# On Low-complexity Lattice Reduction Algorithms for Large-scale MIMO Detection: the Blessing of Sequential Reduction 

Shanxiang Lyu, Jinming Wen, Jian Weng and Cong Ling


#### Abstract

Lattice reduction is a popular preprocessing strategy in multiple-input multiple-output (MIMO) detection. In a quest for developing a low-complexity reduction algorithm for largescale problems, this paper investigates a new framework called sequential reduction ( SR ), which aims to reduce the lengths of all basis vectors. The performance upper bounds of the strongest reduction in SR are given when the lattice dimension is no larger than 4 . The proposed new framework enables the implementation of a hash-based low-complexity lattice reduction algorithm, which becomes especially tempting when applied to large-scale MIMO detection. Simulation results show that, compared to other reduction algorithms, the hash-based SR algorithm exhibits the lowest complexity while maintaining comparable error performance.


Index Terms-lattice reduction, MIMO, large-scale, hashbased.

## I. Introduction

The number of antennas has been scaled up to tens or hundreds in multiple-input multiple-output (MIMO) systems to fulfill the performance requirements needed by the next generation communication systems [1]. A critical challenge that comes with very large arrays is to design reliable and computationally efficient detectors. Though the well-known maximum likelihood detector (MLD) provides optimal error performance, it suffers from exponential complexity that grows with the number of transmit antennas [2]. In the past two decades, lattice-reduction-aided suboptimal detection techniques have been well investigated [3]-[5], whose instantaneous complexity does not depend on constellation size and noise realizations, but collect the same diversity as the MLD for MIMO systems [6]-[8]. Although conventional lattice reduction algorithms suffice for small-scale MIMO systems, there is still an avenue to pursue a more practical

[^0]low-complexity reduction algorithm for large-scale systems. Moreover, an efficient reduction algorithm for large-scale problems may also find its applications to cryptanalysis [9] and image processing [10].

The principle of designing a reduction algorithm varies depending on the desired basis properties: to make all the basis vectors short, or to make the condition number of the reduced basis small. There are several popular types of lattice reduction strategies, such as Minkowski reduction, KorkineZolotareff reduction (KZ) [11], Gauss reduction [12], Lenstra-Lenstra-Lovász (LLL) reduction [13], Seysen reduction [14], etc. They yield reduced bases with shorter or more orthogonal basis vectors, and provide a trade-off between the quality of the reduced basis and the computational effort required for finding it. In essence, a reduction algorithm aims to find a unimodular matrix to transform an input basis into another one with better property. The process involves a series of elementary operations noted as reflection, swapping, and translation. These operations vary for distinct algorithms.

Much work has been done to advance conventional reduction algorithms. Regarding KZ, refs. [15]-[17] give some practical implementations and improve the performance bounds. As for blockwise KZ, its faster implementations and the expected basis properties are given in [18]. Researchers have also been constructing and analyzing the variants of LLL with great effort. For instance, the size reduction step is optimized in [16], [19], [20], the implementation order of swaps is simplified in [21]-[23], and the fixed complexity versions of LLL are given in [19], [22], [24]. In contrast, the direction on Seysen reduction has few follow-up studies [25], [26], partly because of the fact that Seysen reduction has unsatisfactory performance in high dimensions.

While LLL and blockwise KZ are still the default choices in cryptography to reduce a basis in hundreds of dimensions, the element-based lattice reduction (ELR) proposed in [27] has become more attracting in large MIMO, which preprocesses a large basis with even lower complexity than LLL. Later ref. [27] has been generalized to $\mathrm{ELR}^{+}$[28] for small-scale problems, but the theoretical characterization of ELR and $\mathrm{ELR}^{+}$has not been given a rigorous treatment, even for small dimensions. It is noteworthy that ELR and ELR ${ }^{+}$have totally different structures with LLL variants, and one might be lured into the belief that ELR and $\mathrm{ELR}^{+}$can be tuned to arrive at more sophisticated methods. Nevertheless, no analytical skills can be inherited from LLL/KZ literature [12], [13], which makes the performance analysis of the new algorithms
complicated.
In this work, we investigate a general form of ELR and $\mathrm{ELR}^{+}$which we refer to as sequential reduction (SR). We derive the objective function from a MIMO detection task, and present the general form of an SR algorithm which can solve/approximate the smallest basis problem. Unlike KZ or Minkowski reduction, SR reduces the basis vectors by using sub-lattices so as to avoid a basis expansion process. The strongest algorithm in SR tries to minimize the length of basis vectors with the aid of a closest vector problem (CVP) oracle. We show bounds on the basis lengths and orthogonal defects for small dimensions. After that, the feasibility of applying SR to reduce a large dimensional basis is analyzed, and we actually construct a hash-based algorithm for this task. Our simulation results then show the plausibility of using SR in large-scale MIMO systems.

Preliminary results of this work have been partly presented in a conference paper [29]. Compared with [29], this work contains the following new contributions:

- The performance bounds on small dimensional bases are rigorously analyzed (Theorems 2 and 3). Unlike the results in [29] that rely on an assumption about covering radius, these bounds hold for all input bases.
- Comparisons with other types of strong\&weak reduction are made (Section III-B, Section IV-D), including $\eta$ Greedy reduction, KZ and its variants, Minkowski reduction, LLL and its variants, and Seysen reduction.
- A Hash-based SR algorithm is constructed (Section IV). More specifically, the nearest neighbor search problem is approximately solved with the aid of hashing, and not through a brute-force search.
- The theoretical studies are supported with more simulation results (Section V), these include the comparisons with major lattice-reduction-aided MIMO detection algorithms, and the BER performance tested for various channels.
- The types of bases feasible for using SR-Hash is discussed (Appendix A). We numerically show that the dual of large-scale Gaussian random bases have dense pairwise angles.
It is worth mentioning that SR is emerging as a new building block in lattice-reduction-aided MIMO detection. Thus, the proposed SR variants may also benefit list sphere decoding [30] and Klein's sampling algorithm [31].

The rest of this paper is organized as follows. Backgrounds about lattices and lattice reduction in MIMO are reviewed in Section II. The SR framework is subsequently introduced in Section III. The low-complexity version of SR based on hashing is given in Section IV. After that, Section V presents the simulation results. Conclusions and possible future research are presented in the last section.

Notation: Matrices and column vectors are denoted by uppercase and lowercase boldface letters. The $i$ th column and $(j, i)$ th entry of $\mathbf{B}$ are respectively denoted as $\mathbf{b}_{i}$ and $b_{i, j} . \mathbf{I}_{n}$ and $\mathbf{0}_{n}$ respectively denote the $n \times n$ identity matrix and $n \times 1$ zero vector, and the operation $(\cdot)^{\top}$ denotes matrix transposition. $[n]$ denotes the set $\{1, \ldots, n\}$. For a set $\Gamma, \mathbf{B}_{\Gamma}$ denotes the columns of $\mathbf{B}$ indexed by $\Gamma . \operatorname{span}\left(\mathbf{B}_{\Gamma}\right)$ denotes the
vector space spanned by vectors in $\mathbf{B}_{\Gamma} \cdot \pi_{\mathbf{B}_{\Gamma}}(\mathbf{x})$ and $\pi_{\mathbf{B}_{\Gamma}}^{\perp}(\mathbf{x})$ denote the projection of $\mathbf{x}$ onto $\operatorname{span}\left(\mathbf{B}_{\Gamma}\right)$ and the orthogonal complement of $\operatorname{span}\left(\mathbf{B}_{\Gamma}\right)$, respectively. $\lfloor x\rceil$ denotes rounding $x$ to the nearest integer, $|x|$ denotes getting the absolute value of $x$, and $\|\mathbf{x}\|$ denote the Euclidean norm of vector $\mathbf{x} . \mathbb{N}$ and $\mathbb{Z}$ respectively denotes the set of natural numbers and integers. The set of $n \times n$ integer matrices with determinants $\pm 1$ is denoted by $\mathrm{GL}_{n}(\mathbb{Z})$.

## II. Preliminaries

## A. Lattices

An $n$-dimensional lattice $\Lambda$ is a discrete additive subgroup in the real field $\mathbb{R}^{n}$. Similarly to the fact that any finitedimensional vector space has a basis, a lattice has a basis. To consider a square matrix for simplicity, a lattice generated by basis $\mathbf{B}=\left[\mathbf{b}_{1}, \ldots, \mathbf{b}_{n}\right] \in \mathbb{R}^{n \times n}$ is defined as

$$
\Lambda(\mathbf{B})=\left\{\mathbf{v} \mid \mathbf{v}=\sum_{i \in[n]} c_{i} \mathbf{b}_{i} ; c_{i} \in \mathbb{Z}\right\}
$$

The dual lattice of $\Lambda$ is defined as $\Lambda^{\dagger}=$ $\left\{\mathbf{u} \in \mathbb{R}^{n} \mid\langle\mathbf{u}, \mathbf{v}\rangle \in \mathbb{Z}, \forall \mathbf{v} \in \Lambda\right\}$. One basis of $\Lambda^{\dagger}$ is given by $\mathbf{B}^{-\top}$.

The Gram-Schmidt (GS) basis of $\mathbf{B}$, referred to as $\mathbf{B}^{*}$, is found by using $\mathbf{b}_{i}^{*}=\pi_{\left\{\mathbf{b}_{1}^{*}, \ldots, \mathbf{b}_{i-1}^{*}\right\}}\left(\mathbf{b}_{i}\right)=\mathbf{b}_{i}-\sum_{j=1}^{i-1} \mu_{i, j} \mathbf{b}_{j}^{*}$, where $\mu_{i, j}=\left\langle\mathbf{b}_{i}, \mathbf{b}_{j}^{*}\right\rangle /\left\|\mathbf{b}_{j}^{*}\right\|^{2}$.

The $i$ th successive minimum of an $n$ dimensional lattice $\Lambda(\mathbf{B})$ is the smallest real positive number $r$ such that $\Lambda$ contains $i$ linearly independent vectors of length at most $r$ :

$$
\lambda_{i}(\mathbf{B})=\inf \{r \mid \operatorname{dim}(\operatorname{span}((\Lambda \cap \mathcal{B}(\mathbf{0}, r))) \geq i\}
$$

in which $\mathcal{B}(\mathbf{t}, r)$ denotes a ball centered at $\mathbf{t}$ with radius $r$.
The orthogonality defect (OD) can alternatively quantify the goodness of a basis

$$
\begin{equation*}
\eta(\mathbf{B})=\frac{\prod_{i=1}^{n}| | \mathbf{b}_{i}| |}{\sqrt{\left|\operatorname{det}\left(\mathbf{B}^{T} \mathbf{B}\right)\right|}} \tag{1}
\end{equation*}
$$

From Hadamard's inequality, we know that $\eta(\mathbf{B}) \geq 1$. As the determinant of a given basis is fixed, the parameter is proportional to the product of the lengths of the basis vectors. A necessary condition for reaching the smallest orthogonality defect is to have a short basis length defined as $l(\mathbf{B})=$ $\max _{i}\left\|\mathbf{b}_{i}\right\|$.

