Submission to STOC 2014 on 11/11/2013.

A quantum algorithm for computing the unit group of an arbitrary degree number field

Kirsten Eisenträger
Department of Mathematics
Penn State University
eisentra@math.psu.edu

Sean Hallgren
Department of Computer Science
and Engineering
Penn State University
hallgren@cse.psu.edu

Alexei Kitaev
Kavli Institute for Theoretical Physics
University of California, Santa Barbara
kitaev@kitp.ucsb.edu
and California Institute of Technology

Fang Song
Department of Combinatorics
& Optimization
and Institute for Quantum Computing
University of Waterloo
fang.song@uwaterloo.ca

Abstract

There are several interesting computational problems coming from number fields whose complexity makes them interesting both for quantum algorithms and also for cryptography. Some problems, such as computing the unit group, class group and solving the principal ideal problem already seem computationally hard classically when the number field has growing discriminant but constant degree. There are efficient quantum algorithms for solving these problems in the constant degree case. Other problems, such as solving the lattice problems in ideal lattices that come from number fields of growing degree, are assumed to be hard in recent homomorphic encryption constructions. No quantum algorithms are known for these lattice problems.

In this paper we give a quantum algorithm for computing the unit group in time polynomial in the degree of the number field. We show that there is a quantum reduction from the problem to solving an abelian hidden subgroup problem over $\mathbb{R}^n$. In doing so we define a natural way for extending the definition of the HSP to a continuous group. As an application we show that the previously solved abelian hidden subgroup instances reduce to this new case. We then give an efficient quantum algorithm for the HSP over $\mathbb{R}^n$.

The paper contains three results. The first is a classical algorithm for computing a basis for certain principal ideals with doubly exponentially large generators in a number field. The second shows that a Gaussian weighted superposition of lattice points, with an appropriate encoding, can be used to provide a unique representation of a real-valued lattice, and satisfy the continuous HSP properties. The third is a quantum algorithm for solving the HSP over $\mathbb{R}^n$.

1 Introduction

The problems where quantum algorithms have exponential speedups over the best known classical algorithm have mostly been of number theoretic origin. Shor found quantum algorithms for factoring and discrete log [Sho97] and Hallgren found a quantum algorithm for Pell's equation [Hal07]. These algorithms were further generalized for finding the unit group of a number field and related problems [Hal05, SV05]. The running time is measured in terms of the discriminant and the degree of the number field. In the latter case, the algorithm is only efficient for constant degree number fields. In this paper we address the arbitrary degree case and give an algorithm that efficient in both the discriminant and the degree.

A number field K can be defined as a subfield of the complex numbers $\mathbb{C}$ which is generated over the rational numbers $\mathbb{Q}$ by an algebraic number, i.e. $K = \mathbb{Q}(\theta)$ where θ is the root of a polynomial with rational coefficients. If K is a number field, then the subset of K consisting of all elements that are roots of monic polynomials with integer coefficients, forms a ring $\mathcal{O}$, called the ring of integers of K. The ring $\mathcal{O} \subseteq K$ can be thought of as a generalization of $\mathbb{Z} \subset \mathbb{Q}$. In particular, we can ask whether $\mathcal{O}$ is a principal ideal domain, whether numbers in $\mathcal{O}$ have unique factorization, and what the set of invertible elements is. The unit group $\mathcal{O}^*$ is the set of invertible algebraic integers inside K, that is, elements $\alpha \in \mathcal{O}$ such that $\alpha^{-1} \in \mathcal{O}$.

Computing the unit group of a number field is an important problem in computational number theory. By Dirichlet's Theorem the group of units $\mathcal{O}^*$ is isomorphic to $\mathcal{O}^* \cong \mu(K) \times \mathbb{Z}^{s+t-1}$, where $\mu(K)$ are the roots of unity contained in K and K has s real embeddings and t pairs of complex conjugate embeddings. An elementary version of the problem is Pell's equation: given a positive non-square integer d, find x and y such that $x^2 - dy^2 = 1$. Solutions to this equation are parametrized by the formula $x_k + y_k\sqrt{d} = \big(x_1 + y_1\sqrt{d}\big)^k$; the numbers $\pm\big(x_k + y_k\sqrt{d}\big)$ are exactly the units of the quadratic ring $\mathbb{Z}\big[\sqrt{d}\big]$ (or a subgroup of index 2 if there is a unit that has norm -1). The fundamental solution (x_1, y_1) is difficult to find, or even to write down because it may be exponential in d (i.e., doubly-exponential). Moreover, the computation of the real number $R = \ln\big(x_1 + y_1\sqrt{d}\big)$ with a polynomial number of precision digits is believed to be a hard problem classically.

A polynomial time quantum algorithm for the computation of R was given in [Hal07]. The approach is to reduce the problem to a hidden subgroup problem (HSP) over the real numbers $\mathbb{R}$, and then to give a quantum algorithm for that hidden subgroup problem. In this context, the HSP amounts to having a periodic function on $\mathbb{R}$ which is 1-1 within the period. The goal is to approximate the period.

For the unit group the corresponding oracle function g takes a real number u to a lattice $g(u) \subset \mathbb{R}^2$. More specifically, we can embed $\mathcal{O}$ as a lattice and then $g(u)$ is obtained by stretching by factor of $e^{\boldsymbol{u}}$ in one direction and squeezing in the other:

$$g(u) = (e^u, e^{-u})\,\mathcal{O} := \big\{\big(e^u z^{(1)}, e^{-u} z^{(2)}\big) : (z^{(1)}, z^{(2)}) \in \mathcal{O}\big\},$$
$$\text{where} \quad \mathcal{O} = \big\{\big(x + y\sqrt{d},\, x - y\sqrt{d}\big) : x, y \in \mathbb{Z}\big\}.$$

The function g is periodic with period R and 1-1 within the period. However, exponential stretching and squeezing of lattices is not computationally trivial. Furthermore, the standard quantum algorithm for the hidden subgroup problem requires a unique representation of the oracle function value (a representation up to an equivalence relation will not work). In [Hal07] these issues were addressed using an intricate notion of "reduced ideals". This method was extended to constant degree number fields [Hal05, SV05], but it is difficult to generalize this method to rings of higher degree. At a minimum, computing the required reduced ideals seems to require solving the shortest

vector problem in ideal lattices of dimension n, and enumerating lattice points also seems necessary. Cryptosystems whose security relies on the hardness of solving problems in ideal lattices have been suggested for cryptography [PR07, LPR10]. Another problem is running the hidden subgroup algorithm for the continuous group $G = \mathbb{R}^m$, where rounding causes errors. Such errors are tolerable when m is fixed but worsen in higher dimensions.

We propose a different scheme, leading to a quantum reduction from computing the unit group of a number field of arbitrary degree n to solving an abelian hidden subgroup problem over $\mathbb{R}^m$, where $m = O(n)$. It involves several important ingredients. First, we represent a lattice by a reduced basis (up to some precision). The exponential transformation is performed using repeated squaring of lattices. These lattices can be multiplied because they are also ideals. Having obtained some basis of the lattice $L = f(u)$, we construct a canonical *quantum representation* of L, namely the Gaussian-weighted superposition of lattice points with a sufficiently large dispersion. To ensure stability against rounding errors, each lattice point is represented by a superposition of nearby points of a fine grid. (For example, in one dimension, such a superposition straddles two adjacent grid points.) The initial motivation for this was using double Gaussian states as in [GKP01], required a different representation of points. In addition to showing how to classically compute approximate bases for the stretched lattices, we prove that the inner product of Gaussian lattice states has a hidden subgroup property.

One byproduct of this work is a generalization of the HSP to uncountable topological groups such as $\mathbb{R}$. Most exponential speedups by quantum algorithms either use or try to use the HSP [FIM$^+$03, HMR$^+$10]. In the HSP a function $f : G \to S$ is given on a group G to some set S. For an unknown subgroup $H \subseteq G$, the function is constant on cosets of H and distinct on different cosets. The goal is to find a set of generators for H in time polynomial in the appropriate input size, e.g. $\log |G|$. When G is finite abelian or $\mathbb{Z}^m$ there is an efficient quantum algorithm to solve the problem.

Using the usual definition of the HSP for the group $G = \mathbb{R}$ does not work as can be seen by the following illustration. When the group is discrete the function can be evaluated on any group element. For example, it's possible to verify that a given element h is in H, by testing if $g(0) = g(h)$. Over the reals, if the period is some transcendental number x, then no algorithm could ever even query $g(x)$, and then see that it matches $g(0)$. It is possible to address this by giving an ad-hoc technical definition if we replace $\mathbb{R}$ by a discrete set with rounding, as in the case of constant degree number fields [Hal07, Hal05, SV05]. However, it is not known how to solve the HSP with such a definition. Here we give a cleaner definition using continuous functions which aids us in finding an algorithm to solve the general problem.

Definition 1.1 (*The continuous HSP over* $\mathbb{R}^m$). The unknown subgroup $L \subseteq \mathbb{R}^m$ is a full-rank lattice satisfying some promise: the norm of the shortest vector is at least λ and the unit cell volume is at most d. The oracle has parameters (a, r, ε). Let $f : \mathbb{R}^m \to S$ be a function, where S is the set of unit vectors in some Hilbert space. We assume that f hides L in the following way.

1. f is periodic on L: for all $v \in L$, $x \in \mathbb{R}^m$, $f(x) = f(x + v)$;
2. $\big\| |f(x)\rangle - |f(y)\rangle \big\| \leqslant a \cdot \mathrm{dist}(x, y)$ for all $x, y \in \mathbb{R}^m$ (Lipschitz);
3. If $\min_{v \in L} \|x - y - v\| \geqslant r$, then $\big|\langle f(x)|f(y)\rangle\big| \leqslant \varepsilon$.

Given an efficiently computable function with this property, compute a basis for L.

We show that computing the unit group of an arbitrary degree number field can be (quantum) reduced to this definition of the HSP, and we also give a quantum algorithm for solving it. We prove the following main theorem

Theorem 1.2. *There is an efficient algorithm to compute the unit group of a number field K that is polynomial in the degree of K and polynomial in log of the discriminant of K .*

This follows from Theorem 4.1, Theorem 5.5, Section E.4, and Theorem 6.2.

Computing the unit group is one of the main computational tasks in algebraic number theory [Coh93]. Two of the others are solving the principal ideal problem and computing the class group. Based on the previous quantum algorithms for solving these three problems in the constant degree case, the unit group seems to be the most difficult part. The other two problems can be solved using the unit group algorithm and general hidden subgroup techniques. We leave the other two problems open for arbitrary degree. The main issue will be proving that the HSP functions constructed to solve them will be Lipschitz.

In the context of cryptography, even over degree two number fields, the problem of computing the unit group and solving the Principal Ideal Problem are considered to be hard classically. It was used as a basis in the Buchmann-Williams key exchange problem in an effort to find a system that is harder to break than factoring based systems. On the other hand, the typical ideal lattice problem, such as finding short vectors over degree two number fields, is easy because the degree is constant.

In the last few years, since the discovery of homomorphic encryption and the following efforts to make the systems more efficient and more secure, assumptions related to number fields have been used. These systems are set up based on high degree number fields. In [GH11], a version of the PIP where a special generator is the secret was used as the hardness assumption. The Ring-LWE problem which forms the basis in [LPR10, BV11] assumes that finding short vectors in ideal lattices of high degree number fields is hard.

To summarize, the constant degree assumptions are broken by quantum algorithms. The relatively recent high degree number field assumptions about computing short vectors are still open in terms of security against quantum computers. However, in this paper we show that it is now possible to efficiently compute the unit group in these number fields, which could move towards understanding whether the new homomorphic cryptosystems really are secure against quantum computers.

2 Number-theoretic background

In the following K will denote a number field of degree n over $\mathbb{Q}$ and $\mathcal{O}$ will denote its ring of integers. When we want to consider $\mathcal{O}$ as a lattice in $E = \mathbb{R}^s \times \mathbb{C}^t$ with $s + 2t = n$ (see below), we will write $\underline{\mathcal{O}}$. We use bold letters to designate elements of E and vectors in general.

If $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ is a basis for a lattice $\Lambda \subseteq \mathbb{R}^n$, let B be the matrix $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ which is composed of column vectors $\boldsymbol{b}_k$. Then $d(\Lambda) := |\det(B)|$ is the unit cell volume of the lattice Λ generated by the basis.

Elements of K can be conveniently represented by using the embeddings of K into the field of complex numbers. In general, there are n such embeddings, which break into s real ones and t complex-conjugate pairs:

$$\tau_1, \ldots, \tau_s : K \to \mathbb{R}, \qquad \tau_{s+1}, \ldots, \tau_{s+t}, \overline{\tau_{s+1}}, \ldots, \overline{\tau_{s+t}} : K \to \mathbb{C} \qquad (s + 2t = n).$$

Each element $z \in K$ is mapped to the corresponding *conjugate vector*

$$\boldsymbol{\tau}(z) = \left(z^{(1)}, \ldots, z^{(s)}, z^{(s+1)}, \ldots, z^{(s+t)}, \overline{z^{(s+1)}}, \ldots, \overline{z^{(s+t)}}\right)^T \in \mathbb{R}^s \times \mathbb{C}^{2t},$$

where the last t coordinates are redundant. Thus, K is embedded into $E = \mathbb{R}^s \times \mathbb{C}^t$. Conjugate vectors are added and multiplied coordinate-wise. Many useful functions on K extend naturally to E. For example, the algebraic trace and norm are defined for arbitrary conjugate vectors:

$$\mathrm{tr}(\boldsymbol{z}) = \sum_{j=1}^{s} z^{(j)} + \sum_{j=s+1}^{s+t} \bigl(z^{(j)} + \overline{z^{(j)}}\bigr), \qquad \mathcal{N}(\boldsymbol{z}) = \prod_{j=1}^{s} z^{(j)} \prod_{j=s+1}^{s+t} |z^{(j)}|^2.$$

Both these functions take real values.

As far as the additive structure is concerned, the ring E is simply an n-dimensional real space. We can define a Euclidean inner product on E by letting

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle \;=\; \mathrm{tr}(\boldsymbol{x}\overline{\boldsymbol{y}}) \;=\; \sum_{j=1}^{s} x^{(j)} y^{(j)} + 2 \sum_{j=s+1}^{s+t} \Bigl(\bigl(\mathrm{Re}\, x^{(j)}\bigr)\bigl(\mathrm{Re}\, y^{(j)}\bigr) + \bigl(\mathrm{Im}\, x^{(j)}\bigr)\bigl(\mathrm{Im}\, y^{(j)}\bigr) \Bigr). \tag{2.1}$$

The length of a vector with respect to this inner product is denoted by $\|\boldsymbol{z}\|$.

Now let $\{\omega_1, \ldots, \omega_n\}$ be some basis (over $\mathbb{Z}$) for the ring $\mathcal{O}$ of integral elements in K. One way to characterize $\mathcal{O}$ is by the "multiplication table", i.e., the decomposition of $\omega_j \omega_k$ into ω_l with integer coefficients. In the conjugate vector representation, $\mathcal{O}$ becomes a lattice $\underline{\mathcal{O}} \subseteq E$ with basis $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$. From the computational perspective, it is important to have some upper bound on the length of the basis vectors or, equivalently, on the coefficients in the multiplication table. To this end, we use the notion of *discriminant*, which is defined as the determinant of the matrix G with entries $G_{jk} = \mathrm{tr}(\omega_j \omega_k)$. The discriminant $D = D(\mathcal{O})$ depends only on the ring but not the basis. The extension degree, n, and the discriminant, D, constitute a natural set of parameters characterizing the "complexity" of the ring. Our algorithm for finding the group of units is polynomial in $n + \log |D|$.

For various algorithmic tasks, e.g. the computation of the lattice $e^{\boldsymbol{u}}\underline{\mathcal{O}}$, the basis vectors of $\underline{\mathcal{O}}$ must be known with sufficient precision. We will use the fact that the embedding of elements of K can be found to any polynomial number of precision bits in polynomial time [Thi95].

More background is given in Appendix B.

3 Overview of the algorithm

The group of units $\mathcal{O}^*$ consists of elements $z \in \mathcal{O}$ such that $\mathcal{N}(z) = \pm 1$, and they are represented by conjugate vectors of the form $\boldsymbol{z} = e^{\boldsymbol{u}}\boldsymbol{v}$. Here $\boldsymbol{u} = (u^{(1)}, \ldots, u^{(s+t)}) \in \mathbb{R}^{s+t}$ satisfies the condition $\sum_j u^{(j)} = 0$, and the components of $\boldsymbol{v} = (v^{(1)}, \ldots, v^{(s+t)}) \in E = \mathbb{R}^s \times \mathbb{C}^t$ are real or complex numbers of absolute value 1. Thus, the group of units $\mathcal{O}^*$ is contained in

$$G = \mathbb{R}^{s+t-1} \times \bigl((\mathbb{Z}_2)^s \times (\mathbb{R}/\mathbb{Z})^t\bigr).$$

More specifically, $\mathcal{O}^*$ is the hidden subgroup in G which corresponds to the following oracle:

$$g : G \to \text{lattices in } E \;:\; (\boldsymbol{u}, \boldsymbol{v}) \mapsto e^{\boldsymbol{u}}\boldsymbol{v}\underline{\mathcal{O}}. \tag{3.1}$$

We give an efficient classical realization of this function, where the output (i.e. a lattice $L \subset E$) is represented by some basis with a certain precision. Unfortunately, such a representation is not unique, and therefore g cannot be used as an oracle for a quantum HSP algorithm. To deal with this issue, we compose the function g with another function:

$$\tilde{f} : \text{lattices in } E \to \text{quantum states} \;:\; L \mapsto |\tilde{f}(L)\rangle. \tag{3.2}$$

where $|\tilde{f}(L)\rangle$ is a uniquely defined quantum superposition that encodes the lattice L. Thus, we obtain a usable quantum oracle

$$f = \tilde{f} \circ g : G \to \text{quantum states} : (\boldsymbol{u}, \boldsymbol{v}) \mapsto \left|\tilde{f}\left(e^{\boldsymbol{u}}\boldsymbol{v}\underline{\mathcal{O}}\right)\right\rangle. \tag{3.3}$$

Note: By abuse of notation, we will later denote $\tilde{f}$ as f. For example, we will refer to the quantum state $|f(L)\rangle$ for the quantum state representing lattice L.

Finally, we reduce the HSP problem for G to that for $\mathbb{R}^n$ and apply a general algorithm for finding the hidden subroup in $\mathbb{R}^m$.

Thus, our algorithm for finding the group of units splits into three self-contained parts:

- A classical algorithm for the function $g : (\boldsymbol{u}, \boldsymbol{v}) \mapsto \boldsymbol{v}e^{\boldsymbol{u}}\underline{\mathcal{O}}$. Note that we cannot compute $e^{\boldsymbol{u}}$ because it is an exponentially long number. Instead, we begin with representing $\boldsymbol{u}$ as $2^l\boldsymbol{u}_0$, where $\boldsymbol{u}_0$ is sufficiently small, and compute the lattice $e^{\boldsymbol{u}_0}\underline{\mathcal{O}}$ directly. Then we apply the following procedure l times: given a basis $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$ of the lattice $\Lambda = e^{\boldsymbol{w}}\underline{\mathcal{O}}$ (for some $\boldsymbol{w}$ that does not need to be known), we compute some basis of the lattice $\Lambda^2 = e^{2\boldsymbol{w}}\underline{\mathcal{O}}$. The repeated squaring yields a basis of the lattice $e^{\boldsymbol{u}}\underline{\mathcal{O}}$; then we multiply it by $\boldsymbol{v}$. The lattice squaring procedure is not trivial. We need to compute all products $\boldsymbol{z}_j\boldsymbol{z}_k$ and find some basis of the lattice they generate. This requires the detection of linear dependencies with integer coefficients as well as some way to prevent the vector lengths from growing. The algorithm for computing a basis for $e^u\underline{\mathcal{O}}$ is new as far as we know.

- A quantum procedure for the creation of the state $|\tilde{f}(L)\rangle$ representing a lattice $L \subset \mathbb{R}^n$. Assuming a lower bound on the length of a shortest vector, $\lambda_n(L) \geqslant \lambda$ and an upper bound on the unit cell volume, $d(L) \leqslant d$, we find a Lipschitz constant of the function $\tilde{f}$ and estimate the inner product $\langle\tilde{f}(L)|\tilde{f}(L')\rangle$ when the lattices L and L' are far apart. The function $\tilde{f}$ is defined and implemented as follows. We first create a Gaussian-weighted superposition of points $\boldsymbol{z} \in L$. Then for each coefficient vector $\boldsymbol{x} \in \mathbb{Z}^n$ representing a lattice point $\boldsymbol{z}$ we use a straddle encoding (to be defined in Sect. 5) to account for rounding errors. The original value of $\boldsymbol{x}$ may now be erased (in a reversible way, which requires the reconstruction of $\boldsymbol{x}$ from $\widetilde{\boldsymbol{z}} \approx \boldsymbol{z}$). Finally, we need to "uncompute" the previously computed basis to guarantee quantum coherence.

- An efficient quantum algorithm for finding a hidden subgoup in $\mathbb{R}^m$, as discussed in the introduction.

All of our results are stated for the ring of integers $\mathcal{O}$ of a given number field K, but they can easily be extended to general *orders* of K, i.e. to rings contained in K whose additive group is isomorphic to $\mathbb{Z}^n$.

4 Computing a basis for $e^t\mathcal{O}$

In this section we show how to compute an approximate basis for the lattice $\boldsymbol{e^t}\underline{\mathcal{O}}$. Because $\boldsymbol{e^t}$ is in general doubly exponential in size and we have to use floating point computations, this is a non-trivial operation. The basic steps are to alternate ideal multiplication with size reduction (see Sections C.1, C.1.3) to compute a short basis for the product of the two ideals that were multiplied. The algebraic numbers that appear in this computation would take exponentially many bits to represent exactly. Instead we show that a polynomial number of bits of precision is sufficient. The idea is to use the fact that we are always using ideal lattices which lower bounds and upper

bounds on the vector lengths appearing throughout the computation, so that the precision loss can be bounded at each step. With this we can pick a precision high enough, some polynomial number of bits, so that we still have high precision at the end. The precision we need is that for any vector of length at most $s\sqrt{n}$, the computed vector is within $1/(2N)$ of the actual vector. Here s and N are parameters chosen such that the lattice Gaussian superposition in Section 5 will be a good approximation to the lattice.

Given $\boldsymbol{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n$ such that $\sum t_i = 0$ we will show how to compute a basis of the lattice $f(\boldsymbol{t}) = (e^{t_1}, \ldots, e^{t_n}) \cdot \underline{\mathcal{O}}$. The function $f : \mathbb{R}^n \to \{\text{real-valued lattices}\}$ is constant and distinct on cosets of the Log embedding of (the free part of) the units. We will later handle the fact that we only have approximations of these lattices, in particular, how to create useful superpositions using the approximations.

The main subroutine needed for computing f computes a basis of the product of two lattices. Lattices A and B can be multiplied in this case since they are always of the form $(a_1, \ldots, a_n) \cdot \underline{\mathcal{O}}$, where $a_i \in \mathbb{R}$, and $\mathcal{O}^2 = \mathcal{O}$. In particular, given the bases of two lattices $A = \langle \boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \rangle$ and $B = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \rangle$, all pairwise products $\boldsymbol{w}_i \boldsymbol{v}_j$ of basis vectors are computed giving a generating set of the lattice AB, to which we apply the algorithm from Section C.1 to obtain a size reduced basis.

To ensure that the entire computation can be done in polynomial time we must upper bound the determinant of each lattice and lower bound the shortest vector, which bounds the basis sizes.

4.1 Splitting up the computation

The computation must be split up carefully to avoid ending up with doubly exponential size coefficients. First it is split into two parts. The first part handles the integer part of the t_i's which is complicated by the fact that e^{t_i} will be doubly-exponential in general. The solution will be to compute a sequence of ideals $A_{-1}, A_0, \ldots, A_m$ of bounded determinant such that $A_{-1} \times \prod_{i=0}^{m} A_i^{2^i} = f(\boldsymbol{t})$.

For $1 \leq i \leq n-1$ let $t_i = r_i + s_i$, where $r_i \in \mathbb{Z}$ and $0 \leq s_i < 1$. Let $r_n = -\sum_{i=1}^{n-1} r_i$ and $s_n = t_n - r_n$. Using the fact that $(e^{t_1}, \ldots, e^{t_n}) \cdot \underline{\mathcal{O}} = (e^{r_1}, \ldots, e^{r_n}) \cdot \underline{\mathcal{O}} \cdot (e^{s_1}, \ldots, e^{s_n}) \cdot \underline{\mathcal{O}}$ we will compute these two pieces separately. Define

$$A_j = (e^{r_{1j}}, \ldots, e^{r_{(n-1)j}}, (e^{-1})^{\sum_{i=0}^{n-1} r_{ij}}) \cdot \underline{\mathcal{O}},$$

where r_{ij} is $\text{sign}(r_i)$ times bit j of $|r_i|$. From the determinant formula it follows that the determinant of A_j is the determinant of $\mathcal{O}$ times $(e^{-1})^{\sum_i r_{ij}} \prod_{i=0}^{n-1} e^{r_{ij}} = e^{\sum_i r_{ij} - \sum_i r_{ij}} = 1$. This also bounds the powers of A_j. The log of the determinant of $\mathcal{O}$ and n define the input size to the problem.

The second part handles the fractional part of the t_i's by directly computing the ideal $A_{-1} = (e^{s_1}, \ldots, e^{s_n}) \cdot \underline{\mathcal{O}}$ using the first poly many terms in the formula $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$ to get the desired approximation. From the determinant formula it follows that $\det A_{-1}$ is $\prod_i e^{s_i}$ times the determinant of $\mathcal{O}$. The product $\prod_i e^{s_i}$ is between e^{-2n} and e^{2n}.

