A Note on Discrete Gaussian Combinations of Lattice Vectors

Divesh Aggarwal∗ Oded Regev†

May 11, 2014

Abstract

We analyze the distribution of $\sum_{i=1}^{m} v_i x_i$ where $x_1, \ldots, x_m$ are fixed vectors from some lattice $\mathcal{L} \subset \mathbb{R}^n$ (say $\mathbb{Z}^n$) and $v_1, \ldots, v_m$ are chosen independently from a discrete Gaussian distribution over $\mathbb{Z}$. We show that under a natural constraint on $x_1, \ldots, x_m$, if the $v_i$ are chosen from a wide enough Gaussian, the sum is statistically close to a discrete Gaussian over $\mathcal{L}$. We also analyze the case of $x_1, \ldots, x_m$ that are themselves chosen from a discrete Gaussian distribution (and fixed).

Our results simplify and qualitatively improve upon a recent result by Agrawal, Gentry, Halevi, and Sahai [AGHS13].

∗Courant Institute of Mathematical Sciences, New York University. Email: divesha@cs.nyu.edu.
†Courant Institute of Mathematical Sciences, New York University. This material is based upon work supported by the National Science Foundation under Grant No. CCF-1320188. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.
1 Introduction

The study of discrete Gaussian distributions has been a crucial step in most of the work in lattice-based cryptography in the last few years. A discrete Gaussian distribution over some fixed lattice \( \mathcal{L} \), denoted as \( \mathcal{D}_{\mathcal{L},s} \) for some parameter \( s > 0 \), is a distribution in which each lattice point is sampled with probability proportional to the probability density function of a continuous Gaussian distribution of width \( s \) evaluated at that point. There have been several results, e.g. [MR07, GPV08, Pei10, BF11, MP13], showing that these functions have many useful properties similar to the corresponding continuous functions. However, some aspects of the discrete Gaussian distribution are not as well understood as those of its continuous counterpart.

**Our Result.** We analyze the following distribution. Let \( x_1, \ldots, x_m \) be \( m \) fixed points in \( \mathbb{Z}^n \) for some \( m > n \). Let \( X \) be the \( n \times m \) matrix formed by the vectors \( x_1, \ldots, x_m \) as the column vectors. For \( r > 0 \), define the distribution \( \mathcal{E}_{X,r} := \{ X \cdot v : v \leftarrow \mathcal{D}_{\mathbb{Z}^m,r} \} \).

We show in Theorem 3.2 in Section 3 that if \( X \) satisfies a certain constraint and if \( r \) is large enough, then the distribution \( \mathcal{E}_{X,r} \) is statistically close to the discrete Gaussian distribution with appropriate covariance (denoted as \( \mathcal{D}_{\mathbb{Z}^n,rX^T} \)). The constraint on \( X \) that we need is that \( X \) has small entries and additionally there are short vectors \( u_1, \ldots, u_n \in \mathbb{Z}^m \) such that \( X \cdot u_i = e_i \) for \( 1 \leq i \leq n \), where \( e_i \) are unit vectors in \( n \) dimensions.

We then show that if the vectors \( x_1, \ldots, x_m \) are chosen according to the discrete Gaussian distribution, then \( X \) satisfies the aforementioned constraint, and hence the following result holds:

**Theorem 1.1 (Informal).** For integers \( n \geq 1 \), \( m = \text{poly}(n) \), \( s = \Omega(\sqrt{\log(n/\varepsilon)}) \), and \( r = \Omega(ns\sqrt{\log(1/\varepsilon)}) \), and \( X \leftarrow (\mathcal{D}_{\mathbb{Z}^n,s})^m \), we have with probability \( 1 - 2^{-n} \), the statistical distance between \( \mathcal{E}_{X,r} \) and \( \mathcal{D}_{\mathbb{Z}^n,rX^T} \) is at most \( \varepsilon \).

The formal statement and proof appear as Theorem 5.1 in Section 5. We also mention there how to obtain a more general version for the case in which the columns of \( X \) are chosen from a discrete Gaussian distribution over an arbitrary lattice \( \mathcal{L} \subset \mathbb{R}^n \); this follows easily from the case of \( \mathbb{Z}^n \).

The motivation for studying the distribution \( \mathcal{E}_{X,r} \) for this choice of \( X \) comes from a recent breakthrough construction of multilinear maps from ideal lattices by Garg, Gentry, and Halevi [GGH13] (see [AGHS13] for details). Also, more recently, our result has been used by Ling et al. [LPSS13] to give a lattice-based traitor tracing scheme.

Our result improves on the main result of Agrawal, Gentry, Halevi, and Sahai [AGHS13] who showed a statement similar to Theorem 1.1 with the bound on \( r \) depending linearly on \( m \). Improving this dependence on \( m \) was left by them as an open question, which we answer here.

**Our Techniques.** As was done in [AGHS13], the first part of our proof is to reduce the question to a question about bounding the successive minima of a so-called orthogonal lattice. Namely, let \( A \subset \mathbb{Z}^m \) be the orthogonal lattice of \( X \), i.e., the lattice formed by all integer vectors in \( \mathbb{Z}^m \) that are orthogonal to all \( n \) rows of \( X \), which is (typically) a lattice of rank \( m - n \). Our goal then is to find \( m - n \) linearly independent vectors in it whose maximum length is as small as possible, i.e., we need to bound the last successive minimum \( \lambda_{m-n} \) of the lattice \( A \). We feel that the techniques used
in this part are rather standard, and instead of repeating the proofs, we cite the relevant lemmas from [AGHS13].