The $\varepsilon$ CVP problem is, given a vector $\mathbf{y} \in \mathbb{R}^{n}$ and a lattice $\Lambda(\mathbf{B})$, find a vector $\mathbf{v} \in \Lambda(\mathbf{B})$ such that:

$$
\|\mathbf{y}-\mathbf{v}\|^{2} \leq \varepsilon\|\mathbf{y}-\mathbf{w}\|^{2}, \forall \mathbf{w} \in \Lambda(\mathbf{B}) .
$$

An algorithm that solves an $\varepsilon$ CVP problem is referred to as an $\varepsilon \operatorname{CVP}$ oracle. We write $\mathbf{v}=\varepsilon \operatorname{CVP}(\mathbf{y}, \mathbf{B})$ or $\mathbf{v}=\operatorname{CVP}(\mathbf{y}, \mathbf{B})$ if $\varepsilon=1$.

## B. Lattice-reduction-aided MIMO detection

We considered an uplink multiuser large MIMO system, in which $n_{T}$ single-antenna users send data to a base station with $n_{R}$ antennas, and both $n_{T}, n_{R}$ are in the order of tens
or hundreds. A received complex-valued signal vector at the base station is written as:

$$
\begin{equation*}
\mathbf{y}_{c}=\mathbf{B}_{c} \mathbf{x}_{c}+\mathbf{w}_{c}, \tag{2}
\end{equation*}
$$

where $\mathbf{B}_{c} \in \mathbb{C}^{n_{R} \times n_{T}}$ denotes a channel matrix perfectly known at the base station, $\mathbf{x}_{c} \in \mathbb{C}^{n_{T}}$ refers to a signal vector with entries drawn from a QAM constellation, and $\mathbf{w}_{c} \in \mathbb{C}^{n_{R}}$ denotes a zero-mean additive noise vector with entries independently and identically following the complex normal distribution $\mathcal{C N}\left(0, \sigma_{w}^{2}\right)$.

To simplify the analysis we will focus on representations in the real field, so (2) is transformed to an equivalent real value system with

$$
\begin{equation*}
\mathbf{y}=\overline{\mathbf{B}} \mathbf{x}+\mathbf{w} \tag{3}
\end{equation*}
$$

where

$$
\overline{\mathbf{B}}=\left[\begin{array}{cc}
\Re\left(\mathbf{B}_{c}\right) & -\Im\left(\mathbf{B}_{c}\right)  \tag{4}\\
\Im\left(\mathbf{B}_{c}\right) & \Re\left(\mathbf{B}_{c}\right)
\end{array}\right]
$$

and $\mathbf{y}=\left[\Re(\mathbf{y})^{\top}, \Im(\mathbf{y})^{\top}\right]^{\top}, \mathbf{x}=\left[\Re(\mathbf{x})^{\top}, \Im(\mathbf{x})^{\top}\right]^{\top}, \mathbf{w}=$ $\left[\Re(\mathbf{w})^{\top}, \Im(\mathbf{w})^{\top}\right]^{\top}$ are all real and imaginary compositions. Here the noise variance of $\mathbf{w}$ becomes $\sigma^{2}=\sigma_{w}^{2} / 2$.

Lattice reduction is essentially multiplying a given basis with a unimodular matrix $\mathbf{U} \in \mathrm{GL}_{n}(\mathbb{Z})$ to get a reduced basis $\tilde{\mathbf{B}} \triangleq \overline{\mathbf{B}} \mathbf{U}$. For a lattice-reduction-aided detector, we first rewrite Eq. (3) as:

$$
\mathbf{y}=\tilde{\mathbf{B}}\left(\mathbf{U}^{-1} \mathbf{x}\right)+\mathbf{w}
$$

To make the unimodular transform compact for the QAM constellation, we need to scale and shift signal vector $\mathbf{x}$ to get $\mathbf{x} \leftarrow\left(\mathbf{x}+\mathbf{1}_{n \times 1}\right) / 2$, so that the constraint on $\mathbf{x}$ become a consecutive integer set $\Xi^{n}$. Let $\mathbf{y} \leftarrow\left(\mathbf{y}+\tilde{\mathbf{B}} \mathbf{U}^{-1} \mathbf{1}_{n \times 1}\right) / 2$, then the inferred signal vector is given by:

$$
\begin{equation*}
\hat{\mathbf{x}}=2 \mathcal{Q}_{\Xi^{n}}\left(\mathbf{U} \mathcal{Q}_{\mathbb{Z}^{n}}(\mathcal{E}(\mathbf{y}, \tilde{\mathbf{B}}))\right)-\mathbf{1}_{n \times 1} \tag{5}
\end{equation*}
$$

where $\mathcal{E}(\mathbf{y}, \tilde{\mathbf{B}})$ denotes a low-complexity detector that could be zero-forcing (ZF) or successive-interference-cancellation (SIC), and $Q(\cdot)$ denotes a quantization function with respect to its subscript. Given certain information about the signal vector, the detectors can be implemented under an minimum-mean-square-error (MMSE) principle. The MMSE-based ZF/SIC detectors are similarly given by extending the size of the system: $\mathbf{y} \leftarrow\left[\mathbf{y}^{\top}, \mathbf{0}_{1 \times n}\right]^{\top}, \overline{\mathbf{B}} \leftarrow\left[\overline{\mathbf{B}}^{\top}, \sigma / \sigma_{s} \mathbf{I}_{n}\right]^{\top}$, with $\sigma_{s}^{2}$ referring to the variance of a signal symbol.

## C. The objective in lattice reduction

Hereby we explain the design criteria of lattice reduction used in MIMO detection. For a set of linearly independent vectors $\overline{\mathbf{B}}=\left[\overline{\mathbf{b}}_{1}, \ldots, \overline{\mathbf{b}}_{n}\right]$, we define its fundamental parallelepiped as

$$
\mathcal{P}(\overline{\mathbf{B}})=\left\{\sum_{i=1}^{n} c_{i} \overline{\mathbf{b}}_{i} \mid-1 / 2 \leq c_{i} \leq 1 / 2\right\}
$$

Choosing $\mathcal{E}(\mathbf{y}, \tilde{\mathbf{B}})$ as the SIC [32] detector, then the pairwise error probability $P_{e}$ based on (5) becomes

$$
\begin{align*}
P_{e} & =1-\operatorname{Pr}\left(\mathbf{w} \in \mathcal{P}\left(\overline{\mathbf{B}}^{*}\right)\right) \\
& =1-\prod_{i=1}^{n} \operatorname{Pr}\left(\left|\mathbf{w}^{\top} \overline{\mathbf{b}}_{i}^{*}\right|<\left\|\overline{\mathbf{b}}_{i}^{*}\right\|^{2} / 2\right) \\
& =1-\prod_{i=1}^{n} \operatorname{erf}\left(\frac{\left\|\overline{\mathbf{b}}_{i}^{*}\right\|}{2 \sqrt{2} \sigma}\right) \\
& \leq 1-\prod_{i=1}^{n} \operatorname{erf}\left(\frac{1}{2 \sqrt{2} \sigma\left\|\mathbf{d}_{i}\right\|}\right) \tag{6}
\end{align*}
$$

where the last inequality comes from $\left\|\overline{\mathbf{b}}_{i}^{*}\right\|=$ $\left\|\pi_{\overline{\mathbf{b}}_{1}, \ldots, \overline{\mathbf{b}}_{i-1}}\left(\overline{\mathbf{b}}_{i}\right)\right\| \geq\left\|\pi_{\overline{\mathbf{b}}_{1}, \ldots, \overline{\mathbf{b}}_{i-1}, \overline{\mathbf{b}}_{i}, \ldots, \overline{\mathbf{b}}_{n}}\left(\overline{\mathbf{b}}_{i}\right)\right\|=1 /\left\|\mathbf{d}_{i}\right\|$, and $\mathbf{d}_{i}$ is the $i$ th vector in the dual basis of $\Lambda^{\dagger}$.
From (6), it becomes clear that the upper bound on $P_{e}$ is mainly controlled by the lengths of vectors in the dual basis, i.e., $\left\|\mathbf{d}_{1}\right\|, \ldots,\left\|\mathbf{d}_{n}\right\|$. Based on this observation, we can solve/approximate the following problem in the dual lattice to attain better error rate performance for the above lattice-reduction-aided SIC detector.

Definition 1 (SBP). The smallest basis problem (SBP) is, given a lattice $\Lambda$, find the basis with the smallest orthogonality defect.

To address SBP, a designed reduction algorithm should make all basis vectors as short as possible. Moreover, since the basis dimension is in the order of tens or hundreds in large MIMO, we need a low-complexity lattice reduction algorithm that reduces the basis with satisfactory performance.

## III. SEQUENTIAL REDUCTION FRAMEWORK

The fundamental principle of sequential reduction is to reduce a basis vector by using all other vectors that span a sublattice. In the new method, given an input basis $\mathbf{B}^{1}$, we sequentially solve $\mathbf{s}_{i}=\varepsilon \operatorname{CVP}\left(\mathbf{b}_{i}, \mathbf{B}_{[n] \backslash i}\right)$ with $[n] \backslash i=$ $\{1, \ldots, n\} \backslash i$. For each $\mathbf{s}_{i}$, we test whether the residue distance is shorter: $\left\|\mathbf{b}_{i}-\mathbf{s}_{i}\right\|^{2}<\tau\left\|\mathbf{b}_{i}\right\|^{2}$, where $\tau \in(0,1]^{2}$ is a parameter to control the complexity. If this holds, we update $\mathbf{b}_{i}$ by $\mathbf{b}_{i} \leftarrow \mathbf{b}_{i}-\mathbf{s}_{i}$. Here both $\mathbf{s}_{i}=\mathbf{0}$ and the $\mathbf{s}_{i}$ that makes $\left\|\mathbf{b}_{i}-\mathbf{s}_{i}\right\|=\left\|\mathbf{b}_{i}\right\|$ are declared as ineffective attempts. A threshold parameter $m$ is set to count these useless trials. The algorithm terminates if $m>n$, which means no more vectors can be further reduced. The general form of sequential reduction is summarized in Algorithm 1.
An SR algorithm maintains a lattice basis due to the following reason. In round $m$, suppose $\sum_{k \in[n] \backslash i} c_{k} \mathbf{b}_{k}$ is a valid reduction on $\mathbf{b}_{i}$, then the lattice basis updating process becomes $\mathbf{B} \leftarrow \mathbf{B T}^{m}$, with $T_{k, k}^{m}=1 \forall k \in[n], T_{k, i}^{m}=$ $-c_{k} \forall k \in[n] \backslash i$ and all other entries are zeros. Since $\mathbf{T}^{m}$ is an integer matrix with determinant $1, \mathbf{T}^{m}$ is unimodular, and the composition of the transform matrices from different rounds maintains a unimodular matrix.
If an exact CVP oracle is chosen in line 4 of Algorithm 1, we call the algorithm SR-CVP. By choosing other approximate