To see that $f(\boldsymbol{t}) = A_{-1} \cdot \prod_j A_j^{2^j}$, we have

$$\begin{aligned}
A_{-1} \cdot \prod_j A_j^{2^j} &= A_{-1} \cdot \prod_j \left((e^{r_{1j}}, \ldots, e^{r_{(n-1)j}}, (e^{-1})^{\sum_i r_{ij}}) \cdot \underline{\mathcal{O}} \right)^{2^j} \\
&= A_{-1} \cdot (e^{\sum_j r_{1j} 2^j}, \ldots, e^{\sum_j r_{(n-1)j} 2^j}, (e^{-1})^{\sum_{j,i} r_{ij} 2^j}) \cdot \underline{\mathcal{O}} \\
&= (e^{s_1}, \ldots, e^{s_n}) \cdot (e^{r_1}, \ldots, e^{r_{n-1}}, (e^{-1})^{\sum_i r_i}) \cdot \underline{\mathcal{O}} \\
&= (e^{t_1}, \ldots, e^{t_{n-1}}, (e^{-1})^{\sum_i t_i}) \cdot \underline{\mathcal{O}} = f(\boldsymbol{t})
\end{aligned}$$

The algorithm now works as follows. First compute a $\mathbb{Z}$-basis $\omega_1, \ldots, \omega_n$ of $\mathcal{O}$. Next compute the conjugate vector representation $\boldsymbol{z}_i = \underline{\omega_i}$. Compute A_{-1} as described above. Next compute each

A_j by first computing $(e^{r_{1j}}, \ldots, e^{r_{nj}})$ and then $(e^{r_{1j}}, \ldots, e^{r_{nj}}) \cdot \boldsymbol{z}_i$ for each i. Next use repeated squaring of ideals to compute a basis for $A_j = ((e^{t_{1j}}, \ldots, e^{t_{nj}}) \cdot \underline{\mathcal{O}})^{2^j}$. Finally, multiply the $A_j^{2^j}$'s and A_{-1}.

4.2 The algorithm for computing $e^{\boldsymbol{t}}\mathcal{O}$

Given t, compute a basis for $\boldsymbol{e^t\mathcal{O}}$:

1. Choose a polynomial q.
2. For each bit index j do the following:
3. Compute the diagonal matrix T_j, where $(T_j)_{i,i} = e^{r_{i,j}}$ for $i < n$, and $(T_j)_{n,n} = (e^{-1})^{\sum_{i=1}^{n-1} r_{ij}}$.
4. Compute $A_j := T_j \cdot \underline{\mathcal{O}}$ and compute a short basis for it using the algorithm in Section C.1.
5. Square A_j j times, using j applications of ideal multiplication below.
6. Multiply the resulting ideals together. To multiply two ideals B and C proceed as follows: Let the ideal B have basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ and ideal C have basis $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$.
   (a) Multiply pairwise columns to get n^2 vectors $\boldsymbol{c}_1\boldsymbol{b}_1, \boldsymbol{c}_1\boldsymbol{b}_2, \ldots, \boldsymbol{c}_1\boldsymbol{b}_n, \boldsymbol{c}_2\boldsymbol{b}_1, \ldots \boldsymbol{c}_n\boldsymbol{b}_n$.
   (b) Use the algorithm from Section C.1 to compute a short basis for BC.
   (c) Truncate the precision to q bits.

Together with the results from Section C.1 we have the following theorem.

Theorem 4.1. *There is an algorithm that on input $\boldsymbol{t} \in \mathbb{Q}^n$ and δ computes a basis $\hat{\boldsymbol{b}}_1, \ldots, \hat{\boldsymbol{b}}_n$ that is δ-close to a basis for $\boldsymbol{e^t\mathcal{O}}$, has bounded size, and runs in time polynomial in $\log \Delta$, n, $\log ||\boldsymbol{t}||$ and $\log 1/\delta$.*

Proof sketch. We analyze the complexity of the algorithm from Section 4.2. By Thiel [Thi95] we can take the initial precision as high as we need in Step 3 and Step 4. The main step in the algorithm multiplying ideals B and C, so it is enough to show that each multiplication step BC can be done efficiently and that we can bound the loss of precision. By Lemma C.1 we can compute n^2 generators for BC from n generators for B and n from C and bound the loss of precision. By Theorem C.5, we can compute a basis approximating $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ for BC in polynomial time, with

$$||\boldsymbol{b}_j|| \leq (\sqrt{n^3} + 2)2^{\frac{n^2-1}{2}} \cdot \lambda_j(BC).$$

The loss of precision for the squaring step is bounded in Theorem C.5. Therefore we may choose the initial number of precision bits q to be high enough to satisfy δ at the end. □

In the proof of the above theorem we needed to compute a bounded-length basis from a generating set. This is also used in the hidden subgroup problem algorithm when computing the basis for the unit lattice from a generating set for its dual. The input and output vectors for this are approximate and we need an algorithm that can find integer dependencies among the rounded vectors and also to bound the errors that result from the transformation to the reduced basis. An algorithm for computing a lattice basis from a set of generators was given by Buchmann and Pohst [BP87] and Buchmann and Kessler [BK93]. They do not bound all of the errors, though. In order to bound the errors that result from the transformation to the reduced basis we need better bounds on the coefficient sizes in the transformation than they give. For that reason we will present their

algorithm and improve their analysis. Both [BP87] and [BK93] analyze the same algorithm for computing a basis, so we give an outline of their algorithm and bounds before giving our analysis and proving the error bounds. This is done in Section C.1 in the appendix. More details for this section can be found in Appendix C.

5 Quantum representation of lattices

The representation of a lattice by a basis is not unique, which makes in unsuitable for use in an algorithm that deals with quantum superpositions of lattices. To avoid this issue, we will represent each lattice by a unique quantum state, namely, a superposition of the lattice points with certain weights. Let us first discuss some desired properties of such a representation. We want that a small deformation of the lattice result in a small change in the quantum state, whereas substantially different lattices be mapped to almost orthogonal vectors. These requirements can be formalized as follows.

Definition 5.1. Let $\mathrm{dist}(x, y)$ denote the distance between two points in a metric space X, and let $\mathcal{H}$ be some Hilbert space. A map $f : X \to \mathcal{H}$ is called an (a, r, ε) *quantum encoding* if the following conditions are met:

1. $\langle f(x)|f(x)\rangle = 1$ for all $x \in X$;
2. $\big\||f(x)\rangle - |f(y)\rangle\big\| \leqslant a \cdot \mathrm{dist}(x, y)$ for all $x, y \in X$;
3. If $\mathrm{dist}(x, y) \geqslant r$, then $\big|\langle f(x)|f(y)\rangle\big| \leqslant \varepsilon$.

Given such an encoding, the vector $|f(x)\rangle$ is called the *signature state* for x.

The number a in condition 2 is called a *Lipschitz constant* of function f. When $X = \mathbb{R}^n$ (or, more generally, when X is a Riemannian manifold), f can be approximated by a smooth function to an arbitrary precision in the sup-norm at cost of an arbitrary small parameter change in the above definition. For smooth functions, a is simply an upper bound on the first derivative of f.

Lemma 5.2. *Let $f_1 : X_1 \to \mathcal{H}_1$ and $f_2 : X_2 \to \mathcal{H}_2$ take values in unit vectors and satisfy the Lipschitz condition with constants a_1 and a_2, respectively.*

a) If $X_1 = X_2 = X$, then the function $f_1 \otimes f_2 : X \to \mathcal{H}_1 \otimes \mathcal{H}_2$ is $(a_1 + a_2)$-Lipschitz.

b) If X_1 and X_2 are Euclidean spaces, then the function $g : X_1 \times X_2 \to \mathcal{H}_1 \otimes \mathcal{H}_2$ defined as $g(x_1, x_2) = f_1(x_1) \otimes f_2(x_2)$ is a-Lipschitz, where $a = \sqrt{a_1^2 + a_2^2}$.

Example 5.3 (Straddle encoding). A representation of real numbers by quantum superposition of integers can be defined as follows:

$$\big|\mathrm{str}_\nu(x)\big\rangle = \cos\left(\tfrac{\pi}{2}t\right)|k\rangle + \sin\left(\tfrac{\pi}{2}t\right)|k+1\rangle, \qquad \text{where } \; k = \lfloor x/\nu \rfloor, \quad t = x/\nu - k. \tag{5.1}$$

The map $\mathrm{str}_\nu : \mathbb{R} \to \mathbb{C}^{\mathbb{Z}}$ is an $\left(\frac{\pi}{2\nu}, 2\nu, 0\right)$ quantum encoding. Applying this map to each coordinate on an n-dimensional real vector, we obtain an $\left(\frac{\pi}{2\nu}\sqrt{n}, 2\nu\sqrt{n}, 0\right)$ encoding $\mathrm{str}_{\nu,n} : \mathbb{R}^n \to \left(\mathbb{C}^{\mathbb{Z}}\right)^{\otimes n}$. (Here we have used statement (b) of Lemma 5.2.)

We usually set $\nu = 2^{-q}$. The transformation $|x\rangle \mapsto |x\rangle \otimes |\operatorname{str}_\nu(x)\rangle$ can be implemented efficiently if we assume that x is represented as $2^{-l}\tilde{x}$, where $l \geqslant q$ and $\tilde{x}$ is an integer, which is actually stored in the quantum memory. (In practice, l should be substantially greater than q so that the rounding error in x does not matter.) To construct $|x\rangle \otimes |\operatorname{str}_\nu(x)\rangle$ from $|x\rangle$, we compute k and t, create the state $\cos\left(\frac{\pi}{2}t\right)|0\rangle + \sin\left(\frac{\pi}{2}t\right)|1\rangle$, add k, and reverse the computation of k and t.

A full-rank lattice $L \subseteq \mathbb{R}^n$ may be represented by a superposition of its points with Gaussian amplitudes. Each lattice point is in turn represented using the straddle encoding. Thus,

$$\boxed{|f(L)\rangle = \gamma^{-1/2} \sum_{x\in L} e^{-\pi\|x\|^2/s^2} \left|\operatorname{str}_{n,\nu}(x)\right\rangle} \qquad \text{where} \quad \gamma = \sum_{x\in L} e^{-2\pi\|x\|^2/s^2}. \tag{5.2}$$

We must prove the HSP properties for this state when the lattice L is of the form $e^{\boldsymbol{t}}\mathcal{O}$. To prove that the states over different ideal lattices have small inner product when $\boldsymbol{t}_1 - \boldsymbol{t}_2$ is not near a lattice point we need the following lemma showing that the two corresponding ideal lattices do not have too many points close to each other. The proof is in Appendix E.1.

Lemma 5.4. *Let $e^{\boldsymbol{t}_1}$ and $e^{\boldsymbol{t}_2}$ be vectors of algebraic norm 1. If some nonzero point of $e^{\boldsymbol{t}_1}\mathcal{O}$ is equal to some point of $e^{\boldsymbol{t}_2}\mathcal{O}$ and has length at most R, then the distance between any unequal pair of points, one from $e^{\boldsymbol{t}_1}\mathcal{O}$ and one from $e^{\boldsymbol{t}_2}\mathcal{O}$, is at least $\sqrt{n}/R^n$.*

In Appendix E.3 we prove the main HSP property which is stated next. Since the inner product is at most a constant < 1, a tensor product of n copies will reduce the inner product to exponentially small. Here we state the theorem in terms of the free part of the unit group $\operatorname{Log}\mathcal{O}^* \leq \mathbb{R}^{s+t-1}$.

Theorem 5.5. *Choose $s = 2^{2n}\sqrt{nD}$, $\nu = 1/(4n(s\sqrt{n})^{2n})$. Let $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathbb{R}^n$. Let $\gamma = (\gamma_1, \ldots, \gamma_n)$ be $\boldsymbol{t}_2 - \boldsymbol{t}_1 - u$, where u is the unit closest to $\boldsymbol{t}_2 - \boldsymbol{t}_1$ in $\operatorname{Log}\mathcal{O}^*$. Then the function f is Lipschitz with the constant specified in Theorem D.4. The inner product $\langle e^{\boldsymbol{t}_1}\mathcal{O}|e^{\boldsymbol{t}_2}\mathcal{O}\rangle$ is at most 3/4 if for some i, either $\ln(1 - (s\sqrt{n})^{n-1}2\nu\sqrt{n}) \geq \gamma_i$ or $\gamma_i \geq \ln(1 + (s\sqrt{n})^{n-1}2\nu\sqrt{n})$.*

6 The hidden subgroup problem on a continuous group

First we show that known cases of the abelian HSP can be reduced to the new continuous case in Definition 1.1 over the new HSP instance $\mathbb{R}^m$.

6.1 Application: Reduction to the case $G = \mathbb{R}^m$

Our HSP algorithm is applicable to Abelian groups of the form $\mathbb{R}^k \times \mathbb{Z}^l \times (\mathbb{R}/\mathbb{Z})^s \times H$, where H is finite. We call such groups "elementary". The reduction to $G = \mathbb{R}^k \times \mathbb{Z}^l$ is straightforward. In the case of interest, the hidden subgroup L is a full-rank lattice in $G \subseteq \mathbb{R}^k \times \mathbb{R}^l$ such that $\lambda_1(L \cap \mathbb{R}^k) \geqslant \lambda$ and $d(L) \leqslant d$ for some fixed numbers λ and d. We now describe the further reduction to the group $\tilde{G} = \mathbb{R}^{k+l}$.

The main idea can be illustrated in the one-dimensional case, where the parameter λ has no meaning. We embed $G = \mathbb{Z}$ into $\tilde{G} = \mathbb{R}$ in the standard way, set $\nu = 2^{-q}$ for some $q \geqslant 2$, and define the $\mathbb{R}$-oracle g in terms of the $\mathbb{Z}$-oracle f as follows:

$$\begin{aligned} &\left|g(x)\right\rangle = c_0 \left|\operatorname{str}_\nu(t)\right\rangle \otimes \left|f(s)\right\rangle + c_1 \left|\operatorname{str}_\nu(t-1)\right\rangle \otimes \left|f(s+1)\right\rangle, \\ &\text{where} \quad s = \lfloor x \rfloor, \quad t = x - s, \quad c_0 = \cos\left(\tfrac{\pi}{2}t\right), \quad c_1 = \sin\left(\tfrac{\pi}{2}t\right). \end{aligned} \tag{6.1}$$

It is clear that g is a continuous function. If f is a periodic function, then g is also periodic with the same period. We will estimate the parameters of the new oracle later in the more general setting.

To construct the state $|g(x)\rangle$ using the original oracle f, we compute s and t, use them as parameters in the following sequence of operations, and "uncompute" s and t:

$$|0\rangle \mapsto \sum_z c_z|z\rangle \mapsto \sum_z c_z\,|z\rangle \otimes \big|f(s+z)\big\rangle \mapsto \sum_z c_z\,\big|\mathrm{str}_\nu(t-z)\big\rangle \otimes \big|f(s+z)\big\rangle,$$

where $z \in \{0,1\}$. The last step, $|z\rangle \mapsto |\,\mathrm{str}_\nu(t-z)\rangle$ requires that we discriminate between the states $|\,\mathrm{str}_\nu(t)\rangle$ and $|\,\mathrm{str}_\nu(t-1)\rangle$. This is easy because the supports of those states on the ν-grid do not overlap.

Let us now consider the general case, $G = \mathbb{R}^k \times \mathbb{Z}^l$. The group G is embedded in $\tilde{G} = \mathbb{R}^{k+l}$ by scaling the $\mathbb{Z}$ factors by λ. This is to guarantee that $\lambda_1(\tilde{L}) \geqslant \lambda$, where $\tilde{L}$ is the image of L under the embedding. The other condition on the new hidden subgroup reads: $d(\tilde{L}) \leqslant \tilde{d}$, where $\tilde{d} = d\lambda^l$. The generalization of Eq. (6.1) is straightforward:

$$\big|g(\boldsymbol{x}, x_1, \ldots, x_l)\big\rangle = \sum_{z_1,\ldots,z_l \in \{0,1\}} \left(\bigotimes_{j=1}^{l} |\psi(x_j, z_j)\rangle\right) \otimes \big|f(\boldsymbol{x},\, s(x_1,z_1), \ldots, s(x_l,z_l)\big\rangle, \tag{6.2}$$

$$\text{where} \quad s(x,z) = \lfloor x/\lambda \rfloor + z, \qquad |\psi(x,z)\rangle = \cos\left(\tfrac{\pi}{2}t\right)\big|\mathrm{str}_\nu(t)\big\rangle \quad \text{with } t = x/\lambda - s(x,z).$$

Note that the terms in the above sum are mutually orthogonal vectors.

Theorem 6.1. *If the oracle function f has parameters (a, r, ε) (see Definition F.1), then the new function g has the following parameters:*

$$\tilde{a}^2 = a^2 + l\left(\frac{\pi}{2\nu\lambda}(1+\nu)\right)^2, \qquad \tilde{r}^2 = r^2 + l(2\nu\lambda)^2, \qquad \tilde{\varepsilon} = \varepsilon.$$

The proof is in the Appendix.

6.2 An HSP algorithm for the new group $\mathbb{R}^m$

The HSP algorithm for $\mathbb{R}^m$ can be thought of in the usual structure, however the analysis is difficult and relies on our new continuous definition of the HSP. The analysis works in continuous space and is discretized in a general way to derive the algorithm. Imagine creating a superposition of points in $\mathbb{R}^m$ with a sufficiently broad wavefunction w, applying the oracle, and measuring in the Fourier basis. We show that Fourier sampling generates a point u of the reciprocal lattice L^* with the probability distribution

$$q_u = \frac{1}{d(L)^2} \int_{(\mathbb{R}^m/L)^2} \langle f_{x'}|f_x\rangle\, e^{2\pi i\langle x-x', u\rangle}\, dx\, dx'. \tag{6.3}$$

This distribution is not close to uniform but we are able to show that the probability of staying in any sublattice is bounded. For rounding, the Fourier samples deviate from the lattice points by, roughly, the inverse width of the wavefunction w. We show that reconstruction of a lattice from an approximate generating set can be done using an improved analysis of [BK93] in Section C.1. In Section C.2 we show that the condition number of a reduced basis is bounded so that the dual lattice, which is the hidden subgroup, can be computed.

Theorem 6.2. *There is a polynomial time quantum algorithm solving the HSP over $\mathbb{R}^m$.*

Proof sketch. First we create the superposition $|\psi_\delta\rangle = \delta^{m/2} \sum_{\tilde{x}\in\mathbb{Z}^m} w(x)\,|\tilde{x}\rangle \otimes |f(x)\rangle$ with $x = \delta\tilde{x}$.. This is a valid superposition because the window function w is zero outside a finite range and the sum is over $\mathbb{Z}^m$ inside that range. By Theorem F.9 we can sample points close to L^* points. By Theorem F.13 a polynomial number of samples generate L^*. By Section C.2 we can compute a basis for L. □

References

[Ban93] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(1):625–635, 1993.

[BK93] Johannes Buchmann and Volker Kessler. Computing a reduced lattice basis from a generating system, 1993. Preprint, August 4, 1993.

[BP87] Johannes Buchmann and Michael Pohst. Computing a lattice basis from a system of generating vectors. In *Eurocal'87*, volume 378 of *LNCS*, pages 54–63. Springer-Verlag, June 1987.

[BV11] Zvika Brakerski and Vinod Vaikuntanathan. Fully homomorphic encryption from ring-LWE and security for key dependent messages. In *Advances in cryptology—CRYPTO 2011*, volume 6841 of *LNCS*, pages 505–524. Springer, 2011.

[Coh93] Henri Cohen. *A course in computational algebraic number theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1993.

[FIM+03] Katalin Friedl, Gabor Ivanyos, Frederic Magniez, Miklos Santha, and Pranab Sen. Hidden translation and orbit coset in quantum computing. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, San Diego, CA, 9–11June 2003.

[GH11] C. Gentry and S. Halevi. Implementing gentry's fully-homomorphic encryption scheme. *Eurocrypt 2011*, pages 132–150, 2011.

[GKP01] Daniel Gottesman, Alexei Kitaev, and John Preskill. Encoding a qubit in an oscillator. *Phys. Rev. A*, 64:012310, Jun 2001.

[GR02] Lov Grover and Terry Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions, August 2002. arxiv:quant-ph/0208112.

[GVL96] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, 3rd edition, 1996.

[Hal05] Sean Hallgren. Fast quantum algorithms for computing the unit group and class group of a number field. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 468–474, 2005.

[Hal07] Sean Hallgren. Polynomial-time quantum algorithms for Pell's equation and the principal ideal problem. *Journal of the ACM*, 54(1):1–19, 2007.

[HMR+10] Sean Hallgren, Cristopher Moore, Martin Rötteler, Alexander Russell, and Pranab Sen. Limitations of quantum coset states for graph isomorphism. *J. ACM*, 57:34:1–34:33, November 2010.

[Kit95] Alexei Kitaev. Quantum measurements and the abelian stabilizer problem, 1995. quant-ph/9511026.

[KW08] Alexei Kitaev and William A. Webb. Wavefunction preparation and resampling using a quantum computer, January 2008. arXiv:0801.03422.

[LPR10] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In *Advances in cryptology—EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 1–23. Springer, 2010.

[Mic08] Daniele Micciancio. Efficient reductions among lattice problems. In *Proceedings of SODA 2008*, pages 84–93, New York, 2008. ACM.

[PR07] Chris Peikert and Alon Rosen. Lattices that admit logarithmic worst-case to average-case connection factors. In *STOC '07: Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 478–487, New York, NY, USA, 2007. ACM Press.

[San91] Jonathan W. Sands. Generalization of a theorem of Siegel. *Acta Arith.*, 58(1):47–57, 1991.

[Sho97] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[SV05] Arthur Schmidt and Ulrich Vollmer. Polynomial time quantum algorithm for the computation of the unit group of a number field. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 475–480, 2005.

[Thi95] Christoph Thiel. *On the complexity of some problems in algorithmic algebraic number theory*. PhD thesis, Universität des Saarlandes, Saarbrücken, Germany, 1995.

A Background on lattices

A.1 Notation and elementary facts

A basis in $\mathbb{R}^n$ can be represented by a matrix $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ whose columns are the basis vectors. The dual basis $B^* = (\boldsymbol{b}_1^*, \ldots, \boldsymbol{b}_n^*)$ is defined by the condition $\langle \boldsymbol{b}_j^*, \boldsymbol{b}_k \rangle = \delta_{jk}$. It is clear that

$$B^* = (B^{-1})^T. \tag{A.1}$$

If $L \subset \mathbb{R}^n$ is a full-rank lattice with basis vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$, then $\boldsymbol{b}_1^*, \ldots, \boldsymbol{b}_n^*$ form a basis of the reciprocal lattice L^*.

The unit cell volume of lattice L is denoted by $d(L)$. It is clear that $d(L) = |\det B|$.

We denote by $\boldsymbol{b}_k^\perp$ the k-th Gram-Schmidt vector, i.e. the projection of $\boldsymbol{b}_k$ on the orthogonal complement to the subspace spanned by $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{k-1}$. Note that

$$|\det B| = \prod_{k=1}^{n} \|\boldsymbol{b}_k^\perp\|. \tag{A.2}$$

Since $\|\boldsymbol{b}_k^\perp\| \leqslant \|\boldsymbol{b}_k\|$, the above equation implies the *Hadamard inequality,*

$$|\det B| \leqslant \prod_{k=1}^{n} \|\boldsymbol{b}_k\|. \tag{A.3}$$

For an arbitrary set S in a real linear space, $\overline{S}$ denotes the linear span of S. If L is a rank n lattice, then $T = \overline{L}/L$ is an n-dimensional torus. The *fundamental parallelepiped* corresponding to

some basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ is

$$\mathcal{P}(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) = \left\{ \sum_{k=1}^{n} \mu_k \boldsymbol{b}_k : \ \mu_k \in \left[-\tfrac{1}{2}, \tfrac{1}{2}\right) \right\}. \tag{A.4}$$

A.2 Successive minima

Definition A.1. Associated to any rank n lattice L and a symmetric convex body S are the *successive minima* $\lambda_1, \ldots, \lambda_n$. The *$k$-th minimum* is defined as follows:

$$\lambda_k(S, L) = \inf\{r : \ rS \text{ contains } k \text{ linearly independent vectors of } L\}.$$

We usually take S to be the unit ball and drop it from the notation. Thus, $\lambda_1(L)$ is the length of a shortest nonzero vector in L.