The main part of our proof (and that of [AGHS13]) is therefore bounding $\lambda_{m-n}(A)$. The approach we take is using a simple, yet powerful, idea from a recent result by Kuperberg, Lovett, and Peled [KLP12] (who were interested in an entirely different application, namely, showing the existence of some combinatorial structures like $t$-wise independent permutations). Using this idea, we get that in order to prove a bound on the successive minima of the orthogonal lattice, it suffices to satisfy two constraints: (1) that $X$ has small entries, and (2) that there are short vectors $u_1, \ldots, u_n$ such that $X \cdot u_i = e_i$ for $1 \leq i \leq n$, where $e_i$ are unit vectors in $n$ dimensions.

This already gives us sufficient constraints on $X$ and $r$ under which $E_{X,r}$ is guaranteed to be close to a discrete Gaussian. In order to complete the proof of Theorem 1.1, it suffices to show that the $X$ chosen there satisfies these two constraints with high probability. The first constraint, namely, that $X$ has small entries follows easily by definition. Proving the second constraint is trickier and we show this in Section 4 using a careful application of the pigeon-hole principle.

Comparison with [AGHS13]. As mentioned above, our Theorem 1.1 is similar to the main result of [AGHS13], the main difference being the quantitative improvement in the bound on $r$. Namely, the bound on $r$ obtained by [AGHS13] is $\tilde{\Omega}(mn \log (1/\varepsilon))$ which could be much worse than our bound for large $m$. We note that our result gives a worse bound if $s \gg m$. However, for applications like [GGH13, LPSS13], one chooses $s = \Theta(\sqrt{\log(n/\varepsilon)})$ in which case our result gives a better bound for all $m, n, \varepsilon$.

Another advantage of our approach is that our Theorem 3.2 (which we use to prove Theorem 1.1) applies to a general choice of $X$, and not just to a random $X$. This might be useful in future applications.

In terms of techniques, as mentioned above, the most challenging step in both our proof and the proof in [AGHS13] is to bound the last successive minimum $\lambda_{m-n}$ of the orthogonal lattice $A$, which is a lattice of rank $m - n$ in dimension $m$. The way this is done in [AGHS13] is by first defining a superlattice $A_q$ of $A$ which is of full rank, and then considering the dual lattice $M_q$ of $A_q$. They then obtain a lower bound on $\lambda_{n+1}(M_q)$, which using Banaszczyk’s transference theorem [Ban93], implies an upper bound on $\lambda_{m-n}(A_q)$. Finally, they argue that this is also an upper bound on $\lambda_{m-n}(A)$. This somewhat indirect proof results in the bound on $r$ depending linearly on $m$. Improving this dependence was left as an open question in [AGHS13]. In comparison, our method to bound $\lambda_{m-n}(A)$ is more direct.

Other Related Work. Micciancio and Peikert [MP13] recently showed hardness results for the main lattice-based cryptographic problems. One of the key ingredients in the proof was a new convolution theorem which was a strengthening of a previous similar result by Peikert [Pei10]. This theorem, like our result, also looks at sums of discrete Gaussian samples. The difference is that in their statement the combination vector $v$ is fixed and the matrix $X$ comprises of (spherical) discrete Gaussian with parameter $s$ significantly bigger than the smoothing parameter of the underlying lattice. In our case, the matrix is fixed “once and for all”, and only $v$ varies. In our setting we essentially analyze the sum of 1-dimensional discrete Gaussians in $n$-dimensional space.

We note that by combining the result of [MP13] with that of [AGHS13], it might be possible to derive a result similar to our Theorem 1.1 i.e., to improve the dependence of $r$ on $m$. The idea would be to partition the sum $\sum_{i=1}^m v_i X_i$ into small blocks, argue that each block is close to
a discrete Gaussian (using [AGHS13]) and then arguing that the overall sum must therefore also
be close to a discrete Gaussian (using [MP13]). As this approach would almost certainly lead to
worse parameters and a far more complicated proof, we did not attempt to pursue it.

2 Preliminaries

All logarithms, unless otherwise stated, are to the base 2. Natural logarithms, i.e., to the base e,
are denoted by ln. The norm \( \| \cdot \| \) considered in this paper is the \( l_2 \) norm, unless otherwise stated.
The transpose and inverse of a matrix \( A \) are denoted as \( A^T \) and \( A^{-1} \), respectively. The inverse
transpose of a matrix \( A \) is denoted as \( A^{-T} \).

Let \( D \) be a discrete distribution. We denote by \( D[x] \) the probability it assigns to \( x \), and by
\( X \sim D \) a random variable distributed according to \( D \). For two distributions \( D, D' \) their statistical
distance is \( \Delta(D; D') = \frac{1}{2} \sum_x |D[x] - D'[x]| \).

