On the Closest Vector Problem with a Distance Guarantee

Daniel Dadush* dadush@cwi.nl
Oded Regev†‡
Noah Stephens-Davidowitz† noahsd@cs.nyu.edu

Abstract

We present a substantially more efficient variant, both in terms of running time and size of preprocessing advice, of the algorithm by Liu, Lyubashevsky, and Micciancio [LLM06] for solving CVPP (the preprocessing version of the Closest Vector Problem, CVP) with a distance guarantee. For instance, for any $\alpha < 1/2$, our algorithm finds the (unique) closest lattice point for any target point whose distance from the lattice is at most α times the length of the shortest nonzero lattice vector, requires as preprocessing advice only $N \approx \tilde{O}(n \exp(\alpha^2 n/(1-2\alpha)^2))$ vectors, and runs in time $\tilde{O}(nN)$.

As our second main contribution, we present reductions showing that it suffices to solve CVP, both in its plain and preprocessing versions, when the input target point is within some bounded distance of the lattice. The reductions are based on ideas due to Kannan [Kan87] and a recent sparsification technique [DK13]. Combining our reductions with the LLM algorithm gives an approximation factor of $O(n/\sqrt{\log n})$ for search CVPP, improving on the previous best of $O(n^{1.5})$ due to Lagarias, Lenstra, and Schnorr [LLS90]. When combined with our improved algorithm we obtain, somewhat surprisingly, that only $O(n)$ vectors of preprocessing advice are sufficient to solve CVPP with (the only slightly worse) approximation factor of $O(n)$.

1 Introduction

A *lattice* is the set of all integer combinations of n linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in $\mathbb{R}^n$. These vectors are known as a *basis* of the lattice. In the last couple of decades, lattices became a central object of investigation in theoretical computer science due to their wide range of algorithmic and cryptographic applications.

The two most fundamental lattice problems are the Shortest Vector Problem (SVP) and the Closest Vector Problem (CVP). Given an n-dimensional lattice $\mathcal{L}$ (specified using an arbitrary basis), the SVP is to find a shortest non-zero vector in $\mathcal{L}$, and, given in addition a target point $\mathbf{t} \in \mathbb{R}^n$, the CVP is to find a closest vector to $\mathbf{t}$ in $\mathcal{L}$. For their approximation versions, the goal is to compute solutions whose length or distance is within some factor of optimal, and in the associated decisional versions, one must estimate the length or distance to within the desired factor.

From a computational complexity point of view, lattice problems are quite fascinating. For the nearly exponential approximation factor of $2^{O(n \log\log n/\log n)}$, efficient algorithms are known [LLL82,

*Centrum Wiskunde & Informatica (CWI), Amsterdam.
†Courant Institute of Mathematical Sciences, New York University.
‡Supported by the National Science Foundation (NSF) under Grant No. CCF-1320188. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

Bab86, Sch87, AKS01, MV13]. On the other hand, for solving the exact problems (or even for approximating to within $\mathrm{poly}(n)$ factors) the best known algorithms run in time $2^{O(n)}$ [AKS01, MV13]. It is known that for some $c > 0$, approximating CVP to within $n^{c/\log\log n}$ is NP-hard (see [DKRS03] as well as [Kho10] and references therein). Under reasonable complexity assumptions, SVP is also known to be hard for the same approximation factor [Mic01b, Kho04, HR12]. Finally, for approximation factor $\sqrt{n}$ both problems are known to be in NP∩coNP and hence unlikely to be NP-hard [GG00, AR05]. For an introduction to the area see, e.g., [MG02, Reg10].

In this paper we also consider a natural variant of CVP known as the *Closest Vector Problem with Preprocessing (CVPP)*. The motivation comes from applications in coding theory and cryptography where the lattice is often fixed once and for all, and the input only consists of the target point $\mathbf{t}$. In CVPP, the algorithm is allowed to spend an unlimited amount of time *preprocessing* the given lattice and output at the end a polynomial-size description of the lattice. Then, given that description and a target point $\mathbf{t}$, our goal is to efficiently solve $\mathrm{CVP}(\mathcal{L}, \mathbf{t})$. As usual, one can consider either the search or the decision versions.