[^1]```
Algorithm 1: The general form of an SR algorithm.
    Input: lattice basis \(\mathbf{B}=\left[\mathbf{b}_{1}, \ldots, \mathbf{b}_{n}\right]\), complexity
            threshold \(\tau\);
    Output: reduced lattice basis B.
    \(i=0, m=1\);
    while \(m \leq n\) do
        \(i \leftarrow(i \bmod n)+1 ; \quad \triangleright\) The column index;
        \(\mathbf{s}_{i}=\varepsilon \operatorname{CVP}\left(\mathbf{b}_{i}, \mathbf{B}_{[n] \backslash i}\right) ; \quad\) Exact/approximated CVP
        solvers;
        if \(\left\|\mathbf{b}_{i}-\mathbf{s}_{i}\right\|^{2}<\tau\left\|\mathbf{b}_{i}\right\|^{2}\) then
            \(\mathbf{b}_{i} \leftarrow \mathbf{b}_{i}-\mathbf{s}_{i} ;\)
            \(m=1\);
        else
            \(m \leftarrow m+1 ;\)
```

CVP solvers, we can obtain other variants that have lower complexity. As shown in Fig. 1, SR encompasses SR-CVP, SRPair and SR-Hash. The red box in the figure denotes SR, whose SR-CVP, SR-Pair and SR-Hash algorithms feature decreasing complexity. Their analogies in the conventional KZ framework are shown in the black box.


Fig. 1: SR contains algorithms with different performancecomplexity trade-offs.

## A. Basis properties of $S R-C V P$

We need to understand the performance limits of SR by first analyzing SR-CVP. Hereby we set $\tau=1$ in the analysis for brevity. When no more attempts using $\mathbf{s}_{i}=\operatorname{CVP}\left(\mathbf{b}_{i}, \mathbf{B}_{[n] \backslash i}\right)$ can further reduce the basis in the algorithm, for all $\mathbf{s}_{i} \in$ $\Lambda\left(\mathbf{B}_{[n] \backslash i}\right)$ we have

$$
\begin{align*}
\underbrace{\left\|\mathbf{b}_{i}\right\|^{2}}_{\text {term } 1} & \leq\left\|\mathbf{b}_{i}-\mathbf{s}_{i}\right\|^{2} \\
& =\left\|\pi_{\mathbf{B}_{[n] \backslash i}}^{\perp}\left(\mathbf{b}_{i}\right)+\pi_{\mathbf{B}_{[n] \backslash i}}\left(\mathbf{b}_{i}\right)-\mathbf{s}_{i}\right\|^{2} \\
& =\underbrace{\left\|\pi_{\mathbf{B}_{[n] \backslash i}}^{\perp}\left(\mathbf{b}_{i}\right)\right\|^{2}}_{\text {term } 2}+\underbrace{\left\|\pi_{\mathbf{B}_{[n] \backslash i}}\left(\mathbf{b}_{i}\right)-\mathbf{s}_{i}\right\|^{2}}_{\text {term } 3}, \tag{7}
\end{align*}
$$

where the second equality is due to Pythagoras' theorem. Note that SR-CVP provides the tightest constraint on term 3, and approximations of CVP oracles also distinguish themselves on the same term. Based on (7), we can prove the following
theorem that consists of upper bounds for the basis length and the orthogonality defect.

Theorem 2. For any dimension $n \leq 4$, an $S R-C V P$ reduced basis satisfies:

$$
\begin{align*}
l(\mathbf{B}) & \leq \sqrt{\frac{4 n}{5-n}} \lambda_{n}(\mathbf{B})  \tag{8}\\
\eta(\mathbf{B}) & \leq\left(\frac{4}{5-n}\right)^{n / 2} \frac{\lambda_{n}^{n}(\mathbf{B})}{\lambda_{1}^{n}(\mathbf{B})} \tag{9}
\end{align*}
$$

Proof. We first show upper bounds for terms 2 and 3 in (7), respectively. By constructing a sublattice $\Lambda\left(\mathbf{B}^{\prime}\right)$ from vectors with lengths $\lambda_{1}(\mathbf{B}), \ldots, \lambda_{n}(\mathbf{B})$ in $\Lambda(\mathbf{B})$, the covering radius satisfies

$$
\begin{aligned}
\rho(\mathbf{B}) & =\max _{\mathbf{x}} \operatorname{dist}(\mathbf{x}, \Lambda(\mathbf{B})) \\
& \leq \max _{\mathbf{x}} \operatorname{dist}\left(\mathbf{x}, \Lambda\left(\mathbf{B}^{\prime}\right)\right) \\
& \leq 1 / 2 \sqrt{\sum_{i=1}^{n} \lambda_{i}^{2}(\mathbf{B})}
\end{aligned}
$$

where the last inequality is obtained after applying Babai's nearest plane algorithm [33]. Since the residue distance of CVP is upper bounded by the covering radius of the sublattice $\Lambda\left(\mathbf{B}_{[n] \backslash i}\right)$, we have for term 3 that

$$
\begin{align*}
\left\|\pi_{\mathbf{B}_{[n] \backslash i}}\left(\mathbf{b}_{i}\right)-\mathbf{s}_{i}\right\|^{2} & \leq \rho^{2}\left(\mathbf{B}_{[n] \backslash i}\right) \\
& \leq \frac{1}{4} \sum_{j \neq i} \lambda_{j}^{2}\left(\mathbf{B}_{[n] \backslash i}\right) \\
& \leq \frac{1}{4} \sum_{j \neq i}\left\|\mathbf{b}_{j}\right\|^{2} . \tag{10}
\end{align*}
$$

Regarding term 2, we use QR decomposition to get $\left[\mathbf{B}_{[n] \backslash i}, \mathbf{b}_{i}\right]=\mathbf{Q R}$, from which we obtain $\left\|\pi_{\mathbf{B}_{[n] \backslash i}}^{\perp}\left(\mathbf{b}_{i}\right)\right\|=\left|r_{n, n}\right|$. Now w.l.o.g. assume that the successive minima $\lambda_{1}(\mathbf{B}), \ldots, \lambda_{n}(\mathbf{B})$ come from vectors $\mathbf{v}_{1} \triangleq\left[\mathbf{B}_{[n] \backslash i}, \mathbf{b}_{i}\right] \mathbf{c}_{1}, \ldots, \mathbf{v}_{n} \triangleq\left[\mathbf{B}_{[n] \backslash i}, \mathbf{b}_{i}\right] \mathbf{c}_{n}$. To produce $n$ linearly independent vectors, there exists at least one vector denoted as $\mathbf{c}_{k}$ whose $n$th entry $c_{k, n}$ is nonzero. Then we have $\left\|\mathbf{R c}_{k}\right\|^{2}=\lambda_{k}^{2}(\mathbf{B}) \leq \lambda_{n}^{2}(\mathbf{B})$. Together with $\left\|\mathbf{R c}_{k}\right\|^{2}=c_{k, n}^{2} r_{n, n}^{2}+\sum_{j=1}^{n-1} v_{n, j}^{2} \geq\left\|\pi_{\mathbf{B}_{[n] \backslash i}^{\perp}}\left(\mathbf{b}_{i}\right)\right\|^{2}$, it arrives at

$$
\begin{equation*}
\left\|\pi_{\mathbf{B}_{[n] \backslash i}}^{\perp}\left(\mathbf{b}_{i}\right)\right\|^{2} \leq \lambda_{n}^{2}(\mathbf{B}) . \tag{11}
\end{equation*}
$$

By substituting (10) and (11) to (7) for all basis vectors, we have

$$
\left\{\begin{array}{l}
\left\|\mathbf{b}_{1}\right\|^{2} \leq \lambda_{n}^{2}(\mathbf{B})+\frac{1}{4} \sum_{j \neq 1}\left\|\mathbf{b}_{j}\right\|^{2} \\
\vdots \\
\left\|\mathbf{b}_{n}\right\|^{2} \leq \lambda_{n}^{2}(\mathbf{B})+\frac{1}{4} \sum_{j \neq n}\left\|\mathbf{b}_{j}\right\|^{2}
\end{array}\right.
$$

The sum of these $n$ inequalities yields

$$
\sum_{i=1}^{n}\left\|\mathbf{b}_{i}\right\|^{2} \leq n \lambda_{n}^{2}(\mathbf{B})+\frac{n-1}{4} \sum_{i=1}^{n}\left\|\mathbf{b}_{i}\right\|^{2}
$$

If $n \leq 4$, we have

$$
\begin{equation*}
\sum_{i=1}^{n}\left\|\mathbf{b}_{i}\right\|^{2} \leq \frac{4 n}{5-n} \lambda_{n}^{2}(\mathbf{B}) \tag{12}
\end{equation*}
$$

Based on Eq. (12), the longest vector in the basis can be trivially bounded as

$$
l(\mathbf{B}) \leq \sqrt{\frac{4 n}{5-n}} \lambda_{n}(\mathbf{B})
$$

To analyze the orthogonal defect, we apply the arithmetic mean-geometric mean inequality on (12) to get

$$
\begin{equation*}
\prod_{i=1}^{n}\left\|\mathbf{b}_{i}\right\| \leq\left(\frac{1}{n} \sum_{i=1}^{n}\left\|\mathbf{b}_{i}\right\|^{2}\right)^{n / 2} \leq\left(\frac{4}{5-n}\right)^{n / 2} \lambda_{n}^{n}(\mathbf{B}) \tag{13}
\end{equation*}
$$

Clearly the volume of the lattice is lower bounded by $\lambda_{1}^{n}(\mathbf{B})$ for $n \leq 4$, so along with (13) we obtain (9).

If we alternatively set $\tau<1$, then upon termination of SR we have $\left\|\mathbf{b}_{i}\right\|^{2} \leq 1 / \tau\left\|\mathbf{b}_{i}-\mathbf{s}_{i}\right\|^{2}$. Along with the techniques used in Theorem 2, we obtain

$$
\begin{equation*}
l(\mathbf{B}) \leq \sqrt{\frac{4 n}{4 \tau-n+1}} \lambda_{n}(\mathbf{B}) \tag{14}
\end{equation*}
$$

Since we have to ensure that the denominator $4 \tau-n+1$ is larger than 0 , we claim that inequality (14) holds if $n<4 \tau+1$. If the CVP oracle is replaced by another suboptimal solver referred to as $\varepsilon \mathrm{CVP}$, then when bounding term 3 we have

$$
\left\|\pi_{\mathbf{B}_{[n] \backslash i}}\left(\mathbf{b}_{i}\right)-\mathbf{s}_{i}\right\|^{2} \leq \varepsilon \rho^{2}\left(\mathbf{B}_{[n] \backslash i}\right) .
$$

Similarly to the above, it yields

$$
\begin{equation*}
l(\mathbf{B}) \leq \sqrt{\frac{4 n}{4-\varepsilon n+\varepsilon}} \lambda_{n}(\mathbf{B}) \tag{15}
