Minkowski and KZ reduction of nearly orthogonal lattice bases

Sanjeeb Dash, Ramesh Neelamani, and Gregory B. Sorkin*

November 9, 2011

Abstract

We prove that if a lattice basis is *nearly orthogonal* (the angle between any basis vector and the linear subspace spanned by the other basis vectors is at least $\frac{\pi}{3}$ radians), then a KZ-reduced basis can be obtained from it in polynomial time. We also show that if a nearly orthogonal lattice basis has *nearly equal* vector lengths (within a certain constant factor of one another), then the basis is Minkowski reduced. We use these results to show that m i.i.d. random vectors drawn from a uniform distribution over the unit ball in $\mathbb{R}^n$ form a Minkowski-reduced basis of the lattice generated by the vectors asymptotically almost surely as n tends to infinity, if $m \leq cn$ for any constant $c < 1/4$, and form a KZ-reduced basis for $c < 1/5$. The result on Minkowski reduction in lattices generated by random vectors extends a result of Donaldson (1979) who proved this property for fixed m as n tends to infinity.

Keywords

lattices, shortest lattice vector, random lattice, Minkowski reduction, KZ reduction

1 Introduction

An m-dimensional lattice $\mathcal{L}$ in $\mathbb{R}^n$ is a set $\mathcal{L} = \{u_1 b_1 + u_2 b_2 + \cdots + u_m b_m \,|\, u_i \in \mathbb{Z}\}$ of all integer linear combinations of m linearly independent vectors $b_1, \ldots, b_m$ in $\mathbb{R}^n$, and the set $B = \{b_1, \ldots, b_m\}$ is called a *basis* of $\mathcal{L}$. Neelamani, Dash and Baraniuk [14] define a lattice basis to be θ-*orthogonal* if the angle between any basis vector and the linear subspace spanned by the remaining basis vectors is at least θ. They call a θ-orthogonal basis *nearly orthogonal* if θ is at least $\frac{\pi}{3}$ radians. They prove that a $\frac{\pi}{3}$-orthogonal basis always contains a shortest non-zero lattice vector. Thus, the shortest lattice vector problem (SVP) becomes trivial for a $\frac{\pi}{3}$-orthogonal basis, just as it is trivial for an orthogonal basis.

We prove additional properties of $\frac{\pi}{3}$-orthogonal bases in this paper. We show that such a basis is *nearly KZ-reduced*, in the sense that a $\frac{\pi}{3}$-orthogonal basis can be transformed into a KZ-reduced basis in polynomial time. Secondly, we show that if all vectors of a θ-orthogonal ($\theta \geq \frac{\pi}{3}$) basis have lengths no more than $2/\cos\theta$ times the length of the shortest basis vector, then the basis is *Minkowski reduced* for some ordering of the vectors.

Various authors have studied lattice problems in the context of a random lattice, i.e., a lattice generated by (whose basis is) a set of random vectors. Donaldson [6] proved that for fixed m, a random $n \times m$ matrix chosen uniformly from the unit sphere in $\mathbb{R}^{nm}$ defines a Minkowski reduced basis of the lattice generated

*S. Dash and G.B. Sorkin are with the Department of Mathematical Sciences, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598; Email: {sanjeebd,sorkin}@us.ibm.com. R. Neelamani is with the ExxonMobil Exploration Company, 222 Benmar, Houston , TX 77060; Email: ramesh.neelamani@exxonmobil.com.

by its columns a.a.s. as $n \to \infty$. (An event is true almost surely if it is true with probability 1; it is true asymptotically almost surely, or a.a.s., if it is true with probability tending to 1 as $n \to \infty$.) It follows that as $n \to \infty$, one of the columns of the random matrix is a shortest lattice vector a.a.s. In earlier work, Daudé and Vallée [5] showed that in a random lattice generated by n vectors chosen independently and uniformly from the unit ball in $\mathbb{R}^n$, the expected number of iterations of the LLL algorithm [10] is $O(n^2 \log n)$. Now consider an m-dimensional lattice in $\mathbb{R}^n$ generated by m vectors chosen independently and uniformly from the unit ball in $\mathbb{R}^n$; these vectors form a basis because they are linearly independent almost surely. Recently, Akhavi, Marckert and Rouault [2] proved that as $n - m \to \infty$, the probability that this basis is "nearly LLL-reduced" [10] tends to 1. More precisely, they show that such a basis satisfies the *size condition* in the LLL algorithm (condition (i) in Definition 3). Neelamani, Dash and Baraniuk [14] show that for $c = 0.071$, if $m \leq cn$, then as $n \to \infty$, this basis is a.a.s. $\frac{\pi}{3}$-orthogonal. In Section 4, we improve the value of c to 1/4. We show that if c is any constant less than 1/4, then as $n \to \infty$, $m \leq cn$ i.i.d. vectors chosen from a uniform distribution over the unit ball in $\mathbb{R}^n$ form a Minkowski reduced basis a.a.s. (for some re-ordering of the vectors). If $c < 1/5$, then $m \leq cn$ vectors from the above distribution form a KZ-reduced basis a.a.s. These results imply Donaldson's result cited above.