The computational hardness of CVPP was investigated in a sequence of works [Mic01a, FM04, Reg04, AKKV11], culminating in a hardness factor of $2^{\log^{1-\varepsilon} n}$ for any $\varepsilon > 0$ by Khot, Popat, and Vishnoi under reasonable complexity assumptions [KPV14]. Behind the latest two hardness results is a preprocessing version of the PCP theorem.

The situation in terms of positive results, which is the focus of this work, is even more interesting. It follows from the early work of Lagarias, Lenstra, and Schnorr [LLS90] on so-called Korkine-Zolotarev bases that there exists an $n^{3/2}$ approximation algorithm for CVPP. Somewhat surprisingly, prior to this work, their algorithm was still the best known approximation algorithm for CVPP.

Improved algorithms were known only for the *decision* variant of CVPP in which the task is to approximate the distance of the target point to the lattice. An $O(n)$ approximation algorithm was given in [Reg04] and then improved by Aharonov and Regev [AR05] to an $O(\sqrt{n/\log n})$ approximation algorithm, a natural approximation factor that seems very difficult to beat. We are therefore in the (somewhat absurd!) situation that we know that there is a close vector but we somehow can't find it! We note that an equivalence between the search and decision versions of CVP holds for the exact case [MG02], but is not known to hold for the approximate case.

Since the latter algorithm is very natural and closely related to our work, we describe it here briefly. The main idea is to define for any lattice $\mathcal{L} \subset \mathbb{R}^n$ the *periodic Gaussian function* $f : \mathbb{R}^n \to \mathbb{R}^+$, given by

$$f(\mathbf{t}) = \frac{\rho(\mathcal{L} + \mathbf{t})}{\rho(\mathcal{L})}, \tag{1}$$

where $\rho(A) = \sum_{\mathbf{x} \in A} \exp(-\pi \|\mathbf{x}\|^2)$. See Figure 1 for an illustration. The algorithm now follows from two observations. The first is that for points $\mathbf{t}$ at distance greater than $\sqrt{n}$ from the lattice, $f(\mathbf{t})$ is essentially zero, whereas for $\mathbf{t}$ at distance less than $\sqrt{\log n}$, $f(\mathbf{t})$ is non-negligible, so being able to compute f would suffice to solve the decision problem. The second crucial idea is that the function f, despite being defined in terms of a sum over infinitely many lattice points, can be approximated to within any $\pm 1/\mathrm{poly}(n)$ by a function with a polynomial-size circuit. *Finding* that circuit seems hard, but since it only depends on the lattice, we can do it in the preprocessing phase. To show that such an estimator exists, they first observe that the Poisson summation formula gives the identity

$$f(\mathbf{t}) = \mathop{\mathbb{E}}_{\mathbf{w} \sim D_{\mathcal{L}^*}} \left[\cos(2\pi \langle \mathbf{w}, \mathbf{t} \rangle)\right], \tag{2}$$

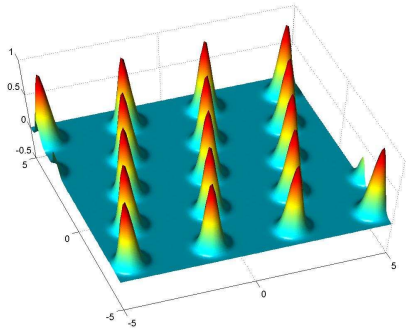


Figure 1: The periodic Gaussian function

where $\mathbf{w}$ is a vector of the dual lattice $\mathcal{L}^*$ sampled from $D_{\mathcal{L}^*}$, the so-called *discrete Gaussian distribution* over $\mathcal{L}^*$. This naturally leads to the definition of the estimator

$$f_W(\mathbf{t}) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i=1}^{N} \cos(2\pi \langle \mathbf{w}_i, \mathbf{t} \rangle) \,, \tag{3}$$

where $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N) \in \mathcal{L}^*$ are i.i.d. samples from $D_{\mathcal{L}^*}$, which one can show satisfies $f_W \approx f$ with high probability over the choice of W assuming N is a large enough $\text{poly}(n)$. Once the vectors in W are given as preprocessing advice, computing f_W is clearly efficient. This completes the description of the decision CVPP algorithm from [AR05].