**Gaussian function.** For any \( s > 0 \), the (spherical) Gaussian function on \( \mathbb{R}^n \) with parameter \( s \) is
defined as \( \rho_s(x) = \exp(-\pi \| x \|^2/s^2) \) for all \( x \in \mathbb{R}^n \). If \( s = 1 \), then it is omitted. For a rank-\( n \) matrix
\( S \in \mathbb{R}^{m \times n} \), the ellipsoidal Gaussian function on \( \mathbb{R}^n \) with covariance matrix \( \Sigma = S^T S \) is defined as:
\[
\forall x \in \mathbb{R}^n, \quad \rho_S(x) = \exp \left( -\pi x^T (S^T S)^{-1} x \right).
\]
When \( S = sI_n \), \( \rho_S \) is the same as \( \rho_s \). Also, if \( m = n \), then \( \rho_S(x) = \rho(S^{-T} x) \).

**Lattices.** A lattice is a discrete additive subgroup of \( \mathbb{R}^m \). A set of linearly independent vectors
that generates a lattice is called a basis and is denoted by \( B = \{ b_1, \ldots, b_n \} \subset \mathbb{R}^m \) for integers
\( m \geq n \geq 1 \). The lattice generated by the basis \( B \) is
\[
\mathcal{L} = \mathcal{L}(B) = \left\{ Bz = \sum_{i=1}^n z_i b_i : z \in \mathbb{Z}^n \right\}.
\]
We say that the rank of this lattice is \( n \) and its dimension is \( m \). For \( i = 1, \ldots, n \), the *successive
minimum* \( \lambda_i(\mathcal{L}) \) is defined as the smallest value such that a ball of radius \( \lambda_i(\mathcal{L}) \) centered around
the origin contains at least \( i \) linearly independent lattice vectors.

The dual lattice \( \mathcal{L}^\ast \) of \( \mathcal{L} \) is defined as \( \mathcal{L}^\ast = \left\{ x \in \mathbb{R}^n : \langle \mathcal{L}, x \rangle \subseteq \mathbb{Z} \right\} \). For a lattice \( \mathcal{L} \) and positive real \( \varepsilon > 0 \), the smoothing parameter \( \eta_{\varepsilon}(\mathcal{L}) \) is the smallest real \( s > 0 \) such that
\( \rho_{1/s}(\mathcal{L}^\ast \setminus \{0\}) \leq \varepsilon \), where \( \rho_{1/s}(A) \) for a set \( A \) denotes \( \sum_{x \in A} \rho_{1/s}(x) \).

For a rank-\( n \) lattice \( \mathcal{L}, S \in \mathbb{R}^{m \times n} \), and \( c \in \mathbb{R}^n \), the ellipsoidal Gaussian distribution with
parameter \( S \) and support \( \mathcal{L} + c \) is defined as:
\[
\forall x \in \mathcal{L} + c, \quad D_{\mathcal{L}+c,S}(x) = \frac{\rho_S(x)}{\rho_S(\mathcal{L} + c)},
\]
where \( \rho_S(A) \) for a set \( A \) denotes \( \sum_{x \in A} \rho_S(x) \).

**Matrices and Singular Values.** For a rank-\( n \) matrix \( S \in \mathbb{R}^{m \times n} \), there exist orthogonal matrices
\( U, V \) (i.e., \( U^T U = I \), and \( V^T V = I \)) and a diagonal matrix
\[
\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{m \times n}, \quad \text{with} \quad \sigma_1 \geq \cdots \geq \sigma_n \geq 0,
\]
such that \( S = U \Sigma V^T \). The values \( \sigma_1, \ldots, \sigma_n \) are called the singular values of \( S \). The largest singular
value of \( S \) is denoted as \( \sigma_1(S) \), and the least singular value is \( \sigma_n(S) \).
Some Known Results.

Lemma 2.1 ([AGHS13 Lemma 3]). For any \( c \geq 1/\sqrt{2\pi} \), rank-\( n \) lattice \( \mathcal{L} \), \( 0 < \varepsilon < 1 \), and matrix \( S \) s.t. \( \sigma_n(S) \geq \eta_\varepsilon(\mathcal{L}) \), we have

\[
\Pr_{\mathcal{L},S} (\|v\| \geq \sigma_1(S)c\sqrt{n}) \leq \frac{1 + \varepsilon}{1 - \varepsilon} \left( c\sqrt{2\pi e} \cdot e^{-\pi c^2} \right)^n.
\]

Lemma 2.2 ([AGHS13 Lemma 4]). For any rank-\( m \) lattice \( \mathcal{L} \), \( 0 < \varepsilon < 1 \), vector \( c \in \mathbb{R}^m \), and full-rank matrix \( S \in \mathbb{R}^{m \times m} \), such that \( \sigma_m(S) \geq \eta_\varepsilon(\mathcal{L}) \), we have

\[
\rho_S(\mathcal{L} + c) \in \left[ \frac{1 - \varepsilon}{1 + \varepsilon}, 1 \right] \cdot \rho_S(\mathcal{L}).
\]

Lemma 2.3 ([AGHS13 Corollary 2]). For any full-rank lattice \( \mathcal{L} \in \mathbb{R}^n \), \( \varepsilon > 0 \), \( c > 2 \), and a rank-\( n \) matrix \( S \) such that \( \sigma_n(S) \geq (1 + c)\eta_\varepsilon(\mathcal{L}) \), the following holds. For any \( T \subset \mathcal{L} \), and any \( v \in \mathcal{L} \),

\[
\mathcal{D}_{\mathcal{L},S}(T) - \mathcal{D}_{\mathcal{L},S}(T - v) \leq \frac{\text{erf}(q/2 + 2q/c)}{\text{erf}(2q)} \cdot \frac{1 + \varepsilon}{1 - \varepsilon},
\]

where \( q = \frac{\|v\|\sqrt{\pi}}{\sigma_n(S)} \), and \( \text{erf} \) is the Gaussian error function defined as \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \).