2 Definitions

Consider an m-dimensional lattice $\mathcal{L}$ in $\mathbb{R}^n$, $m \leq n$. By an *ordered basis* of $\mathcal{L}$, we mean a basis with a given ordering of the basis vectors. We represent such a basis by an ordered set, for example $(b_1, \ldots, b_m)$, or as the columns of a matrix B. We represent an unordered basis as $\{b_1, \ldots, b_m\}$. We refer to the Euclidean norm of a vector v as its length, and denote it by $\|v\|$. We use $\lambda(\mathcal{L})$ to denote the length of a shortest non-zero vector in a lattice $\mathcal{L}$. The *(linear) span* of a set of vectors $b_1, \ldots, b_m$ is the set $\{a_1 b_1 + \cdots + a_m b_m \mid a_i \in \mathbb{R}\}$ of all (real) linear combinations of the set of vectors (not just integer linear combinations as for a lattice).

Definition 1 (Minkowski reduced)**.** *An ordered basis $(b_1, \ldots, b_m)$ is* Minkowski reduced *if b_1 is a shortest lattice vector, and for $i = 2, 3, \ldots, m$, b_i is a shortest vector among lattice vectors not in the span of $\{b_1, \ldots, b_{i-1}\}$.*

Every lattice has a basis which is Minkowski reduced. In definitions like the above, in our randomized setting we could replace "a shortest vector" with "the shortest vector", as it is almost surely unique up to sign, but we will stay with the more general phrasing.

The QR decomposition, or thin QR factorization, of an $n \times m$ matrix B is its representation as $B = QR$ where Q is an $n \times m$ orthonormal matrix ($Q^T Q = I$) and R is an $m \times m$ upper-triangular matrix.

The following definition comes from Kannan [8].

Definition 2 (Proper)**.** *An ordered basis $B = QR$, $R = (r_{ij})$, is* proper *if $r_{ij} \in (-|r_{ii}|/2, |r_{ii}|/2]$ for all $i, j \in 1, \ldots, m$.*

Lovász [12] refers to a proper basis as a *weakly reduced* basis. Any lattice has a proper basis; indeed, any basis can be converted to a proper basis (and this is part of the LLL algorithm). We will show how to make an upper-triangular basis R proper by adding appropriate multiples of earlier columns to the ith column; with $B = QR$ this corresponds to similar column operations on B (since column operations can be expressed as right-multiplication). Taking $i = 2, \ldots, n$ in turn, and for each i taking $j = i - 1, \ldots, 1$, simply subtract from column i the multiple of column j bringing r_{ij} into the range $(-|r_{ii}|/2, |r_{ii}|/2]$.

Definition 3 (LLL-reduced)**.** *For $\delta > 0$, a basis $B = QR$, $R = (r_{ij})$, is* δ-LLL-reduced *if*

(i) for $i = 1, \ldots, m-1$, $|r_{i+1,i+1}| \geq \delta |r_{ii}|$ *("size condition"), and*
(ii) it is proper.

For any fixed δ in the interval $(0, \sqrt{3}/2)$, for any lattice, in polynomial time the LLL algorithm [10] obtains a δ-LLL-reduced basis with the property that the first basis vector has length at most $\lambda(\mathcal{L})/\delta^{n-1}$. (For $\delta > \sqrt{3}/2$ a δ-LLL-reduced basis need not exist; with a bit of checking it can be seen that $\sqrt{3}/2$ is best possible for the lattice generated by $(1, 0)$ and $(1/2, \sqrt{3}/2)$.)

We define $\mathrm{proj}_{\perp b_1, \ldots, b_i}(v)$ to be the projection of the vector v on the subspace orthogonal to the span of $b_1, \ldots, b_i$.

The next definition is from Kannan [8], who attributes the notion to Korkine and Zolotarev [7].

Definition 4 (KZ reduced)**.** *An ordered basis* $B = (b_1, \ldots, b_n)$ *is* KZ-reduced *if:*
(i) b_1 *is a shortest lattice vector;*
(ii) the basis is proper;
(iii) in the lattice generated by $B' = \{\mathrm{proj}_{\perp b_1}(b_2), \ldots, \mathrm{proj}_{\perp b_1}(b_n)\}$, B' *is a KZ-reduced basis.*

Kannan also describes a slightly weaker condition without giving it a name.