\end{equation*}
$$

in which $n<4 / \varepsilon+1$.
Let $\theta_{i}$ be the angle between $\mathbf{b}_{i}$ and the subspace $\operatorname{span}\left(\mathbf{B}_{[n] \backslash i}\right)$, and define $\theta_{\max } \triangleq \max _{i} \theta_{i}$. Such a maximum angle between basis vectors and subspaces can also be bounded, as shown in the following theorem.
Theorem 3. An $S R-C V P$ reduced basis satisfies $\cos ^{2} \theta_{\max } \leq$ $\frac{n-1}{4}$.

Proof: Based on (7) and (10) we have

$$
\begin{equation*}
\left\|\mathbf{b}_{i}\right\|^{2}-\left\|\pi_{\mathbf{B}_{[n] \backslash i}}^{\perp}\left(\mathbf{b}_{i}\right)\right\|^{2} \leq\left\|\pi_{\mathbf{B}_{[n] \backslash i}}\left(\mathbf{b}_{i}\right)-\mathbf{s}_{i}\right\|^{2} \leq \frac{1}{4} \sum_{j \neq i}\left\|\mathbf{b}_{j}\right\|^{2} \tag{16}
\end{equation*}
$$

It then follows from $\left\|\mathbf{b}_{i}\right\|^{2} \cos ^{2} \theta_{i}=\left\|\mathbf{b}_{i}\right\|^{2}-\left\|\pi_{\mathbf{B}_{[n] \backslash i}^{\perp}}^{\perp}\left(\mathbf{b}_{i}\right)\right\|^{2}$ that

$$
\left\|\mathbf{b}_{i}\right\|^{2} \cos ^{2} \theta_{\max } \leq\left\|\mathbf{b}_{i}\right\|^{2} \cos ^{2} \theta_{i} \leq \frac{1}{4} \sum_{j \neq i}\left\|\mathbf{b}_{j}\right\|^{2}
$$

Similarly to the techniques used in proving Theorem 2, we sum (16) for $i=1, \ldots, n$ to get

$$
\cos ^{2} \theta_{\max } \leq \frac{n-1}{4}
$$

Clearly the above theorem is non-trivial when $n \leq 4$, and this will come in handy when attacking a counter example in subsection IV-D.

## B. Discussions

1) Comparison with $\eta$-Greedy reduction [34, Fig.5] (also noted as $\mathrm{ELR}^{+}$-SLV in [28]). Rather than applying CVP for all vectors, $\eta$-Greedy reduction only performs CVP for the longest basis vector. According to its definition [34], it is only a special case of SR-CVP and all SR-CVP reduced basis must be greedy-reduced. For example, consider the following basis

$$
\left[\begin{array}{lllll}
2 & 0 & 0 & 0 & 1 \\
0 & 2 & 0 & 0 & 1 \\
0 & 0 & 2 & 0 & 1 \\
0 & 0 & 0 & 2 & 1 \\
0 & 0 & 0 & 0 & \varepsilon
\end{array}\right]
$$

with parameter $\varepsilon \in(0,1)$. The shortest vector $\left[\begin{array}{lllll}0 & 0 & 0 & 0 & \pm 2 \varepsilon\end{array}\right]^{\top}$ cannot be reached by greedy reduction. Specifically, by using $1 \times \mathbf{b}_{5}$ as the query point, $\eta$-Greedy cannot find $2 \mathbf{b}_{5}-\mathbf{b}_{1}-\mathbf{b}_{2}-\mathbf{b}_{3}-\mathbf{b}_{4}$ and $-2 \mathbf{b}_{5}+\mathbf{b}_{1}+\mathbf{b}_{2}+\mathbf{b}_{3}+\mathbf{b}_{4}$. In contrast, SR-CVP additionally considers the cases of using $\mathbf{b}_{1}, \mathbf{b}_{2}, \mathbf{b}_{3}$, and $\mathbf{b}_{4}$ as query points. A shortest vector $\mathbf{v}=\sum_{i=1}^{n} c_{i} \mathbf{b}_{i}$ with at least one coefficient $c_{k}= \pm 1$ must be contained in the SR-CVP reduced basis.
2) Comparison with KZ and its variants [16], [35]. Recall that a basis $\mathbf{B}$ is called KZ reduced if it satisfies the size reduction conditions, and $\pi_{\mathbf{B}_{[i-1]}}^{\perp}\left(\mathbf{b}_{i}\right)$ is the shortest vector of the projected lattice $\pi_{\mathbf{B}_{[i-1]}}^{\perp}\left(\left[\mathbf{b}_{i}, \ldots, \mathbf{b}_{n}\right]\right)$ for $1 \leq i \leq n$ [35]. For a KZ reduced basis, it satisfies [35] $\left\|\mathbf{b}_{i}\right\| \leq \frac{\sqrt{i+3}}{2} \lambda_{i}(\mathbf{B}), 1 \leq i \leq n$. Though boosted KZ [16] can solve the length increasing issue caused by size reduction, tuning $\pi_{\mathbf{B}_{[i-1]}}^{\perp}\left(\mathbf{b}_{i}\right)$ to be the shortest vector in the projected lattice can still make the basis longer. On the contrary, this issue is totally avoided in SR-CVP.
3) Comparison with Minkowski reduction. Recall that a lattice basis $\mathbf{B}$ is called Minkowski reduced if for any integers $c_{1}, \ldots, c_{n}$ such that $c_{i}, \ldots, c_{n}$ are altogether coprime, it has $\left\|\mathbf{b}_{1} c_{1}+\cdots+\mathbf{b}_{n} c_{n}\right\| \geq\left\|\mathbf{b}_{i}\right\|$ for $1 \leq$ $i \leq n$ [15]. For a Minkowski reduced basis, it satisfies $[15]\left\|\mathbf{b}_{i}\right\| \leq \max \left\{1,(5 / 4)^{(i-4) / 2}\right\} \lambda_{i}(\mathbf{B}), 1 \leq i \leq n$. Whereas Minkowski reduction is optimal as it reaches all the successive minima when $n \leq 4$, our results in Theorem 2 only show the SR-CVP reduced basis is not far from the optimal one. Here we argue that SRCVP has simpler structure. While Minkowski reduction requires solving integer least squares problems with GCD constraints and delicate basis expansion, SR-CVP only involves unconditional CVP solvers and its basis expansion process is trivial. Moreover, the SR-CVP algorithm can be approximately implemented by its many low-complexity siblings in the SR family.

## C. Complexity of $S R$ and $S R-C V P$

We argue that even when the threshold parameter $\tau=$ 1 , the decrease from $\left\|\mathbf{b}_{i}\right\|$ to $\left\|\mathbf{b}_{i}-\mathbf{s}_{i}\right\|$ can be finitely counted because a lattice is discrete. Therefore we define $\epsilon=\sup _{\mathbf{b}_{i}, \mathbf{s}_{i}} \frac{\left\|\mathbf{b}_{i}-\mathbf{s}_{i}\right\|}{\left\|\mid \mathbf{b}_{i}\right\|_{n}}$ which satisfies $\epsilon<1$. As $\sum_{i=1}^{n}\left\|\mathbf{b}_{i}\right\|^{2}$ is no smaller than $\sum_{i=1}^{n} \lambda_{i}^{2}(\mathbf{B})$ while this metric keeps decreasing for every $n$ iterations, the number of calls to the CVP oracle is not larger than $n \log \left(\frac{\|\mathbf{B}\|_{\mathrm{F}}^{2}}{\sum_{i=1}^{n} \lambda_{i}^{2}(\mathbf{B})}\right)$ with $\tau \leq 1$, where $\|\mathbf{B}\|_{\mathrm{F}}^{2}$ denotes the Frobenius norm of the input basis, and the $\log$ function is over $\min \{1 / \tau, 1 / \epsilon\}$. Therefore we conclude that the number of iterations in SR is polynomial.

Regarding SR-CVP, since the reduction in each round is quite strong, we can use the following heuristic implementation to minimize the number of iterations: first reduce the longest vector (similarly to $\eta$-Greedy), then reduce other basis vectors in descending order with $n-1$ rounds of CVP. Our simulation results show that this version of SR-CVP is competitive with Minkowski reduction and boosted KZ reduction.

While we can employ a state-of-the-art implementation for CVP, its complexity for a random basis is exponential [36]-[38]. In the next section, we will focus on approximate versions of CVP.

## IV. Hash-based Approximation: SR-Hash

## A. The nearest neighbor problem in $S R$-Pair

When the $\varepsilon$ CVP subroutine is not implemented with an exact CVP algorithm but rather a pairwise cancellation with the following form:

$$
\begin{align*}
\mathbf{b}_{i} & =\arg \min \left\|\mathbf{b}_{i}^{(j)}\right\|, j=\{1, \ldots, N\} \backslash i,  \tag{17}\\
\mathbf{b}_{i}^{(j)} & =\mathbf{b}_{i}-\left\lfloor\left\langle\mathbf{b}_{i}, \mathbf{b}_{j}\right\rangle /\left\langle\mathbf{b}_{j}, \mathbf{b}_{j}\right\rangle\right\rceil \mathbf{b}_{j},
\end{align*}
$$

we refer to the whole algorithm as SR-Pair. This algorithm coincides with the element-based reduction in [27]. Although this sub-routine only has a complexity in the order of $O(n N)$, reaching another variant with lower complexity is possible.

Recall the nearest neighbor problem in the field of large dimensional data processing is: given a list of $n$-dimensional vectors $L=\left\{\mathbf{v}_{1}, \mathbf{v}_{2}, \ldots, \mathbf{v}_{N}\right\} \in \mathbb{R}^{n}$, preprocess $L$ in such a way that, when later given a target vector $\mathbf{q} \notin L$, one can efficiently find an element $\mathbf{v} \in L$ which is almost the closest to $\mathbf{q}$. Since Eq. (17) exactly defines a search for the nearest neighbor of $\mathbf{b}_{i}$ among the vectors in $\mathbf{B}$, then it motivates us to reduce this complexity to $O(n \log N)$ based on locality-sensitive-hashing (LSH) [39], [40].