Lemma 2.4 ([AS04 Azuma’s inequality, Chapter 7]). Let \( X_0, X_1, \ldots \) be a set of random variables that form a discrete-time sub-martingale, i.e., for all \( n \geq 0 \),

\[
E[X_{n+1} \mid X_1, \ldots, X_n] \geq X_n.
\]

If for all \( n \geq 0 \), \( |X_n - X_{n-1}| \leq c \), then for all integers \( N \) and positive real \( t \),

\[
\Pr(X_N - X_0 \leq -t) \leq \exp \left( \frac{-t^2}{2Nc^2} \right).
\]

3 Sufficient Condition for \( X \)

Let \( X \in \mathbb{Z}^{n \times m} \) be a full row-rank matrix. For a full-rank matrix \( R \in \mathbb{R}^{m \times m} \), and \( c \in \mathbb{R}^m \), we consider the distribution \( \mathcal{E}_{X,R,c} \), which is a generalization of \( \mathcal{E}_{X,r} \), defined in Section 1,

\[
\mathcal{E}_{X,R,c} := \{ X \cdot v : v \in \mathcal{D}_{\mathbb{Z}^m + c, R} \}.
\]

In this section, we establish a sufficient condition on \( X \) such that \( \mathcal{E}_{X,R,c} \) is statistically close to a discrete Gaussian distribution, namely, to \( \mathcal{D}_{\mathbb{Z}^n + X,c,R^T} \).

Definition 3.1. Let \( X \) be an \( n \times m \) matrix and let \( x_1, \ldots, x_n \in \mathbb{Z}^m \) be the rows of \( X \). For any positive \( q_1 = q_1(n), q_2 = q_2(n) \), the matrix \( X \) is said to have quality \( (q_1, q_2) \) if all column-vectors of \( X \) have \( \ell_2 \) norm at most \( q_1 \) and there exist pairwise orthogonal \( u_1, \ldots, u_n \in \mathbb{Z}^m \), such that for all \( i, j \in [n] \), \( u_i \cdot x_j = \delta_{ij} \), and \( \|u_i\|_2 \leq q_2 \).

Note that the fact that \( X \) has quality \( (q_1, q_2) \) for \( q_1, q_2 \in \mathbb{R}^+ \) implies that \( X \) is a full row-rank matrix, and that \( X \mathbb{Z}^m = \mathbb{Z}^n \). In this section, we will prove the following theorem.
Theorem 3.2. For any \( m > n \geq 1 \), and \( 0 < \varepsilon < 1/3 \), \( c \in \mathbb{R}^m \), if \( X \) is an \( n \times m \) matrix of quality \((q_1,q_2)\), and \( R \in \mathbb{R}^{m \times m} \) is full-rank such that

\[
\sigma_m(R) \geq (1 + q_1 q_2) \cdot \sqrt{\frac{\ln(2(m-n)(1+1/\varepsilon))}{\pi}},
\]

then

\[
\Delta(\mathcal{E}_{X,R,c}, \mathcal{D}_{\mathbb{Z}^m + X_c, RX^T}) \leq 2\varepsilon.
\]

Define a lattice \( A = A(X) \) as the lattice containing integer vectors in \( \mathbb{Z}^m \) orthogonal to all the row vectors of \( X \),

\[
A = A(X) := \{v \in \mathbb{Z}^m : X \cdot v = 0\}.
\]

Note that the rank of \( A \) is \( m - n \) since \( X \) has rank \( n \) and the vectors in \( A \) span the vector space orthogonal to the \( n \) row vectors of \( X \). The following result is a slight extension of \cite{AGHST13}, Lemma 10 (mentioned already in an unpublished version of \cite{LPSS13}). It shows (in the case of centered spherical distribution) that if \( r \) is bigger than the smoothing parameter of \( A \), then \( \mathcal{E}_{X,r} \) is statistically indistinguishable from \( \mathcal{D}_{\mathbb{Z}^m, RX^T} \).

Lemma 3.3. Let \( X \) be an \( n \times m \) full row-rank matrix and let \( A = A(X) \) be as defined above. For any full-rank matrix \( R \in \mathbb{R}^{m \times m} \), \( c \in \mathbb{R}^m \), such that \( \sigma_m(R) > \eta_c(A) \) for some \( \varepsilon < 1/3 \),

\[
\Delta(\mathcal{E}_{X,R,c}, \mathcal{D}_{\mathbb{Z}^m + X_c, RX^T}) \leq 2\varepsilon.
\]

Proof. The support of \( \mathcal{E}_{X,R,c} \) is \( \mathbb{Z}^m + X_c \). Fix some \( z \in \mathbb{Z}^m + X_c \). The probability mass assigned to \( z \) by \( \mathcal{E}_{X,R,c} \) is proportional to \( \rho_R(A_z) = \rho(R^T A_z) \), where

\[
A_z = A_z(X) := \{v \in \mathbb{Z}^m + c : X \cdot v = z\}.
\]

Note that \( A_z = A + w_z \), for any arbitrary element \( w_z \in A_z \) (since \( A_z \) is non-empty).