Moving on to the search problem, a natural approach is to perform some sort of hill-climbing or gradient ascent on the periodic Gaussian function f (using our estimator f_W) starting from the target point. As can be seen in Figure 1, f attains its maxima in lattice points, and so one would expect this process to converge to the nearest lattice point. Indeed, this is the approach followed by Liu, Lyubashevsky, and Micciancio [LLM06]: they showed (improving on earlier work of Klein [Kle00]) how, given the estimator f_W, to efficiently find the nearest lattice point to any target that is within distance $O(\sqrt{\log n / n}) \cdot \lambda_1(\mathcal{L})$ of the lattice, where $\lambda_1(\mathcal{L})$ denotes the length of the shortest non-zero vector in $\mathcal{L}$. Notice, however, that this falls short of solving CVPP since the algorithm is only guaranteed to work for target points that are close to the lattice. This problem is known as the *Bounded Distance Decoding problem* (BDD), or BDDP in its preprocessing version.

Extending these ideas to the search version of CVPP, or even just to BDDP for a larger decoding radius, has proved to be elusive. The bound $O(\sqrt{\log n / n}) \cdot \lambda_1(\mathcal{L})$ arises as a result of the following tension. On one hand, we would like to choose the width of the Gaussians in f as wide as possible in order to increase the radius in which f is detectable and in which we can apply gradient ascent. On the other hand, making them too wide causes "interference" between the various peaks so we no longer have as clean a picture as in Figure 1. We demonstrate this interference in Section 7 by presenting a simple example in which f has a local maximum at distance $\lambda_1 / \sqrt{2}$ from the lattice whose value is exponentially close to the global maximum of 1.

1.1 Our contributions

Solving BDDP by hill climbing. Our main technical contribution, given in Sections 3 and 4, is an improvement of the hill-climbing algorithm of LLM. While the basic approach is the same, our algorithm uses a more natural gradient ascent procedure, compared with LLM's "discrete" version. Namely, at each step we replace the current point $\mathbf{t}$ with an approximation of

$$\mathbf{t} + \nabla f(\mathbf{t})/(2\pi f(\mathbf{t})) \; . \tag{4}$$

Letting $\mathbf{y}$ be the closest lattice point to $\mathbf{t}$ and ignoring the interference coming from other peaks, we can think of f as $\exp(-\pi\|\mathbf{t}-\mathbf{y}\|^2)$, in which case (4) is easily seen to equal $\mathbf{y}$, our desired output. For comparison, LLM uses small axis-aligned steps, replacing $\mathbf{t}$ with $\mathbf{t} \pm \delta\mathbf{e}_i$ for some $\delta > 0$ and $i \in [n]$. Combining our more natural algorithm with a rather detailed analysis of the periodic Gaussian function f (which is of independent interest, see Section 4.1) and of its estimator f_W (Section 4.2), we obtain improvements on several fronts.

Firstly, we are able in some cases to extend the decoding radius. Namely, instead of $\sqrt{(\log n)/n} \cdot \lambda_1(\mathcal{L})$, we can handle targets at distance of up to $\Omega(\sqrt{\log(1/\varepsilon)}/\eta_\varepsilon(\mathcal{L}^*))$ for $\varepsilon = 1/\mathrm{poly}(n)$, or slightly above the inverse of the smoothing parameter of the dual lattice (see Section 2.3 for the definition). This is never worse and is sometimes significantly better than the bound in LLM. For instance, already for $\mathbb{Z}^n$, we get distance $\Omega(1)$, which is a constant factor of $\lambda_1(\mathcal{L})$, compared with $\sqrt{(\log n)/n}$ in LLM. This improvement is a result of our refined analysis of f, and highlights the fact that $1/\eta_\varepsilon(\mathcal{L}^*)$ is the right measure of the interference between the peaks of f.

A second improvement is in the size of the advice required from preprocessing, which apart from being inherently interesting, is a good proxy for the efficiency of the algorithm. In LLM, the advice consisted of an unspecified polynomial number of dual lattice vectors. In our algorithm, we require only $O(n\log(1/\varepsilon)/\sqrt{\varepsilon})$ dual lattice vectors.