Remark 4. If we choose SIC as the $\varepsilon$ CVP subroutine, then along with LLL preprocessing we have [33]

$$
\begin{equation*}
\left\|\mathbf{b}_{i}-\pi_{\mathbf{B}_{[n] \backslash i}}\left(\mathbf{b}_{i}\right)-\mathbf{s}_{i}\right\| \leq 2(2 / \sqrt{3})^{n-1} \rho(\mathbf{B}) \tag{18}
\end{equation*}
$$

for such an SR-SIC algorithm. However, the computation complexity of this algorithm is still too high as it requires the pre-processing by LLL.

## B. Angular LSH

LSH roughly works as follows: first all $N$ candidates are dispatched to different buckets with labels, then when searching the nearest neighbor of a query point $\mathbf{q}$, we can alternatively do this only for $N^{\prime}$ candidates that have the same label with $\mathbf{q}$, where $N^{\prime} \ll N$. There are label functions $f$ which map an $n$-dimensional vector $\mathbf{v}$ to a low-dimensional sketch of $\mathbf{v}$. For certain distance function $D$, vectors which are nearby in the sense of $D$ have a high probability of having the same sketch, while vectors which are far away have a low probability of having the same image under $f$.
To reach this property, we introduce the definition of an LSH family $\mathcal{F}$.

Definition 5. A family $\mathcal{F}=\left\{f: \mathbb{R}^{n} \rightarrow \mathbb{N}\right\}$ of hash functions is said to be $\left(r_{1}, r_{2}, p_{1}, p_{2}\right)$-sensitive for a similarity measure $D$ if for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{n}$, we have i) If $D(\mathbf{u}, \mathbf{v}) \leq r_{1}$, then $\operatorname{Pr}_{f \in \mathcal{F}}(f(\mathbf{u})=f(\mathbf{v})) \geq p_{1}$; ii)If $D(\mathbf{u}, \mathbf{v}) \geq r_{2}$, then $\operatorname{Pr}_{f \in \mathcal{F}}(f(\mathbf{u})=f(\mathbf{v})) \leq p_{2}$.
For the sake of constructing a hash family with $p_{1} \approx 1$ and $p_{2} \approx 0$, normally one first constructs $p_{1} \approx p_{2}$ and then uses the so called AND- and OR-compositions to turn it into an $\left(r_{1}, r_{2}, p_{1}^{\prime}, p_{2}^{\prime}\right)$-sensitive hash family $\mathcal{F}^{\prime}$ with $p_{1}^{\prime}>p_{1}$ and $p_{2}^{\prime}<p_{2}$, thereby amplifying the gap between $p_{1}$ and $p_{2}$. Specifically, by combining $k$ AND-compositions and $t$ OR-compositions, we can turn an $\left(r_{1}, r_{2}, p_{1}, p_{2}\right)$-sensitive hash family $\mathcal{F}$ into an $\left(r_{1}, r_{2}, 1-\left(1-p_{1}^{k}\right)^{t}, 1-\left(1-p_{2}^{k}\right)^{t}\right)$ sensitive hash family $\mathcal{F}^{\prime}$. As long as $p_{1}>p_{2}$, we can always find values of $k$ and $t$ such that $1-\left(1-p_{1}^{k}\right)^{t} \rightarrow 1$ and $1-\left(1-p_{2}^{k}\right)^{t} \rightarrow 0$.

Note that if given a hash family $\mathcal{H}$ which is $\left(r_{1}, r_{2}, p_{1}, p_{2}\right)$ sensitive with $p_{1} \gg p_{2}$, then we can use $\mathcal{F}$ to distinguish between vectors which are at most $r_{1}$ away from $\mathbf{v}$, and vectors which are at least $r_{2}$ away from $\mathbf{v}$ with non-negligible probability, by only looking at their hash values. Although large values of $k$ and $t$ can amplify the gap between $p_{1}$ and $p_{2}$, large parameters come at the cost of having to compute many hashes and having to store many hash tables in memory. To minimize the overall time complexity, we need the following lemma that shows how to balance $k$ and $t$. In practice, we can further tune $k$ and $t$ to have the best performance.

Lemma 6 ([41], [42]). Suppose there exists an ( $r_{1}, r_{2}, p_{1}, p_{2}$ )sensitive family $\mathcal{F}$. For a list $L$ of size $N$, let

$$
\rho=\frac{\log p_{1}^{-1}}{\log p_{2}^{-1}}, k=\frac{\log N}{\log p_{2}^{-1}}, t=O\left(N^{\rho}\right) .
$$

Then given a query point $\mathbf{q}$, with high probability we can either find an element $\mathbf{v} \in L$ such that $D(\mathbf{q}, \mathbf{v}) \leq r_{2}$, or conclude that with high probability, no element $\mathbf{v} \in L$ with $D(\mathbf{q}, \mathbf{v})>$ $r_{1}$ exist, with the following costs: i) Time for preprocessing the list: $O\left(k N^{1+\rho}\right)$; ii) Space complexity of the preprocessed data: $O\left(N^{1+\rho}\right)$; iii) Time for answering a query: $O\left(N^{\rho}\right)$.

In the sequel, we examine the implementation of LSH based on angular hashing. Angular hashing means generating random hyperplanes $\mathbf{h}_{1}, \ldots, \mathbf{h}_{k}$, such that the whole space is sliced into $2^{k}$ regions. After that, to find the nearest neighbor of $\mathbf{q}$,
one only compares $\mathbf{q}$ to points in the same region $\mathcal{R}$. Here we introduce the angular distance similarity function

$$
D(\mathbf{u}, \mathbf{v})=\arccos \left(\frac{\mathbf{u}^{\top} \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\right)
$$

With this measure two vectors are nearby if their common angle is small. Its corresponding hash family is defined by

$$
\mathcal{F}=\left\{f_{\mathbf{a}}: \mathbf{a} \in \mathbb{R}^{n},\|\mathbf{a}\|=1\right\}, f_{\mathbf{a}}(\mathbf{v})= \begin{cases}1 & \text { if } \mathbf{a}^{\top} \mathbf{v} \geq 0 \\ 0 & \text { if } \mathbf{a}^{\top} \mathbf{v}<0\end{cases}
$$

Intuitively, the space that is orthogonal to a defines a hyperplane, and $f_{\mathbf{a}}$ maps the two regions separated by this hyperplane to different bits. In particular, for any two angles $\theta_{1}<\theta_{2}$, the family $\mathcal{F}$ is $\left(\theta_{1}, \theta_{2}, 1-\frac{\theta_{1}}{\pi}, 1-\frac{\theta_{2}}{\pi}\right)$ sensitive. Further with $k$ AND- and $t$ OR- compositions, we have $\left(\theta_{1}, \theta_{2}, 1-\left(1-\left(1-\frac{\theta_{1}}{\pi}\right)^{k}\right)^{t}, 1-\left(1-\left(1-\frac{\theta_{2}}{\pi}\right)^{k}\right)^{t}\right)$ sensitive hash family.

To illustrate LSH and in particular the angular LSH method described above, Fig. 2 shows how hyperplane hashing might work in a 2-D setting. In the figure, we have a list of 8 candidates: $L=\left\{\mathbf{b}_{1}, \ldots, \mathbf{b}_{8}\right\}$, and we use $k=2$ hyperplanes for $t=2$ hash tables. Each table stores the hash keys (labels) along with elements being placed in buckets, where elements having the same keys will be placed in the same buckets. In the two tables, the AND-compostions of 11 respectively correspond to $\mathbf{b}_{1}, \mathbf{b}_{2}, \mathbf{b}_{3}$ and $\mathbf{b}_{1}$. Based on OR-composition, the nearest neighbor of $\mathbf{b}_{1}$ is found inside $\left\{\mathbf{b}_{2}, \mathbf{b}_{3}\right\}$.


Fig. 2: Demonstration of LSH.

## C. LSH-Based Reduction

Now we show how to incorporate LSH into the sequential reduction algorithm. The pseudo-codes of SR-Hash are presented in Algorithm 2. It first hashes all vectors in lines $2-3$. Then inside the loop of the element-based reduction, the search for finding the nearest neighbor of $\mathbf{b}_{i}$ is within $\mathbf{C}=\cup_{l=1}^{t} T_{l}\left(f_{l}\left( \pm \mathbf{b}_{i}\right)\right)$. Every time when a shorter $\mathbf{b}_{i}$ is found, its hash labels and positions in buckets are updated in lines 11 and 13.

In summary, SR-Hash can be seen as a generalization of the naive brute-force search inside SR-Pair for finding nearest neighbors, as $k=0, t=1$ corresponds to checking all other basis vectors for nearby vectors, while increasing both $k$ and $t$ leads to fewer comparisons but a higher cost of computing hash keys and checking buckets.