Consider now the quantity \( \rho(R^T A_z) \). Let \( X' = X \cdot R^T \in \mathbb{R}^{n \times m} \), and note that \( \ker X' = R^{-T} \ker X \). Let \( Y' = (X'X'^T)^{-1}X' \) be the pseudo-inverse of \( X' \) (i.e., \( X'Y'^T = I_n \)) and the rows of \( Y' \) span the same linear subspace of \( \mathbb{R}^m \) as the rows of \( X' \).

Define \( u_z := Y'^T z \). Note that \( u_z \) is the point in the affine subspace \( \ker X' + R^{-T} w_z \) closest to the origin: it is clearly in this affine subspace, and moreover, being in the row space of \( Y' \), it must be orthogonal to \( \ker Y' = \ker X' \). Therefore, for any \( r \in R^{-T} A_z \), we have that \( r - u_z \) is orthogonal to \( u_z \), and thus, \( \rho(r) = \rho(u_z) \cdot \rho(r - u_z) \). Hence,

\[
\rho(R^T A_z) = \rho(u_z) \cdot \rho(R^T A_z - u_z).
\]

Since \( u_z \in R^{-T} \ker X + R^{-T} w_z \), it follows that

\[
R^{-T} A_z - u_z = R^{-T}(A - c'),
\]

for some \( c' \) in the span of the lattice \( A \). Thus, using Lemma 2.2, we have that

\[
\rho(R^T A_z) = \rho(u_z) \cdot \rho_R(A - c') \in \left[ \frac{1 - \varepsilon}{1 + \varepsilon}, 1 \right] \cdot \rho_R(A) \cdot \rho(u_z) = \left[ \frac{1 - \varepsilon}{1 + \varepsilon}, 1 \right] \cdot \rho_R(A) \cdot \rho_{X^T}(z).
\]

This implies that the statistical distance between \( \mathcal{E}_{X,R,c} \) and \( \mathcal{D}_{\mathbb{Z}^m + X_c, RX^T} \) is at most

\[
1 - \frac{1 - \varepsilon}{1 + \varepsilon} \leq 2\varepsilon.
\]

\( \square \)
Thus, in order to bound the required statistical distance, we need to bound the smoothing parameter $\eta_{\varepsilon}(A)$, for which we use the following result.

**Lemma 3.4** ([MR07] Lemma 3.3). For any rank-$n$ lattice $\mathcal{L}$ and any $\varepsilon > 0$,

$$
\eta_{\varepsilon}(\mathcal{L}) \leq \lambda_{n}(\mathcal{L}) \cdot \sqrt{\frac{\ln(2n(1+1/\varepsilon))}{\pi}}.
$$

In particular, $\eta_{\varepsilon}(\mathbb{Z}^n) \leq \sqrt{\frac{\ln(2n(1+1/\varepsilon))}{\pi}}$.

Therefore, we now only need to bound $\lambda_{m-n}(A)$, and then using Lemma 3.4 we get a bound on $\eta_{\varepsilon}(A)$. For this, we use the following result from [KLP12], which we state and prove in our notation.

**Lemma 3.5.** Let $X \in \mathbb{Z}^{m \times n}$ and let $x_1, \ldots, x_n \in \mathbb{Z}^m$ be the rows of $X$. If $X$ has quality $(q_1, q_2)$, then there exist linearly independent $v_1, \ldots, v_{m-n} \in \mathbb{Z}^m$ such that $v_k \cdot x_j = 0$ and $\|v_k\|_2 \leq 1 + q_1 q_2$ for all $j, k$. In particular $\lambda_{m-n}(A) \leq 1 + q_1 q_2$.

**Proof.** Let $u_1, \ldots, u_n$ be as in Definition 3.1. Define vectors $v_1, \ldots, v_m$ as:

$$
v_k = e_k - \sum_{i=1}^{n} x_{ik} u_i,
$$

where $x_{ik}$ is the $k$-th coordinate of $x_i$, and $e_k$’s are standard unit vectors in $m$ dimensions. Thus,

$$
v_k \cdot x_j = x_{jk} - \sum_{i=1}^{n} x_{ik} (u_i \cdot x_j) = x_{jk} - x_{jk} = 0,
$$

using the fact that $u_i \cdot x_j = \delta_{ij}$.

Using the fact that $u_i$’s are orthogonal and the triangle inequality, we have that

$$
\|v_k\| \leq 1 + \sqrt{\sum_{i=1}^{n} |x_{ik}|^2 \|u_i\|^2} \leq 1 + q_2 \sqrt{\sum_{i=1}^{n} |x_{ik}|^2} \leq 1 + q_1 q_2.
$$

Also, clearly there are at least $m-n$ linearly independent vectors in $v_1, \ldots, v_m$, since, by definition, the set of vectors $\{v_1, \ldots, v_m, u_1, \ldots, u_n\}$ together generate all the unit vectors $e_k$, and hence span $\mathbb{R}^m$. \qed

Combining Lemmata 3.3, 3.4, and 3.5 immediately leads to a proof of Theorem 3.2.