Definition 5 (nearly KZ reduced)**.** *A basis is* nearly KZ-reduced *if it satisfies*
(i) b_1 *is a shortest lattice vector;*
(ii) in the lattice generated by $B' = \{\mathrm{proj}_{\perp b_1}(b_2), \ldots, \mathrm{proj}_{\perp b_1}(b_n)\}$, B' *is a nearly KZ-reduced basis.*

Note that the recursive requirement of properness in KZ-reducedness is implied by properness of the original basis B. As usual it suffices to consider an upper-triangular basis. The recursive statement simply refers to a submatrix with the first i rows and columns eliminated, for which properness follows instantly from that of the whole matrix. By the same token, and as shown by Kannan [8], a nearly KZ-reduced basis can be transformed to a KZ-reduced basis (i.e., properness attained) in polynomial time, via the procedure used in LLL-reduction to make a basis proper, as outlined below Definition 2; it is easy to check that this procedure does not spoil the "shortest lattice vector" condition (i).

Kannan [8] shows how to construct a KZ-reduced basis for any lattice (naturally this is an exponential-time algorithm, since even finding a shortest vector b_1 is not known to be in the polynomial-time class P). It is also an easy observation from [8] that a KZ-reduced basis is also $(\sqrt{3}/2)$-LLL-reduced.

The next definitions come from Neelamani, Dash and Baraniuk [14].

Definition 6 (θ-orthogonal; nearly orthogonal)**.** *A basis* $\{b_1, \ldots, b_m\}$ *is* θ-orthogonal *if, for each* i*, the angle between* b_i *and the linear subspace spanned by the other basis vectors is at least* θ*. It is* nearly orthogonal *if it is* $\frac{\pi}{3}$*-orthogonal (we will describe all angles in radians).*

Neelamani, Dash and Baraniuk [14] prove the following result.

Theorem 1. *A* $\frac{\pi}{3}$*-orthogonal basis contains a shortest non-zero lattice vector.*

3 Main results

For any set of vectors $\{b_1, \ldots, b_m\}$, define $b'_i = \mathrm{proj}_{\perp b_1}(b_i)$.

Lemma 2. *Let* B*, with columns* $b_1, \ldots, b_m$*, be a* θ*-orthogonal collection of vectors for some* $\theta > 0$*. Then* B'*, with columns* $\{b'_2, \ldots, b'_m\}$*, is also a* θ*-orthogonal collection.*

Proof. Without loss of generality, it suffices to show that the angle between b'_m and the subspace spanned by $\{b'_2, \ldots, b'_{m-1}\}$ lies in the range $[\theta, \frac{\pi}{2}]$. Taking the QR decomposition $B = QR$ of the basis, it is equivalent to prove the same for R (the orthogonal transformation Q preserves all angles); equivalently, we may assume without loss of generality that B is upper triangular. Note that B' is simply B with its first column and row deleted. Writing $\theta_m \geq \theta$ for the angle between b_m and the subspace spanned by $b_1, \ldots, b_{m-1}$,

$$\sin^2(\theta_m) = \frac{r_{mm}^2}{\sum_{k=1}^{m} r_{km}^2}.$$

Writing θ'_m for the angle between b'_m and the subspace spanned by $b'_2, \ldots, b'_{m-1}$,

$$\sin^2(\theta'_m) = \frac{r_{mm}^2}{\sum_{k=2}^{m} r_{km}^2} \geq \sin^2(\theta_m).$$

We conclude that $\theta'_m \geq \theta_m \geq \theta$. Therefore $\{b'_2, \ldots, b'_m\}$ is θ-orthogonal. □

Definition 7 (Projection ordered)**.** *An ordered basis $B = b_1, \ldots, b_n$ is* projection-ordered *if (i) b_1 is a shortest basis vector, and (ii) $B' = \{b'_2, \ldots, b'_m\}$ is a projection-ordered basis of the lattice it generates.*

Any basis can be projection-ordered in polynomial time: just choose a shortest basis vector as the first, then recursively projection-order the projections of the others.

Theorem 3. *A projection-ordered $\frac{\pi}{3}$-orthogonal basis is nearly KZ-reduced.*

Proof. We prove this by induction on the dimension m of the lattice. The statement is vacuously true for $m = 1$. Let $B = (b_1, \ldots, b_m)$ be a projection-ordered $\frac{\pi}{3}$-orthogonal basis of an m-dimensional lattice. For the first property of KZ-reducedness, Theorem 1 shows that b_1 is a shortest lattice vector (not merely a shortest basis vector). For the second property, by definition, $B' = (b'_2, \ldots, b'_m)$ is projection-ordered, by Lemma 2 it is $\frac{\pi}{3}$-orthogonal, and of course it is of dimension $m - 1$, so by the inductive hypothesis it is nearly KZ-reduced. □

Therefore, a KZ-reduced basis can be obtained from a $\frac{\pi}{3}$-orthogonal basis in polynomial time, by projection-ordering it and then applying the algorithm to make a matrix proper.