Third, we show that our gradient ascent converges in just two steps (after which we apply a simple rounding procedure) compared to $\mathrm{poly}(n)$ steps for LLM. In both algorithms, the time complexity of each step is $O(n)$ times the number of preprocessing vectors, which is also significantly lower in our algorithm. This fast convergence is due to the fact that a single step of our algorithm reduces the distance to the nearest lattice point by at least a constant factor, starting from any target within the decoding radius, and it reduces this distance by a polynomial factor when the target is closer by a constant factor. In comparison, the LLM algorithm reduces this distance by a factor of only $1 - 1/\mathrm{poly}(n)$.

Finally, we note that our hill-climbing algorithm is quite interesting also in the regime of superpolynomial running time, and provides a smooth tradeoff between running time and decoding radius. For instance, by an appropriate setting of parameters, we obtain for any $\alpha < 1/2$ an algorithm that can handle targets at distance up to $\alpha\lambda_1(\mathcal{L})$ using $N \approx \widetilde{O}(n\exp(\alpha^2 n/(1-2\alpha)^2))$ vectors as advice and runs in time $\widetilde{O}(nN)$. (See Corollary 3.3 for the precise statement.)

Reducing CVP to CVP on close targets. In our second main contribution, appearing in Sections 5 and 6, we show that in order to solve either CVP or CVPP, *it suffices to answer queries on target points that are close to the lattice*.

In Section 5 we focus on the preprocessing setting. We show in Theorem 5.1 that for any non-increasing function $\alpha(n) > 0$, in order to solve $\sqrt{n}/(2\alpha(n))$-approximate CVPP it suffices to answer queries within distance $\alpha\lambda_1(\mathcal{L})$. By combining this reduction with the LLM algorithm (or

our improved algorithm), we immediately obtain an $O(n/\sqrt{\log n})$ approximation algorithm for search CVPP, improving on Lagarias et al.'s algorithm [LLS90]. In terms of techniques, we closely follow Kannan's idea [Kan87] of looking for a projection of the lattice in which the target point is relatively close to the lattice.

By refining the ideas used in Theorem 5.1, we give in Theorem 5.2 a similar reduction with the additional property that it incurs almost no blowup in the amount of preprocessing advice needed. Combining this reduction (as it appears in Corollary 5.3) with our improved BDDP algorithm, we obtain an algorithm for $O(n)$-CVPP that uses only $O(n)$ vectors of advice. This is quite remarkable since $O(n)$ vectors are "not much more" than the n needed to form a basis; and it is an interesting open question whether there exists a basis using which one can obtain even a polynomial approximation for CVPP (see below). Apart from the theoretical interest in minimizing the advice, this might have applications in cryptography or coding theory.

In Section 6, we consider the setting without preprocessing and show for any $\tau = \tau(n) > 0$ and $\gamma = \gamma(n) \geq 1$, a reduction from $\sqrt{1+\tau} \cdot \gamma$-approximate CVP to CVP with the slightly harder approximation factor γ but with a distance bound of $\sqrt{1+\tau^{-1}} \cdot \lambda_1(\mathcal{L})$. We note that this reduction also applies to approximation factors well below $\sqrt{n}$, in contrast to our reduction in the preprocessing setting. Notice that here we also require distance guarantees above $\lambda_1(\mathcal{L})$, while the preprocessing setting can work well below the unique decoding radius of $\lambda_1(\mathcal{L})/2$. When combined with the hardness result of [DKRS03], our reduction shows that approximate CVP is hard even when the target is guaranteed to be close to the lattice. See Corollary 6.2 for the precise hardness result. Our reduction relies on a lattice sparsification technique introduced by Dadush and Kun [DK13] who used it to develop deterministic single-exponential time algorithms for approximate CVP under general norms.

1.2 Open Questions and Discussion

The main open question is whether one can improve our $O(n/\sqrt{\log n})$ approximation factor of search CVPP, and possibly match the best known approximation factor $O(\sqrt{n/\log n})$ for the decision version.

Another open problem is to provide a deeper understanding of the computational complexity of BDD both with and without preprocessing. Liu et al. [LLM06] showed that $\frac{1}{\sqrt{2}}$-BDD is NP-hard in the non-preprocessing version, however nothing is known for smaller distance bounds. This is in contrast to the situation for CVPP and SVP, where *any* constant factor approximation is NP-hard. A natural question is therefore: is α-BDD NP-hard for any constant α?