We note that one can relax Definition 3.1 by not requiring the pairwise orthogonality of the vectors $u_1, \ldots, u_n$. In this case, Theorem 3.2 still holds if we take $q_1$ to be a bound on the $\ell_1$ norm of the columns of $X$ (instead of the $\ell_2$ norm as it is now).

### 4 The matrix $X$ is of good quality

In this section, we show that $X \leftarrow (\mathcal{D}_{\mathbb{Z}^n, S})^m$ has quality $(q_1, q_2)$ for “small” $q_1, q_2$ with overwhelming probability. For proving this result, we need the following claim that easily follows from the pigeonhole principle.
Claim 4.1. Let $B, n$ be any positive integers such that $Bn \geq 16$, and let $x_1, x_2, \ldots, x_\ell \in \mathbb{Z}^n$ be such that $\|x_j\|_\infty \leq B$ for $1 \leq j \leq \lfloor 2n \log(Bn) \rfloor$. Then there exist $\alpha_1, \ldots, \alpha_{\lfloor 2n \log(Bn) \rfloor} \in \{-1, 0, 1\}$, not all zero, such that $\sum_{j=1}^{\lfloor 2n \log(Bn) \rfloor} \alpha_j x_j = 0$.

Proof. Let $\ell := \lfloor 2n \log(Bn) \rfloor$. Any 0/1 combinations of $x_1, \ldots, x_\ell$ is in $\{-B\ell, \ldots, -1, 0, 1, \ldots, B\ell\}^n$, and thus the total number of distinct resulting vectors is at most $(2B\ell + 1)^n < 2^\ell$. By the pigeonhole principle there exist two distinct 0/1 combinations that result in the same vector, which implies the result by taking their difference. $\square$

Lemma 4.2. Let $n \geq 100$ be an integer, and let $\varepsilon = \varepsilon(n) \in (0, 1/1000)$. Let $S \in \mathbb{R}^{n \times n}$ be such that $\sigma_n(S) \geq 9\sqrt{\frac{\ln(2n(1+1/\varepsilon))}{\pi}}$. Denote $\sigma_1 = \sigma_1(S)$ and $\sigma_n = \sigma_n(S)$. Let $m$ be an integer such that $m > 30n \log(\sigma_1 n)$ and let $X \leftarrow (D_{\mathbb{Z}^n,S})^m$ be an $n \times m$ matrix. Then, with probability $1 - 2^{-n}$, $X$ has quality $(\sigma_1 \sqrt{n \log m}, 2\sqrt{30n \log(\sigma_1 n)})$.

Proof. By Lemma 2.1 with $c = \sqrt{\log m}$, we have using the union bound that with probability $1 - 2^{-n-1}$, all column-vectors of $X$ have norm at most $\sigma_1 \sqrt{n \log m}$. It is sufficient to show that with probability $1 - 2^{-n-1}$, there exist pairwise orthogonal $u_1, \ldots, u_n \in \mathbb{Z}^m$ such that $\|u_i\|_2 \leq 2\sqrt{30n \log(\sigma_1 n)}$, and $X \cdot u_i = e_i$ for $1 \leq i \leq n$.

Denote $t := 3n \log(\sigma_1 n)$. For all $i \in [n]$, we show the existence of $u_i$ with probability $1 - 2^{-t/10}$, and then the result follows by the union bound. Let $x_1, \ldots, x_m$ be the columns of $X$. Define sets $S_j \subset \mathbb{Z}^n$ for $1 \leq j \leq m$ as follows:

$$\forall j \in [m], \ S_j = \left\{ \sum_{k=1}^j \alpha_k x_k : \alpha_k \in \{-1, 0, 1\} \right\}.$$ 

From now on we show the existence of $u_1$ and at the end of the proof we explain how to show the existence of the other $u_i$. We will show that with probability $1 - 2^{-t/10}$, $S_{[10t]} \cap (S_{[10t]} + e_1)$ is non-empty, which implies that $e_1$ can be obtained as a $\{-2, -1, 0, 1, 2\}$ combination of the first $[10t]$ vectors. Then we are done since the vector $u_1$ can be defined as this coefficient vector with first $[10t]$ coordinates in $\{-2, -1, 0, 1, 2\}$ and the remaining coordinates zero. It is easy to see that $\|u_1\|$ is at most $2\sqrt{10t}$.

The idea used to prove this is that if this does not happen, then $S_j$ and $S_j + e_1$ must be disjoint for all $j \leq [10t]$, and so at most one of $x_{j+1}$ and $x_{j+1} - e_1$ is in $S_j$ for any $j \leq [10t]$. Using the randomness of $x_j$, we conclude that with constant probability, $x_{j+1} \notin S_j$ and this cannot happen for many values of $j$ by Claim 4.1.