Theorem A.2. *If $L \subset \mathbb{R}^n$ is a full-rank lattice, then $\lambda_k(L)\, \lambda_{n-k+1}(L^*) \geqslant 1$.*

Proof. Let $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n \in L$ be linearly independent vectors such that $\|\boldsymbol{b}_j\| = \lambda_j(L)$, and let $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n \in L^*$ be similar vectors for the reciprocal lattice. Consider the matrix of inner products $\langle \boldsymbol{h}_i, \boldsymbol{b}_j \rangle$, where $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, k\}$. Its columns (indexed by j) are linearly independent, hence there are k linearly independent rows. In particular, the first $n-k+1$ rows cannot all be zero, i.e.

$$\langle \boldsymbol{h}_i, \boldsymbol{b}_j \rangle \neq 0 \quad \text{for some } i \leqslant n-k+1 \text{ and } j \leqslant k.$$

It follows that $\|\boldsymbol{h}_i\| \cdot \|\boldsymbol{b}_j\| \geqslant 1$ because $\langle \boldsymbol{h}_i, \boldsymbol{b}_j \rangle$ is an integer. On the other hand, $\|\boldsymbol{h}_i\| \leqslant \lambda_{n-k+1}(L^*)$ and $\|\boldsymbol{b}_j\| \leqslant \lambda_k(L)$, which implies the desired result. □

Theorem A.3 (Minkowski's Convex Body Theorem). *Let L be a lattice of rank n and $S \subset \overline{L}$ a convex body symmetric about the origin. If $\operatorname{vol}(S) > 2^n d(L)$, then S contains a nonzero lattice point.*

Minkowski's theorem is equivalent to the inequality $\lambda_1(S, L) \leqslant 2\big(d(L)/\operatorname{vol}(S)\big)^{1/n}$ for an arbitrary symmetric convex body S. When S is the unit ball, we have $\operatorname{vol}(S) \geqslant \big(2/\sqrt{n}\big)^n$, hence

$$\boxed{\lambda_1(L) \leqslant \sqrt{n}\, d(L)^{1/n}} \tag{A.5}$$

Theorem A.4 (Minkowski's Second Theorem). *For any rank n lattice L, the successive minima satisfy*

$$\prod_{k=1}^{n} \lambda_k(S, L) \leqslant \frac{2^n d(L)}{\operatorname{vol}(S)}.$$

Minkowski's Second Theorem for the unit ball is expressed by this inequality:

$$\prod_{i=k}^{n} \lambda_k(L) \leqslant n^{n/2} d(L). \tag{A.6}$$

It implies the following bound:

$$\boxed{\lambda_n(L) \leqslant \frac{n^{n/2} d(L)}{\lambda_1(L)^{n-1}}} \tag{A.7}$$

A.3 Norms

Consider two matrices, $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ and $\widetilde{B} = (\widetilde{\boldsymbol{b}}_1, \ldots, \widetilde{\boldsymbol{b}}_n)$, whose columns are the exact and approximate basis vectors, respectively. We define $\operatorname{dist}(B, \widetilde{B})$ to be the geodesic distance on the group $\mathbf{GL}(\mathbb{R}, n)$:

$$\operatorname{dist}(B, \widetilde{B}) \;=\; \inf\bigl\{\|A\|_2 :\; e^A \widetilde{B} = B\bigr\}, \tag{A.8}$$

where $\|A\|_2$ is the Hilbert-Schmidt norm of matrix A,

$$\|(a_{jk})\|_2 \stackrel{\text{def}}{=} \sum_{j,k} a_{jk}^2.$$

The distance function satisfies the standard axioms; it is also preserved by passing to the dual basis:

$$\operatorname{dist}\bigl((B^{-1})^T, (\widetilde{B}^{-1})^T\bigr) \;=\; \operatorname{dist}(B, \widetilde{B}). \tag{A.9}$$

In practice, we are interested in the case where $\operatorname{dist}(B, \widetilde{B})$ is small, so that

$$\operatorname{dist}(B, \widetilde{B}) \approx \bigl\|(\delta B)\widetilde{B}^{-1}\bigr\|_2, \qquad \delta B = B - \widetilde{B}. \tag{A.10}$$

A.4 The LLL algorithm and reduced bases

Let $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k)$ be an arbitrary basis of some lattice Λ (not necessarily sorted, but the order of vectors is important).

Let E_k and Λ_k be the real and integer linear span of $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$, respectively. Let $\boldsymbol{b}_k^\perp$ stand for the Gram-Schmidt vector, i.e., the projection of $\boldsymbol{b}_k$ onto the orthogonal complement to E_{k-1}. We write $\boldsymbol{b}_k^\perp = u_k \boldsymbol{e}_k$, where $\boldsymbol{e}_k$ is a unit vector; these vectors form an orthonormal basis. Note that $u_k = \|\boldsymbol{b}_k^\perp\| = d(\Lambda_k)/d(\Lambda_{k-1})$. Gram-Schmidt coefficients $\mu_{k,j}$ are defined by any of these equivalent equations:

$$\boldsymbol{b}_k \;=\; \boldsymbol{b}_k^\perp + \sum_{j=1}^{k-1} \mu_{k,j} \boldsymbol{b}_j^\perp \;=\; u_k \boldsymbol{e}_k + \sum_{j=1}^{k-1} \mu_{k,j} u_j \boldsymbol{e}_j,$$
$$\mu_{k,j} = (\boldsymbol{e}_j, \boldsymbol{b}_k)/u_j.$$

Definition A.5. Let τ_1 and τ_2 be some constants such that $1/2 < \tau_1 < \tau_2 < 1$. A sequence of linearly independent vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n \in \Lambda$ is called an *LLL-reduced basis* if the following conditions are fulfilled:

$$|\mu_{k,j}| \;\leqslant\; \tau_1 \quad \text{for } 1 \leqslant j < k \leqslant n, \tag{A.11}$$

$$\bigl\|\boldsymbol{b}_k^\perp + \mu_{k,k-1}\boldsymbol{b}_{k-1}^\perp\bigr\| \;\geqslant\; \tau_2 \|\boldsymbol{b}_{k-1}^\perp\| \tag{A.12}$$

$$\Bigl(\text{equivalently, } u_k^2 \geqslant (\tau_2^2 - \mu_{k,k-1}^2) u_{k-1}\Bigr) \quad \text{for } k = 2, \ldots, n. \tag{A.13}$$

(When testing these conditions numerically, one may replace τ_1 by $1/2$ and τ_2 by a slightly larger number τ_2' to give some room for round-off errors.)

The first inequality is easy to satisfy if we add to $\boldsymbol{b}_k$ a suitable linear combination of $\boldsymbol{b}_j$ with integer coefficients. The exact rule is as follows:

$$\textbf{reduce}(k) \quad \text{for} \quad j = k-1, \ldots, 1$$
$$\boldsymbol{b}_k \;\leftarrow\; \boldsymbol{b}_k - \operatorname{round}\bigl((\boldsymbol{e}_j, \boldsymbol{b}_k)/u_j\bigr)\, \boldsymbol{b}_j, \tag{A.14}$$

where "round" means taking the closest integer. Note that this procedure (which we refer to as *reduction step*) does not change the subspaces E_j or the sublattices Λ_j. We also have

$$u_k \geqslant \tau u_{k-1}, \qquad \text{where} \quad \tau = \sqrt{\tau_2^2 - \tau_1^2}. \tag{A.15}$$

An LLL-reduced basis can be computed in time polynomial in the size of a rational input.

B More number theoretic background

B.1 Units and the logarithmic embedding

The *group of units* $\mathcal{O}^*$ is embedded in the group $E^* = (\mathbb{R}^s \times \mathbb{C}^t)^*$ of invertible conjugate vectors. Such vectors can be parametrized using *logarithmic coordinates,* namely

$$\boldsymbol{z} = \Bigl(\underbrace{\dots\,(-1)^{\mu^{(j)}} e^{u^{(j)}}\,\dots}_{j=1,\dots,s}, \underbrace{\dots\,e^{2\pi i\varphi^{(j)}} e^{u^{(s+j)}/2}\,\dots}_{j=1,\dots,t}, \underbrace{\dots\,e^{-2\pi i\varphi^{(j)}} e^{u^{(s+j)}/2}\,\dots}_{j=1,\dots,t}\Bigr)^T,$$

where $u^{(1)},\dots,u^{(s+t)}$ are real numbers, $\mu^{(1)},\dots,\mu^{(s)} \in \{0,1\}$ and $\varphi^{(1)},\dots,\varphi^{(t)} \in \mathbb{R}/\mathbb{Z}$. A natural distance between points in E^* comes from the following considerations. If two points, $\boldsymbol{z}$ and $\widetilde{\boldsymbol{z}}$ are infinitely close to each other and $\delta\boldsymbol{z} := \widetilde{\boldsymbol{z}} - \boldsymbol{z}$, then

$$\bigl\|(\delta\boldsymbol{z})\boldsymbol{z}^{-1}\bigr\|^2 = \sum_{j=1}^{s}(\delta u^{(j)})^2 + \frac{1}{2}\sum_{j=s+1}^{s+t}(\delta u^{(j)})^2 + 8\pi^2\sum_{j=1}^{t}(2\pi\delta\varphi^{(j)})^2. \tag{B.1}$$

The right-hand side of this equation defines an invariant Riemannian metric on the group E^*. In more elementary terms, it defines a Euclidean inner product on $\mathbb{R}^{s+t} \times \mathbb{R}^t$.

Now, $z \in \mathcal{O}$ is a unit if and only if $\mathcal{N}(z) = \pm 1$. Thus, the group of units $\mathcal{O}^*$ is contained in the subgroup $G \subset E^*$ given by the equation $\sum_{j=1}^{s+t} u^{(j)} = 0$. Note that

$$G \cong \mathbb{R}^{s+t-1} \times (\mathbb{Z}_2)^s \times (\mathbb{R}/\mathbb{Z})^t. \tag{B.2}$$

In fact, $\mathcal{O}^*$ is a discrete subgroup of G, therefore the group $\mu(K) = \mathcal{O}^* \cap \bigl((\mathbb{Z}_2)^s \times (\mathbb{R}/\mathbb{Z})^t\bigr)$ is finite. The elements of $\mu(K)$ are exactly the torsion elements of $\mathcal{O}^*$, i.e. the roots of unity in K. The quotient $L = \mathcal{O}^*/\mu(K)$ is a lattice in $\mathbb{R}^{s+t-1}$, and Dirichlet's Theorem says that L has full rank.

An important parameter is the unit cell volume $d(L)$ of lattice L. Let $\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_{s+t-1}\}$ be a basis of L. To calculate $d(L)$, we augment the basis by an orthogonal unit vector, $\boldsymbol{v} = \frac{1}{\sqrt{n}}(1,\dots,1,2,\dots,2)^T$. Thus,

$$\begin{aligned} d(L) &= 2^{-t/2}\bigl|\det(\boldsymbol{u}_1,\dots,\boldsymbol{u}_{s+t-1},\boldsymbol{v})\bigr| \\ &= 2^{-t/2}\sqrt{n}\,R, \end{aligned} \qquad \text{where} \quad R = \left|\det\begin{pmatrix} u_1^{(1)} & \dots & u_{s+t-1}^{(1)} \\ \vdots & \ddots & \vdots \\ u_1^{(s+t-1)} & \dots & u_{s+t-1}^{(s+t-1)} \end{pmatrix}\right|.$$

The number R is called the *regulator* of K.

B.2 Ideals

Definition B.1. A *ideal* of $\mathcal{O}$ is a lattice $\Lambda \subseteq \mathbb{R}^s \times \mathbb{C}^t$ such that $z\Lambda \subseteq \Lambda$ for all $z \in \mathcal{O}$. (This lattice is either zero or has full rank.) An ideal Λ is called *integral* if $\Lambda \subseteq \underline{\mathcal{O}}$. Two ideals are called *equivalent* and we write $\Lambda' \sim \Lambda$ if $\Lambda' = \boldsymbol{v}\Lambda$ for some $\boldsymbol{v} \in (\mathbb{R}^s \times \mathbb{C}^t)^*$. The ideals equivalent to $\underline{\mathcal{O}}$ are called *principal*. The number

$$\mathcal{N}(\Lambda) = d(\Lambda)/d(\underline{\mathcal{O}}) = d(\Lambda)/\sqrt{|D|}$$

is called the *norm* of Λ.

For nonzero $\boldsymbol{x} \in \Lambda$, $|\mathcal{N}(\boldsymbol{x})| \geqslant \mathcal{N}(\Lambda)$. Applying the inequality between the geometric and arithmetic means, we get this result:

Lemma B.2. *The length of a shortest nonzero vector in an arbitrary ideal Λ is at least $\sqrt{n}\,\mathcal{N}(\Lambda)^{1/n}$.*

B.3 Some important bounds

Before we can discuss an algorithm for computing the group of units $\mathcal{O}^*$, we need to know the size of the numbers that would describe this group with a reasonable precision. First, let us estimate the smallest distance between two different units. In fact, Lemma B.2 (applied to $\Lambda = \mathcal{O}$) already gives such a bound; let us express it in terms of the logarithmic coordinates.

Lemma B.3. *The distance between the trivial and a nontrivial unit (in the sense of metric (B.1)) is greater or equal to* $1/2$.

Proof. Let $\boldsymbol{z}$ be given by the vector $\boldsymbol{u} = (u^{(1)}, \dots, u^{(s+t)}, \varphi^{(1)}, \dots, \varphi^{(t)})$ in the logarithmic coordinates. We show that if the length of $\boldsymbol{u}$ is smaller than $1/2$, then each coordinate of the conjugate vector $\boldsymbol{z}-\boldsymbol{1}$ is bounded by 1, implying that $\|\boldsymbol{z}-\boldsymbol{1}\| < \sqrt{n}$. Indeed, if the length of $\boldsymbol{u}$ is less than $1/2$, then $|u^{(j)}| < 1/2$ for $j = 1, \dots, s$, and hence $|z^{(j)} - 1| = |e^{u^{(j)}} - 1| < 1$. Similarly, for $j = 1, \dots, t$ we have the inequalities $|u^{(j)}|/\sqrt{2} < 1/2$ and $2\pi\sqrt{2}|\varphi^{(j)}| < 1/2$, hence $|e^{2\pi\varphi^{(j)}} e^{u^{(s+j)}} - 1|^2 < 1$. Thus, $\|\boldsymbol{z} - \boldsymbol{1}\| < \sqrt{n}$, and therefore $\boldsymbol{z}$ cannot be a nontrivial unit by Lemma B.2. □

As we reduce the problem of finding the subgroup $\mathcal{O}^* \subseteq G$ to that of a subgroup (actually, a full-rank lattice) $M \subset \mathbb{R}^m$, the previous lemma gives a lower bound on the length of a shortest nonzero vector: $\lambda_1(M) \geqslant 1/2$. Another important parameter is the unit cell volume, $d(M)$. In terms of the original problem, it is

$$\operatorname{vol}(G/\mathcal{O}^*) = \frac{2^s (2\pi)^t R}{\omega_K} \sqrt{n},$$

where $\omega_K = |\mu(K)|$ is the size of the torsion subgroup of $\mathcal{O}^*$. An upper bound on $\operatorname{vol}(G/\mathcal{O}^*)$ follows from a more general result. Let $\mathcal{O}$ be an arbitrary order. The definitions of the discriminant, the torsion subgroup, and the regulator are still applicable, as well as Dirichlet's Theorem. Let $h_{\mathcal{O}}$ be the number of equivalence classes of invertible ideals. Sands [San91, Theorem 5.4] showed that

$$\frac{2^s R_{\mathcal{O}}}{\omega_{\mathcal{O}}}\, h_{\mathcal{O}} < 4 \left(\frac{2}{n-1}\right)^{n-1} \sqrt{|D_{\mathcal{O}}|}\, \big(\ln |D_{\mathcal{O}}|\big)^{n-1} \big(2 \ln\ln |D_{\mathcal{O}}|\big)^{n/2}.$$

In our case, we obtain the following bound:

$$\begin{aligned} \operatorname{vol}(G/\mathcal{O}^*) &< 4n^{1/2} \left(\frac{2}{n-1}\right)^{n-1} \sqrt{|D|}\, \big(\ln |D|\big)^{n-1} \big(4\pi \ln\ln |D|\big)^{n/2} \\ &\leqslant O\left(\sqrt{|D_{\mathcal{O}}|}\, \big(\log |D_{\mathcal{O}}|\big)^n\right). \end{aligned} \tag{B.3}$$

C Classical computation of an approximate basis of $e^t\mathcal{O}$

In this section we analyze the algorithm in Section 4.2 for computing an approximate basis of the ideal lattice $e^{\boldsymbol{t}}\mathcal{O}$. A real number x is typically approximated by a rational number $\tilde{x} = \frac{p}{2^m}$ such that $|\tilde{x} - x| \leqslant \varepsilon$. The parameter ε is called *absolute error* and the ratio $\gamma = \varepsilon/|x|$ is called *relative error*. A vector $\boldsymbol{x} = (x^{(1)}, \ldots, x^{(n)})$ is approximated by another $\widetilde{\boldsymbol{x}} = (\tilde{x}^{(1)}, \ldots, \tilde{x}^{(n)})$ such that $\|\widetilde{\boldsymbol{x}} - \boldsymbol{x}\| \leqslant \varepsilon$; the relative error is defined as $\gamma = \varepsilon/\|\boldsymbol{x}\|$. The choice of norm is not important since we will have factors of the form $2^{-n^{\Theta(1)}}$.

We start with a lemma bounding the error in the ideal multiplication step.

Lemma C.1. *Suppose a basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ for a lattice $\boldsymbol{e^{t_1}}\mathcal{O}$ and a basis $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ for $\boldsymbol{e^{t_2}}\mathcal{O}$ are given with relative precision γ, and each $\boldsymbol{e^{t_i}}$ has norm 1, then the resulting products $\boldsymbol{b}_i\boldsymbol{c}_j$ have relative precision $4\gamma\sqrt{n}||\boldsymbol{b}_i||^n$.*

Proof. Suppose we are multiplying vector $\boldsymbol{r}$ by vector $\boldsymbol{s}$ point-wise, and each is given with relative precision ϵ. Let A be a diagonal matrix with $1/r_i$ in the (i,i) position, $\boldsymbol{b} := \boldsymbol{s}$, so that $\boldsymbol{rs} = A^{-1}\boldsymbol{b} =: \boldsymbol{x}$. The matrix A can be computed from $\boldsymbol{r}$ by just inverting the diagonal elements, with small precision loss. By Lemma 2.7.1 and Theorem 2.7.2 in [GVL96], if A has relative precision ϵ, and $\boldsymbol{b}$ has relative precision ϵ, then the solution $\boldsymbol{x} = \boldsymbol{rs}$ to the equation $\hat{A}\boldsymbol{x} = \hat{\boldsymbol{b}}$, where $\hat{A}$ and $\hat{\boldsymbol{b}}$ are the rounded versions, must have relative precision bounded by $2\epsilon\kappa(A)/(1-\epsilon\kappa(A)) \leq 4\epsilon\kappa(A) \leq 4\epsilon||A||_2 \cdot ||A^{-1}||_2 \leq 4\epsilon||\boldsymbol{r}||_\infty^{n-1} \cdot ||r|| \leq 4\epsilon\sqrt{n}||\boldsymbol{r}||^n$. Here we use the fact that $\boldsymbol{r}$ has an integer norm, so $\prod_i r_i \geq 1$, and so the maximum $1/r_i$ component is at most the product of the rest. □

In the next section we show how to compute a basis from a generating set.

C.1 Computing a basis of a lattice from an approximate generating set

Suppose that vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \in \mathbb{R}^n$ generate an r-dimensional lattice L. We are given approximations $\hat{\boldsymbol{a}}_1, \ldots, \hat{\boldsymbol{a}}_k \in \mathbb{Z}^n$ such that $||\boldsymbol{a}_i - \hat{\boldsymbol{a}}_i/2^q|| \leq \sqrt{n}/2^{q+1}$. We want to compute a set of vectors $\hat{\boldsymbol{c}}_1, \ldots, \hat{\boldsymbol{c}}_r$ which approximate a basis $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_r$ of L with some precision q', which will necessarily be smaller than q. The problem of computing a basis for a lattice from a set of generators is addressed in [BP87] and [BK93], and both papers use the same algorithm which is a reduction to LLL. We will use their algorithm to solve this problem, but we need a better analysis to make sure that the output accuracy q' is not too much smaller than q. This is done in Theorem C.5 below.

C.1.1 The Buchmann-Pohst algorithm

The reduction to LLL takes the input vectors $\hat{\boldsymbol{a}}_1, \ldots, \hat{\boldsymbol{a}}_k$ and creates a new lattice with basis defined from the concatenated vectors

$$\tilde{\boldsymbol{a}}_j = (\boldsymbol{e}_j, \hat{\boldsymbol{a}}_j) \quad (1 \leq j \leq k),$$

where $\boldsymbol{e}_j$ denotes the j-th unit vector in $\mathbb{Z}^k$. These vectors $\tilde{\boldsymbol{a}}_1, \ldots \tilde{\boldsymbol{a}}_k \in \mathbb{Z}^{n+k}$ are clearly linearly independent and form a basis of the lattice $\tilde{L} = \bigoplus_{j=1}^k \mathbb{Z}\tilde{\boldsymbol{a}}_j$. Note that with this setup, the bottom of the matrix has vectors $\lfloor 2^q \boldsymbol{a}_i \rceil$ as the basis vectors, where $\lfloor \cdot \rceil$ rounds each coordinate of the vector. The LLL-algorithm is then applied to the lattice basis $\tilde{\boldsymbol{a}}_1, \ldots \tilde{\boldsymbol{a}}_k$ to obtain an LLL-reduced basis $\tilde{\boldsymbol{b}}_1, \ldots, \tilde{\boldsymbol{b}}_k$. For the first $k - r$ of these vectors we will denote the top and bottom components as

$$\tilde{\boldsymbol{b}}_j = (\boldsymbol{m}_j, \hat{\boldsymbol{z}}_j)^t. \quad (1 \leq j \leq k - r),$$

and for the last r vectors as

$$\tilde{\boldsymbol{b}}_{k-r+j} = (\boldsymbol{m}'_j, \hat{\boldsymbol{b}}_j)^t. \quad (1 \leq j \leq r).$$

Note that the vectors $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_{k-r} \in \mathbb{Z}^k$ are the coefficient vectors transforming the vectors $\hat{\boldsymbol{a}}_1, \ldots, \hat{\boldsymbol{a}}_k$ to the vectors $\hat{\boldsymbol{z}}_1, \ldots, \hat{\boldsymbol{z}}_{k-r}$, respectively, and the vector $\boldsymbol{m}'_j$ takes $\hat{\boldsymbol{a}}_1, \ldots, \hat{\boldsymbol{a}}_k$ to $\hat{\boldsymbol{b}}_j$ (for $1 \leq j \leq r$).

In [BP87] and [BK93] it is shown that the resulting basis has the following two properties. First, the top left $k-r$ columns contain a linearly independent set of relations $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_{k-r} \in \mathbb{Z}^k$ for the exact vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$. A relation vector $\boldsymbol{m}_j = (m_{1j}, \ldots, m_{kj})^t$ satisfies $\sum_{i=1}^{k} m_{ij}\boldsymbol{a}_i = 0$. In the exact matrix the vectors $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{k-r}$ would be zero, so the approximate vectors $\hat{\boldsymbol{z}}_i$ will be small. The second property of the resulting basis is that the bottom right of the matrix contains approximations $\hat{\boldsymbol{b}}_1, \ldots, \hat{\boldsymbol{b}}_r$ to a basis for the lattice generated by $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$.

In [BK93] the following theorem is shown.

Theorem C.2 (Theorem 4.1 in [BK93]). *Assume that q is chosen such that*

$$2^q > \frac{(\sqrt{nk}+2) \cdot \lambda \cdot 2^{\frac{k-3}{2}}}{\mu}$$

with $\lambda := (\frac{k\sqrt{n}}{2} + \sqrt{k}) \cdot \frac{\alpha^r}{d(L)}$, $\alpha = \max ||\boldsymbol{a}_i||$, and μ a lower bound for $\lambda_1(L)$.

The vectors $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_{k-r}$ are linearly independent relations for the $\boldsymbol{a}_i$ and satisfy

$$||\boldsymbol{m}_j|| \leq 2^{\frac{k-1}{2}} \cdot \lambda,$$

and the vectors $\boldsymbol{m}'_j$ satisfy

$$||\boldsymbol{m}'_j|| \leq 2^q 2^{(k+1)/2} \lambda_j(L_r) \ (1 \leq j \leq r).$$

Moreover, the vectors

$$\boldsymbol{b}_j = \sum_{i=1}^{k} m'_{i,j}\boldsymbol{a}_i \ (1 \leq j \leq r)$$

form a basis for L and satisfy

$$||\boldsymbol{b}_j|| \leq (\sqrt{kn}+2)2^{\frac{k-1}{2}} \lambda_j(L_r) \quad (1 \leq j \leq r). \tag{C.1}$$

This holds for every sublattice L_r which is spanned by a subset of r linearly independent vectors of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$.

C.1.2 New bounds on the coefficients of the basis transform

From the theorem above we can efficiently compute a basis from a generating set of a lattice. The first thing to note is that the resulting basis is not as small as we need it to be because its size is given in terms of the successive minima of a sublattice L_r of L rather than the lattice L itself. If we have enough precision left in the output vectors $\hat{\boldsymbol{b}}_1, \ldots, \hat{\boldsymbol{b}}_r$ we could run the procedure and Theorem C.2 a second time, namely on the vectors $(\boldsymbol{e}_1, \hat{\boldsymbol{b}}_1), \ldots, (\boldsymbol{e}_r, \hat{\boldsymbol{b}}_r)$ where $\boldsymbol{e}_i$ is the i-th unit vector in $\mathbb{Z}^r$ and the $\hat{\boldsymbol{b}}_i$ $(1 \leq i \leq r)$ are the approximations to the basis vectors $\boldsymbol{b}_i$ obtained from the first application of the algorithm. We would use $\alpha' = (\sqrt{kn}+2)2^{\frac{k-1}{2}}\alpha$, to obtain a small basis whose size is bounded in terms of the successive minima of L.

However, the 2^q factor for the coefficient vectors $\boldsymbol{m}'_j$ mapping the $\hat{\boldsymbol{a}}_i$'s to the basis vectors $\hat{\boldsymbol{b}}_1, \ldots, \hat{\boldsymbol{b}}_r$ destroys the precision guarantees that we started with. Furthermore, during our repeated squaring routine we need to start with a very high precision and have some relatively small loss on each iteration. For that reason we must prove a better bound on the coefficients resulting from

this algorithm. Once we upper bound the coefficient sizes, the following lemma lets us compute how much precision we lose at each step. After we have the lemma, we will bound the coefficient vectors $\boldsymbol{m}'_j$ $(1 \leq j \leq r)$ in Lemma C.4 below. Then we prove the main theorem, Theorem C.5, that bounds the loss of precision when computing a short basis for L that is described in terms of the successive minima for L.