```
Algorithm 2: The SR-Hash algorithm.
    Input: original lattice basis \(\mathbf{B}=\left[\mathbf{b}_{1}, \ldots, \mathbf{b}_{n}\right]\),
            complexity threshold \(\tau\), LSH parameters \(t, k\).
    Output: Reduced basis \(\left[\mathbf{b}_{1}, \ldots, \mathbf{b}_{n}\right]\) of \(\Lambda\).
    1 Initialize \(t\) empty hash tables \(\left(T_{l}\right)_{l=1}^{t}\), each has \(k\) random
    hash functions \(f_{l, 1}, \ldots f_{l, k} \in \mathcal{F}\);
    for \(i=1 \cdots n\) do
        Add \(\mathbf{b}_{i}\) to all hash tables \(\left(T_{l}\right)_{l=1}^{t}\), with hash values
            \(\left(f_{l}\left(\mathbf{b}_{i}\right)\right)_{l=1}^{t}\) and vectors in the same bucket noted as
            \(T_{l}\left(f_{l}\left(\mathbf{b}_{i}\right)\right)\);
\(4 i=0, m=1\);
    while \(m \leq n\) do
        \(i \leftarrow(i \bmod n)+1 ; \quad \triangleright\) The column index;
        Obtain the set of candidates \(\mathbf{C}=\cup_{l=1}^{t} T_{l}\left(f_{l}\left( \pm \mathbf{b}_{i}\right)\right)\);
        \(\mathbf{c}_{l}=\arg \min _{\mathbf{c}_{l} \in \mathbf{C}}\left\|\mathbf{b}_{i}-\left\lfloor\left\langle\mathbf{b}_{i}, \mathbf{c}_{l}\right\rangle /\left\langle\mathbf{c}_{l}, \mathbf{c}_{l}\right\rangle\right\rangle \mathbf{c}_{l}\right\|^{2} ;\)
        \(\mathbf{s}_{i}=\left\lfloor\left\langle\mathbf{b}_{i}, \mathbf{c}_{l}\right\rangle /\left\langle\mathbf{c}_{l}, \mathbf{c}_{l}\right\rangle\right\rangle \mathbf{c}_{l} ;\)
        if \(\left\|\mathbf{b}_{i}-\mathbf{s}_{i}\right\|^{2}<\tau\left\|\mathbf{b}_{i}\right\|^{2}\) then
            Remove \(\mathbf{b}_{i}\) from all hash tables;
            \(\mathbf{b}_{i} \leftarrow \mathbf{b}_{i}-\mathbf{s}_{i} ;\)
            Add \(\mathbf{b}_{i}\) to all hash tables;
            \(m=1\);
        else
            \(m \leftarrow m+1 ;\)
```


## D. Discussions

1) Comparison with SR-CVP. Here we emphasize that SR-Pair/SR-Hash is only a weak approximation for SR-CVP, and these low complexity algorithms may have quite inferior performance. Consider the counter example given for ELR [27] (the same as SR-Pair). Clearly SR-Pair/SRHash is unable to reduce a basis whose Gram matrix is

$$
\mathbf{G}=\left[\begin{array}{ccc}
1 & 0.5-\nu & 0.5-\nu \\
0.5-\nu & 1 & -0.5+\nu \\
0.5-\nu & -0.5+\nu & 1
\end{array}\right]
$$

with $\nu \rightarrow 0$. Under spherical coordinate system of $(r, \varrho, \varphi)$ with $r=1, \varrho=\pi / 3$, and $\varphi=\pi / 2-\nu$, the lattice basis $\mathbf{A}$ corresponded to $\mathbf{G}$ (up to a unitary transform) is given by

$$
\mathbf{A}=\left[\begin{array}{ccc}
\sin \varphi \cos (\pi / 3) & -\sin \varphi \cos (\pi / 3) & 1  \tag{19}\\
\sin \varphi \sin (\pi / 3) & \sin \varphi \sin (\pi / 3) & 0 \\
\cos \varphi & -\cos \varphi & 0
\end{array}\right]
$$

This basis has an angle $\theta_{i}<\nu \rightarrow 0$ between any $\mathbf{a}_{i}$ and $\operatorname{span}\left(\mathbf{A}_{[3] \backslash i}\right)$, and $\eta(\mathbf{A})=\infty$ if $\nu \rightarrow 0$. If $\boldsymbol{A}$ is reduced by using SR-CVP, we have $\theta_{\max } \geq \pi / 4$ according to Theorem 3, so $\mathbf{A}$ is not a stable basis for SR-CVP. Moreover, the actual reduced basis has the following form:

$$
\tilde{\mathbf{A}}=\left[\begin{array}{ccc}
2 \sin \varphi \cos (\pi / 3)-1 & -\sin \varphi \cos (\pi / 3) & 1 \\
0 & \sin \varphi \sin (\pi / 3) & 0 \\
2 \cos \varphi & -\cos \varphi & 0
\end{array}\right]
$$

Its OD is

$$
\eta(\tilde{\mathbf{A}})=\frac{\sqrt{4 \cos ^{2} \varphi+\sin ^{2} \varphi-2 \sin \varphi+1}}{\sqrt{3} \sin \varphi \cos \varphi}
$$

when given $\varphi=\pi / 2-10^{-4}$, we have $\left.\eta(\tilde{\mathbf{A}})\right|_{\varphi=\pi / 2-10^{-4}}=1.1547$. Therefore, the proposed low-complexity $S R$ algorithms are only feasible for bases whose input vectors are dense in some directions. Our simulation results and Appendix A will show that the dual lattice basis in MIMO detection is one example of this.
2) Comparison with LLL and its variants [19], [22], [24]. Note that the worst case complexity of LLL for bases in the real field is unbounded [43], and the variants that control the order of swaps or a selective implementation of size reduction cannot remove this curse. On the contrary, SR variants with a polynomial time $\varepsilon$ CVP routine can enjoy the overall polynomial-time complexity. Regarding performance bounds, LLL and its variants (the maintains the Siegel condition and size reduction condition) often have bounds of the form $l(\mathbf{B}) \leq 2^{n-1} \lambda_{n}(\mathbf{B})$, while SR-Pair and SR-Hash are heuristic.
3) Comparison with Seysen reduction [14]. Rather than minimizing the orthogonality defect of a basis, a metric called Seysen's measure can reflect whether both the primal and dual bases are short: $\sum_{i=1}^{n}\left\|\mathbf{b}_{i}\right\|^{2}\left\|\mathbf{d}_{i}\right\|^{2}$. Seeking for the global minimum of this metric is extremely hard; when referring to Seysen's algorithm [25], it is the one that finds a local minimum of $\sum_{i=1}^{n}\left\|\mathbf{b}_{i}\right\|^{2}\left\|\mathbf{d}_{i}\right\|^{2}$ without any theoretical performance guarantee. Similarly to SR-Pair, Seysen's algorithm performs basis updates in a pair-wise manner:

$$
\mathbf{b}_{j}=\mathbf{b}_{j}+c_{i, j} \mathbf{b}_{i}, i \neq j
$$

with $c_{i, j}=\left\lfloor\frac{1}{2}\left(\frac{\left\langle\mathbf{d}_{i}, \mathbf{d}_{j}\right\rangle}{\left\|\mathbf{d}_{i}\right\|^{2}}-\frac{\left\langle\mathbf{b}_{i}, \mathbf{b}_{j}\right\rangle}{\left\|\mathbf{b}_{i}\right\|^{2}}\right)\right\rceil$. Due to the additional inner product calculation in the dual basis, Seysen's algorithm is more complicated than SR-Pair, and it does not support the hash-based implementation. Moreover, in large ( $\geq 35$ ) dimensions Seysen's algorithm often halts at a local minimum [14, P.375]. Since the error rate performance is only controlled by the length of the dual basis, our empirical results also show that Seysen's algorithm is not competitive for large dimensions.

## V. Simulation results

## A. Performance of SR-CVP

Hereby we employ the OD's to compare SR-CVP with other strong lattice reduction algorithms, including the boosted Korkin-Zolotarev reduction noted as "bKZ", the Minkowski reduction noted as "Minkowski", and the $\eta$-Greedy reduction [34, Fig.5] noted as " $\eta$-Greedy". Results are averaged over $1 \times 10^{4}$ Monte-Carlo runs, and SR-CVP is implemented by the heuristic version in subsection III-C.

Fig. 3 plots dimension versus OD for distinct algorithms for the primal and dual of a Gaussian random matrix with


Fig. 3: The orthogonal defects of different types of strong reduction.
entries from $\mathcal{N}(0,1)$, respectively. The figure shows that ODs of SR-CVP, bKZ and Minkowski reduced bases are almost indistinguishable. $\eta$-Greedy has the worst performance as expected, because it is not designed to minimized all basis vectors. Since Minkowski reduction is the state-of-theart algorithm for generating the shortest basis in practice, our results show that SR-CVP practically reaches optimality as well.

Fig. 4 plots the averaged number of CVP runs in Fig. 3 when using $\eta$-Greedy and SR-CVP. It is known that both Minkowski and bKZ cost around $n$ oracles for the shortest vector problem (SVP) or CVP. Fig. 4 reflects that SR-CVP actually needs fewer than $n$ rounds of CVP, and is only slightly more complicated than $\eta$-Greedy.

## B. SR-Hash vs. SR-Pair and LLL variants

In this subsection, we study the complexity/performance tradeoffs of different types of weak lattice reduction. The modulation is set as 16 QAM, and the results are obtained from $1 \times 10^{4}$ Monte Carlo runs. We denote the zero-forcing detector by "ZF", the successive interference cancellation detector by "SIC", and lattice-reduction-aided detectors with prefixes: "LLL-SIC/ZF" [32], "bLLL-SIC/ZF" [16], "SR-PairSIC/ZF" (this paper), "SR-Hash-SIC/ZF" (this paper), and "Seysen-SIC/ZF" [26]. Here comparisons are made for major lattice-reduction-aided methods in large-scale MIMO systems, because they represent pre-processing based methods that may attain the diversity order of ML detection [6], [44].

1) i.i.d. channels: Assume that each entry of the channel matrix is chosen from a standard normal distribution $\mathcal{C N}(0,1)$. Fig. 5 plots the bit error rate (BER) performance of different uncoded MIMO detectors in a real domain $2 n_{T} \times$ $2 n_{R}=60 \times 60 \mathrm{MIMO}$ system. Here the linear detectors are


Fig. 4: The number of effective CVP runs in $\eta$-Greedy and SR-CVP.
implemented with the MMSE criterion. Parameters in LSH are chosen as $t=\left\lfloor n^{0.585}\right\rceil=11, k=\lfloor\log n\rceil=6$.

In the high SNR region of Fig. 5-(a), we observe that, in addition to the well-known fact that ZF, SIC and Seysen-SIC fail to achieve the full diversity order, b-LLL-SIC, SR-PairSIC and SR-Hash-SIC all attain approximately 1dB gain over LLL-SIC. As for Fig. 5-(b), the variants of SR both outperform conventional and boosted LLL algorithms. Both sub-figures indicate that SR-Hash gets very close to SR-Pair.

The complexity of implementing the lattice reduction algorithms is plotted in Fig. 6, where sub-figure (a) is for the effective channel matrix under the MMSE criterion, and subfigure (b) under the ZF criterion. Considering the difficulty in analyzing the number of floating-point operations for hash operations, here we measure the complexity by the number of vector comparisons. This equals to the number of iterations times: the size of the basis for SR-Pair, the number of vectors in the same buckets for SR-Hash, and to the size of vectors for doing size-reductions for both LLL and bLLL. From Fig. 6-(a), we observe that the LLL variants are not affected by SNR in the MMSE matrix, and Seysen, SR-Pair and SRHash gradually increase with the rise of SNR. This shows the complexity of Seysen, SR-Pair and SR-Hash are dependent on the quality of the input bases. Regarding the stationary lines in Fig. 6-(b), the numbers of comparisons of Seysen, SR-Pair and SR-Hash reflect the asymptotic values of their counterparts in Fig. 6-(a). Both subfigures reveal that the hash method helps to reduce the complexity of SR-Pair significantly. A natural question that arises here is whether the complexity dependency of SR-Pair\&SR-Hash on input bases may lead to inferior performance at low SNR. To address this question, we plot the SNR versus OD relations of different reduction algorithms in Fig. 7. We observe from the figure that even at low SNR, SR-Pair\&SR-Hash featuring low complexity still