Theorem 4. *Let $B = \{b_1, \ldots, b_m\}$ be a θ-orthogonal basis for a lattice $\mathcal{L}$ with $\theta \geq \frac{\pi}{3}$. Further, suppose that*

$$\frac{\min_i \|b_i\|}{\max_i \|b_i\|} \geq 2\cos\theta. \tag{1}$$

Then some ordering of the basis is Minkowski reduced.

Note that as θ increases above $\frac{\pi}{3}$, the first hypothesis becomes stronger and the second one weaker. The theorem is easy to show in the extreme cases. If $\theta = \frac{\pi}{2}$, the basis is orthogonal, and when ordered by increasing length is clearly Minkowski-reduced. If $\theta = \frac{\pi}{3}$, $2\cos\theta = 1$, equation (1) and Theorem 1 imply that all the basis vectors have lengths equal to $\lambda(\mathcal{L})$, and the basis is trivially Minkowski-reduced.

Proof. Let B be a basis satisfying the hypotheses of Theorem 4. Assume without loss of generality that B is scaled and ordered such that

$$1 = \|b_1\| \leq \|b_2\| \leq \cdots \leq \|b_m\|, \tag{2}$$

so that (1) implies $\|b_1\| / \|b_m\| \geq 2\cos\theta$.

We claim that any lattice vector $w = \sum u_i b_i$ is at least as long as the longest basis vector appearing with nonzero coefficient, i.e., $\|w\| \geq \|b_k\|$ where $k = \max\{i\colon u_i \neq 0\}$. Then for any $i \in \{1, \ldots, m\}$, b_i is a shortest nonzero lattice vector not contained in the span of $b_1, \ldots, b_{i-1}$, because any such lattice vector w has $k \geq i$, and thus (by the claim) $\|w\| \geq \|b_k\| \geq \|b_i\|$. Therefore B is Minkowski-reduced.

We now prove the claim. Let $v = \sum_{i<k} u_i b_i$, so that $w = v + u_k b_k$. Let θ' be the angle between b_k and v, and let $l_1 = \|v\|$ and $l_2 = \|b_k\|$. Theorem 1 implies that $l_1 \geq \lambda(\mathcal{L}) = \|b_1\|$. Observe that

$$\begin{aligned}\|w\|^2 = \|v + u_k b_k\|^2 &= (l_1 \pm u_k l_2 \cos\theta')^2 + (u_k l_2 \sin\theta')^2 \\ &= l_1^2 + u_k^2 l_2^2 \pm 2u_k l_1 l_2 \cos\theta' \\ &\geq l_1^2 + u_k^2 l_2^2 - |u_k| l_1,\end{aligned}$$

using $2l_2\cos\theta' \leq \|b_1\| = 1$ from (1) and (2). Therefore

$$\|w\|^2 - \|b_k\|^2 \geq l_1^2 + u_k^2 l_2^2 - |u_k| l_1 - l_2^2. \tag{3}$$

If $|u_k| = 1$ then the right hand side of (3) is $l_1^2 - l_1 \geq 0$ as $l_1 \geq 1$. Otherwise $|u_k| \geq 2$, and the right hand side of (3) satisfies

$$\begin{aligned}l_1^2 + u_k^2 l_2^2 - |u_k| l_1 - l_2^2 &\geq l_1^2 + u_k^2 l_2^2 - |u_k| l_1 l_2 - l_2^2 \\ &= \left(l_1 - \tfrac{1}{2}|u_k| l_2\right)^2 + \tfrac{3}{4} u_k^2 l_2^2 - l_2^2 \\ &\geq \left(l_1 - \tfrac{1}{2}|u_k| l_2\right)^2 + 2l_2^2 \\ &\geq 0.\end{aligned}$$

Either way, this proves that $\|w\|^2 - \|b_k\|^2 \geq 0$, establishing the claim.

□

4 Random lattices

We define a random lattice as one generated by (the columns of) a random matrix, considering several distributions over $n \times m$ matrices B:

(i) entries of B are standard normal random variables $N(0, 1)$;
(ii) each column of B is an independent random vector uniform over the unit ball in $\mathbb{R}^n$;
(iii) B as a whole is drawn uniformly from the unit ball in $\mathbb{R}^{nm}$;
(iv) each entry of B is an independent symmetric Bernoulli random variable ("symmetric" meaning taking values $+1$ or -1) with success probability $1/2$.