Define binary random variables $Y_1, \ldots, Y_m$ as follows: $Y_j = 1$ if and only if either

- $S_j \cap (S_j + e_1)$ is non-empty, or
- $x_{j+1} \notin S_j$ and $\|x_{j+1}\|_\infty \leq \sigma_1 \sqrt{n}$.

Consider the following claims about the distribution of $Y_1, \ldots, Y_m$.

Claim 4.3. For any values of $y_1, \ldots, y_{j-1}$, and for all $j$, $\Pr(Y_j = 1 \mid Y_1 = y_1, \ldots, Y_{j-1} = y_{j-1}) \geq 0.3$. 

Proof. Throughout this proof, we condition on arbitrary values of \( x_1, \ldots, x_j \), and our goal is to show that the probability of \( Y_j = 1 \) is at least 0.3.

If \( S_j \cap (S_j + e_1) \) is non-empty (an event that depends only on \( x_1, \ldots, x_j \)), then we are done. Assume therefore that our choice of \( x_1, \ldots, x_j \) is such that \( S_j \) and \( S_j + e_1 \) are disjoint. By Lemma 2.1 with all but exponentially small probability in \( n \), \( \|x_{j+1}\|_\infty \leq \|x_{j+1}\|_2 \leq \sigma_1 \sqrt{n} \). It therefore suffices to show that \( x_{j+1} \in S_j \) with probability at most, say, 0.695. This follows since from Lemma 3.4 we get that \( \sigma_n \geq 9\eta_e(Z^n) \), and thus we have by Lemma 2.3 that for any set \( S \subset Z^n \)

\[
\Pr(x_{j+1} \in S) - \Pr(x_{j+1} \in S + e_1) \leq \frac{\operatorname{erf}(3\sqrt{n}/(4\sigma_n))}{\operatorname{erf}(2\sqrt{n}/\sigma_n)} \cdot \frac{1 + \varepsilon}{1 - \varepsilon} < 0.39 ,
\]

where the last inequality follows because \( \operatorname{erf} \) is nearly linear for small arguments. Finally, the claim follows by noticing that the sum of the two probabilities in the left hand side is at most 1. \( \square \)

Claim 4.4.

\[
\Pr(Y_1 + \cdots + Y_{10t} \geq 2t) \geq 1 - 2^{-t/10} .
\]

Proof. Define random variables \( Z_0, Z_1, \ldots \) recursively as \( Z_0 = 0 \) and for \( i \geq 1 \), \( Z_i = Z_{i-1} + Y_i - \frac{3}{10} \), so for all \( j \),

\[
Y_1 + \cdots + Y_j = Z_j - Z_0 + \frac{3j}{10} .
\]

Clearly \( Z_0, Z_1, \ldots \) is a sub-martingale, and for all \( j \geq 1 \), \( |Z_j - Z_{j-1}| \leq \frac{7}{10} \). Thus, using Lemma 2.4 we get that

\[
\Pr(Y_1 + \cdots + Y_{10t} \leq 2t) = \Pr(Z_{10t} - Z_0 \leq -t) \leq \exp \left( \frac{-t^2}{2(10t)(0.7)^2} \right) = 2^{-t/10} ,
\]

which implies the result. \( \square \)

Claim 4.5. If \( Y_1 + \cdots + Y_{10t} \geq 2t \), then \( S_{10t} \cap (S_{10t} + e_1) \) is non-empty.

Proof. Assume \( Y_{j_1} = \cdots = Y_{j_{10t}} = 1 \) and \( S_{10t} \cap (S_{10t} + e_1) \) is empty. Consider the vectors \( x_{j_1+1}, \ldots, x_{j_{10t}+1} \). We have that \( \|x_{j_{10t}+1}\|_\infty \leq \sigma_1 \sqrt{n} \) for \( 1 \leq k \leq \lfloor 2t \rfloor \). Since \( 2t = 6n \log(\sigma_1 n) \geq 2n \log(\sigma_1 n^{3/2}) \), by Claim 4.1 there exist \( \alpha_1, \ldots, \alpha_{\lfloor 2t \rfloor} \in \{-1, 0, 1\} \), not all zero, such that

\[
\sum_{k=1}^{\lfloor 2t \rfloor} \alpha_k x_{j_k+1} = 0 .
\]

Let \( k \) be the largest index such that \( \alpha_k \neq 0 \). This implies \( x_{j_k+1} \in S_{j_k} \), contradicting the fact that \( Y_{j_k} = 1 \). \( \square \)

Claim 4.3, 4.4 and 4.5 together prove the existence of \( u_1 \). To obtain a vector \( u_2 \) that is orthogonal to \( u_1 \), we do the following. Let \( X' \) be the \((n+1) \times m\) matrix formed by adding another row, \( u_1 \), to \( X \). We follow the same argument as above with \( X \) replaced by \( X' \) and \( e_1 \) replaced by \( e_2 \). Let \( x_j' \in Z^{n+1} \) be \( x_j \) concatenated with \( u_{ij} \), the \( j \)-th coordinate of \( u_1 \). By construction \( \|u_1\|_\infty \leq 2 \), and hence \( \|x_j'\|_\infty \leq \sigma_1 \sqrt{n} \) if and only if \( \|x_j\|_\infty \leq \sigma_1 \sqrt{n} \). For completing the proof,
we need to bound \( \Pr(x'_{j+1} \in S') - \Pr(x'_{j+1} \in S' + e_2) \) for an arbitrary set \( S' \subset \mathbb{Z}^{n+1} \). Since the \((n+1)\)-th coordinate \( x'_{j+1} \) is fixed to be \( u_{1(j+1)} \), we have that