To obtain a better bound on the coefficients in the LLL reduction we will consider the coefficient space $\mathbb{R}^k$ of these vectors and decompose it as

$$\mathbb{R}^k = N \oplus T,$$

where N is the subspace of $\mathbb{R}^k$ generated by the relation vectors $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_{k-r}$ and T is the orthogonal complement to N in $\mathbb{R}^k$. We denote the projections from $\mathbb{R}^k$ to N and T by π_N and π_T, respectively. We denote by α the length of the longest vector among the $\boldsymbol{a}_i$,

$$\alpha = \max\{||\boldsymbol{a}_1||, \ldots, ||\boldsymbol{a}_k||\}.$$

Lemma C.3. *Let L be a lattice of rank r generated by $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \in \mathbb{R}^n$. The map $f : Span(L) \to T$ given by*

$$f\left(\sum_{i=1}^{k} x_i \boldsymbol{a}_i\right) = \pi_T(x_1, \ldots, x_k)$$

is well-defined and an isomorphism between $Span(L)$ and T. For $\boldsymbol{b} \in Span(L)$ we have

$$||f(\boldsymbol{b})|| \leq \sqrt{r}\frac{\alpha^{r-1}}{d(L)}||\boldsymbol{b}||.$$

Proof. It is clear that the map is well-defined since given any two representations of $\boldsymbol{b}$ as $\boldsymbol{b} = \sum_{i=1}^{k} x_i \boldsymbol{a}_i$, $\boldsymbol{b} = \sum_{i=1}^{k} x'_i \boldsymbol{a}_i$, they have the property that $(x_1 - x'_1, \ldots, x_k - x'_k) \in N$. Hence $\pi_T(x_1, \ldots, x_k) = \pi_T(x'_1, \ldots, x'_k)$. The map is clearly injective and both vector spaces have dimension r, so it's an isomorphism.

Given $\boldsymbol{b} \in Span(L)$, we bound $||f(\boldsymbol{b})||$ by considering all vectors $(x_1, \ldots, x_k)$ in $\mathbb{R}^k$ whose image under π_T equals $f(\boldsymbol{b})$. (This will be an equivalence class of vectors in $\mathbb{R}^k/N$.) These vectors $(x_1, \ldots, x_k)$ in $\mathbb{R}^k$ correspond to the different ways in which it is possible to represent $\boldsymbol{b}$ as a linear combination of the $\boldsymbol{a}_i$. If we can bound the length of one such vector $(x_1, \ldots, x_k)$ with $\boldsymbol{b} = \sum x_i \boldsymbol{a}_i$, then we also have an upper bound for $||f(\boldsymbol{b})|| = \pi_T(x_1, \ldots, x_k)$, since the projection onto T can only decrease the length.

To bound the length of one such coefficient vector, w.l.o.g. we can arrange the $\boldsymbol{a}_i$ so that the first r of them are linearly independent. Let L_r denote the lattice they span. Since L_r is a full-rank sublattice of L, we can write

$$\boldsymbol{b} = y_1 \boldsymbol{a}_1 + \cdots + y_r \boldsymbol{a}_r$$

with $y_1, \ldots, y_r \in \mathbb{Q}$. We have to bound the length of $(y_1, \ldots, y_r, 0, \ldots, 0)$. Let D be a common denominator for $y_1, \ldots, y_r$. Then

$$D\boldsymbol{b} = (Dy_1)\boldsymbol{a}_1 + \cdots + (Dy_r)\boldsymbol{a}_r.$$

We can bound the length of the vector $(Dy_1, \ldots, Dy_r) \in \mathbb{Z}^r$ by

$$||(Dy_1, \ldots, Dy_r)||_\infty \leq ||D\boldsymbol{b}|| \cdot \frac{\alpha^{r-1}}{d(L_r)}.$$

This follows from Proposition 2.2(b) in [BK93], where we replace $\boldsymbol{v}_j$ with $D\boldsymbol{b}$ and hence x_{ij} with Dy_i and where we replace $||\boldsymbol{v}_j|| = \lambda_j(L_r)$ with $||D\boldsymbol{b}||$. The proof is word for word the same except in the last displayed equation on page 4 of [BK93] where we bound

$$|Dy_i| = \left| \frac{\det(\boldsymbol{a}'_1, \ldots, \boldsymbol{a}'_{i-1}, \varrho(\boldsymbol{b}), \boldsymbol{a}'_{i+1}, \ldots, \boldsymbol{a}'_r)}{\det(\boldsymbol{a}'_1, \ldots, \ldots, \boldsymbol{a}'_r)} \right|$$

by $||D\boldsymbol{b}|| \cdot \alpha^{r-1}/d(L_r)$, since each of the $\boldsymbol{a}'_i$ has length at most α, $D\boldsymbol{b}$ has length $||D\boldsymbol{b}||$, and the quantity in the denominator is the determinant of the lattice $\varrho(L_r)$ which equals the determinant of L_r. (Here L_r is the lattice generated by $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$.) Since

$$||(Dy_1, \ldots, Dy_r)||_\infty \leq ||D\boldsymbol{b}|| \cdot \frac{\alpha^{r-1}}{d(L_r)},$$

this gives us

$$||(y_1, \ldots, y_r)||_\infty \leq ||\boldsymbol{b}|| \cdot \frac{\alpha^{r-1}}{d(L_r)}.$$

Now, putting everything together we obtain

$$||f(\boldsymbol{b})|| \leq ||(y_1, \ldots, y_r, 0, \ldots, 0)|| \leq \sqrt{r} \cdot ||\boldsymbol{b}|| \cdot \frac{\alpha^{r-1}}{d(L_r)} \leq \sqrt{r} \cdot ||\boldsymbol{b}|| \cdot \frac{\alpha^{r-1}}{d(L)}.$$

□

Lemma C.4. *The square length of the coefficient vectors $\boldsymbol{m}'_j$ $(1 \leq j \leq r)$ that transform the generators $\boldsymbol{a}_i$ of L into a basis vector $\boldsymbol{b}_j$ of L as in Theorem C.2 above is bounded by $(\frac{\alpha^{r-1}}{d(L)} \cdot \sqrt{r}\Delta_{1,j})^2 + \Delta_2^2$. Here $\Delta_{1,j}$ is the right-hand-side of Equation (10) in [BK93],*

$$\Delta_{1,j} = (\sqrt{kn} + 2) \cdot 2^{\frac{k-1}{2}} \cdot \lambda_j(L_r).$$

Here L_r is any sublattice of L which is spanned by a subset of r linearly independent vectors of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$. The quantity Δ_2 is

$$\Delta_2 = \sqrt{k-r} \cdot 2^{\frac{k-1}{2}} \left(\frac{k\sqrt{n}}{2} + \sqrt{k} \right) \cdot \frac{\alpha^r}{d(L)}.$$

Proof. Given a coefficient vector $\boldsymbol{m}'_j$ as in the lemma (for $1 \leq j \leq r$), we will show that $\pi_T(\boldsymbol{m}'_j) \leq \alpha^{r-1}/d(L)\sqrt{r}\Delta_{1,j}$ and that $\pi_N(\boldsymbol{m}'_j) \leq \Delta_2$. Since N and T are orthogonal this will imply that

$$||\boldsymbol{m}'_j||^2 = ||\pi_T(\boldsymbol{m}'_j)||^2 + ||\pi_N(\boldsymbol{m}'_j)||^2 \leq \left(\frac{\alpha^{r-1}}{d(L)} \Delta_{1,j} \right)^2 + \Delta_2^2.$$

To bound $\pi_T(\boldsymbol{m}'_j)$ we observe that by the definition of f we have

$$f(\boldsymbol{b}_j) = f\left(\sum_{i=1}^{k} m'_{i,j} \boldsymbol{a}_i \right) = \pi_T(m'_{1,j}, \ldots, m'_{k,j}) = \pi_T(\boldsymbol{m}'_j).$$

By Lemma C.3, this implies that

$$||\pi_T(\boldsymbol{m}'_j)|| \leq ||\boldsymbol{b}_j|| \cdot \sqrt{r} \cdot \frac{\alpha^{r-1}}{d(L)}.$$

By Theorem C.2, $||\boldsymbol{b}_j|| \leq \Delta_{1,j}$, proving the bound on $\pi_T(\boldsymbol{m}'_j)$.

To bound $\pi_N(\boldsymbol{m}'_j)$, we use the fact that the matrix with M in the top and the coefficients of the $\hat{b}_i$ is a size reduced basis for the lattice $\tilde{L}$. The Gram-Schmidt coefficients are at most 1/2, and hence

$$\begin{aligned}||\pi_N(\boldsymbol{m}'_j)||^2 &\leq (||\boldsymbol{m}_1^{\perp}||/2)^2 + (||\boldsymbol{m}_2^{\perp}||/2)^2 + \cdots + (||\boldsymbol{m}_{k-r}^{\perp}||/2)^2 \\ &\leq ||\boldsymbol{m}_1||^2 + \cdots + ||\boldsymbol{m}_{k-r}||^2 \\ &\leq (k-r)\max\{||\boldsymbol{m}_1||^2, \ldots, ||\boldsymbol{m}_{k-r}||^2\}.\end{aligned}$$

By Theorem C.2, the linearly independent relations $m_1, \ldots, m_{k-r}$ satisfy

$$\max\{||\boldsymbol{m}_1||, \ldots, ||\boldsymbol{m}_{k-r}||\} \leq 2^{\frac{k-1}{2}}\left(\frac{k\sqrt{n}}{2} + \sqrt{k}\right) \cdot \frac{\alpha^r}{d(L)}.$$

Therefore $||\pi_N(\boldsymbol{m}'_j)|| \leq \sqrt{k-r} \cdot 2^{\frac{k-1}{2}}\left(\frac{k\sqrt{n}}{2} + \sqrt{k}\right) \cdot \frac{\alpha^r}{d(L)}$. □

Now we can prove the main theorem.

C.1.3 Computing a short basis

Theorem C.5. *Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \in \mathbb{R}^n$ generate a lattice L of rank r, let μ be a lower bound on the shortest vector, and let q be such that $2^q \geq (k2^{(k+1)/2} \cdot \max||a_i||)^r/(\mu \det(L)^2)$. Given approximations of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \in \mathbb{R}^n$ with q bits of precision, a basis $\hat{\boldsymbol{c}}_1, \ldots, \hat{\boldsymbol{c}}_r$ for L can be computed in time polynomial in q, where the exact vectors $\boldsymbol{c}_j$ satisfy*

$$||\boldsymbol{c}_j|| \leq (\sqrt{kn} + 2)2^{\frac{k-1}{2}} \cdot \lambda_j(L).$$

The absolute error on each output vector $\hat{\boldsymbol{c}}_i$ is bounded by $rk\gamma_1\gamma_3\sqrt{n}/2^{q+1}$, where γ_1 and γ_3 are defined below.

Proof. In the first step we apply algorithm from [BP87] which runs the LLL-algorithm on the matrix $(\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_k)$ where $\tilde{\boldsymbol{a}}_i = (\boldsymbol{e}_i, \hat{\boldsymbol{a}}_i)$. By Lemma C.4 and by Theorem C.2, we obtain a basis

$$\hat{\boldsymbol{b}}_j = \sum_{i=1}^{k} m'_{i,j}\hat{\boldsymbol{a}}_i,$$

approximating exact vectors $\boldsymbol{b}_j$ whose lengths are bounded by

$$||\boldsymbol{b}_j|| \leq (\sqrt{kn} + 2)2^{\frac{k-1}{2}} \cdot \alpha = \gamma_2,$$

and where the coefficient vectors $\boldsymbol{m}'_1, \ldots, \boldsymbol{m}'_r$ satisfy

$$||\boldsymbol{m}'_j|| \leq \sqrt{\left(\frac{\alpha^{r-1}}{d(L)}\sqrt{r}\Delta_{1,n}\right)^2 + \Delta_2^2} = \gamma_1$$

with $\Delta_{1,n}, \Delta_2$ as in Lemma C.4.

Then apply LLL again, this time to the matrix whose columns are the vectors $(\boldsymbol{e}_i, \hat{\boldsymbol{b}}_i)_{1\leq i\leq r} \in \mathbb{Z}^{r+n}$, we obtain basis vectors $\hat{\boldsymbol{c}}_1, \ldots, \hat{\boldsymbol{c}}_r$ approximating exact vectors $\boldsymbol{c}_j$ whose lengths are bounded by Theorem C.2 as

$$||\boldsymbol{c}_j|| \leq (\sqrt{kn} + 2)2^{\frac{k-1}{2}} \cdot \lambda_j(L).$$

After the first run of LLL the output vectors have absolute error $||\sum_{i=1}^{k} m'_{ij}(\boldsymbol{a}_i - \hat{\boldsymbol{a}}_i/2^q)|| \leq k\gamma_1\sqrt{n}/2^{q+1}$. We now verify that there is enough precision to run it a second time.

For the second run we will apply Theorem C.2 with:

- $\alpha' = (\sqrt{kn}+2)2^{(k-1)/2}\alpha$
- $\lambda' = (r\sqrt{n}/2+\sqrt{r})(\alpha')^r/\det(L)$ (here $k=r$), and
- $2^q/k\gamma_1 \geq (\sqrt{nr}+2)2^{(r-3)/2}\lambda'/\mu$.

The last inequality is satisfied by the choice of q and the following calculation. We have

$$\Delta_{1,j} = (\sqrt{kn}+2)2^{(k-1)/2}\lambda_j(L_r) \leq k2^{k/2}\max||a_i||,$$

and

$$\Delta_2 = \sqrt{k-r}2^{(k-1)/2}(k\sqrt{n}/2+\sqrt{k})\frac{\alpha^r}{\det L} \leq k^2 2^{k/2}\frac{(\max||a_i||)^r}{\det(L)},$$

so $||b_i - \hat{b}_i/2^q||$ is bounded by $k\gamma_1\sqrt{n}/2^{q+1}$ with

$$\begin{aligned}\gamma_1^2 = \left(\frac{\alpha^{r-1}}{\det(L)}\sqrt{r}\Delta_{1,n}\right)^2 + \Delta_2^2 &\leq (\frac{(\max||a_i||)^r}{\det(L)}\sqrt{r}k2^{k/2})^2 + (k^2 2^{k/2}\frac{(\max||a_i||)^r}{\det(L)})^2 \\ &\leq (k^2 2^{(k+1)/2}\frac{(\max||a_i||)^r}{\det(L)})^2.\end{aligned}$$

By Lemma C.4, the length of the coefficient vectors transforming $\hat{\boldsymbol{b}}_i$ into $\hat{\boldsymbol{c}}_i$ is bounded by

$$\gamma_3 = \sqrt{\left(\frac{(\alpha')^r}{d(L)}\cdot\sqrt{r}\cdot||\boldsymbol{b}_j||\right)^2 + \left(\sqrt{k-r}\sqrt{k}\cdot\frac{(\alpha')^r}{d(L)}\right)^2}$$

with $\alpha' = \gamma_2 \leq k\cdot 2^{\frac{k+1}{2}}\cdot\alpha$.

We have

$$\begin{aligned}\gamma_3^2 &\leq (k(k2^{k/2}\max||a_i||)^4/\det(L)^2 + (k(k2^{k/2}\max||a_i||)/\det(L))^2 \\ &\leq 4k(k2^{k/2}\max||a_i||)^r/\det(L)^2.\end{aligned}$$

The absolute error for the output is bounded by

$$||\sum_{i=1}^r m'_{ij}(\boldsymbol{b}_i - \hat{\boldsymbol{b}}_i)|| \leq r\gamma_3 k\gamma_1\sqrt{n}/2^{q+1}.$$

This is at most $4k^{5r}2^{2kr}(\max||a_i||)^{3r}/\det(L)^4 rk\sqrt{n}/2^{q+1}$.

□

C.2 Computing the dual basis

Suppose we have an LLL-reduced basis $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_n$ in the columns of B for a lattice L^*. We want to compute the basis B^{-T} for $L = \mathrm{Log}\,\mathcal{O}^*$. Then:

Lemma C.6. *An LLL-reduced basis B has condition number $\kappa(B) = ||B||\cdot||B^{-1}|| \leq n^4 2^{3n}\frac{\lambda_n(L^*)}{\lambda_1(L^*)}$.*

Proof. Since B is LLL-reduced we have $||B|| \leq n\max||\boldsymbol{b}_i|| \leq n2^{(n-1)/2}\lambda_n(L^*)$.

For the inverse, write $B = QR = Q\cdot\mathrm{diag}(||\boldsymbol{b}_1^\perp||,\ldots,||\boldsymbol{b}_n^\perp||)\cdot M$, where Q contains the normalized columns of the Gram-Schmidt decomposition $\boldsymbol{b}_i^\perp/||\boldsymbol{b}_i^\perp||$, and M is upper unit triangular (i.e. 1's on the diagonal) and $M_{i,j} = \mu_{j,i} = \langle\boldsymbol{b}_j,\boldsymbol{b}_i^\perp\rangle/\langle\boldsymbol{b}_i^\perp,\boldsymbol{b}_i^\perp\rangle$. Then $B^{-1} = M^{-1}\cdot\mathrm{diag}(1/||\boldsymbol{b}_1^\perp||,\ldots,1/||\boldsymbol{b}_n^\perp||)\cdot$

Q^{-1}. For the diagonal we have $1/||\boldsymbol{b}_i^\perp||^2 \leq 2^{i-1}/||\boldsymbol{b}_1^\perp||^2 \leq 2^{i-1}/\lambda_1(L^*)^2$. The first inequality is because the basis is LLL-reduced and the second is because $\boldsymbol{b}_1^\perp \in L^*$. So the maximum diagonal entry is bounded by $2^{n/2}/\lambda_1(L^*)$.

The matrix M^{-1} has entries that are sums of products of the $\mu_{i,j}$'s. In a given entry $(M^{-1})_{i,j}$ the summands are at most 1 and the number of summands is at most 2^{j-i}, so the entry is at most 2^n. Therefore $||B^{-1}|| \leq n^2 \max_{i,j} |(B^{-1})_{i,j}| \leq n^2 2^{2n} \lambda_n(L)$. □

In the case of the unit group we have L as the unit lattice and we are sampling from its dual L^*. By Lemma B.3, $\lambda_1(L) \geq 1/2$. Using the inequalities $\lambda_1(L)\lambda_n(L^*) \leq n$ and $1 \leq \lambda_1(L^*)\lambda_n(L)$, we have $\lambda_n(L^*)/\lambda_1(L^*) \leq n\lambda_n(L)/\lambda_1(L)$ which can be bounded by the unit lattice parameters.

Using this bound on the condition number we can compute the accuracy of the dual basis B^{-T}. We compute the inverse of B^T by solving $B^T \boldsymbol{b}_i' = e_i$ for each standard basis vector to get a primal basis vector $\boldsymbol{b}_i'$.

By [GVL96, Lemma 2.71] we need to have ϵ accuracy for B where $\epsilon\kappa(B) = r < 1$. Then by [GVL96, Theorem 2.7.2] the output vector $\boldsymbol{b}_i'$ will have relative precision $\frac{2\epsilon}{1-r}\kappa(B) = \frac{2r}{1-r}$.

D The Lipschitz bound on the lattice Gaussians

In this section we bound the Lipschitz constant for the HSP function.

D.1 Sums over lattice points with Gaussian weights

As previously mentioned, the HSP oracle in our algorithm uses a representation of a lattice L by the Gaussian superposition $|f(L)\rangle$, see Eq. (5.2). To estimate the oracle parameters, we will employ some properties of the following function:

$$\rho(L) = \sum_{x \in L} e^{-\pi \|x\|^2}. \tag{D.1}$$

More generally, we define

$$\rho(S, u) = \sum_{x \in S} e^{2\pi i \langle u, x \rangle} e^{-\pi \langle x, x \rangle}, \tag{D.2}$$

where S is a set of points in $\mathbb{R}^n$ and $u \in \mathbb{R}^n$. These functions were used by Banaszczyk [Ban93] to answer some purely geometric questions about lattices. A few results in this subsection are borrowed from Banaszczyk's paper.

The function ρ satisfies a duality relation:

$$d(L)\, \rho(L) = \rho(L^*). \tag{D.3}$$

It is obtained by applying the Poisson summation formula (F.13) to the function $f(x) = e^{-\pi \langle x, x \rangle}$. Note that the Fourier transform $\widehat{f}$ coincides with f. A slightly more general result is as follows:

$$d(L)\, \rho(L, u) = \rho(L^* + u). \tag{D.4}$$

In this case, we apply the Poisson summation formula to $f(x) = e^{2\pi i \langle u, x \rangle} e^{-\pi \langle x, x \rangle}$ and $\widehat{f}(y) = e^{-\pi \langle y+u, y+u \rangle}$.

Lemma D.1. *Let W be a linear transformation of $\mathbb{R}^n$. If $\|W\| \leqslant 1$, then*

$$1 \leqslant \frac{\rho(WL)}{\rho(L)} \leqslant (\det W)^{-1}.$$

Proof. The inequality $\rho(WL) \geqslant \rho(L)$ follows from the fact that $e^{-\|Wx\|^2} \geqslant e^{-\|x\|^2}$ for all x. On the other hand, if $\Lambda = WL$, then $L^* = W^T\Lambda^*$ and hence $\rho(L^*) \geqslant \rho(\Lambda^*)$. Let us rewrite this inequality using the duality relation (D.3):

$$d(L)\rho(L) \geqslant d(\Lambda)\rho(\Lambda) = d(WL)\rho(WL).$$

Thus, $\rho(WL)/\rho(L) \leqslant d(L)/d(WL) = (\det W)^{-1}$. □

Lemma D.2. *For any vector $v \in \mathbb{R}^n$,*

$$e^{-\pi\|v\|^2} \leqslant \frac{\rho(L+v)}{\rho(L)} \leqslant 1.$$

Proof. The lower bound is equivalent to Eq. (7) in Ref. [Ban93]. It is derived as follows:

$$\begin{aligned}\rho(L+v) &= \sum_{x\in L} e^{-\pi\langle x+v, x+v\rangle} = e^{-\pi\langle v,v\rangle}\sum_{x\in L} e^{-2\pi\langle v,x\rangle}e^{-\pi\langle x,x\rangle} = e^{-\pi\langle v,v\rangle}\sum_{x\in L}\cosh\bigl(2\pi\langle v,x\rangle\bigr)\,e^{-\pi\langle x,x\rangle}\\ &\geqslant e^{-\pi\langle v,v\rangle}\rho(L).\end{aligned}$$

The last equality was obtained by replacing each term in the sum, $a_x = e^{-2\pi\langle v,x\rangle}e^{-\pi\langle x,x\rangle}$ with $\frac{1}{2}(a_x + a_{-x})$. To prove the upper bound, we first apply the duality relation (D.4) and then proceed as above:

$$\begin{aligned}\rho(L+v) &= d(L)^{-1}\rho(L^*, v) = d(L)^{-1}\sum_{y\in L^*} e^{2\pi i\langle y,v\rangle}e^{-\pi\langle y,y\rangle} = d(L)^{-1}\sum_{y\in L^*}\cos\bigl(2\pi\langle y,v\rangle\bigr)\,e^{-\pi\langle y,y\rangle}\\ &\leqslant d(L)^{-1}\rho(L^*) = \rho(L).\end{aligned}$$

□

Let us now consider the moments of a Gaussian distribution on a lattice: each point $x \in L$ is taken with probability $w(x) = \rho(L)^{-1}\,e^{-\pi\|x\|^2}$. In particular, we are interested in the second and the fourth moments:

$$M^{(2)}_{jk}(L) = \rho(L)^{-1}\sum_{x\in L} x_j x_k\, e^{-\pi\|x\|^2} = -(2\pi)^{-2}\rho(L)^{-1}\,\frac{\partial^2\rho(L,u)}{\partial u_j \partial u_k}\bigg|_{u=0}, \tag{D.5}$$

$$M^{(4)}_{jklm}(L) = \rho(L)^{-1}\sum_{x\in L} x_j x_k x_l x_m\, e^{-\pi\|x\|^2} = (2\pi)^{-4}\rho(L)^{-1}\,\frac{\partial^4\rho(L,u)}{\partial u_j \partial u_k \partial u_l \partial u_m}\bigg|_{u=0}. \tag{D.6}$$

Expressing the partial derivatives with the help of Eq. (D.4), we find duality relations for the moments:

$$M^{(2)}_{jk}(L) = (2\pi)^{-1}\delta_{jk} - M^{(2)}_{jk}(L^*), \tag{D.7}$$

$$\begin{aligned}M^{(4)}_{jklm}(L) = {}& (2\pi)^{-2}\bigl(\delta_{jk}\delta_{lm} + \delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl}\bigr)\\ &- (2\pi)^{-1}\bigl(\delta_{jk}M^{(2)}_{lm}(L^*) + \delta_{jl}M^{(2)}_{km}(L^*) + \delta_{jm}M^{(2)}_{kl}(L^*)\\ &\qquad + \delta_{kl}M^{(2)}_{jm}(L^*) + \delta_{km}M^{(2)}_{jl}(L^*) + \delta_{lm}M^{(2)}_{jk}(L^*)\bigr)\\ &+ M^{(4)}_{jklm}(L^*).\end{aligned} \tag{D.8}$$

Since the matrix $M^{(2)}(L^*)$ is positive-semidefinite, Eq. (D.7) implies that

$$M^{(2)}(L) \leqslant \frac{1}{2\pi} \tag{D.9}$$

(which means that $\frac{1}{2\pi} I - M^{(2)}(L)$ is positive-semidefinite). This result is equivalent to Lemma 1.3 from Ref. [Ban93].