We call distributions (i), (ii), and (iii) *simple Gaussian distributions*; as we will discuss later, they are almost exactly equivalent for our purposes. We consider distributions (i)-(iii) (and not other similar distributions such as vectors drawn uniformly from the unit sphere in $\mathbb{R}^n$) as they are discussed in [14], [2] and [6], respectively, and we wish to compare our results with those in the above papers. We note that (i) is equivalent to choosing the column vectors of B from the standard Gaussian measure over $\mathbb{R}^n$, or B itself from that over $\mathbb{R}^{mn}$. Trivially, the columns of B are almost surely linearly independent (making B a lattice basis) in the first three cases if $m \leq n$. In the last case, the columns of B are a.a.s. linearly independent: indeed [3] shows a much stronger result that for any fixed c with $0 < c < 1$, as $n \to \infty$ with $m = cn$, the smallest

singular value of $B/\sqrt{n}$ is a.a.s. almost exactly $1-\sqrt{c}$. (This result of [3] is further strengthened in [11], which allows $m \leq cn$ and gives a concentration inequality for the smallest singular value of $B/\sqrt{n}$.)

The ideas and techniques we use here are similar to those in Daudé and Vallée [5] and Akhavi, Marckert and Rouault [2]. However, our results are for the case $m < n$, applying only to non full-dimensional lattices.

We rely on the following concentration inequality of Cirelson, Ibragimov and Sudakov [4], as given by Massart [13, Thm 3.4]. (We give just one of two inequalities from [13, Thm 3.4].) It is far more general than we need, but simple and convenient.

Theorem 5. *Consider some Lipschitz function ζ on the Euclidean space $\mathbb{R}^N$ with Lipschitz constant L. If P denotes the canonical Gaussian measure on $\mathbb{R}^N$, and M the corresponding mean, or median, of ζ, then for every $\lambda \geq 0$,*

$$\mathbb{P}\big(|\zeta - M| \geq \lambda\big) \leq 2\exp\left(-\tfrac{\lambda^2}{2L^2}\right). \tag{4}$$

In the context of the theorem above we will need a convenient asymptotic formula for the mean or median of a Chi-distributed random variable.

Lemma 6. *A random variable $X \sim \chi_k$ has mean $\mu = \sqrt{k}(1 + O(1/k))$.*

Proof. It is well known that

$$\mu = \sqrt{k}\frac{\Gamma((k+1)/2)}{\Gamma(k/2)}.$$

Substituting

$$\Gamma(z) = \sqrt{2\pi/z}(z/e)^z(1 + O(1/z))$$

gives

$$\mu = \sqrt{k}(1+1/k)^{k/2}e^{-1/2}(1 + O(1/k)).$$

The logarithm of the middle term is

$$\tfrac{k}{2}\ln(1+\tfrac{1}{k}) = \tfrac{k}{2}(\tfrac{1}{k} - \tfrac{1}{2k^2} + \cdots) = \tfrac{1}{2} + O(\tfrac{1}{k}),$$

and for $x < 1$ we have $\exp(x) = 1 + O(x)$ so that the term itself is

$$(1+1/k)^{k/2} = \exp(1/2 + O(1/k)) = \exp(1/2)\exp(O(1/k)) = \sqrt{e}(1 + O(1/k)).$$

Thus,

$$\mu = \sqrt{k}(1 + O(1/k)).$$

□

(As an alternative to Lemma 6, an adequate estimate of the median can be obtained even more easily if a bit more convolutedly: X^2 has mean k^2 and variance $2k$, by Markov's inequality this implies that its median is within $2\sqrt{2}k$ of the mean, and the median of X is the square root of that of X^2.)

Neelamani, Dash and Baraniuk [14] showed that lattice bases defined by distributions (i), (ii) and (iv) above are a.a.s. $\frac{\pi}{3}$-orthogonal as n tends to infinity and $m \leq 0.071n$. For the first two distributions, we improve the ratio of m to n based on the following lemma.

Lemma 7. *Let* $\mathbf{X}$ *be a vector* $(X_1, \ldots, X_n)$ *whose entries are independent standard normal random variables,* $X_i \sim N(0,1)$*. Let* $c \in (0,1)$*. As* $n \to \infty$*, the (random) angle between* $\mathbf{X}$ *and the subspace* $\mathbf{L}$ *spanned by* m *similar, independent random vectors* $\mathbf{X}_1, \ldots, \mathbf{X}_m$ *with* $m = \lfloor cn \rfloor$ *is within* $n^{-0.1}$ *radians of*

$$\tan^{-1}\left(\sqrt{\tfrac{1}{c} - 1}\right)$$

with probability $1 - \exp(-\Omega(n))$*.*

Proof. $\mathbf{L}$ is almost surely m-dimensional. It is well known that the squared length of $\mathbf{X}$ has chi-squared distribution χ^2_n, and its orientation is uniform over a unit sphere in $\mathbb{R}^n$. By the latter property, the (random) angle between $\mathbf{X}$ and $\mathbf{L}$ is identically distributed to the angle between $\mathbf{X}$ and any fixed m-dimensional subspace $\mathbf{L}'$. For convenience, let $\mathbf{L}'$ be the subspace spanned by the m unit vectors $(1,0,0,\ldots)$, $(0,1,0,\ldots)$, *etc*.

