Electrical Engineering. He was a visiting researcher with Philips Research Laboratories Eindhoven, The Netherlands, ENST Paris, France, and the Heinrich Hertz Institute Berlin, Germany. His research interests are in information theory, harmonic analysis, and signal processing.

Dr. Bölcskei received the 2001 IEEE Signal Processing Society Young Author Best Paper Award, the 2006 IEEE Communications Society Leonard G. Abraham Best Paper Award, the ETH "Golden Owl" Teaching Award, and was an Erwin Schrödinger Fellow of the Austrian National Science Foundation. He was a plenary speaker at several IEEE conferences and served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the EURASIP *Journal on Applied Signal Processing*. He is currently on the editorial board of "Foundations and Trends in Networking" and serves as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY. He was TPC Co-Chair of the 2008 IEEE International Symposium on Information Theory. He also serves on the Board of Governors of the IEEE Information Theory Society.