It is more difficult to estimate $M^{(4)}(L)$. We will do that for sparse and dense lattices as defined below. Let us first consider the contribution from the tail of the Gaussian function. The following lemma generalizes the first part of Lemma 1.5 from Ref. [Ban93], and the proof is almost the same.

Lemma D.3. *If* $2\pi r^2 \geqslant n + p$, *then*

$$\rho(L)^{-1} \sum_{\substack{x\in L \\ \|x\|\geqslant r}} \|x\|^p e^{-\pi\|x\|^2} \leqslant \bigl(\alpha\sqrt{n}\bigr)^p \left(\sqrt{2\pi e}\,\alpha\, e^{-\pi\alpha^2}\right)^n, \qquad \textit{where} \quad \alpha = r/\sqrt{n}.$$

Proof. The idea is to write the expression under the sum as $\left(\|x\|^p e^{-(1-t)\pi\|x\|^2}\right) e^{-t\pi\|x\|^2}$ and replace the first factor by a constant. Let

$$0 < t \leqslant 1 - \frac{p}{2\pi r^2}.$$

Under these assumptions, the function $x \mapsto x^p e^{-(1-t)\pi x^2}$ is monotonically decreasing for $x \geqslant r$. Thus,

$$\sum_{\substack{x\in L \\ \|x\|\geqslant r}} \|x\|^p e^{-\pi\|x\|^2} \leqslant r^p e^{-(1-t)\pi r^2} \sum_{\substack{x\in L \\ \|x\|\geqslant r}} e^{-t\pi\|x\|^2} \leqslant r^p e^{-(1-t)\pi r^2} \rho\bigl(t^{1/2}L\bigr).$$

Lemma D.1 says that $\rho(t^{1/2}L) \leqslant t^{-n/2}\rho(L)$, hence

$$\rho(L)^{-1} \sum_{\substack{x\in L \\ \|x\|\geqslant r}} \|x\|^p e^{-\pi\|x\|^2} \leqslant r^p e^{-(1-t)\pi r^2} t^{-n/2}.$$

Setting $t = \frac{n}{2\pi r^2}$, we obtain the required inequality. □

For each given p, Lemma D.3 can be written as an asymptotic bound:

$$\rho(L)^{-1} \sum_{\substack{x\in L \\ \|x\|\geqslant r}} \|x\|^p e^{-\pi\|x\|^2} \leqslant e^{-\Omega(\alpha^2 n)} \qquad \text{if} \quad \alpha = r/\sqrt{n} \gg 1.$$

More exactly, there are some constant α_0 and c (depending on p) such that for all $\alpha \geqslant \alpha_0$ the left-hand side is less or equal to $e^{-c\alpha^2 n}$. Setting $r = \lambda_1(L)$ and considering the $p = 0$ and $p > 0$ cases separately, we get two bounds that are applicable when L is sparse, i.e. $\alpha = \lambda_1(L)/\sqrt{n} \gg 1$:

$$1 - \rho(L)^{-1} \leqslant e^{-\Omega(\alpha^2 n)}, \qquad \rho(L)^{-1} \sum_{x\in L} \|x\|^p e^{-\pi\|x\|^2} \leqslant e^{-\Omega(\alpha^2 n)} \quad \text{for } p > 0. \tag{D.10}$$

If p is even, the left-hand side of the second inequality is a certain combination of the moment elements, e.g. $\sum_{j,l} M^{(4)}_{jjll}(L)$ for $p = 4$. It also serves as an upper bound for each individual element. Furthermore, any linear combination of $M^{(p)}_{jk\ldots}$ is likewise bounded because polynomial factors (such as n^2) are absorbed by $e^{-\Omega(\alpha^2 n)}$.

Finally, let us consider the opposite limit:

$$\text{Dense lattices:} \qquad \beta = \sqrt{n}\,\lambda_n(L) \ll 1. \tag{D.11}$$

If L is dense, then L^* is sparse because $\lambda_n(L)\,\lambda_1(L^*) \geqslant 1$. Thus, the duality relations for $\rho(L)$ and the moments, together with the previous bounds applied to L^*, yield these results:

$$\rho(L) = d(L)^{-1}\big(1 + e^{-\Omega(\beta^{-2}n)}\big), \tag{D.12}$$

$$M^{(2)}_{jk}(L) = (2\pi)^{-1}\delta_{jk} + e^{-\Omega(\beta^{-2}n)}, \tag{D.13}$$

$$M^{(4)}_{jklm}(L) = (2\pi)^{-2}\big(\delta_{jk}\delta_{lm} + \delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl}\big) + e^{-\Omega(\beta^{-2}n)}. \tag{D.14}$$

D.2 The encoding of lattices by Gaussian states

In this subsection, we show that the quantum encoding $L \mapsto |f(L)\rangle$ given by Eq. (5.2) satisfies Definition 5.1 for certain values of a, r, and ε. We restrict the function f to lattices with the unit cell volume $d(L) \leqslant d$ and the shortest vector length $\lambda_1(L) \geqslant \lambda$. In addition, we assume that $\nu \leqslant \frac{\lambda}{2\sqrt{n}}$ so that the vectors $\big|\mathrm{str}_{n,\nu}(x)\big\rangle$ representing the lattice points are mutually orthogonal.

To find a Lipschitz constant of f, we consider two nearby (infinitely close) lattices L and $\widetilde{L}$. Let L be represented by a basis matrix B, i.e. $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$, where the column vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ form a basis of L. The second lattice is given by a basis matrix $\widetilde{B} = WB$ that is infinitely close to B. By definition, the distance between L and $\tilde{L}$ is

$$\mathrm{dist}(L, \tilde{L}) = \|\delta W\|_2 = \sqrt{\mathrm{Tr}(\delta W^T \delta W)}, \qquad \text{where} \quad \delta W = (\delta B)B^{-1}, \quad \delta B = \tilde{B} - B. \tag{D.15}$$

It is convenient to write the function f as follows:

$$|f(L)\rangle = \sum_{x\in\Lambda} \frac{e^{-\pi\|x\|^2/2}}{\sqrt{\rho(\Lambda)}} \Big|\mathrm{str}_{n,\nu}\Big(\frac{s}{\sqrt{2}}x\Big)\Big\rangle, \qquad \text{where} \quad \Lambda = \frac{\sqrt{2}}{s}L. \tag{D.16}$$

Note that the lattices Λ and $\tilde{\Lambda}$ are related by the same matrix δW. The infinitesimal variation of the function value, $|\delta f(L)\rangle$ is a sum of two terms:

$$|\delta f_1(L)\rangle = \sum_{x\in\Lambda} \frac{e^{-\pi\|x\|^2/2}}{\sqrt{\rho(\Lambda)}} \Big|\delta\,\mathrm{str}_{n,\nu}\Big(\frac{s}{\sqrt{2}}x\Big)\Big\rangle, \tag{D.17}$$

$$|\delta f_2(L)\rangle = \sum_{x\in\Lambda} \delta\left(\frac{e^{-\pi\|x\|^2/2}}{\sqrt{\rho(\Lambda)}}\right) \Big|\mathrm{str}_{n,\nu}\Big(\frac{s}{\sqrt{2}}x\Big)\Big\rangle. \tag{D.18}$$

Thus,

$$\big\||\delta f(L)\rangle\big\| \leqslant \big\||\delta f_1(L)\rangle\big\| + \big\||\delta f_2(L)\rangle\big\|. \tag{D.19}$$

We first estimate $\big\||\delta f_1(L)\rangle\big\|$. Recall that the straddle encoding $\mathrm{str}_{n,\nu}$ has Lipschitz constant $\frac{\pi\sqrt{n}}{2\nu}$, hence

$$\left\|\Big|\delta\,\mathrm{str}_{n,\nu}\Big(\frac{s}{\sqrt{2}}x\Big)\Big\rangle\right\| \leqslant \frac{s}{\sqrt{2}}\,\frac{\pi\sqrt{n}}{2\nu}\,\|\delta x\|.$$

On the other hand,

$$\delta x = (\delta W)x, \qquad \|\delta x\|^2 = \sum_{j,k} x_j\big(\delta W^T \delta W\big)_{jk} x_k.$$

It follows that

$$\big\||\delta f_1(L)\rangle\big\|^2 \leqslant \frac{\pi^2 n s^2}{8\nu^2}\rho(\Lambda)^{-1}\sum_{x\in\Lambda} e^{-\pi\|x\|^2}\|\delta x\|^2 = \frac{\pi^2 n s^2}{8\nu^2}\sum_{j,k} M^{(2)}_{jk}(\Lambda)\,\big(\delta W^T\delta W\big)_{jk} \leqslant \frac{\pi n s^2}{16\nu^2}\|\delta W\|_2^2,$$

where we have used the bound (D.9) for the second moment of the lattice Gaussian. Thus,

$$\big\||\delta f_1(L)\rangle\big\| \leqslant \frac{\sqrt{\pi n}\,s}{4\nu}\,\|\delta W\|_2. \tag{D.20}$$

Let us now estimate $\big\||\delta f_2(L)\rangle\big\|$. Replacing the constant orthonormal vectors $\big|\mathrm{str}_{n,\nu}\big(\frac{s}{\sqrt{2}}x\big)\big\rangle$ with $|x\rangle$, we find that

$$\big\||\delta f_2(L)\rangle\big\|^2 = \left\langle \delta\frac{\psi}{\sqrt{\langle\psi|\psi\rangle}}\middle|\delta\frac{\psi}{\sqrt{\langle\psi|\psi\rangle}}\right\rangle,$$

$$\text{where}\quad |\psi\rangle = \sum_{x\in\Lambda} e^{-\pi\|x\|^2/2}|x\rangle, \qquad |\delta\psi\rangle = -\pi\sum_{x\in\Lambda}\big(x^T(\delta W)x\big)e^{-\pi\|x\|^2/2}|x\rangle.$$

We proceed with a somewhat tedious calculation:

$$\left|\delta\frac{\psi}{\sqrt{\langle\psi|\psi\rangle}}\right\rangle = \frac{1}{\sqrt{\langle\psi|\psi\rangle}}|\delta\psi\rangle - \frac{\langle\psi|\delta\psi\rangle+\langle\delta\psi|\psi\rangle}{2\langle\psi|\psi\rangle^{3/2}}|\psi\rangle,$$

$$\begin{aligned}
\left\langle \delta\frac{\psi}{\sqrt{\langle\psi|\psi\rangle}}\middle|\delta\frac{\psi}{\sqrt{\langle\psi|\psi\rangle}}\right\rangle &= \frac{\langle\delta\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} - \left(\frac{\langle\psi|\delta\psi\rangle+\langle\delta\psi|\psi\rangle}{2\langle\psi|\psi\rangle}\right)^2 \\
&= \pi^2\rho(\Lambda)^{-1}\sum_{x\in\Lambda}\big(x^T(\delta W)x\big)^2 e^{-\pi\|x\|^2} - \left(\pi\rho(\Lambda)^{-1}\sum_{x\in\Lambda}\big(x^T(\delta W)x\big)e^{-\pi\|x\|^2}\right)^2 \\
&= \pi^2\sum_{j,k,l,m} M^{(4)}_{jklm}(\delta W)_{jk}(\delta W)_{lm} - \left(\pi\sum_{j,k} M^{(2)}_{jk}(\delta W)_{jk}\right)^2 \\
&= \pi^2\sum_{j,k,l,m}\Big(M^{(4)}_{jklm} - M^{(2)}_{jk}M^{(2)}_{lm}\Big)(\delta W)_{jk}(\delta W)_{lm} \\
&= \frac{1}{4}\sum_{j,k,l,m}\Big(\delta_{jl}\delta_{km}+\delta_{jm}\delta_{kl}+e^{-\Omega(\beta^{-2}n)}\Big)(\delta W)_{jk}(\delta W)_{lm} \leqslant \left(\frac{1}{2}+e^{-\Omega(\beta^{-2}n)}\right)\|\delta W\|_2^2\,.
\end{aligned}$$

When passing to the last line, we applied the asymptotic bounds (D.13) and (D.14) to the lattice Λ. These bounds are valid if $\beta = \sqrt{n}\,\lambda_n(\Lambda) \gg 1$, i.e. if $\sqrt{n}\,\lambda_n(\Lambda)$ is smaller than a certain constant. This condition is equivalent to $s \gg \sqrt{n}\,\lambda_n(L)$. A upper bound for $\lambda_n(L)$ follows from equation (A.7) and the assumptions about L:

$$\lambda_n(L) \leqslant \frac{n^{n/2}d(L)}{\lambda_1(L)^{n-1}} \leqslant \sqrt{n}\,\big(\sqrt{n}/\lambda\big)^{n-1}\,d.$$

We may also replace $e^{-\Omega(\beta^{-2}n)}$ by $1/2$ so that $\big\||\delta f_2(L)\rangle\big\| \leqslant \|\delta W\|_2$. To summarize, we have obtained the following result.

Theorem D.4. *Let $\nu \leqslant \frac{\lambda}{2\sqrt{n}}$ and $s \geqslant c\,n\,(\sqrt{n}/\lambda)^{n-1}\,d$ for a certain constant c, and let us restrict the encoding $L \mapsto |f(L)\rangle$ to lattices with $d(L) \leqslant d$ and $\lambda_1(L) \geqslant \lambda$. On such lattices, f has a Lipschitz constant*

$$a = \frac{\sqrt{\pi n}\,s}{4\nu} + 1.$$

E The HSP property of the lattice Gaussian

E.1 Lattice Gaussians

In this section we define the lattice Gaussian superpositions $|f(L)\rangle$ to represent the lattices $L = e^{\boldsymbol{t}}\mathcal{O}$. We will pick parameters s and ν, an accuracy ϵ for the given basis vectors, and show that the superpositions satisfy the desired hidden subgroup condition. That is, if $\boldsymbol{t}_2 - \boldsymbol{t}_1$ is close to a unit, then the superpositions $|e^{\boldsymbol{t_1}}\mathcal{O}\rangle$ and $|e^{\boldsymbol{t_2}}\mathcal{O}\rangle$ have a large inner product, and that the inner product falls off exponentially as $\boldsymbol{t}_2 - \boldsymbol{t}_1$ gets farther from a unit.

Since we do not have a unique classical representation (e.g. a unique basis) for the lattice, which would be required to define the usual type of hidden subgroup problem, we define a quantum superposition over the points in the lattice to represent it. To make the superposition have norm one over the infinite set of lattice points, the superposition will have an outside Gaussian with parameter s which will specify the set of lattice points we will consider. The vectors in this range (up to length $s\sqrt{n}$) will have some rounding error ε. The straddle encoding with parameter ν used to represent each point to handle the rounding. The straddle encoding at two different vectors representing the same point, i.e. that are separated by at most 2ϵ, will have large inner product. Finally, we must ensure that different points of $e^{\boldsymbol{t_1}}\mathcal{O}$ and $e^{\boldsymbol{t_2}}\mathcal{O}$ that are not the same are far enough apart so that the inner Gaussians for two different points are very far apart and have small inner product. The lattice Gaussian used to represent the lattice $L = e^{\boldsymbol{t}}\mathcal{O}$ is

$$|f(L)\rangle = \gamma(L) \sum_{v \in L} g_s(v) \left|\mathrm{str}_{n,\nu}(v)\right\rangle. \tag{E.1}$$

Only a finite piece out to radius $\sqrt{n}$ times the parameter is used, as specified in Section E.2. We start with a lemma showing how points of two different lattices are distributed.

If the width s of the Gaussian is large, the sum in the expression for γ can be approximated by an integral, i.e.

$$\gamma \approx \frac{1}{d(L)} \int_{\mathbb{R}^n} e^{-2\pi\|x\|^2/s^2}\, dx = \frac{\left(s/\sqrt{2}\right)^n}{d(L)}.$$

In fact, s must be large enough for the state $|f(L)\rangle$ to be efficiently constructible and for different lattices to be distinguishable by the corresponding states. Likewise, ν must be sufficiently small. The exact conditions depend on the lattice parameters $d(L)$ and $\lambda_1(L)$; they will be discussed in Section E.1.

While the straddle encoding achieves $\varepsilon = 0$ (i.e. the signature vectors of distant points are strictly orthogonal), the orthogonality properties of lattice Gaussians are poor. Indeed, consider two lattices L_1, L_2 with $d(L_1) = d(L_2)$ that intersect over an index 2 sublattice. Then $\langle f(L_1)|f(L_2)\rangle$ is approximately $1/2$. This situation can be ameliorated using the following corollary of Lemma 5.2, part (a).

Lemma E.1 (Orthogonality amplification). *Let $f : X \to \mathcal{H}$ be an (a, r, ε) quantum encoding. Then $f^{\otimes n} : X \to \mathcal{H}^{\otimes n}$ is an (na, r, ε^n) encoding.*

Proof of Lemma 5.4 Suppose $\boldsymbol{e^{t_1}a} = \boldsymbol{e^{t_2}b}$ $(\boldsymbol{a}, \boldsymbol{b} \in \underline{\mathcal{O}})$ is a point in the intersection of $\boldsymbol{e^{t_1}}\mathcal{O}$ and $\boldsymbol{e^{t_2}}\mathcal{O}$ and has length at most R. Since the length of $\boldsymbol{e^{t_1}a}$ is at most R, each coordinate of $\boldsymbol{e^{t_1}a}$ is bounded by R, and hence the norm of $\boldsymbol{e^{t_1}a}$, which is the product over all coordinates, is bounded by R^n. Since $\boldsymbol{e^{t_1}}$ has norm 1, the norm of $\boldsymbol{a}$ is then bounded by R^n as well. To see how close points in $\boldsymbol{e^{t_1}}\mathcal{O}$ and $\boldsymbol{e^{t_2}}\mathcal{O}$ can be (without being equal) we consider the minimum distance of points in the lattice $\boldsymbol{e^{t_1}}\mathcal{O} + \boldsymbol{e^{t_2}}\mathcal{O}$.

To compute this lattice we first compute $\boldsymbol{e^{t_1-t_2}}\mathcal{O}+\mathcal{O}$: Since $\boldsymbol{e^{t_1}a} = \boldsymbol{e^{t_2}b}$ we have $\boldsymbol{b/a} = \boldsymbol{e^{t_1-t_2}}$. So $\boldsymbol{e^{t_1-t_2}}\mathcal{O} + \mathcal{O} = \boldsymbol{(b/a)}\mathcal{O} + \mathcal{O}$. Let $N(\boldsymbol{a})$ denote the norm of $\boldsymbol{a}$. (I.e., $N(\boldsymbol{a}) = N(a_1) = \prod a_i$, where $\boldsymbol{a} = (a_1, \ldots, a_n)$.)

Claim: $(\boldsymbol{b/a})\mathcal{O} + \mathcal{O} \subseteq \frac{1}{N(\boldsymbol{a})}\mathcal{O}$.

Proof of claim: Clearly $\mathcal{O} \subseteq \frac{1}{N(\boldsymbol{a})}\mathcal{O}$. Let $\boldsymbol{a} = (a_1, \ldots, a_n) \in \underline{\mathcal{O}}$. Since each a_i satisfies the same minimal polynomial as a_1 they are all algebraic integers, and hence so is the product $\prod_{i=2}^{n} a_i$. Since $a_2 \cdot a_3 \cdot \cdots \cdot a_n = N(a)/a_1$, the product is also in K and hence it is in $\mathcal{O}$. Then $b_1 \cdot a_2 \cdot \cdots \cdot a_n$ is in $\mathcal{O}$ as well. Thus

$$(b_1/a_1)\mathcal{O} = \frac{b_1 \cdot a_2 \cdot \cdots \cdot a_n}{N(a_1)}\mathcal{O} \subseteq \frac{1}{N(a_1)}\mathcal{O}.$$

Hence $(\boldsymbol{b/a})\underline{\mathcal{O}} + \underline{\mathcal{O}} \subseteq \frac{1}{N(\boldsymbol{a})}\underline{\mathcal{O}}$. This proves the claim and shows that $\boldsymbol{e^{t_1-t_2}}\mathcal{O} + \mathcal{O} \subseteq \frac{1}{N(\boldsymbol{a})}\mathcal{O}$.

The shortest vector in $\mathcal{O}$ is $(1, \ldots, 1)$ which has length $\sqrt{n}$. Since we showed that $N(\boldsymbol{a}) \leq R^n$, this implies that the shortest vector in $\frac{1}{N(\boldsymbol{a})}\mathcal{O}$ has length at least $\sqrt{n}/R^n$.

Now we consider $\boldsymbol{e^{t_1}}\mathcal{O} + \boldsymbol{e^{t_2}}\mathcal{O}$. By the above argument

$$\boldsymbol{e^{t_1}}\mathcal{O} + \boldsymbol{e^{t_2}}\mathcal{O} = \boldsymbol{e^{t_2}}(\boldsymbol{e^{t_1-t_2}}\mathcal{O} + \mathcal{O}) \subseteq \frac{1}{N(\boldsymbol{a})}\boldsymbol{e^{t_2}}\mathcal{O}.$$

Since the shortest vector in $\boldsymbol{e^{t_2}}\mathcal{O}$ has length at least $\sqrt{n}$, the shortest vector in $\boldsymbol{e^{t_1}}\mathcal{O} + \boldsymbol{e^{t_2}}\mathcal{O}$ has length at least $\sqrt{n}/R^n$. Hence points in $\boldsymbol{e^{t_1}}\mathcal{O}$ and $\boldsymbol{e^{t_2}}\mathcal{O}$, which are not equal, are at least $\sqrt{n}/R^n$ apart. □

E.2 Inner product computations

One useful fact states that summation of Gaussian function over a lattice is actually concentrated on a "small'' region of the lattice. More precisely, let $L \subseteq \mathbb{R}^n$ be a lattice, then $g_s(L|_{\sqrt{n}s}) \geq (1 - 2^{-2n})g_s(L)$. This implies that the states we use approximate the case where we use the whole infinite lattice.

Now we study the inner product for two lattices $L, L' \subseteq \mathbb{Q}^n$. These are rational representations of some real valued lattices. If two real-valued lattices intersect in some sublattice, then the rational representations of the lattices will have a sublattice points close, but maybe not equal.

Lemma E.2. *Let $u, v \in \mathbb{R}^n$ with $\|v\| \leq \ell$ and $\|u\| \leq d$. Then $g_s(v+u) \in [e^{-\pi\frac{2\ell d+\ell^2}{s^2}}, e^{\pi\frac{2\ell d}{s^2}}] \cdot g_s(u)$. In particular, if $s > 2\pi n \cdot \max\{\sqrt{\ell d}, \ell\}$, $g_s(v+u) \in (1 \pm \frac{1}{n^2}) \cdot g_s(u)$.*

Proof. $g_s(v+u) = e^{-\pi\frac{\|u\|^2+2\langle u,v\rangle+\|v\|^2}{s^2}} = g_s(u) \cdot e^{-\pi\frac{2\langle u,v\rangle+\|v\|^2}{s^2}}$. By the Cauchy-Schwarz inequality $|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$, we have

$$g_s(u)e^{-\pi\frac{2\|u\|\cdot\|v\|+\|v\|^2}{s^2}} \leq g_s(v+u) \leq g_s(u)e^{-\pi\frac{\|v\|^2-2\|u\|\cdot\|v\|}{s^2}}.$$

Since $\|v\| \leq \ell$ and $\|u\| \leq d$, we have $e^{-\pi\frac{2\|u\|\cdot\|v\|+\|v\|^2}{s^2}} \geq e^{-\pi\frac{2\ell d+\ell^2}{s^2}}$, and $e^{-\pi\frac{\|v\|^2-2\|u\|\cdot\|v\|}{s^2}} \leq e^{\pi\frac{2\ell d}{s^2}}$. Finally, if we pick $s \geq 2\pi n \cdot \max\{\sqrt{\ell d}, \ell\}$, then $e^{-\pi\frac{2\ell d+\ell^2}{s^2}} \geq 1 - \frac{1}{n^2}$, and $e^{\pi\frac{2\ell d}{s^2}} \leq 1 + \frac{1}{n^2}$. □

Define the truncated sets $\tilde{L} = L|_{\sqrt{n}s}$, $\tilde{L'} = L'|_{\sqrt{n}s}$, $\tilde{I} = I|_{\sqrt{n}s}$, $\tilde{G} = G|_{\sqrt{n}t}$, where $S|_r := S \cap \mathbf{B}(0, r)$ means the elements of S inside the ball $\mathbf{B}(0, r) \subseteq \mathbb{R}^n$ around the origin of radius r.