The projection of $\mathbf{X}$ onto $\mathbf{L}'$ is

$$\mathbf{X}' = (X_1, \ldots, X_m, 0, \ldots, 0), \tag{5}$$

and the orthogonal component of $\mathbf{X}$ with respect to $\mathbf{L}'$ is

$$\mathbf{X}'' = (0, \ldots, 0, X_{m+1}, \ldots, X_n). \tag{6}$$

The angle ϕ between $\mathbf{X}$ and $\mathbf{L}'$ is precisely the angle between $\mathbf{X}$ and $\mathbf{X}'$, namely

$$\tan^{-1}\left(\left\|X''\right\| / \left\|X'\right\|\right), \tag{7}$$

where $\|X''\| \sim \chi_{n-m}$ and $\|X'\| \sim \chi_m$ (and, though we will not use this fact, the two variables are independent). Since n and m are both large ($m = \lfloor cn \rfloor$ and $n \to \infty$), both $\|X''\|$ and $\|X'\|$ are concentrated random variables, and their ratio is also concentrated, as we now show.

For a random vector $\mathbf{X} = (X_1, \ldots, X_k)$ playing the role of $\mathbf{X}'$ or $\mathbf{X}''$ above depending on the choice of k, in Theorem 5, let $\zeta(X) = \|X\|$. Then ζ is a Lipschitz function with Lipschitz constant $L = 1$. By Lemma 6, $\mathbb{E}[\zeta(X)] = \mu = \sqrt{k}(1 + O(1/k))$, and for n sufficiently large, $|\mathbb{E}[\zeta] - \sqrt{k}| \leq 1$. Then Theorem 5 yields

$$\mathbb{P}\big(|\zeta - \sqrt{k}| \geq 1 + \lambda\big) \leq 1 \exp\left(-\tfrac{1}{2}\lambda^2\right).$$

Take $\lambda = \frac{1}{2}k^{0.4}$. For n sufficiently large, $\frac{1}{2}k^{0.4} \geq 1$, whereupon

$$\mathbb{P}\big(|\zeta - \sqrt{k}| \geq k^{0.4}\big) \leq 2 \exp\left(-\tfrac{1}{8}k^{0.8}\right).$$

Since $k = \Theta(n)$, another way to put it is that with probability $1 - \exp(-\Omega(n^{0.8}))$,

$$\|X\| = \zeta = \sqrt{k} + O(k^{0.4}) = \sqrt{k}\,(1 + O(n^{-0.1})). \tag{8}$$

By the union bound, (8) holds for both X' and X'' with probability $1 - \exp(-\Omega(n^{0.8}))$, in which case

$$\begin{aligned}\frac{\|X''\|}{\|X'\|} &= \frac{\sqrt{n-m}}{\sqrt{m}}\,(1 + O(n^{-0.1})) = \sqrt{1/c - 1}\,(1 + O(n^{-0.1})) \\ &= \sqrt{1/c - 1} + O(n^{-0.1}).\end{aligned}$$

Since $\tan^{-1}$ has bounded derivative, in the same event as above,

$$\begin{aligned}\phi &= \tan^{-1}\left(\|X''\| / \|X'\|\right) \\ &= \tan^{-1}\left(\sqrt{1/c-1}\right) + O(n^{-0.1}).\end{aligned} \tag{9}$$

□

The following lemma draws a trivial corollary to Lemma 7 and extends it to simple Gaussian matrices of types (ii) and (iii).

Lemma 8. *For any $c \in (0,1)$, let $m = \lfloor cn \rfloor$ and let B be an $n \times m$ matrix drawn from a simple Gaussian distribution* (i)*,* (ii)*, or* (iii)*. As $n \to \infty$, the (random) angle between any column of B and the subspace spanned by the remaining columns is within $n^{-0.1}$ radians of $\tan^{-1}\left(\sqrt{1/c-1}\right)$ with probability $1 - \exp(-\Omega(n))$.*

Proof. For the standard Gaussian distribution (i), we know that (8) holds for each column with failure probability $\exp(-\Omega(n^{0.8}))$; by the union bound over the n columns, it holds for *every* column, with probability $1 - \exp(-\Omega(n^{0.8}))$.

The rest we derive from this, the well known fact that all these distributions are isotropic, and the fact that the column vector lengths are predictable in an almost-sure, almost exact sense. Specifically, we will now show that for any $c \in (0,1]$, with $n \to \infty$ and $m = \lfloor cn \rfloor$, in a matrix B drawn from any of these distributions, a.a.s. every column has almost exactly the same length. The details are routine, but we give them for the sake of completeness.