Fig. 5: The BER performance of different detectors in large MIMO.
outperform Seysen and LLL in terms of OD.
2) Correlated Channels: Results in the last example were obtained for i.i.d. frequency-flat Rayleigh fading channels. The performance of MIMO systems in realistic radio environments however sometimes depends on spatial correlation. Therefore, we investigated the effect of channel correlation on the performance of the new reduction algorithms. Based on [45], the spatially correlated channel is modeled as

$$
\widetilde{\mathbf{B}_{c}}=\Psi \mathbf{B}_{c},
$$

where $\Psi \in \mathbb{R}^{n_{R} \times n_{R}}$ is the correlation matrix defined by

$$
\Psi=\left[\begin{array}{cccc}
1 & \rho & \cdots & \rho^{n_{R}-1} \\
\rho & 1 & \cdots & \rho^{n_{R}-2} \\
\vdots & \vdots & \ddots & \vdots \\
\rho^{n_{R}-1} & \rho^{n_{R}-2} & \cdots & 1
\end{array}\right]
$$

and $\rho$ refers to the spatial correlation coefficient.
With the same chosen parameters in the algorithm as those for i.i.d. channels, Fig. 8 demonstrates the BER performances against SNR in correlated channels respectively with $\rho=0.1$ and $\rho=0.3$. It reveals that, as $\rho$ increases, the SR aided detectors suffer from more severe performance degradation


Fig. 6: The complexity of different lattice reduction algorithms in i.i.d. channels.


Fig. 7: The ODs of MMSE matrices reduced by different algorithms.
than the LLL aided methods, although the BER gaps between SR variants and LLL variants are very small. This is not unexpected because we do have examples showing SR-Pair/SRHash cannot reduce certain matrices (e.g., the matrix in (19)). Lastly, as plotted in Fig. 9, the complexity of SR-Pair/SR-Hash is still much lower than those of LLL variants and Seysen, and the proposed SR-Hash has much lower complexity than SRPair.

## VI. Conclusions and Future Work

To summarize, we have unveiled a new lattice reduction family called sequential reduction, which enjoys a polynomial number of iterations. Theoretical bounds on basis lengths and orthogonality defects are derived under the premise that an exact CVP subroutine has been invoked. Though we only manage to prove these results for small dimensions, they still provide insights on understanding the performance of such a class of


Fig. 8: The BER performance of lattice-reduction-aided SIC detectors in correlated channels with $\rho=0.1$ and $\rho=0.3$.
algorithms. Within the SR framework, the SR-Hash method can serve as an effective subprogram, and simulation results show that the complexity-performance trade-off outperforms those of SR-Pair and LLL variants in large MIMO detection.

We believe that the studies we initiated here, only scratch the tip of the iceberg about the new lattice reduction family. Many important questions remain to be answered. Research on the interactions and combinations of SR with other techniques such as floating-point arithmetic [46], [47], randomized detection algorithms [31], [48], success probability analysis [49], [50], and numerous other topics is now being pursued.

## Appendix A <br> ON THE TYPE OF BASES FEASIBLE FOR SR-PAIR\&SR-HASH

We first argue that a large dimensional Gaussian random basis is always SR-CVP reduced, and thus being SR-Pair and SR-Hash reduced. The inability to change such bases is however not a problem because these bases are close to being orthogonal.


Fig. 9: The complexity of lattice reduction algorithms in correlated channels with $\rho=0.1$ and $\rho=0.3$.

Proposition 7. For a Gaussian random basis whose entries follow the distribution $\mathcal{N}(0,1)$, the probability that it is not $S R-C V P$ reduced goes to zero as $n \rightarrow \infty$.

Proof: We need to show that for all choices of coefficients $a_{i}^{\prime} s$ in $\mathbb{Z}$ with at least one nonzero $a_{i}$, the probability

$$
\operatorname{Pr}\left(\left\|\mathbf{b}_{1}+\sum_{i=2}^{n} \mathbf{b}_{i} a_{i}\right\|^{2} \leq\left\|\mathbf{b}_{1}\right\|^{2}\right)
$$

vanishes as the problem size $n$ increases. Since $\sum_{i=2}^{n} \mathbf{b}_{i} a_{i}$ is an isotropic Gaussian random vector with covariance $\mathbb{E}\left(\left(\sum_{i=2}^{n} \mathbf{b}_{i} a_{i}\right)\left(\sum_{i=2}^{n} \mathbf{b}_{i} a_{i}\right)^{\top}\right)=\left(\sum_{i=2}^{n} a_{i}^{2}\right) \mathbf{I}_{n}$, then for any $\beta>0$,

$$
\begin{align*}
& \operatorname{Pr}\left(\left\|\mathbf{b}_{1}+\sum_{i=2}^{n} \mathbf{b}_{i} a_{i}\right\|^{2} \leq\left\|\mathbf{b}_{1}\right\|^{2}\right) \\
& \leq \mathbb{E}\left(e^{-\beta\left(\left\|\mathbf{b}_{1}+\sum_{i=2}^{n} \mathbf{b}_{i} a_{i}\right\|^{2}-\left\|\mathbf{b}_{1}\right\|^{2}\right)}\right) \\
& =\int \frac{d \mathbf{x} d \mathbf{v}}{(2 \pi)^{n}} \\
& \left.e^{-\frac{1}{2}\left[\mathbf{v}^{\top}, \mathbf{x}^{\top}\right][ } \begin{array}{cc}
2 \sqrt{\sum_{i=2}^{n} a_{i}^{2}} \beta \mathbf{I}_{n} & \left(1+2 \beta \sum_{i=2}^{n} a_{i}^{2}\right) \mathbf{I}_{n}
\end{array}\right]\left[\begin{array}{c}
\mathbf{v} \\
\mathbf{x}
\end{array}\right] \\
& =\operatorname{det}\left(\left[\begin{array}{cc}
2 \sqrt{\sum_{i=2}^{n} a_{i}^{2}} \beta \mathbf{I}_{n} \\
=\left(\begin{array}{cc}
\mathbf{I}_{n}^{n} a_{i}^{2} & \mathbf{I}_{n} \\
1+2 \beta(1-2 \beta) \sum_{i=2}^{n} a_{i}^{2}
\end{array}\right)^{-1 / 2}
\end{array}\right.\right. \\
& \left.\left.\begin{array}{ll}
1 & 2 \sqrt{\sum_{i=2}^{n} a_{i}^{2}} \beta \mathbf{I}_{n} \\
\left.1+2 \sum_{i=2}^{n} a_{i}^{2}\right) \\
\mathbf{I}_{n}
\end{array}\right]\right)^{n / 2} \tag{20}
\end{align*}
$$

By optimizing over $\beta$ in the denominator, we have $1+$ $2 \beta(1-2 \beta) \sum_{i=2}^{n} a_{i}^{2} \leq 1+\frac{1}{4} \sum_{i=2}^{n} a_{i}^{2}$. This means we can use $\beta=\frac{1}{4}$ to reach the tightest bound for inequality (20). Therefore, for any $\varepsilon>0$, we have

$$
\begin{aligned}
& \lim _{n \rightarrow \infty} \operatorname{Pr}\left(\left\|\mathbf{b}_{1}+\sum_{i=2}^{n} \mathbf{b}_{i} a_{i}\right\|^{2} \leq\left\|\mathbf{b}_{1}\right\|^{2}\right) \\
& \leq \lim _{n \rightarrow \infty}\left(\frac{1}{1+2 \beta(1-2 \beta) \sum_{i=2}^{n} a_{i}^{2}}\right)^{n / 2} \\
& <\varepsilon
\end{aligned}
$$

Next, we investigate the reduction on the dual of a Gaussian random basis, which arises in our detection problem. For an input basis $\mathbf{B}$ we define

$$
\theta_{i, j}=\arccos \left(\frac{\left|\left\langle\mathbf{b}_{i}, \mathbf{b}_{j}\right\rangle\right|}{\left\|\mathbf{b}_{i}\right\|\left\|\mathbf{b}_{j}\right\|}\right), 1 \leq i \neq j \leq n
$$

The following lemma says that the SR-Pair method can provide a Gauss-reduced basis for all pairs of vectors with pairwise angles $\theta_{i, j}>\pi / 3$.
Lemma 8. For an SR-Pair reduced basis, we have $\theta_{i, j}>\pi / 3$ for all $i \neq j$.
Proof. If a lattice basis B is non-reducible by SR-Pair, we have $\left\lfloor\left\langle\mathbf{b}_{i}, \mathbf{b}_{j}\right\rangle /\left\langle\mathbf{b}_{j}, \mathbf{b}_{j}\right\rangle\right\rceil=0 \forall i \neq j$. Therefore the lemma follows from

$$
\cos \theta_{i, j}<\frac{1}{2} \frac{\left\|\mathbf{b}_{i}\right\|}{\left\|\mathbf{b}_{j}\right\|}<\frac{1}{2} \frac{\min \left(\left\|\mathbf{b}_{i}\right\|,\left\|\mathbf{b}_{j}\right\|\right)}{\max \left(\left\|\mathbf{b}_{i}\right\|,\left\|\mathbf{b}_{j}\right\|\right)} \leq 1 / 2
$$

Here we argue that the pairwise angles are dense in the dual of a Gaussian random basis. Fig. 10-(a) plots the histogram of such random matrices. It shows that so a large number of vectors satisfy $\theta_{i, j}<\pi / 3$, and these vectors will trigger the reduction in SR-Pair/SR-Hash. On the contrary, as predicted by Proposition 7, Fig. 10-(b) shows that the primal basis will not be reduced by SR-Pair/SR-Hash. Bases with dense angles also feature large orthogonality defects. In Fig. 11, we plot the OD versus dimension $n$ relations respectively for the dual and primal Gaussian random matrices. The figure shows the dual bases approximately have a growth rate of $O\left(20^{n^{1.5}}\right)$, while that of the primal basis is extremely small. The above confirms that the objective lattice bases in MIMO detection are easily reducible by tuning the pairwise angles.

## REFERENCES

[1] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: opportunities and challenges with very large arrays," IEEE Signal Process. Mag., vol. 30, no. 1, pp. 40-60, 2013.
[2] 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, 2002.
[3] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in Proc. Global Telecommunications Conference (GLOBECOM), Taipei, Taiwan, pp. 424-428, 2002.
[4] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-reductionaided broadcast precoding," IEEE Trans. Commun., vol. 52, no. 12, pp. 2057-2060, 2004.
[5] Z. Wang, Y. Huang, and S. Lyu, "Lattice-reduction-aided gibbs algorithm for lattice gaussian sampling: Convergence enhancement and decoding optimization," IEEE Trans. Signal Processing, vol. 67, no. 16, pp. 4342-4356, 2019.


Fig. 10: The histogram of the pairwise angles for the dual and primal of $60 \times 60$ Gaussian random bases.


Fig. 11: The ODs of square Gaussian random bases.
[6] M. Taherzadeh, A. Mobasher, and A. K. Khandani, "LLL reduction achieves the receive diversity in MIMO decoding," IEEE Trans. Information Theory, vol. 53, no. 12, pp. 4801-4805, 2007.
[7] -_, "Communication over MIMO broadcast channels using latticebasis reduction," IEEE Trans. Inf. Theory, vol. 53, no. 12, pp. 45674582, 2007.
[8] X. Ma and W. Zhang, "Performance analysis for MIMO systems with lattice-reduction aided linear equalization," IEEE Trans. Commun., vol. 56, no. 2, pp. 309-318, 2008.
[9] P. Q. Nguyen and B. Vallée, Eds., The LLL Algorithm. Springer Berlin Heidelberg, 2010.
[10] R. Neelamani, R. L. de Queiroz, and R. G. Baraniuk, "Compression color space estimation of JPEG images using lattice basis reduction," in Proc. IEEE Int. Conf. Image Process. (ICIP), Thessaloniki, Greece, pp. 890-893, 2001.
[11] A. Korkinge and G. Zolotareff, "Sur les formes quadratiques positives," Math. Ann., vol. 11, no. 2, pp. 242-292, 1877.
[12] D. Micciancio and S. Goldwasser, Complexity of Lattice Problems. Boston, MA: Springer, 2002.
[13] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," Mathematische Annalen, vol. 261, no. 4, pp. 515-534, 1982.
[14] M. Seysen, "Simultaneous reduction of a lattice basis and its reciprocal
basis," Combinatorica, vol. 13, no. 3, pp. 363-376, sep 1993.
[15] W. Zhang, S. Qiao, and Y. Wei, "HKZ and minkowski reduction algorithms for lattice-reduction-aided MIMO detection," IEEE Trans. Signal Process., vol. 60, no. 11, pp. 5963-5976, 2012.