Corollary E.3 (Sets). *Let $v \in \mathbb{R}^n, S \subseteq \mathbb{R}^n$ with $\|v\| \leq \ell$ and $\|x\| \leq d\ \forall x \in S$, then $g_s(v+S) \in (1 \pm \frac{1}{n^2})g_s(S)$ whenever $s \geq 2\pi n \cdot \max\{\sqrt{\ell d}, \ell\}$. If I is a lattice, then $g_s(v+\tilde{I}) \geq (1 - \frac{1}{n^2})g_s(\tilde{I})$ whenever $s > 4\pi^2\ell n^{5/2}$.*

Proof. By Lemma E.2, $g_s(v+S) = \sum_{x\in S} g_s(v+x) = \sum_{x\in S}(1\pm\frac{1}{n^2})g_s(x) = (1\pm\frac{1}{n^2})g_s(S)$. □

Corollary E.4. *Let I and I' be lattices satisfying that there is a 1-1 correspondence $h: I \to I'$ s.t. $\|x - h(x)\| \le \ell, \forall x \in \tilde{I}$. Then $g_s(I) \in (1\pm\frac{2}{n^2})g_s(I')$ whenever $s \ge (4\pi^2 n^{5/2}+1)\ell$.*

Proof.

$$g_s(I) \ge g_s(\tilde{I}) = \sum_{x\in\tilde{I}} g_s(h(x)+(x-h(x))) \overset{(1)}{\ge} (1-\frac{1}{n^2})g_s(\tilde{I}')$$
$$\ge (1-\frac{1}{n^2})(1-2^{-2n})g_s(I') \ge (1-\frac{2}{n^2})g_s(I').$$

On the other hand,

$$(1-2^{-2n})g_s(I) \le g_s(\tilde{I}) = \sum_{x\in\tilde{I}} g_s(h(x)+(x-h(x))) \overset{(2)}{\le} (1+\frac{1}{n^2})g_s(\tilde{I}') \le (1+\frac{1}{n^2})g_s(I')$$

and hence $g_s(I) \le (1+\frac{2}{n^2})g_s(I')$.

In (1) & (2) we applied Lemma E.2 with $u := h(x)$, $v := x - h(x)$. Here $\|h(x)\| \le \sqrt{n}s + \ell$ and $\|x - h(x)\| \le \ell$, so we need to pick $s \ge 2\pi n \max\{\sqrt{\ell(\sqrt{n}s+\ell)}, \ell\}$. It is easy to verify that setting $s \ge (4\pi^2 n^{5/2}+1)\ell$ suffices. □

We then derive a lemma characterizing Gaussian summations on a lattice and its proper sublattices.

Lemma E.5. *Let I be a proper sublattice of L, then $\frac{g_s(I)}{g_s(L)} \le 2/3$, for any $n \ge 2$ and $s \ge 4\pi^2 n^3 \lambda_n(L)$.*

Proof. First we argue that we can pick a $v \in L$ but not contained in I, with length $\|v\| \le \sqrt{n}\lambda_n(L)$. The existence follows from the fact that for any lattice L, there is a basis $\mathbf{B} = \{b_i : b_i \in \mathbb{Z}^m\}_{i=1}^n$ of L such that $\|b_i\| \le \sqrt{i}\lambda_i(L), i = 1,\dots,n$, where $\lambda_i(L)$ is the ith successive minimum of L. If there is no such v, that would imply I contains all vectors in L of length $\le \sqrt{n}\lambda_n(L)$, and hence contains a basis of L, contradicting the assumption that I is a proper sublattice of L. Then observe that L at least contains $I \cup (v+I)$, and therefore $g_s(L) \ge g_s(I) + g_s(v+I) \ge g_s(I) + g_s(v+\tilde{I}) \ge g_s(I) + (1-\frac{1}{n^2})g_s(\tilde{I})$ by Corollary E.3, if $s \ge 4\pi^2 \cdot n^{5/2} \cdot \sqrt{n}\lambda_n(L) = 4\pi^2 n^3 \lambda_n(L)$. Since $g_s(\tilde{I}) \ge (1-2^{-2n})g_s(I)$, we have that $g_s(L) \ge (1 + (1-\frac{1}{n^2})(1-2^{-2n}))g_s(I) \ge \frac{3}{2}g_s(I)$ when $n \ge 2$. □

Lemma E.6. *Given lattices L and L', sublattices $I \subseteq L$ and $I' \subseteq L'$. Assume there is a 1-1 correspondence $h: I \to I'$ s.t., for any $(x, x') \in \tilde{L} \times \tilde{L}'$,*

1. *if $(x, x') \in C := \{(u, v) \in I \times I' : v = h(u)\}$, then $\|x - x'\| \le \varepsilon$,*
2. *otherwise $\|x - x'\| \ge 2\nu\sqrt{n}$.*

Then $\langle f(L)|f(L')\rangle \le \left(1+\frac{2}{n^2}\right)^{1/2} \cdot \frac{g_s(I)}{\sqrt{g_s(L)g_s(L')}}$ whenever $s \ge 4\pi^2 n^{5/2}\varepsilon$.

Proof. In the following, x will always vary over $\tilde{L}$, while x' will vary over points in $\tilde{L}'$.

$$\begin{aligned}
\langle f(L)|f(L')\rangle &= \alpha \sum_{v\in L, v'\in L'} \sqrt{g_s(v)g_s(v')}\langle \mathrm{str}_{n,\nu}(v)|\,\mathrm{str}_{n,\nu}(v')\rangle \\
&= \alpha \sum_{\|x-x'\|\le\varepsilon} \sqrt{g_s(x)g_s(x')}\langle \mathrm{str}_{n,\nu}(x)|\,\mathrm{str}_{n,\nu}(x')\rangle \\
&+ \alpha \sum_{\|x-x'\|\ge 2\nu\sqrt{n}} \sqrt{g_s(x)g_s(x')}\langle \mathrm{str}_{n,\nu}(x)|\,\mathrm{str}_{n,\nu}(x')\rangle \\
&+ \alpha \sum_{v\in L-\tilde{L}, v'\in L'-\tilde{L}'} \sqrt{g_s(v)g_s(v')}\langle \mathrm{str}_{n,\nu}(v)|\,\mathrm{str}_{n,\nu}(v')\rangle \\
&=: \alpha(A+B+C),
\end{aligned}$$

where $\alpha := \left(\sqrt{g_s(L)g_s(L')}\right)^{-1}$. Then

$$\begin{aligned}
A &= \sum_{\|x-x'\|\le\varepsilon} \sqrt{g_s(x)g_s(x')}\langle \mathrm{str}_{n,\nu}(x)|\,\mathrm{str}_{n,\nu}(x')\rangle \\
&= \sum_{x\in\tilde{I}, x'=h(x)} \sqrt{g_s(x)g_s(x')}\langle \mathrm{str}_{n,\nu}(x)|\,\mathrm{str}_{n,\nu}(x')\rangle \\
&\le \sum_{x\in\tilde{I}, x'=h(x)} \sqrt{g_s(x)g_s(x')} \\
&\le \sqrt{(1+1/n^2)}g_s(\tilde{I}).
\end{aligned}$$

The last inequality uses Lemma E.2 with $u = x$, $v = x - x'$, $s \ge 4\pi^2 n^{5/2}\varepsilon$.

The second term $B = 0$ since the straddle encoding has inner product zero when two values are more than $2\nu\sqrt{n}$ apart.

The third term $\alpha|C|$ is at most $3\cdot\left\|\sqrt{g_s(L)}^{-1}\sum_{v\in L\setminus\tilde{L}}\sqrt{g_s(v)}|\,\mathrm{str}_{n,\nu}(v)\rangle\right\| \le 3\sqrt{\sum_{v\in L\setminus\tilde{L}} g_s(v)/g_s(L)} \le 3\cdot 2^{-n}$, which is negligibly small. Here we need to assume that $\min\{\lambda_1(L),\lambda_1(L')\} \ge 2\nu\sqrt{n}$. □

We now use Lemma E.6 and consider the case $I \subsetneq L$ and $I' \subsetneq L'$.

Lemma E.7. *Assume the conditions in Lemma E.6. If $I \subsetneq L$ and $I' \subsetneq L'$, then for any $n \ge 5$, $\langle f(L)|f(L')\rangle \le 3/4$ if in addition $s \ge 4\pi n^3 \max\{\lambda_n(L),\lambda_n(L')\}$.*

Proof. If $I \subsetneq L$ and $I' \subsetneq L'$, we apply Lemma E.6 and obtain

$$\begin{aligned}
\langle f(L)|f(L')\rangle &\le \left(1+\frac{2}{n^2}\right)\cdot\sqrt{\frac{g_s(I)}{g_s(L)}}\cdot\sqrt{\frac{g_s(I)}{g_s(L')}} \\
&\le \left(1+\frac{2}{n^2}\right)\sqrt{\frac{2}{3}}\cdot\sqrt{(1+\frac{2}{n^2})\cdot\frac{2}{3}} \le \frac{3}{4},
\end{aligned}$$

where we applied Lemma E.5 to get $g_s(L) \ge \frac{3}{2}g_s(I)$ and $g_s(L') \ge \frac{3}{2}g_s(I')$, and applied Corollary E.4 to get $g_s(I) \le (1+\frac{2}{n^2})g_s(I')$. These inequalities hold for any $n \ge 5$ and $s \ge \max\{4\pi^2 n^3 \cdot \max\{\lambda_n(L),\lambda_n(L')\}, (4\pi^2 n^{5/2}+1)\varepsilon\} = 4\pi^2 n^3 \cdot \max\{\lambda_n(L),\lambda_n(L')\}$. □

E.3 Main Theorem

Next we can choose the parameters to show that a hidden subgroup property holds for the states.

Proof of Theorem 5.5 Theorem D.4 bounds the Lipschitz constant.

Fix an ideal $e^{\boldsymbol{t_1}}\mathcal{O}$ and we will consider its inner product with ideals $e^{\boldsymbol{t_2}}\mathcal{O}$ where $t_2 = s + t$ and $e^s\mathcal{O} \cap e^{t_1}\mathcal{O} \cap B(0, s\sqrt{n}) \neq \{0\}$, i.e., $e^s\mathcal{O}$ has a nonzero exact point of intersection with $e^{\boldsymbol{t_1}}\mathcal{O}$ of length at most $s\sqrt{n}$. By Lemma 5.4 there is a minimum distance of $\sqrt{n}/(s\sqrt{n})^n$ between any equal points in $e^{\boldsymbol{t_1}}\mathcal{O}$ and $e^s\mathcal{O}$. Consider vectors t in a box defined by the condition $\ln(1-(s\sqrt{n})^{n-1}2\nu\sqrt{n}) \leq t_i \leq \ln(1-(s\sqrt{n})^{n-1}2\nu\sqrt{n})$ for all i. Let $I_t = e^{\boldsymbol{t_1}}\mathcal{O} \cap e^{s+t}\mathcal{O}$. Then $I_0 \cap B(0, s\sqrt{n}) \neq \{0\}$ by assumption. By using $\varepsilon = 0$ and assuming that $t_1 - t_2 \notin \mathcal{O}^*$, Lemma E.7 implies that the inner product is at most $3/4$. Unequal points are at least $\sqrt{n}/(s\sqrt{n})^n$ apart which is at least $2\nu\sqrt{n}$ by the choice of ν.

As t varies, a point $v \in e^s\mathcal{O}$ moves by at most $\|e^t v - v\| \leq \|e^t - I\| \cdot \|v\| \leq n(s\sqrt{n})^{n-1}2\nu\sqrt{n} \cdot s\sqrt{n} \leq \sqrt{n}/(2(s\sqrt{n})^n)$ by the choice of ν. First, this implies that a point $v \in e^s\mathcal{O}$ cannot be within $2\nu\sqrt{n}$ of a new point $w \in e^{\boldsymbol{t_1}}\mathcal{O}$, i.e., $\|e^t v - w\| \geq 2\nu\sqrt{n}$, unless $v = w$. So the pairing of points between $e^{\boldsymbol{t_1}}\mathcal{O}$ and $e^s\mathcal{O}$ stays the same for $e^{\boldsymbol{t_1}}\mathcal{O}$ and $e^{s+t}\mathcal{O}$. Second, we can use Lemma E.7 by setting $\varepsilon = 2n^2(\sqrt{n})^{n-1}\nu$ and $I = e^{\boldsymbol{t_1}}\mathcal{O} \cap e^s\mathcal{O}$.

Now do not assume $t_1 - t_2 \notin \mathcal{O}^*$ (so $e^{\boldsymbol{t_1}}\mathcal{O} = e^s\mathcal{O}$ is possible). We argue that at the boundary of the box the inner product only has overlap at the origin of the lattices. Assume that for some i, $\ln(1 - (s\sqrt{n})^{n-1}2\nu\sqrt{n}) \geq t_i$ or $t_i \geq \ln(1 + (s\sqrt{n})^{n-1}2\nu\sqrt{n})$. For this choice of $\boldsymbol{t} = (t_1, \ldots, t_n)$, $\max_i |e^{t_i} - 1| \geq (s\sqrt{n})^{n-1}2\nu\sqrt{n}$. Since $v \in \mathcal{O}$ satisfies $N(v) \geq 1$, each component v_i satisfies $|v_i| \geq 1/(s\sqrt{n})^{n-1}$. Then

$$\begin{aligned}\|e^{\boldsymbol{t}} v - v\| &= \sqrt{(e^{t_1}v_1 - v_1)^2 + \cdots + (e^{t_n}v_n - v_n)^2} \geq \max\{|e^{t_i} - 1| \cdot |v_i|\} \\ &\geq 1/(s\sqrt{n})^{n-1}((s\sqrt{n})^{n-1}2\nu\sqrt{n}) = 2\nu\sqrt{n}.\end{aligned}$$

Therefore on the boundary of the box, the straddle encoding of any nonzero point v has no overlap with $e^t v$. The same holds true for any value of t_2 that is not inside one of the boxes surrounding the points of exact overlap in the domain.

□

E.4 Computing the lattice Gaussian

We show that, assuming $s \geq 2^{2n}n^{n/2+1}\lambda^{-n+1}$, there is an efficient quantum algorithm to generate $|f(L)\rangle = \gamma(L) \sum_{v \in L} g_s(v) \left|\mathrm{str}_{n,\nu}(v)\right\rangle$. For notational ease, we omit normalization below.

We start off creating a Gaussian superposition over L: $\sum_{v \in L} g_s(v)|v\rangle$. This is adapted from [KW08]. Let B be an LLL-reduced basis of L and let $A := B^T B$ be its Gram matrix. Then A can be decomposed into $A = (M^{-1})^T D M^{-1}$ where $D = \mathrm{diag}\{d_i\}$ is a diagonal matrix and M is an upper triangular *shearing* matrix with diagonal entries being one. M can be further decomposed as $\Pi_{i=1}^m M_i$ where each M_i has diagonal entries being one and exactly one nonzero off-diagonal entry. The generating procedure proceeds as follows:

1. Create a state that approximates $\sum_{x \in \mathbb{Z}^n} e^{-\pi \frac{x^T D x}{s^2}} |x\rangle$. This is done by creating n instances of 1-dimensional Gaussian superpositions, using standard techniques as in [GR02], with parameter $s/\sqrt{d_i}$ in each respective dimension.

2. By change of variable $x = M^{-1}y$, we get

$$\sum_{M^{-1}y \in \mathbb{Z}^n} e^{-\pi \frac{y^T (M^{-1})^T D M^{-1} y}{s^2}} |M^{-1}y\rangle = \sum_{M^{-1}y \in \mathbb{Z}^n} e^{-\pi \frac{\|By\|^2}{s^2}} |M^{-1}y\rangle.$$

3. Apply transformation $U : |z\rangle \mapsto |z'\rangle, \forall z \in \mathbb{Z}^n$, where z' is Mz followed by coordinate-wise floor. This gives $\sum_{z \in \mathbb{Z}^n} g_s(B(z + \delta_z))|z\rangle$, with $\delta_z := Mz - z'$.

4. Multiply by basis B, we get $|\phi\rangle := \sum_{z \in \mathbb{Z}^n} \rho_s(B(z + \delta_z))|Bz\rangle = \sum_{v \in L} g_s(v + \delta_v)|v\rangle$.

Because of the properties of M mentioned above, it is not hard to verify that multiplication by M and taking the floor are efficient reversible classical operations, and therefore can be implemented efficiently on a quantum computer.

We now claim that with proper parameter s, $|\phi\rangle$ is exponentially close to the desired Gaussian superposition $|\psi\rangle = \sum_{v \in L} g_s(v)|Bz\rangle$. Roughly speaking, $|\phi\rangle$ differs from $|\psi\rangle$ only by the "noisy" shift $\delta_v := B\delta_z$ in each amplitude. We show below that $\|B\delta z\| \le C := 2^n n^{n/2+1} \lambda^{-n+1}$. Lemma E.2 then tells us, picking large enough s, e.g., $s \ge C \cdot 2^n$, $g_s(v + \delta_v) \approx g_s(v)$ with exponentially small error. As a result, the two states $|\phi\rangle$ and $|\psi\rangle$ will be exponentially close.

We are left to derive an upper bound C on $\|B\delta z\|$. Note that each coordinate of δz lies in $[0, 1)$. That implies that $\|B\delta_z\|$ is at most the diameter of the fundamental parallelepiped, which is upper bounded by $n2^n\lambda_n(L)$ since B is LLL-reduced. Minkowski's theorem tells us that $\Pi_i \lambda_i \le \sqrt{n^n} \det(B)$. By our HSP definition, $\lambda_i \ge \lambda_1 \ge \lambda, i = 1, \ldots, n-1$ and $\det(B) \le d$, we get $\lambda_n \le \sqrt{n^n} d/\lambda^{n-1}$. Therefore $\|B\delta_x\| \le C := 2^n n^{n/2+1} \lambda^{-n+1}$.

Finally we apply straddle encoding for each v on $\sum_{v \in L} g_s(v)|v\rangle$, and we obtain $|f(L)\rangle = \gamma(L) \sum_{v \in L} g_s(v) \left|\mathrm{str}_{n,\nu}(v)\right\rangle$.

F Algorithm for the hidden subgroup problem

In this section we give an algorithm for solving the HSP over $\mathbb{R}^m$.

As explained earlier, the group of units $\mathcal{O}^*$ is contained in $G = \mathbb{R}^{s+t-1} \times \mathbb{Z}_2^s \times (\mathbb{R}/\mathbb{Z})^t$, and we have an explicit upper bound for $\mathrm{vol}(G/\mathcal{O}^*)$. Our HSP algorithm is applicable to Abelian groups G that factor into a continuous and a discrete part, where the first (e.g. $\mathbb{R}^{s+t-1} \times (\mathbb{R}/\mathbb{Z})^t$) is a quotient of $\mathbb{R}^k$ over some lattice, and the second (e.g. $\mathbb{Z}_2^s$) is finitely generated. We call such groups *elementary*. Thus, G is a quotient of $\tilde{G} = \mathbb{R}^k \times \mathbb{Z}^l$, and the problem of finding a hidden subgroup $L \subseteq G$ is equivalent to finding $\tilde{L}$ (the inverse image of L) in $\tilde{G}$. The HSP oracle on $\tilde{G}$ is obtained by composing the quotient map $\tilde{G} \to G$ with the G oracle. Note that this reduction preserves the covolume because $\tilde{G}/\tilde{L} \cong G/L$.

Let us now consider the HSP problem for $G = \mathbb{R}^k \times \mathbb{Z}^l$. We use the following metric on G: if x and y belong to the same connected component, then $\mathrm{dist}_G(x, y) = \|x - y\|$ (the Euclidean distance), otherwise $\mathrm{dist}_G(x, y) = \infty$. Let $L \subseteq G$ be a closed subgroup, and $\gamma : G \to G/L$ the quotient map. An arbitrary invariant metric on G induces the *quotient metric* on G/L as follows:

$$\mathrm{dist}_{G/L}(x, y) = \min\bigl\{\mathrm{dist}_G(\tilde{x}, \tilde{y}) : \ x = \gamma(\tilde{x}), \ y = \gamma(\tilde{y})\bigr\}. \tag{F.1}$$

Definition F.1 (*Hidden subgroup problem for* $G = \mathbb{R}^k \times \mathbb{Z}^l$). The hidden subgroup $L \subseteq G$ is assumed to satisfy the following conditions, where the parameters λ and d are known in advance.

1. L is a discrete subgroup in G, and the quotient G/L is compact. (More explicitly, it means that L is a full-rank lattice in $R^m \supseteq G$.)

2. The shortest vector in L has length at least λ. (By the definition of the distance function, the shortest vector belongs to $\mathbb{R}^k$.)

3. $\mathrm{vol}(G/L) \leqslant d$.

An (a, r, ε) *oracle function* associated with L is a function f from G to some Hilbert space that factors into the quotient map $\gamma : G \to G/L$ and an (a, r, ε) quantum encoding of G/L. That is, $|f(x)\rangle$ has unit norm and depends only on $\gamma(x) = x \bmod L$, the function f is a-Lipschitz, and $|\langle f(x)|f(y)\rangle| \leqslant \varepsilon$ whenever $\mathrm{dist}_{G/L}(x, y) \geqslant r$.

The last condition is stated more explicitly (for $G = \mathbb{R}^m$) in Definition 1.1. The oracle available to an HSP algorithm takes $|x\rangle$ to $|x\rangle \otimes |f(x)\rangle$. The algorithm for computing L will require certain assumptions about a, r, ε in terms of λ and d.

Next we prove Lipschitz constant in the reduction from abelian groups to $\mathbb{R}^m$.

Proof of Theorem 6.1 To find the Lipschitz constant $\tilde{a}$, we need to bound the derivative of g, which is piecewise-continuous. The norm of the derivative is given by the expression

$$\left\| \frac{\partial |g(x)\rangle}{\partial \boldsymbol{x}} \right\|^2 + \sum_{j=1}^{l} \left\| \frac{\partial |g(x)\rangle}{\partial x_j} \right\|^2,$$

where the first term is less or equal to a^2. For the second term, we have this bound:

$$\begin{aligned}
\left\| \frac{\partial |g(x)\rangle}{\partial x_j} \right\|^2 &\leqslant \sum_{z_1,\ldots,z_l \in \{0,1\}} \left\| \frac{\partial}{\partial x_j} \left(\bigotimes_{j=1}^{l} |\psi(x_j, z_j)\rangle \right) \right\|^2 = \sum_{z \in \{0,1\}} \left\| \frac{\partial |\psi(x_j, z)\rangle}{\partial x_j} \right\|^2 \\
&\leqslant \sum_{z \in \{0,1\}} \left(\left| \frac{\pi}{2\lambda} \sin\left(\tfrac{\pi}{2} t\right) \right| \cdot 1 + \left| \cos\left(\tfrac{\pi}{2} t\right) \right| \cdot \frac{\pi}{2\nu\lambda} \right)^2 \quad \text{(where } t \text{ depends on } z\text{)} \\
&= \left(\frac{\pi}{2\nu\lambda} \right)^2 \sum_{z \in \{0,1\}} \left(\nu^2 \sin^2\left(\tfrac{\pi}{2} t\right) + 2\nu \left| \sin\left(\tfrac{\pi}{2} t\right) \cos\left(\tfrac{\pi}{2} t\right) \right| + \cos^2\left(\tfrac{\pi}{2} t\right) \right) \leqslant \left(\frac{\pi}{2\nu\lambda} (1 + \nu) \right)^2.
\end{aligned}$$

Now, suppose that the difference between two vectors, $(\boldsymbol{x}, x_1, \ldots, x_l)$ and $(\boldsymbol{y}, y_1, \ldots, y_l)$ is at least $\tilde{r}$ distance away from any point in $\tilde{L}$. We need to prove that

$$\left| \langle g(\boldsymbol{x}, x_1, \ldots, x_l) | g(\boldsymbol{y}, y_1, \ldots, y_l) \rangle \right| \leqslant \varepsilon. \tag{F.2}$$

Let us assume that the inner product in question does not vanish. Then for each $j = 1, \ldots, l$ there are some $z, w \in \{0, 1\}$ such that the vectors $|\psi(x_j, z)\rangle$ and $|\psi(y_j, w)\rangle$ (see Eq. (6.2)) are not orthogonal. Thus,

$$\left| (x_j - y_j)/\lambda - v_j \right| < 2\nu, \qquad \text{where} \quad v_j = s(x_j, z) - s(y_j, w). \tag{F.3}$$

Since $\nu \leqslant 1/4$, the integer v_j does not depend on z and w. Therefore, each term in the expression for $|g(\boldsymbol{x}, x_1, \ldots, x_l)\rangle$ has nonzero inner product with at most one term in $|g(\boldsymbol{y}, y_1, \ldots, y_l)\rangle$, and vice versa. Using the Cauchy-Schwartz inequality, one can easily prove (F.2), provided

$$\left| \langle f(\boldsymbol{x},\, s(x_1, z_1), \ldots, s(x_l, z_l) | f(\boldsymbol{y},\, s(y_1, w_1), \ldots, s(y_l, w_l) \rangle \right| \leqslant \varepsilon$$

for all such pairs of terms. In the opposite case, the hypothesis of the theorem implies that

$$\mathrm{dist}_{G/L}(\boldsymbol{x} - \boldsymbol{y}, v_1, \ldots, v_l) < r.$$

In other words, there is some $\boldsymbol{v} \in \mathbb{R}^k$ such that $(\boldsymbol{v}, v_1, \ldots, v_l) \in L$ and $\|\boldsymbol{x}-\boldsymbol{y}-\boldsymbol{v}\| < r$. This condition together with (F.3) implies that the vectors $(\boldsymbol{x}-\boldsymbol{y}, x_1-y_1, \ldots, x_l-y_l)$ and $(\boldsymbol{v}, \lambda v_1, \ldots, \lambda v_l) \in \tilde{L}$ are less than $\tilde{r} = \sqrt{r^2 + l(2\nu\lambda)^2}$ distance apart from each other, which contradicts our original assumption. □

F.1 Fourier transform and related tools

The Fourier transform of a function on an elementary Abelian group G is a function on another elementary group,

$$\widehat{G} = \mathrm{Hom}(G, \mathbb{R}/\mathbb{Z}), \tag{F.4}$$

whose elements are continuous homomorphisms from G to $\mathbb{R}/\mathbb{Z}$. This $\widehat{G}$ is called the *dual group*. For example, $\widehat{\mathbb{R}} = \mathbb{R}$ and $\widehat{\mathbb{Z}} = \mathbb{R}/\mathbb{Z}$. The Pontryagin duality theorem states that the dual to $\widehat{G}$ is canonically isomorphic to G.