The uniform distribution (ii) over the unit ball puts measure proportional to r^n on the set of vectors of length $r \leq 1$, so the probability a random vector has length $r \leq R$ is $\int_0^R r^n dr / \int_0^1 r^n dr = R^{n+1}$. Taking $R = 1 - n^{-0.2}$, with probability $1 - \exp(-\Omega(n^{0.8}))$ each of the m vectors has length $1 + O(n^{-0.2})$. A matrix drawn from distribution (iii) can be generated by drawing a matrix from distribution (i) and rescaling its entries to give net length distributed over $r \in [0,1]$ with density proportional to r^{nm}. By our argument for (i), the column lengths are a.a.s. $\sqrt{n}(1 + O(n^{-0.1}))$ before rescaling, implying a net length of $\sqrt{nm}(1 + O(n^{-0.1}))$, and by the same argument as for (ii) the target length is a.a.s. $1 + O(n^{-0.2})$, from which it follows that a.a.s. every rescaled column length is $\sqrt{n} \cdot 1/\sqrt{nm} \cdot (1 + O(n^{-0.2}))$. Each "a.a.s. " refers to a failure probability of order $\exp(-\Omega(n^{0.8}))$, and the union bound leaves this unchanged. □

As a corollary of Lemma 8 we obtain the following.

Theorem 9. *Let $c \in (0,1)$. With $m \leq cn$ and $n \to \infty$, a simple Gaussian matrix B of type* (i)*,* (ii) *or* (iii) *is a.a.s. ϕ-orthogonal, with $\phi = \tan^{-1}(\sqrt{1/c-1}) + O(n^{-0.1})$. If $c < 1/4$, then the matrix is a.a.s. $\frac{\pi}{3}$-orthogonal.*

Proof. We first consider $m = \lfloor cn \rfloor$, in which case the matrix's column vectors $\mathbf{X}_1, \ldots, \mathbf{X}_m$ satisfy the hypotheses of Lemma 8 (the distinction between m and $m+1$ is insignificant). Letting ϕ_i be the angle formed between $\mathbf{X}_i$ and the subspace spanned by the other vectors, Lemma 8 says that with probability $1 - \exp(-\Omega(n^{0.8}))$, B is ϕ-orthogonal with $\phi = \min_i \phi_i = \tan^{-1}(\sqrt{1/c-1}) + O(n^{-0.1})$. Further, solving for $\tan^{-1}\left(\sqrt{1/c-1}\right) = \sqrt{3}$, one gets $c = 1/4$. Therefore, for any constant $c' < 1/4$, as n tends to infinity, a matrix with $\lfloor c'n \rfloor$ i.i.d. columns is a.a.s. $\frac{\pi}{3}$-orthogonal.

If we do not have $m = \lfloor cn \rfloor$ but $m \leq cn$, referring to (5) and (6) we see that making m smaller can only shorten any projection $\mathbf{X}_i'$ and can only lengthen any orthogonal component $\mathbf{X}_i''$, and thus each angle ϕ_i can only become larger. □

We now give our main result regarding random lattices.

Theorem 10. *Let B denote an $n \times m$ matrix drawn from a simple Gaussian distribution, with $m \leq cn$. If $c < 1/4$, then a.a.s. as $n \to \infty$, some ordering of B forms a Minkowski reduced basis of the lattice generated by its columns. If $c < 1/5$, then a.a.s. some ordering of B forms a KZ-reduced basis.*

Proof. Assume $c < 1/4$. Fixing any θ with $\frac{\pi}{3} < \theta < \tan^{-1}(\sqrt{1/c - 1}\,)$, by Theorem 9, B is a.a.s. θ-orthogonal. By the discussion above we know that a.a.s. the column lengths $\|b_i\|$ of B are within a factor $1 + o(1)$ of one another. Henceforth we assume that both "a.a.s. " conditions hold. Since $2\cos\theta < 1$, every ratio $\|b_i\| \,/\, \|b_j\| = 1 + o(1) > 2\cos\theta$. Thus, the columns satisfy the hypotheses of Theorem 4, and we conclude that some ordering of B is Minkowski reduced.

Assume $c < 1/5$. Fixing any θ with $\tan^{-1}(2) < \theta < \tan^{-1}(\sqrt{1/c - 1}\,)$, by Theorem 9, B is a.a.s. θ-orthogonal. Also, a.a.s. the column-lengths of B are within a factor $1 + o(1)$ of one another. Henceforth we assume that both "a.a.s." conditions hold. Without loss of generality, suppose that the columns of B are projection-ordered (recall Definition 7). Consider the decomposition $B = QR$, where $R = (r_{ij})$ is an $m \times m$ upper triangular matrix. Let θ_i be the angle between b_i and the subspace spanned by the preceding columns of B.