\[
\Pr(x'_{j+1} \in S') - \Pr(x'_{j+1} \in S' + e_2) = \Pr(x_{j+1} \in S) - \Pr(x_{j+1} \in S + e_2),
\]

where \( S \subset \mathbb{Z}^n \) is the set of vectors formed by projecting the vectors of \( S'' = \{ y \in S : y_{n+1} = u_{1(j+1)} \} \) onto the first \( n \) coordinates. The remainder of the proof is exactly the same. We similarly get \( u_3, \ldots, u_n \), such that for all \( i \leq n \), \( u_i \) is orthogonal to \( u_1, \ldots, u_{i-1} \).

\section{Sum of Discrete Gaussian Distribution}

In this section, we state the main result of this paper, the proof of which follows easily from the results proved in Section 3 and Section 4.

\begin{theorem}
Let \( m > n \geq 100 \) be integers, and let \( \varepsilon = \varepsilon(n) \in (0, 1/1000) \). Let \( R \in \mathbb{R}^{m \times n} \) be a full-rank matrix, and let \( c \in \mathbb{R}^m \). Let \( S \in \mathbb{R}^{n \times n} \) be a full-rank matrix, and let \( X \leftarrow (D_{\mathbb{Z}^n} S)^m \). If \( m \geq 30n \log(\sigma_1(S)n) \), \( \sigma_m(R) \geq 10n\sigma_1(S) \log m \sqrt{\log(1/\varepsilon) \log(n\sigma_1(S))} \), and \( \sigma_n(S) \geq 9 \sqrt{\ln(2(m-n)1+1/\varepsilon)} \pi \), then, with probability \( 1 - 2^{-n} \) over the choice of \( X \), we have that

\[
\Delta(\mathcal{E}_{X,R,c}, D_{\mathbb{Z}^n + Xc, RX^T}) \leq 2\varepsilon.
\]

\end{theorem}

\begin{proof}
Theorem 5.1 follows easily from Lemma 4.2 and Theorem 3.2 using the observation that

\[
\sigma_m(R) \geq 10n\sigma_1(S) \log m \sqrt{\log(1/\varepsilon) \log(n\sigma_1(S))} \geq \left( 1 + 2\sigma_1(S) \sqrt{n \log m} \sqrt{30n \log(\sigma_1(S)n)} \right) \sqrt{\frac{\ln(2(m-n)(1+1/\varepsilon))}{\pi}}.
\]

It is easy to see that if we sample a vector \( x \) from \( D_{\mathbb{Z}^n} S \) then \( Bx \) is distributed like \( D_{\mathcal{L},SB^T} \) where \( \mathcal{L} = \mathcal{L}(B) \) is the lattice generated by the columns of \( B \) (see, e.g., [AGHS13, Fact 2]). As a result, we can extend Theorem 5.1 to the case in which \( X \) is sampled from \( (D_{\mathcal{L},M})^m \) where \( \mathcal{L} = \mathcal{L}(B) \) for some basis \( B \), and \( M = SB^T \). The only change is that the statistical closeness is to the distribution \( D_{\mathcal{L}+Xc, RX^T} \).

\end{proof}

\section*{Acknowledgements:}
We thank Damien Stehlé for useful comments.

\section*{References}

[AGHS13] S. Agrawal, C. Gentry, S. Halevi, and A. Sahai. Discrete Gaussian leftover hash lemma over infinite domains. In ASIACRYPT, pages 97–116. 2013.

[AS04] N. Alon and J. Spencer. The Probabilistic Method. Wiley, 2004.

[Ban93] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. Mathematische Annalen, 296(01):625–635, 1993.
[BF11] D. Boneh and D. Freeman. Homomorphic signatures for polynomial functions. In *Advances in Cryptology-EUROCRYPT*, pages 149–168. Springer, 2011.

[GGH13] S. Garg, C. Gentry, and S. Halevi. Candidate multilinear maps from ideal lattices. In *Advances in Cryptology-EUROCRYPT*, pages 1–17. Springer, 2013.

[GPV08] C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *40th Symposium on Theory of Computing-STOC*, pages 197–206. ACM, 2008.

[KLP12] G. Kuperberg, S. Lovett, and R. Peled. Probabilistic existence of rigid combinatorial structures. In *44th Symposium on Theory of Computing-STOC*, pages 1091–1106. ACM, 2012.

[LPSS13] S. Ling, D. H. Phan, D. Stehlé, and R. Steinfeld. Hardness of k-lwe and applications in traitor tracing, 2013. Unpublished manuscript.

[MP13] D. Micciancio and C. Peikert. Hardness of SIS and LWE with small parameters. In *Advances in Cryptology-CRYPTO*. Springer, 2013.

[MR07] D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM Journal on Computing*, 37(01):267–302, 2007.

[Pei10] C. Peikert. An efficient and parallel Gaussian sampler for lattices. In *Advances in Cryptology-CRYPTO*, pages 80–97. Springer, 2010.