In this paper, we mainly deal with full-rank lattices in $\mathbb{R}^m$ and the corresponding quotients. The following construction is slightly more general. For any closed subgroup $H \subseteq \mathbb{R}^m$, the *reciprocal subgroup* is defined as follows:

$$H^* = \{x \in \mathbb{R}^m : (\forall y \in H)\, \langle x, y\rangle \in \mathbb{Z}\}. \tag{F.5}$$

For example, if H is a full-rank lattice, then H^* is the reciprocal lattice; if $H \cong \mathbb{R}^k$, then $H^* = H^\perp$ is the orthogonal complement of H. (More generally, $H = L \oplus M$, where L is a lattice and M is a subspace that is orthogonal to the linear span $\overline{L}$ of L. In this case, $H^* = L^* \oplus \big(\overline{L} \oplus M\big)^\perp$.)

Proposition F.2. *If $H \subseteq \mathbb{R}^m$ is a closed subgroup, then $\widehat{H} \cong \mathbb{R}^m/H^*$.*

The Fourier transform F_G is usually defined as a unitary map between two Hilbert spaces, $F_G : \mathrm{L}^2(G) \to \mathrm{L}^2(\hat{G})$, where $\mathrm{L}^2(G)$ is the space of square-integrable functions on G. For functions on a lattice L in a Euclidean space, we use this definition of the inner product:

$$\boxed{\langle f|g\rangle_L = d(L) \sum_{v\in L} \overline{f_v} g_v} \tag{F.6}$$

(The expression $d(L)\sum_{v\in L}\cdots$ mimics the integral over $\mathbb{R}^m$.) Thus, the decomposition of the vector $|f\rangle$ in the standard orthonormal basis also contains a normalization factor:

$$|f\rangle = \sqrt{d(L)} \sum_{v\in L} f_v |v\rangle. \tag{F.7}$$

The Fourier transform on three basic types of infinite groups are given by these equations:

$$(F_{\mathbb{R}^m} f)(y) = \int_{\mathbb{R}^m} e^{2\pi i\langle x,y\rangle} f(x)\, dx \qquad \text{for} \quad y \in \mathbb{R}^m; \tag{F.8}$$

$$(F_{\mathbb{R}^m/M} f)_v = \int_{\mathbb{R}^m/M} e^{2\pi i\langle v,u\rangle} f(u)\, du \qquad \text{for} \quad v \in M^*; \tag{F.9}$$

$$(F_L f)(u) = d(L) \sum_{v\in L} e^{2\pi i\langle v,u\rangle} f_v \qquad \text{for} \quad u \in \mathbb{R}^m/L^*. \tag{F.10}$$

The inverse Fourier transform differs by a minus sign in the exponent. Thus,

$$F_G^{-1} = F_{\widehat{G}} P = P F_{\widehat{G}}, \qquad \text{where} \quad (Pf)(x) := f(-x).$$

The Fourier transform, in particular $F_{\mathbb{R}^m}$, is defined on some other spaces besides L^2. For example, one can use the space of infinitely smooth functions with sufficiently rapid decay at infinity (the Schwartz space) or the corresponding space of generalized functions (tempered distributions). We will use generalized functions as a calculational tool, in particular, employing these identities:

$$\boxed{\int_{\mathbb{R}^m} e^{2\pi i\langle x,y\rangle}\,dx = \delta(y)} \tag{F.11}$$

$$\boxed{d(L)\sum_{v\in L} e^{2\pi i\langle v,y\rangle} = \sum_{u\in L^*}\delta(y-u)} \tag{F.12}$$

As an example of such a calculation, consider the Poisson summation formula:

$$d(L)\sum_{v\in L} f(v) = \sum_{u\in L^*}\widehat{f}(u), \qquad \text{where} \quad \widehat{f} = F_{\mathbb{R}^m} f. \tag{F.13}$$

It holds when both f and its Fourier transform decay sufficiently fast at infinity: as $|x|^{-\alpha}$, where $\alpha > m$. Not worrying too much about analytic details, we can derive (F.13) from (F.12):

$$\begin{aligned}\sum_{v\in L} f(v) &= \sum_{v\in L}\left(\int_{\mathbb{R}^m} e^{-2\pi i\langle v,y\rangle}\,\widehat{f}(y)\,dy\right) = \int_{\mathbb{R}^m}\left(\sum_{v\in L} e^{-2\pi i\langle v,y\rangle}\right)\widehat{f}(y)\,dy \\ &= \frac{1}{d(L)}\int_{\mathbb{R}^m}\sum_{u\in L^*}\delta(y-u)\widehat{f}(y)\,dy = \frac{1}{d(L)}\sum_{u\in L^*}\widehat{f}(u).\end{aligned}$$

Proposition F.3. *Restricting a function to a lattice in the real domain is equivalent to wrapping around the torus in the Fourier domain. Specifically, let ψ be a function on $\mathbb{R}^m$ and ψ_L its restriction to some full-rank lattice L. Then*

$$\big(F_L\psi_L\big)(y) = \sum_{u\in L^*}(F_{\mathbb{R}^m}\psi)(\tilde{y}+u), \qquad \textit{where} \quad y\in\mathbb{R}^m/L^*, \quad \tilde{y}\in\mathbb{R}^m, \quad \tilde{y} \bmod L^* = y.$$

Proof. The desired result is equivalent to the Poisson summation formula (F.13) for the function $f(x) = e^{2\pi i\langle\tilde{y},x\rangle}\psi(x)$. □

The convolution of two functions,

$$\boxed{(f*g)(x) = \int_{\mathbb{R}^m} f(x-y)\,g(y)\,dy} \tag{F.14}$$

is defined when f or g decays sufficiently fast. The convolution is Fourier dual to the multiplication:

$$F_{\mathbb{R}^m}(f*g) = (F_{\mathbb{R}^m}f)(F_{\mathbb{R}^m}g), \qquad F_{\mathbb{R}^m}(fg) = (F_{\mathbb{R}^m}f)*(F_{\mathbb{R}^m}g). \tag{F.15}$$

Similar equations hold for the convolution on an arbitrary (elementary) Abelian group.

In the algorithm described below, we apply the Fourier transform to functions of the form $f : G \to \mathcal{H}$, where G is an Abelian group and $\mathcal{H}$ is some Hilbert space. Such f may be regarded as an element of the tensor product $\mathrm{L}^2(G)\otimes\mathcal{H}$, where the Fourier transform acts on the first factor in the usual way. The result may again be interpreted as a vector-valued function, $\widehat{f} : G^* \to \mathcal{H}$.

We will often be interested in the expression $p(y) = \big\langle\widehat{f}(y)\big|\widehat{f}(y)\big\rangle$. If f has unit L^2 norm, then p is a probability distribution on the dual group G^*. Let G be either $\mathbb{R}^m$ or its quotient by a

full-rank lattice M. In both cases, the second moments of $\widehat{f}$ are related to the partial derivatives of the original function f. Indeed,

$$\Big(F_{\mathbb{R}^m}\tfrac{\partial f}{\partial x_j}\Big)(y) = -2\pi i y_j \widehat{f}(y), \qquad \Big(F_{\mathbb{R}^m/M}\tfrac{\partial f}{\partial x_j}\Big)(y) = -2\pi i u_j \widehat{f}_u.$$

(These formulas are obtained using the integration by parts.) Calculating the inner products $\langle y_j\widehat{f}\,|y_k\widehat{f}\,\rangle$ or $\langle u_j\widehat{f}\,|u_k\widehat{f}\,\rangle$ and using the unitarity of the Fourier transform, we obtain the following expressions for the second moments:

$$\int_{\mathbb{R}^m} y_j y_k \,\langle \widehat{f}(y)\big|\widehat{f}(y)\rangle\, dy = \frac{1}{4\pi^2}\int_{\mathbb{R}^m}\Big\langle \tfrac{\partial f}{\partial x_j}\Big|\tfrac{\partial f}{\partial x_k}\Big\rangle\, dx, \tag{F.16}$$

$$\frac{1}{d(M)}\sum_{u\in M^*} u_j u_k \,\langle \widehat{f}_u\big|\widehat{f}_u\rangle = \frac{1}{4\pi^2}\int_{\mathbb{R}^m/M}\Big\langle \tfrac{\partial f}{\partial x_j}\Big|\tfrac{\partial f}{\partial x_k}\Big\rangle\, dx. \tag{F.17}$$

Proposition F.4. *Let $f : \mathbb{R}^m/M \to \mathcal{H}$ be a function with Lipschitz constant a, where $\mathcal{H}$ is some Hilbert space. Then*

$$\frac{1}{d(M)^2}\sum_{u\in M^*}\|u\|^2\,\langle \widehat{f}_u\big|\widehat{f}_u\rangle \leqslant \frac{a^2}{4\pi^2}.$$

Proof. If f is continuously differentiable, we simply use equation (F.17), where we set $j = k$ and sum over k. A general Lipschitz function can be approximated by its convolution with a suitable continuously differentiable kernel. □

F.2 The HSP algorithm for $G = \mathbb{R}^m$

Let f be an (a, r, ε) oracle function for some full-rank lattice $L \subseteq \mathbb{R}^m$ such that $\lambda_1(L) \geqslant \lambda$ and $d(L) \leqslant d$ (see Definition 1.1). The core part of our algorithm is a sampling subroutine that generates an approximation to a random point of the reciprocal lattice L^*. It works under certain assumptions about the oracle parameters.

Let $\omega : \mathbb{R} \to \mathbb{C}$ be some Lipschitz function with unit L^2 norm supported by the interval $[0, 1]$. For example,

$$\omega(x) = \begin{cases} \sqrt{2}\,\sin(\pi x) & \text{for } x \in [0,1], \\ 0 & \text{otherwise.} \end{cases} \tag{F.18}$$

Let us also choose a sufficiently large number $\Delta = 2^{q_1}$ and a sufficiently small number $\delta = 2^{-q_2}$. Define

$$w(x_1, \ldots, x_m) = \frac{1}{\Delta^{m/2}}\prod_{j=1}^{m}\omega\Big(\frac{x_j}{\Delta}\Big), \tag{F.19}$$

$$w_\delta = w|_{\delta\mathbb{Z}^m} \ \big(\text{restriction of } w \text{ on the lattice } \delta\mathbb{Z}^m\big).$$

In our calculations, we will use the following variables:

Real domain:	$x = \delta\tilde{x} \in \delta\mathbb{Z}^m$	or	$\tilde{x} \in \mathbb{Z}^m$;
Fourier domain:	$y = \delta^{-1}\tilde{y} \in \mathbb{R}^m/\delta^{-1}\mathbb{Z}^m$	or	$\tilde{y} \in \mathbb{R}^m/\mathbb{Z}^m$.

We first create the superposition of points x with the wavefunction w_δ. In the quantum computer, x is actually represented by $\tilde{x}$, therefore the initial state may be written as follows:

$$|w_\delta\rangle = \delta^{m/2}\sum_{\tilde{x}\in\mathbb{Z}^m} w(x)|\tilde{x}\rangle \quad \text{with } x = \delta\tilde{x}.$$

(Our choice of the function ω guarantees the correct normalization on the δ-grid; otherwise we would need to multiply the above expression by some factor that tends to 1 as δ tends to 0.) Then we apply the oracle to get the state

$$|\psi_\delta\rangle = \delta^{m/2} \sum_{\tilde{x}\in\mathbb{Z}^m} w(x)\,|\tilde{x}\rangle \otimes |f(x)\rangle \quad \text{with } x = \delta\tilde{x}. \tag{F.20}$$

The quantum register containing $f(x)$ may be ignored, and we measure the other register in the Fourier basis,

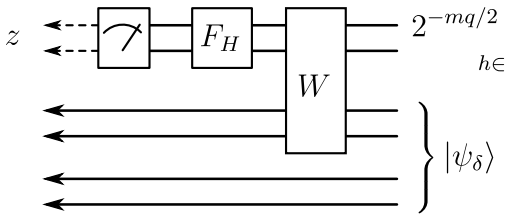
$$|\xi_{\tilde{y}}\rangle = F^{-1}_{\mathbb{Z}^m}|\tilde{y}\rangle = \sum_{\tilde{x}\in\mathbb{Z}^m} e^{-2\pi i\langle\tilde{x},\tilde{y}\rangle}|\tilde{x}\rangle, \qquad \text{where } \tilde{y}\in(\mathbb{R}/\mathbb{Z})^m.$$

An obvious procedure would be to perform the Fourier transform and measure in the standard basis. However, a quantum computer can only do the Fourier transform over a finite group, and the resulting approximation errors are difficult to analyze. Therefore we use a different method.

Note that $|\xi_{\tilde{y}}\rangle$ is an eigenvector of the mutually commuting translation operators $T_{e_1},\dots,T_{e_m}$, where e_j $(j=1,\dots,m)$ are the generators of the group $\mathbb{Z}^m$. The translation by $h\in G$ on an Abelian group G is defined as follows:

$$T_h : \mathrm{L}^2(G)\to\mathrm{L}^2(G), \qquad (T_h f)(x) = f(x-h). \tag{F.21}$$

Thus, the Fourier measurement is equivalent to measuring the eigenvalues of the unitary operators T_{e_j}. A general procedure for the eigenvalue measurement (a.k.a. phase estimation) is described in [Kit95]. To illustrate the difference from the direct use of discrete Fourier transform, let us consider a variant of the phase estimation. In the following circuit, the Fourier transform F_H on the group $H=(\mathbb{Z}_{2^q})^m$ acts on a set of ancillary qubits.



where $W|h,\tilde{x}\rangle = |h,\,\tilde{x}+h\rangle$.

For each value of h, the operator W acts on $\tilde{x}$ as $T_1^{h_1}\cdots T_m^{h_m}$. Note that the "+" in the definition of W means the addition of integer vectors, which are *not* reduced modulo 2^q. Thus, W preserves the decomposition of $|\psi_\delta\rangle$ into the vectors $|\xi_{\tilde{y}}\rangle$. The measurement outcome z may be regarded as a random variable conditioned on $\tilde{y}$, and the latter can be inferred from the former with some precision and confidence. The final result of the sampling subroutine, $Y=-\delta^{-1}2^{-q}z_j$ provides an approximation for y. The error bound for this procedure is pretty standard.

Lemma F.5. *For each j, the probability that the inferred value $\tilde{Y}_j = -2^{-q}z_j$ deviates from $\tilde{y}_j$ by $\geqslant\tilde{\nu}$ is at most $2^{-q}/\tilde{\nu}$. Thus, Y approximates y with precision ν in each coordinate, up to an error probability $\mu_{meas} = m2^{-q}/(\delta\nu)$.*

Proof. For a fixed value of $\tilde{y}$, the components of z are independent random variables with the probability distributions

$$p_j(z_j) = \left|2^{-q}\sum_{k=0}^{2^q-1}\exp\bigl(2\pi ik(2^{-q}z_j+\tilde{y}_j)\bigr)\right|^2 = 2^{-2q}\left|\frac{1-\exp\bigl(2\pi i2^q(2^{-q}z_j+\tilde{y}_j)\bigr)}{1-\exp\bigl(2\pi i(2^{-q}z_j+\tilde{y}_j)\bigr)}\right|^2$$
$$\leqslant 2^{-2q}\bigl(\sin(\pi(2^{-q}z_j+\tilde{y}_j))\bigr)^{-2} \leqslant 2^{-2q}\bigl(2\,\mathrm{dist}_{\mathbb{R}/\mathbb{Z}}(-2^{-q}z_j,\tilde{y}_j)\bigr)^{-2}.$$

In the following calculation, we assume that $\tilde{\nu} \geqslant 2^{-q}$ because otherwise the bound is trivial.

$$\Pr_{z_j}\Big[\mathrm{dist}_{\mathbb{R}/\mathbb{Z}}(-2^{-q}z_j, \tilde{y}_j) \geqslant \tilde{\nu}\Big] \leqslant \frac{2^{-2q}}{4} \sum_{z\in\mathbb{Z}:\, |2^{-q}z+\tilde{y}_j|\geqslant\tilde{\nu}} (2^{-q}z + \tilde{y}_j)^{-2} \leqslant \frac{1}{2}\int_{2^q\tilde{\nu}-1/2}^{+\infty} z^{-2}dz \leqslant (2^q\tilde{\nu})^{-1}.$$

□

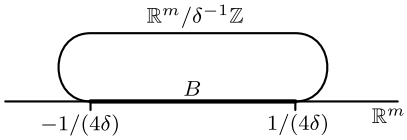
To further analyze the sampling subroutine, we approximate the probability distribution $p_\delta(y)$ of the variable $y = \delta^{-1}\tilde{y}$ by the distribution p that occurs in the $\delta \to 0$ limit. These two distributions are derived from the following quantum states (cf. Eq. (F.20)):

$$|\psi_\delta\rangle = \delta^{m/2} \sum_{\tilde{x}\in\mathbb{Z}^m} |\tilde{x}\rangle \otimes |\psi(x)\rangle, \qquad |\psi\rangle = \int_{\mathbb{R}^m} |x\rangle \otimes |\psi(x)\rangle\, dx, \qquad \text{where} \quad |\psi(x)\rangle = w(x)|f(x)\rangle.$$

(The function $w : \mathbb{R}^m \to \mathbb{C}$ is given by Eq. (F.19); the hidden subgroup oracle f and, thus, the function ψ are vector-valued, i.e. $f, \psi : \mathbb{R}^m \to \mathcal{H}$.) More exactly, p_δ and p are related to the Fourier transform of ψ_δ and ψ, respectively:

$$p_\delta(y) = \big\langle \widehat{\psi}_\delta(y) \big| \widehat{\psi}_\delta(y) \big\rangle \quad \text{with} \quad \widehat{\psi}_\delta = F_{\delta\mathbb{Z}^m}\psi; \qquad p(y) = \big\langle \widehat{\psi}(y) \big| \widehat{\psi}(y) \big\rangle \quad \text{with} \quad \widehat{\psi} = F_{\mathbb{R}^n}\psi. \quad \text{(F.22)}$$

There is a technical problem comparing these distributions because p_δ is defined on the torus $\mathbb{R}^m/\delta^{-1}\mathbb{Z}^m$, whereas p is defined on $\mathbb{R}^m$. We resolve this issue by gluing the torus and the Euclidean space over a "safe domain" $B = \big[-\frac{1}{4\delta}, \frac{1}{4\delta}\big]^m$, where we assume that $\nu \leqslant \frac{1}{4\delta}$. ("Safe" means that the measurement inaccuracy discussed above cannot cause a point $y \in B$ to wrap around the torus.) Now both distributions are defined on the set C obtained by the gluing, albeit with different supports:



$$C = (\mathbb{R}^m/\delta^{-1}\mathbb{Z}^m) \sqcup_B \mathbb{R}^m.$$

Lemma F.6. *The probability distribution p_δ is close to p with respect to the total variation distance:*

$$\frac{1}{2}\int_C |p_\delta(y) - p(y)| \leqslant \mu_{discr}, \qquad \textit{where} \quad \mu_{discr} = \left(4\sqrt{m}\, a + \frac{8\pi m}{\Delta}\right)\delta$$

(The proof is given in Section F.3.)

Let us now focus on the distribution $p(y) = \big\langle \widehat{\psi}(y) \big| \widehat{\psi}(y) \big\rangle$. We have

$$\psi = wf, \qquad \widehat{\psi} = \widehat{w} * (F_{\mathbb{R}^m} f), \qquad \text{where} \quad \widehat{w} = F_{\mathbb{R}^m} w.$$

Since f is a hidden subgroup oracle, we may regard it as a function on $\mathbb{R}^m/L$ and define $\widehat{f} = F_{\mathbb{R}^m/L} f$. That is,

$$\widehat{f}_u = \int_{\mathbb{R}^m/L} e^{2\pi i\langle x,u\rangle} f(x)\, dx \quad \text{ for } u \in L^*, \qquad f(x) = \frac{1}{d(L)} \sum_{u\in L*} e^{-2\pi i\langle x,u\rangle} \widehat{f}_u. \qquad \text{(F.23)}$$

The Fourier transform over $\mathbb{R}^m$ is obtained by the use of Eq. (F.11):

$$(F_{\mathbb{R}^m} f)(y) = \int_{\mathbb{R}^m} e^{2\pi i\langle x,y\rangle} \left(\frac{1}{d(L)} \sum_{u\in L^*} e^{-2\pi i\langle x,u\rangle} \widehat{f}_u\right) dx = \frac{1}{d(L)} \sum_{u\in L^*} \widehat{f}_u\, \delta(y-u).$$

It follows that

$$\widehat{\psi}(y) = \big(\widehat{w} * (F_{\mathbb{R}^m} f)\big)(y) = \frac{1}{d(L)} \sum_{u \in L^*} \widehat{f}_u \, \widehat{w}(y-u), \tag{F.24}$$

$$p(y) = \frac{1}{d(L)^2} \sum_{u,u' \in L^*} \big\langle \widehat{f}_{u'} \big| \widehat{f}_u \big\rangle \, \widehat{w}(y-u) \, \overline{\widehat{w}(y-u')}. \tag{F.25}$$

The last equation is complicated, but we will see that to a good approximation, it is enough to keep the terms with $u = u'$. Let us consider the quantum state $F_{\mathbb{R}^m}|\psi\rangle$ whose wavefunction is given by Eq. (F.24). It consists of identically shaped peaks at the points $u \in L^*$. Each peak has a weight

$$\boxed{q_u = \frac{\langle \widehat{f}_u | \widehat{f}_u \rangle}{d(L)^2} = \frac{1}{d(L)^2} \int_{(\mathbb{R}^m/L)^2} e^{2\pi i \langle x-x', u \rangle} \big\langle f(x') \big| f(x) \big\rangle \, dx \, dx'} \tag{F.26}$$

The numbers q_u can be interpreted as probabilities because they are nonnegative and add up to 1. Indeed, let us normalize f to make a function of unit norm, $g(x) = d(L)^{-1/2} f(x)$. Then

$$\sum_{u \in L^*} q_u = \frac{1}{d(L)} \sum_{u \in L^*} \langle \widehat{g}_u | \widehat{g}_u \rangle = \langle \widehat{g} \, | \widehat{g} \, \rangle = \langle g | g \rangle = 1.$$

We now show that the tails of the peaks can be reduced by a slight modification of the quantum state.

Lemma F.7. *The state $|w\rangle$ is at most $\mu_{tails} = \sqrt{m}/(\Delta\nu)$ distance apart from some state $|w^{(\nu)}\rangle$ whose Fourier transform is supported by $[-\nu, \nu]^m$.*

Proof. Let us consider the unnormalized state $w_*^{(\nu)} = F_{\mathbb{R}^m}^{-1}\big(\theta_{[-\nu,\nu]^m} \widehat{w}\big)$, where θ_S denotes the characteristic function of S. We have

$$\|w_*^{(\nu)} - w\|^2 = \Pr_y\big[y \notin [-\nu, \nu]^m\big],$$

where y is distributed according to the function $|\widehat{w}(y)|^2$. If we write the components of the vector y as $y_j = z_j/\Delta$, then each z_j has the distribution $\big|\widehat{\omega}(z_j)\big|^2$ with $\widehat{\omega} = F_{\mathbb{R}}\omega$. In the following calculation, we denote the first derivative of ω by ω', find the second moment of z_j using equation (F.16), and apply Markov's inequality:

$$\mathbb{E}\big[z_j^2\big] = \int_{\mathbb{R}} z^2 |\widehat{\omega}(z)|^2 \, dz = \frac{1}{4\pi^2} \langle \omega' | \omega' \rangle = \frac{1}{4};$$

$$\Pr\big[|y_j| > \nu\big] = \Pr\big[z_j^2 > (\Delta\nu)^2\big] \leqslant \frac{\mathbb{E}\big[z_j^2\big]}{(\Delta\nu)^2} = \frac{1}{(2\Delta\nu)^2}.$$

The probability of the event $y \notin [-\nu, \nu]^m$ is at most m times greater, hence $\|w_*^{(\nu)} - w\| \leqslant \sqrt{m}/(2\Delta\nu)$. Finally, $w^{(\nu)}$ is defined to be the normalized version of $w_*^{(\nu)}$. Since the quantum state $|w_*^{(\nu)}\rangle$ is the projection of $|w\rangle$ onto a subspace, $\|w_*^{(\nu)} - w\| = \sin\varphi$ and $\|w^{(\nu)} - w\| = 2\sin(\varphi/2)$. It remains to apply the inequality $\sin(\varphi/2) \leqslant \sin\varphi$. □

Corollary F.8. *The probability distribution p is within the total variation distance μ_{tails} from a distribution $p^{(\nu)}$ that is supported by the union of regions $u + [-\nu, \nu]^m$ for $u \in L^*$. If we assume that $\sqrt{m}\,\nu \leqslant \lambda_1(L^*)/2$ so that these regions do not overlap, then the integral of $p^{(\nu)}(y)$ over each region is equal to q_u, see Eq. (F.26).*

Our complete analysis of the sampling subroutine is summarized by the following theorem.