[16] S. Lyu and C. Ling, "Boosted KZ and LLL algorithms," IEEE Trans. Signal Process., vol. 65, no. 18, pp. 4784-4796, Sep. 2017.
[17] J. Wen and X. Chang, "On the KZ reduction," IEEE Trans. Information Theory, pp. 1-1, 2018.
[18] Y. Chen and P. Q. Nguyen, "BKZ 2.0: Better lattice security estimates," in Proc. ASIACRYPT, Seoul, South Korea, 2011, pp. 1-20, 2011.
[19] C. Ling, W. H. Mow, and N. Howgrave-Graham, "Reduced and fixedcomplexity variants of the LLL algorithm for communications," IEEE Trans. Commun., vol. 61, no. 3, pp. 1040-1050, 2013.
[20] W. Zhang, S. Qiao, and Y. Wei, "A diagonal lattice reduction algorithm for MIMO detection," IEEE Signal Process. Lett., vol. 19, no. 5, pp. 311-314, 2012.
[21] H. Vetter, V. Ponnampalam, M. Sandell, and P. A. Hoeher, "Fixed complexity LLL algorithm," IEEE Trans. Signal Process., vol. 57, no. 4, pp. 1634-1637, 2009.
[22] Q. Wen, Q. Zhou, and X. Ma, "An enhanced fixed-complexity LLL algorithm for MIMO detection," in Proc. Global Telecommunications Conference (GLOBECOM), Austin, TX, USA, pp. 3231-3236, 2014.
[23] Q. Wen and X. Ma, "Efficient greedy LLL algorithms for lattice decoding," IEEE Trans. Wirel. Commun., vol. 15, no. 5, pp. 3560-3572, 2016.
[24] J. Wen and X. Chang, "Gfclll: A greedy selection-based approach for fixed-complexity LLL reduction," IEEE Commun. Lett., vol. 21, no. 9, pp. 1965-1968, 2017.
[25] D. Seethaler, G. Matz, and F. Hlawatsch, "Low-complexity MIMO data detection using seysen's lattice reduction algorithm," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Honolulu, Hawaii, USA, pp. 53-56, 2007.
[26] W. Zhang, X. Ma, and A. Swami, "Designing low-complexity detectors based on seysen's algorithm," IEEE Trans. Wireless Communications, vol. 9, no. 10, pp. 3301-3311, 2010.
[27] Q. Zhou and X. Ma, "Element-based lattice reduction algorithms for large MIMO detection," IEEE J. Sel. Areas Commun., vol. 31, no. 2, pp. 274-286, 2013.
[28] -, "Improved element-based lattice reduction algorithms for wireless communications," IEEE Trans. Wirel. Commun., vol. 12, no. 9, pp. 4414-4421, 2013.
[29] S. Lyu and C. Ling, "Sequential lattice reduction," in Proc. 8th Int. Conf. Wirel. Comm. \& Signal Process. (WCSP), Yangzhou, China, pp. 1-5, 2016.
[30] B. M. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," IEEE Trans. Commun., vol. 51, no. 3, pp. 389-399, 2003.
[31] S. Liu, C. Ling, and D. Stehlé, "Decoding by sampling: A randomized lattice algorithm for bounded distance decoding," IEEE Trans. Inf. Theory, vol. 57, no. 9, pp. 5933-5945, 2011.
[32] C. Ling, "On the proximity factors of lattice reduction-aided decoding," IEEE Trans. Signal Process., vol. 59, no. 6, pp. 2795-2808, 2011.
[33] L. Babai, "On lovász' lattice reduction and the nearest lattice point problem," Combinatorica, vol. 6, no. 1, pp. 1-13, 1986.
[34] P. Q. Nguyen and D. Stehlé, "Low-dimensional lattice basis reduction revisited," ACM Trans. Algorithms, vol. 5, no. 4, pp. 46:1-46:48, 2009.
[35] J. C. Lagarias, H. W. Lenstra, and C.-P. Schnorr, "Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice," Combinatorica, vol. 10, no. 4, pp. 333-348, 1990.
[36] Y. Aono and P. Q. Nguyen, "Random sampling revisited: Lattice enumeration with discrete pruning," in Proc. Advances in Cryptology EUROCRYPT, Paris, France, 2017, pp. 65-102, 2017.
[37] Z. Wang and C. Ling, "On the geometric ergodicity of metropolishastings algorithms for lattice gaussian sampling," IEEE Trans. Information Theory, vol. 64, no. 2, pp. 738-751, 2018.
[38] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," IEEE Trans. Signal Process., vol. 53, no. 4, pp. 1474-1484, 2005.
[39] A. Gionis, P. Indyk, and R. Motwani, "Similarity search in high dimensions via hashing," in Proc. 25th International Conference on Very Large Data Bases, VLDB'99, Edinburgh, Scotland, UK, pp. 518-529, 1999.
[40] A. Andoni and P. Indyk, "Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions," in Proc. 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS), Berkeley, California, USA, pp. 459-468, 2006.
[41] P. Indyk and R. Motwani, "Approximate nearest neighbors: Towards removing the curse of dimensionality," in Proc. 30th ACM Symposium on the Theory of Computing (STOC), Dallas, Texas, USA, pp. 604-613, 1998.
[42] T. Laarhoven, "Sieving for shortest vectors in lattices using angular locality-sensitive hashing," in Proc. Advances in Cryptology - CRYPTO, Santa Barbara, CA, USA, 2015, pp. 3-22, 2015.
[43] J. Jaldén, D. Seethaler, and G. Matz, "Worst- and average-case complexity of LLL lattice reduction in MIMO wireless systems," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Caesars Palace, Las Vegas, Nevada, USA, pp. 2685-2688, 2008.
[44] D. Wübben, D. Seethaler, J. Jaldén, and G. Matz, "Lattice reduction," IEEE Signal Process. Mag., vol. 28, no. 3, pp. 70-91, 2011.
[45] J. Choi, S. R. Kim, and I. Choi, "Statistical eigen-beamforming with selection diversity for spatially correlated OFDM downlink," IEEE Trans. Vehicular Technology, vol. 56, no. 5, pp. 2931-2940, 2007.
[46] P. Q. Nguyen and D. Stehlé, "Floating-point LLL revisited," in Proc. Advances in Cryptology - EUROCRYPT, Aarhus, Denmark, 2005, pp. 215-233, 2005.
[47] X. Chang, D. Stehlé, and G. Villard, "Perturbation analysis of the QR factor R in the context of LLL lattice basis reduction," Math. Comput., vol. 81, no. 279, pp. 1487-1511, 2012.
[48] P. Xie, H. Xiang, and Y. Wei, "Randomized algorithms for total least squares problems," Numerical Lin. Alg. with Applic., vol. 26, no. 1, 2019.
[49] J. Wen and X. Chang, "Success probability of the babai estimators for box-constrained integer linear models," IEEE Trans. Information Theory, vol. 63, no. 1, pp. 631-648, 2017.
[50] X. Chang, J. Wen, and X. Xie, "Effects of the LLL reduction on the success probability of the babai point and on the complexity of sphere decoding," IEEE Trans. Information Theory, vol. 59, no. 8, pp. 4915-4926, 2013.

Shanxiang Lyu received the B.S. and M.S. degrees in electronic and information engineering from South China University of Technology, Guangzhou, China, in 2011 and 2014, respectively, and the Ph.D. degree from the Electrical and Electronic Engineering Department, Imperial College London, in 2018. He is currently a lecturer with the College of Cyber Security, Jinan University. His research interests are in lattice theory, algebraic number theory, and their applications.

Jinming Wen received his Bachelor degree in Information and Computing Science from Jilin Institute of Chemical Technology, Jilin, China, in 2008, his M.Sc. degree in Pure Mathematics from the Mathematics Institute of Jilin University, Jilin, China, in 2010, and his Ph.D degree in Applied Mathematics from McGill University, Montreal, Canada, in 2015. He was a postdoctoral research fellow at Laboratoire LIP (from March 2015 to August 2016), University of Alberta (from September 2016 to August 2017) and University of Toronto (from September 2017 to August 2018). He has been a full professor in Jinan University, Guangzhou since September 2018. His research interests are in the areas of lattice reduction and sparse recovery. He has published around 50 papers in top journals (including Applied and Computational Harmonic Analysis, IEEE Transactions on Information Theory/Signal Processing/Wireless Communications) and conferences. He is an Associate Editor of IEEE Access.

Jian Weng received the B.S. and M.S. degrees in computer science from South China University of Technology, Guangzhou, China, in 2001 and 2004, respectively, and the Ph.D. degree in computer science from Shanghai Jiao Tong University, Shanghai, China, in 2008. He is currently a Professor and the Executive Dean with the College of Information Science and Technology, Jinan University, Guangzhou, China. He has authored/coauthored 80 papers in international conferences and journals, such as CRYPTO, EUROCRYPT, ASIACRYPT, TCC, PKC, CT-RSA, IEEE TPAMI, IEEE TDSC, etc. His research areas include public key cryptography, cloud security, blockchain, etc. Prof. Weng was a recipient of the Young Scientists Fund of the National Natural Science Foundation of China in 2018, and the Cryptography Innovation Award from Chinese Association for Cryptologic Research (CACR) in 2015. He served as the General Co-Chair for SecureComm 2016, TPC Co-Chairs for RFIDsec13 Asia, and ISPEC 2011, and program committee members for more than 40 international cryptography and information security conferences. He also serves as an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Cong Ling (S'99-A'01-M'04) received the B.S. and M.S. degrees in electrical engineering from the Nanjing Institute of Communications Engineering, Nanjing, China, in 1995 and 1997, respectively, and the Ph.D. degree in electrical engineering from the Nanyang Technological University, Singapore, in 2005. He had been on the faculties of the Nanjing Institute of Communications Engineering and King's College. He is currently a Reader (Associate Professor) with the Electrical and Electronic Engineering Department, Imperial College London. His research interests are coding, information theory, and security, with a focus on lattices. Dr. Ling has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.


[^0]:    This work was supported in part by the National Natural Science Foundation of China under Grants 61902149, 61932010 and 11871248, in part by the Major Program of Guangdong Basic and Applied Research under Grant 2019B030302008, in part by the Natural Science Foundation of Guangdong Province under Grants 2019B010136003 and 2019B010137005, in part by the Fundamental Research Funds for the Central Universities under Grant 21618329, and in part by the National Key R\&D Plan of China under Grants 2017YFB0802203 and 2018YFB1003701. This paper was presented in part at the 8th International Conference on Wireless Communications and Signal Processing (WCSP), Yangzhou, China, Oct. 2016.
    S. Lyu, J. Wen and J. Weng are with the College of Cyber Security, Jinan University, Guangzhou 510632, China (e-mail: shanxianglyu@gmail.com, jinmingwen1@163.com, cryptjweng@gmail.com). S. Lyu is also with the State Key Laboratory of Cryptology, P.O. Box 5159, Beijing, 100878, China.
    C. Ling is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, United Kingdom (e-mail: cling@ieee.org).

[^1]:    ${ }^{1}$ Unless otherwise specified, B is chosen from the dual of a channel matrix.
    ${ }^{2}$ Choosing $\tau>1$ may make the algorithm diverge.