For any $i, j \in 1, \ldots, m$, $|r_{ii}| \geq \|b_i\| \sin\theta_i$, while $|r_{ij}| \leq \|b_j\| \cos\theta_j$. As $\theta_i, \theta_j \geq \theta$,

$$\frac{|r_{ij}|}{|r_{ii}|} \leq \frac{\|b_j\| \cos\theta_j}{\|b_i\| \sin\theta_i} \leq \frac{\|b_j\| \cos\theta}{\|b_i\| \sin\theta} = (1 + o(1)) \cot\theta < \frac{1}{2}.$$

Therefore B' is a proper matrix. Since it is also projection-ordered, it is KZ-reduced. □

Finally, observe that in a matrix with ± 1 entries drawn from a symmetric Bernoulli distribution, all columns have equal length. If the Bernoullis have success probability $1/2$ and the matrix is $n \times m$ with $m \leq 0.071n$, it is shown in [14] that it is $\frac{\pi}{3}$-orthogonal a.a.s. as $n \to \infty$. The following theorem is then an immediate consequence of Theorem 4.

Theorem 11. *An $n \times m$ matrix with ± 1 entries drawn from a symmetric Bernoulli distribution with success probability 1/2 is a.a.s. Minkowski reduced as $n \to \infty$ with $m \leq 0.071n$.*

To put our results in context, note that a matrix satisfying the hypotheses of Theorem 10, with $c < 1/5$, is a.a.s. LLL-reduced: this follows from the fact that it is KZ-reduced. Donaldson [6] proved that a matrix drawn from the simple Gaussian distribution (iii) is a.a.s. Minkowski reduced as $n \to \infty$ but m is fixed. Thus, Theorem 10 is a significant extension of his result.

Also, as mentioned earlier, Akhavi, Marckert and Rouault [2] proved that a matrix drawn from the simple Gaussian distribution (ii) satisfies the LLL size-reduction property (Definition 3.(i)) as long as $n - m \to \infty$. By reducing the number of columns permitted we are able to derive the stronger results of the present paper, namely Theorem 10, and the stronger assumptions we make are necessary: from (5), (6), (7), and the proof of Lemma 7, the θ-orthogonality we rely on can be seen to fail if $n - m$ grows very slowly.

References

[1] A. Akhavi. Random lattices, threshold phenomena and efficient reduction algorithms, *Theoretical Computer Science* **287** (2002), 359–385.

[2] A. Akhavi, J. F. Marckert, and A. Rouault, On the reduction of a random basis, in Proceedings of SIAM-ALENEX/ANALCO07, New Orleans, January 07.

[3] Z. D. Bai and Y. Q. Yin, Limit of the Smallest Eigenvalue of a Large Dimensional Sample Covariance Matrix, *Ann. Probab.* **21**(3) (1993), 1275–1294.

[4] B. S. Cirelson, I. A. Ibragimov, and V. N. Sudakov, Norms of Gaussian sample functions. In Proceedings of the Third Japan-USSR Symposium on Probability Theory, Lecture Notes in Mathematics **550** 20–41, Springer-Verlag, Berlin (1976).

[5] H. Daudé and B. Vallée, An upper bound on the average number of iterations of the LLL algorithm, *Theoretical Computer Science* **123** (1994), 95–115.

[6] J. L. Donaldson, Minkowski reduction of integral matrices, *Mathematics of Computation* **33**(no. 145) (1979), 201–216.

[7] A. Korkine and G. Zolotarev, Sur les formes quadratiques, *Math. Annalen*, **6**, (1873), 336–389.

[8] R. Kannan, Algorithmic geometry of numbers, *Annual Review of Comp. Sci.* **2** (1987), 231–267.

[9] M. Ledoux, Isoperimetry and Gaussian analysis, Ecole d'Eté de Probabilités de Saint-Flour 1994, Lecture Notes in Mathematics **1648** 165–294, Springer-Verlag, Berlin (1996).

[10] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, Factoring polynomials with rational coefficients, *Mathematische Annalen* **261** (1982), 515–534.

[11] A. Litvak, A. Pajor, M. Rudelson, N. Tomczak-Jaegermann, Smallest singular value of random matrices and geometry of random polytopes, *Adv. Math.* **195**(2) (2005), 491–523.

[12] L. Lovász, An algorithmic theory of numbers, graphs and convexity, CBMS-NSF REgional conference series in applied mathematics, SIAM (1986).

[13] P. Massart, Concentration Inequalities and Model Selection, Ecole d'Eté de Probabilités de Saint-Flour XXXIII - 2003 Lecture Notes in Mathematics **1896**, Springer-Verlag, Berlin (2007).

[14] R. Neelamani, S. Dash, and R. G. Baraniuk, On nearly orthogonal lattice bases and random lattices, *SIAM Journal on Discrete Mathematics* **21**(no. 1) (2007) 199–219.

[15] P. Nguyen, D. Stehle. LLL on the average, *Proceedings of the 7th Algorithmic Number Theory Symposium, ANTS 2006, Berlin*, Lecture Notes in Computer Science **4076** 238–256, Springer-Verlag, Berlin (2008).