Theorem F.9. *Let $\nu \leqslant 1/(4\delta)$ and $\sqrt{m}\,\nu \leqslant \lambda_1(L^*)/2$. Up to an overall error probability $\mu_{meas} + \mu_{discr} + \mu_{tails}$, the sampling subroutine is equivalent to a random process where a point $u \in L^*$ is selected with probability q_u (see Eq. (F.26)) and the output value Y is at most $2\sqrt{m}\,\nu$ distance away from u. Here*

$$\mu_{meas} = \frac{m2^{-q}}{\delta\nu}, \qquad \mu_{discr} = \left(4\sqrt{m}\,a + \frac{8\pi m}{\Delta}\right)\delta, \qquad \mu_{tails} = \frac{\sqrt{m}}{\Delta\nu}.$$

F.3 Discretization errors

This section provides a proof of Lemma F.6. For this particular purpose, it is sufficient to assume that $f : \mathbb{R}^m \to \mathcal{H}$ has Lipschitz constant a and satisfies the condition $\langle f(x)|f(x)\rangle = 1$ for all x. The distribution p in the statement of the lemma is given by the equation $p(y) = \big\langle \widehat{\psi}(y)\big|\widehat{\psi}(y)\big\rangle$, where $\widehat{\psi} = F_{\mathbb{R}^m}\psi$, $|\psi(x)\rangle = w(x)|f(x)\rangle$, and w is given by equations (F.18) and (F.19). The distribution p_δ is constructed in the same way, except that ψ is restricted to the δ-grid. To compare $\widehat{\psi}$ and $\widehat{\psi}_\delta$, we use the fact that the lattice restriction in the real domain is equivalent to wrapping around the torus in the Fourier domain, see Proposition F.3. The lattice in question is $\delta\mathbb{Z}^m$. Thus,

$$\widehat{\psi}_\delta(y) = \sum_{u \in \delta^{-1}\mathbb{Z}^m} \widehat{\psi}(y+u), \tag{F.27}$$

where we may assume that y belongs to the fundamental domain of the torus, i.e. $y \in \left[-\frac{1}{2\delta}, \frac{1}{2\delta}\right]^m$.

Instead of using Eq. (F.27) directly, we will approximate ψ by some function ξ such that its Fourier transform $\widehat{\xi}$ is supported by $B = \left[-\frac{1}{4\delta}, \frac{1}{4\delta}\right]^m$. (The role of the "safe domain" B is explained in the paragraph preceding Lemma F.6.) Then the probability densities $\big\langle \widehat{\xi}(y)\big|\widehat{\xi}(y)\big\rangle$ and $\big\langle \widehat{\xi}_\delta(y)\big|\widehat{\xi}_\delta(y)\big\rangle$ coincide. Thus,

$$\begin{aligned} \frac{1}{2}\int_C |p_\delta(y) - p(y)| &\leqslant \frac{1}{2}\int_{\mathbb{R}^m/\delta^{-1}\mathbb{Z}^m} \Big|p_\delta(y) - \big\langle \widehat{\xi}_\delta(y)\big|\widehat{\xi}_\delta(y)\big\rangle\Big| + \frac{1}{2}\int_{\mathbb{R}^m} \Big|p(y) - \big\langle \widehat{\xi}(y)\big|\widehat{\xi}(y)\big\rangle\Big| \\ &\leqslant \big\|\widehat{\psi}_\delta - \widehat{\xi}_\delta\big\| + \big\|\widehat{\psi} - \widehat{\xi}\big\| \;=\; \|\psi_\delta - \xi_\delta\| + \|\psi - \xi\|. \end{aligned} \tag{F.28}$$

We set

$$\xi = F^{-1}_{\mathbb{R}^m}\big(\widehat{g}\,\widehat{\psi}\,\big) = g * \psi = \int_{\mathbb{R}^m} g(v)\,(T_v\psi)\,dv, \tag{F.29}$$

where T_v is the operator of translation by v and $g : \mathbb{R}^m \to \mathbb{C}$ is a suitable function that will be chosen later. We require that g satisfy the following conditions:

1. $\widehat{g}$ is supported by B (so that ξ is also supported by B);
2. $g(v)$ is real and nonnegative for all v;
3. $\int_{\mathbb{R}^m} g(v)\,dv = 1$.

Conditions (2) and (3) imply that

$$\|\psi - \xi\| \leqslant \int_{\mathbb{R}^m} g(v)\,\|T_v\psi - \psi\|\,dv, \qquad \|\psi_\delta - \xi_\delta\| \leqslant \int_{\mathbb{R}^m} g(v)\,\big\|(T_v\psi - \psi)_\delta\big\|\,dv. \tag{F.30}$$

Lemma F.10. *For all $v \in \mathbb{R}^m$,*

$$\|\psi - T_v\psi\| \leqslant b\|v\|, \qquad \big\|(\psi - T_v\psi)_\delta\big\| \leqslant b\|v\|, \qquad \text{where} \quad b = a + 2\pi\sqrt{m}/\Delta.$$

Proof. Let us first estimate the L^2-norm of the function $\alpha_v : x \mapsto \omega(x) - \omega(x-v)$, where $\omega(x)$ is given by equation (F.18) and $v \in \mathbb{R}$ is fixed. We have $|\omega'(x)| \leqslant \sqrt{2}\,\pi$, hence $|\alpha_v(x)| \leqslant \sqrt{2}\,\pi|v|$. However, $\omega(x) - \omega(x-v)$ actually vanishes unless x or $x-v$ belongs to the interval $[0,1]$. Thus,

$$\|\alpha_v\| \leqslant \left(\int_{[0,1]\cup[v,v+1]} |\alpha(x)|^2\,dx\right)^{1/2} \leqslant 2\pi|v|.$$

Now, we consider the function $\varphi = \psi - T_v\psi$, i.e. $\varphi(x) = \psi(x) - \psi(x-v)$. Since $\psi(x) = w(x)f(x)$, we get the inequality

$$\|\varphi(x)\| \leqslant \|f(x)-f(x-v)\|\cdot|w(x)| + \|\,|f(x-v)\rangle\|\cdot|w(x)-w(x-v)| \leqslant a\|v\|\cdot|w(x)| + |w(x)-w(x-v)|.$$

The L^2-norm of the first term (as a function of x) is equal to $a\|v\|\cdot\|w\| = a\|v\|$ because $\|w\| = 1$. The second term can be written as follows:

$$w(x) - w(x-v) = \sum_{j=1}^{m} \Delta^{-m/2} \left(\prod\nolimits_{k=1}^{j-1} \omega\left(\tfrac{x_k}{\Delta}\right)\right) \left(\omega\left(\tfrac{x_j}{\Delta}\right) - \omega\left(\tfrac{x_j - v_j}{\Delta}\right)\right) \left(\prod\nolimits_{k=j+1}^{m} \omega\left(\tfrac{x_k - v_k}{\Delta}\right)\right).$$

The norm of each term on the right-hand side is bounded by $\|\omega\|^{m-1}\|\alpha_{v_j/\Delta}\| \leqslant 2\pi|v_j|/\Delta$. Summing over j, we get the upper bound $(2\pi\sqrt{m}/\Delta)\,\|v\|$. Thus, $\|\varphi\| \leqslant b\|v\|$, where b is as indicated in the lemma. The inequality $\|\varphi_\delta\| \leqslant b\|v\|$ is proven similarly, using the fact that the restriction of ω to any grid with period δ/Δ has unit L^2-norm. □

We now choose a suitable function g so as to satisfy the above-mentioned conditions as well as obtaining a reasonable upper bound for $\int_{\mathbb{R}^m} g(v)\|v\|\,dv$. Let

$$g(x) = |h(x)|^2, \qquad \text{where} \quad \widehat{h}(y_1,\dots,y_m) = \begin{cases} (8\delta)^{m/2} \prod_{j=1}^{m} \cos(4\pi\delta y_j) & \text{if } y \in \left[-\frac{1}{8\delta}, \frac{1}{8\delta}\right]^m, \\ 0 & \text{otherwise.} \end{cases} \tag{F.31}$$

We will ignore the particular form of h for the moment and only use the fact that it has unit L^2-norm and that $\widehat{h}$ is supported by the region $A = \left[-\frac{1}{8\delta}, \frac{1}{8\delta}\right]^m$. It is clear that $g(x) = |h(x)|^2 \geqslant 0$ and $\int_{\mathbb{R}^m} g(x)\,dx = \|h\|^2 = 1$. Since $\widehat{g} = \widehat{h} * \widehat{\overline{h}}$ and $\widehat{\overline{h}}(y) = \overline{\widehat{h}(-y)}$ (where the bar denotes the complex conjugation), we have

$$\widehat{g}(y) = \int_{\mathbb{R}^m} \widehat{h}(z)\,\overline{\widehat{h}(z-y)}\,dz.$$

Thus, $\widehat{g}$ is supported by the region $A - A \subseteq B$ as required.

To obtain an upper bound for $\int_{\mathbb{R}^m} g(v)\|v\|\,dv$, we first estimate $\int_{\mathbb{R}^m} g(v)\|v\|^2\,dv$ using equation (F.17):

$$\int_{\mathbb{R}^m} g(v)\|v\|^2\,dv = \sum_{j=1}^{m} \int_{\mathbb{R}^m} v_j^2 |h(v)|^2\,dv = \frac{1}{4\pi^2} \sum_{j=1}^{m} \int_{\mathbb{R}^m} \left|\frac{\partial\widehat{h}}{\partial y_j}\right|^2 dy = 4m\delta^2.$$

It follows by Jensen's inequality that

$$\int_{\mathbb{R}^m} g(v)\|v\|\,dv \leqslant 2\sqrt{m}\,\delta.$$

Combining this with Eq. (F.30) and Lemma F.10, we find that

$$\|\psi - \xi\| \leqslant 2\sqrt{m}\,\delta\, b, \qquad \|\psi_\delta - \xi_\delta\| \leqslant 2\sqrt{m}\,\delta\, b, \qquad \text{where} \quad b = a + 2\pi\sqrt{m}/\Delta.$$

Substituting this last result into equation (F.28) completes the proof of Lemma F.6.

F.4 Sampling from lattices

Throughout this section, L is a full-dimensional lattice in $\mathbb{R}^n$, and L^* is the dual lattice. We will analyze the samples from distribution q_u over L^*. Below we first show that $q(\cdot)$ is concentrated on $L^*_R := \{u \in L^* : \|u\| \leq R\}$ for $R = \Omega(a)$, where a is the Lipschitz constant of f. As a result, we can assume that we are sampling from a ball of radius R in $\mathbb{R}^n$.

Proposition F.11. *Assume* $\|\frac{d|f_x\rangle}{dx}\|^2 := \sum_j \|\frac{\partial|f_x\rangle}{\partial x_j}\|^2 \leq a^2$, *then* $\sum_{u\in L^*_R} q_u \geq 1 - (\frac{a}{2\pi R})^2$. *In particular,* $\sum_{u\in L^*_R} q_u \geq 1 - negl(n)$ *for* $R \geq a \cdot 2^n$.

Proof.

From Proposition F.4, we immediately get $\sum_{u\in L^*} q_u\|u\|^2 = \sum_{u\in L^*} \frac{\langle \hat{f}_u|\hat{f}_u\rangle}{d^2(L)}\|u\|^2 \leq \frac{a^2}{4\pi^2}$. Now let $\bar{L}^*_R := L\backslash L^*_R = \{u \in L^* : \|u\| \geq R\}$. We have

$$\sum_{u\in\bar{L}^*_R} q_u \cdot R^2 \leq \sum_{u\in\bar{L}^*_R} q_u\|u\|^2 \leq \sum_{u\in L^*} q_u\|u\|^2 \leq (a/2\pi)^2\,.$$

Hence $\sum_{u\in L^*_R} q_u = 1 - \sum_{u\in\bar{L}^*_R} q_u \geq 1 - (\frac{a}{2\pi R})^2$. □

Next we show a crucial technical tool in Theorem F.12 that will be used extensively. Roughly it states that the probability that a sample according to q_u lands in any maximal sublattice is upper bounded away from 1. Its proof is deferred to Sect. F.5.

Theorem F.12. *If* $M \subset L^*$ *is a maximal sublattice then* $\sum_{u\in M} q_u \leq \max\{\frac{1}{|L^*/M|}, \frac{18r}{\lambda_1(L)}\}$.

Using this theorem, we can prove our main theorem of this section.

Theorem F.13. *Let* $\{u_1, \ldots, u_m\}$ *be* $m := n^2 + cn$ *independent samples according to* q_u *with* $c = \lceil \log(R^n d(L))\rceil$. *Then* $\{u_i\}_{i=1}^m$ *form a generating set with probability* $1 - negl(n)$.

Proof. We know from Theorem F.12 that for any maximal sublattice $M \subsetneq L^*$, $q_M := \sum_{z\in M} q_z \leq 1/2$. We interpret the m samples as two groups and argue that: 1) first n^2 samples generate a full-rank sublattice M except with negligible probability (Prop. F.14 below); and 2) the next cn samples, except with negligible probability, fill in the finite group L^*/M (Prop. F.15 below). Thus in total we get a generating set for L^* with probability $1 - negl(n)$. □

Proposition F.14. $m_1 := n^2$ *independent samples* u_i *from* q_u *generates a sublattice* M *of full rank with prob. at least* $1 - negl(n)$.

Proof. Assume we have already obtained k independent samples $(u_1, \ldots, u_k)$ from q_u. Let $S_k := span\langle u_1, \ldots, u_k\rangle \subsetneq \mathbb{R}^n$ and $M_k := S_k \cap L^*$. We then show that (at least) one of the next n samples will be outside S_k and thus increase the rank by 1. Notice that there exists a basis $\tilde{B} = (v_1, \ldots, v_n)$ of L^*, such that $L(v_1, \ldots, v_d) = M_k$, with $d = rank(S_k)$. (Actually, given $(u_1, \ldots, u_k)$ and a basis B for L^*, such a basis $\tilde{B}$ can be computed efficiently. Cf. [Mic08].) Now consider $M := L(v_1, \ldots, v_{n-1}, p \cdot v_n)$ with an arbitrary prime p. Clearly M is maximal, and $M_k \subseteq M$. Hence within the next n samples, there will be one sample of the form $\ell \cdot v_n$ for $p \nmid \ell$ with probability at least $1 - 1/2^n$. Following this argument, we can conclude that n^2 samples will span a full-rank sublattice M with prob. at least $(1 - 1/2^n)^n \geq 1 - n/2^n = 1 - negl(n)$. □

Proposition F.15. *Assume $u_1, \dots, u_{m_1}$ generate a full-rank sublattice $M \subsetneq L^*$. Let $u'_1, \dots, u'_{m_2}$ be $m_2 = \lceil \log |L^*/M| \rceil \cdot n$ independent samples from q_u. Then $(u_1, \dots, u_{m_1}, u'_1, \dots, u'_{m_2})$ form a generating set of L^* with prob. at least $1 - negl(n)$.*

Proof. The argument is similar to the preceding one. Every n samples double the index of the sublattice generated by $\{u_i\}_{i=1}^{m_1}$ and the samples so far with probability at least $1 - 1/2^n$. Continue this argument we conclude that we fulfill the whole lattice within m_2 samples with prob. at least $(1 - 1/2^n)^{\lceil \log |L^*/M| \rceil}$. By Lemma F.16 below, we conclude that $(1 - 1/2^n)^{\lceil \log |L^*/M| \rceil} \geq 1 - \frac{\log(R^n d(L))}{2^n} = 1 - negl(n)$. □

Finally we derive an upper bound on $|L^*/M|$, which justifies the choice of constant c in Theorem F.13.

Lemma F.16. *Let m be an integer $\geq n$, and suppose m samples span a full-rank sublattice $M \subseteq L^*$. Then $|L^*/M| \leq R^n d(L)$.*

Proof. $|L^*/M| = d(M)/d(L^*) \leq R^n d(L)$ because the samples that generate M are all within distance R from the origin. □

F.5 Proof of Theorem F.12

Proof. We estimate the probability to lie in a sublattice $M \subset L^*$:

$$\begin{aligned}
\sum_{u \in M} q_u &= \sum_{u \in M} \frac{1}{d^2(L)} \int_{x,x' \in \mathbb{R}^n/L} \langle f_{x'} | f_x \rangle e^{2\pi i \langle x - x', u \rangle} dx dx' \\
&= \frac{1}{d^2(L)} \int_{x,x'} \langle f_{x'} | f_x \rangle \sum_{u \in M} e^{2\pi i \langle x - x', u \rangle} dx dx' \\
&\overset{\text{PSF}}{=} \frac{1}{d^2(L)} \int_{x,x'} \langle f_{x'} | f_x \rangle d(M^*) \sum_{v \in M^*} \delta(x - x' - v) dx dx' \\
&= \frac{d(M^*)}{d^2(L)} \int_{x,x'} \langle f_{x'} | f_x \rangle \sum_{v_1 \in M^*/L, v_2 \in L} \delta(x - x' - (v_1 + v_2)) dx dx' \\
&\overset{z := x' + v_2}{=} \frac{1}{|L^*/M|} \frac{1}{d(L)} \int_{x \in \mathbb{R}^n/L} \left(\sum_{v_1 \in M^*/L} \int_{z \in \mathbb{R}^n} \langle f_z | f_x \rangle \delta(x - v_1 - z) dz \right) dx \\
&= \frac{1}{|L^*/M|} \int_{x \in \mathbb{R}^n/L} \frac{1}{d(L)} \sum_{v_1 \in M^*/L} \langle f_{x - v_1} | f_x \rangle dx \\
&\leq \frac{Q_M}{|L^*/M|} \int_{x \in \mathbb{R}^n/L} \frac{1}{d(L)} dx = \frac{Q_M}{|L^*/M|}
\end{aligned}$$

where $Q_M := \sup_{x \in \mathbb{R}^n/L^*} \{ \sum_{y \in M^*/L} \langle f_x | f_{x-y} \rangle \} \leq \sup_{x \in \mathbb{R}^n/L^*} \#\{z \in M^* : \|z - x\| < r\}$.

Let $B_x(r) := \{z \in M^* : \|z - x\| < r\}$. Observe that for any $x \in M^*$, $|B_x(r)| = |B_0(r)|$. On the other hand, let $x \in \mathbb{R}^n/L$ but $x \notin M^*$. Pick $y \in M^* \cap B_x(r)$. We claim that $B_x(r) \subseteq B_y(2r)$. Namely all the M^*-lattice points that are within distance r from x have distance at most $2r$ from y. Because otherwise there exists a $z \in B_x(r)$ such that $\|z - y\| > 2r$. This is a contradiction because y and z are both in $B_x(r)$ and hence $\|z - y\| \leq 2r$. Thus we have

$$\sup_{x \in \mathbb{R}^n/L^*} \#\{z \in M^* : \|z - x\| < r\} \leq \#\{z \in M^* : \|z\| < 2r\}.$$

We show in the next theorem (Theorem F.17) that the number of points of M^* inside any $\tilde{r}$-ball with $\tilde{r} \leq \lambda_1(L)/2$ is bounded by $1+\frac{9\tilde{r}}{\lambda_1(L)} \cdot |M^*/L|$. Therefore $Q_M \leq \max\{1, \frac{18r}{\lambda_1(L)}|M^*/L|\}$ (noticing that our in our HSP function f, $r \leq \lambda/4$), and hence Theorem F.12 holds.

□

Theorem F.17. *Let L, M, M^*, L^*, r as above, and assume $r < \lambda_1(L)/2$. Also assume that $M \subset L^*$ is maximal, i.e. $|L^*/M| = |M^*/L| = p$ with p prime. Then*

$$\#\{x \in M^* : \|x\| \leq r\} \leq 1 + c\frac{r}{\lambda_1(L)} \cdot |M^*/L|.$$

The theorem clearly holds when $\lambda_1(L) = \lambda_1(M^*)$, since in this case the origin is the only point inside the r-ball. Hence in the following we may assume that $\lambda_1(L) > \lambda_1(M^*)$. Then a shortest vector ξ in M^* will generate M^*/L, since M^*/L has prime order.

To prove the theorem we need two lemmas.

Lemma F.18. *Let $G := M^*/L$ be the quotient group. Then $G \cong \mathbb{Z}/p\mathbb{Z}$. Define a metric μ on G by letting*

$$\mu(x, y) = \min\{\|\tilde{x} - \tilde{y}\| : \tilde{x} \mod L = x, \tilde{y} \mod L = y\}.$$

Assume that $\lambda_1(L) \neq \lambda_1(M^*)$. Then $diam_\mu(G) \geq \lambda_1(L)/3$.

Proof. By definition, $diam_\mu(G) = \max_{x,y \in G} \mu(x, y)$. Since $\mu(x+z, y+z) = \mu(x, y)$ (μ is translation invariant), it suffices to consider distances from 0. Let ξ be a shortest vector in M^*. (We assumed that the image of ξ in M^*/L is not zero.) Then M^*/L is generated by the image of ξ.

Claim: There exists a basis for M^* of the form $\xi, b_2, \ldots, b_n$ with $b_2, \ldots, b_n \in L$.

Proof of claim: Since ξ is a shortest vector of M^*, there is a basis of M^* that contains this vector. Now suppose that a basis $\xi, b_2', \ldots, b_n'$ of M^* is given. Since M^*/L is generated by the image of ξ, we have $b_2' + L = k\xi + L$ for some $0 \leq k \leq p-1$. If $k = 0$, then $b_2' \in L$. Otherwise $b_2' = k\xi + b_2$ for some $1 \leq k \leq p-1$ and $b_2 \in L$. Then ξ and $b_2, b_3', \ldots, b_n'$ generate the same lattice as ξ and $b_2', \ldots, b_n'$, so we can replace b_2' with b_2. The same can be done for the other basis vectors. This proves the claim.

Now assume by contradiction that $diam_\mu(G) < \lambda_1(L)/3$. Take the point $\frac{(p-1)}{2}\xi \in M^*$. By the assumption on the diameter,

$$\mu\left(\frac{(p-1)}{2}\xi, 0\right) < \lambda_1(L)/3,$$

so there exists a $y \in L$ with $||\frac{(p-1)}{2}\xi - y|| < \lambda_1(L)/3$. Similarly, since the closest L-point to ξ is the origin, the assumption that $diam_\mu(G) < \lambda_1(L)/3$ implies that $||\xi|| < \lambda_1(L)/3$. But since $G = M^*/L \cong \mathbb{Z}/p\mathbb{Z}$, we have $p\xi \in L$, and hence $p\xi - 2y$ is a vector in L of length

$$\begin{aligned} ||p\xi - 2y|| &= ||2 \cdot \left(\frac{(p-1)}{2}\xi - y\right) + \xi|| \\ &\leq 2 \cdot ||\frac{(p-1)}{2}\xi - y|| + ||\xi|| \\ &< 2/3\lambda_1(L) + 1/3\lambda_1(L) = \lambda_1(L). \end{aligned}$$

Since $p\xi, b_2, \ldots, b_n$ is a basis for L, we can't have $p\xi = 2y$ with $y \in L$. Hence $p\xi - 2y$ is a nonzero vector of L of length $< \lambda_1(L)$, contradiction.

□

Lemma F.19. *Let ξ be a shortest vector in M^*, and assume $\|\xi\| < r$. Again assume that $L \subset M^*$ is maximal. Then there exists a set $A \subseteq M^*/L$ with the following properties:*

- *for all $x, y \in A$, with $x \neq y$ we have $\mu(x, y) > 2r$;*
- *$|A| \geq c^{-1}\lambda_1(L)/r$.*

Proof. Fix $\ell = 3 \cdot r$. Let ξ be a shortest vector in M^*. Define a sequence of elements a_j (for $j \in \{0, \ldots, \lceil \frac{\lambda_1(L)}{3\ell} \rceil\}$) of elements in M^*/L as follows: we let a_j be the first point in the sequence $0, \xi, 2\xi, \ldots$ such that $\mu(0, a_j) \geq j \cdot \ell$. Since $|M^*/L| = p$ with p prime, the image of the element ξ in M^*/L generates M^*/L. Hence Lemma F.18 and our choice for the range of j imply that such elements a_j exist.

Now fix j. Then $a_j = k \cdot \xi$ for some $k \in \mathbb{N}$. By definition of a_j this means that

$$\mu(0, (k-1)\xi) < j\ell.$$

Then the triangle inequality gives

$$\mu(0, k\xi) < j\ell + \|\xi\| \leq j\ell + r.$$

Hence we obtain

$$j\ell \leq \mu(0, a_j) \leq j\ell + r. \tag{F.32}$$

We claim that the $A = \{a_j : 0 \leq j \leq \lceil \frac{\lambda_1(L)}{3\ell} \rceil\}$ is the desired set. By Equation F.32 and since $\ell = 3 \cdot r$, the elements in A are distinct and so A has the correct size if we let $c = 9$. It remains to show that $\mu(a_i, a_j) > 2r = \ell - r$. Assume by contradiction that $\mu(a_i, a_j) \leq \ell - r$. We may assume that $j = i + 1$ since these two points are closest. If $\mu(a_i, a_{i+1}) \leq \ell - r$, then since $\mu(0, a_i) < i\ell + r$ we have

$$\mu(0, a_{i+1}) \leq \mu(0, a_i) + \mu(a_i, a_{i+1}) < (i\ell + r) + (\ell - r) < (i+1)\ell,$$

contradicting our choice of a_{i+1}. This finishes the proof. □

Now we can prove Theorem F.17:

Proof of Theorem F.17 Since $r < \lambda_1(L)/2$, we have that

$$\#\{x \in M^* : \|x\| \leq r\} = \#\{x \in M^*/L : \mu(0, x) \leq r\}.$$

Now suppose A satisfies the conditions as in Lemma F.19. For each $a \in A$ let $B_a = \{x \in M^*/L : \mu(a, x) \leq r\}$. Since $\mu(x, y) > 2r$ for distinct points $x, y \in A$ it follows that the sets B_a are disjoint. Also, since μ is translation invariant, the sets B_a all have the same size (which is $|B_{a_0}| = |B_0|$).

Hence

$$|B_0| = \#\{x \in M^*/L : \mu(0, x) \leq r\} \leq \frac{|M^*/L|}{|A|}.$$

Since $|A| \geq \frac{1}{9}\lambda_1(L)/r$, we obtain

$$\#\{x \in M^* : \|x\| \leq r\} \leq |M^*/L| \cdot \frac{9r}{\lambda_1(L)}.$$

This gives the desired bound. □