An open question already mentioned briefly above is whether there exists a *basis* that one can use to obtain a polynomial approximation for CVPP. A natural approach is to use Babai's algorithm (see Section 2.6), whose approximation factor can be shown to be

$$\max_{1 \leq i \leq n} \frac{\sqrt{\sum_{j=1}^{i} \|\tilde{\mathbf{b}}_j\|^2}}{\|\tilde{\mathbf{b}}_i\|},$$

where the $\tilde{\mathbf{b}}_i$ are the Gram-Schmidt orthogonalization of the given basis. The open question, once specialized to Babai's algorithm, is therefore equivalent to asking whether every lattice has a basis with $\max_{i \leq j} \|\tilde{\mathbf{b}}_i\| / \|\tilde{\mathbf{b}}_j\| < \text{poly}(n)$. The best known upper bound is $n^{O(\log n)}$ [LLS90] using a Korkine-Zolotarev basis.

Finally, we note that our reductions in Section 5 and Section 6 are to the *approximate* bounded-distance problem. That is, we are guaranteed to be close to the lattice and are required to output a nearby lattice point, but *not necessarily the closest*. In contrast, the LLM algorithm and our improvement both have the property that they actually output the closest lattice point, an apparently harder problem. So, we are seemingly unable to use the full strength of our reduction. This leads to the following intuitive question: can the presence of one very close lattice point help in finding a *different* relatively close lattice point? Alternatively, is there a reduction from the approximate distance-bounded problem to its exact version? The current gap between the search and decision versions of CVPP seems to suggest that being very close to the lattice may provide useful information that is still insufficient to find the nearest vector.

2 Preliminaries

2.1 Lattices

A rank d lattice $\mathcal{L} \subset \mathbb{R}^n$ is the set of all integer linear combinations of d linearly independent vectors $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_d)$. $\mathbf{B}$ is called a basis of the lattice and is not unique. We sometimes write $\mathcal{L}(\mathbf{B})$ to signify the lattice generated by $\mathbf{B}$. The length of the shortest nonzero vector, also known as the first successive minimum, is denoted by $\lambda_1(\mathcal{L}) := \min_{\mathbf{0} \neq \mathbf{x} \in \mathcal{L}} \|\mathbf{x}\|$.

For any point $\mathbf{x} \in \mathbb{R}^n$, we define $\mathrm{dist}(\mathbf{x}, \mathcal{L})$ as the minimum of $\|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{y} \in \mathcal{L}$. The covering radius $\mu(\mathcal{L})$ is the supremum of $\mathrm{dist}(\mathbf{x}, \mathcal{L})$ for all $\mathbf{x} \in \mathrm{span}(\mathcal{L})$.

For any lattice $\mathcal{L}$, the dual lattice, denoted $\mathcal{L}^*$, is defined as the set of all points in $\mathrm{span}(\mathcal{L})$ that have integer inner products with all lattice points,

$$\mathcal{L}^* = \{\mathbf{w} \in \mathrm{span}(\mathcal{L}) : \forall \mathbf{y} \in \mathcal{L}, \langle \mathbf{w}, \mathbf{y} \rangle \in \mathbb{Z}\} .$$

Similarly, for a lattice basis $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_d)$, we define the dual basis $\mathbf{B}^* = (\mathbf{b}_1^*, \ldots, \mathbf{b}_d^*)$ to be the unique set of vectors in $\mathrm{span}(\mathcal{L})$ satisfying $\langle \mathbf{b}_i^*, \mathbf{b}_j \rangle = \delta_{i,j}$. It is easy to show that $\mathcal{L}^*$ is itself a rank d lattice and $\mathbf{B}^*$ is a basis of $\mathcal{L}^*$.

In what follows, we typically consider only lattices $\mathcal{L} \subset \mathbb{R}^n$ whose rank is n (lattices of full rank). We note that all of our results apply to more general lattices, as we can simply think of the lattice as embedded in $\mathrm{span}(\mathcal{L})$. We sometimes make use of this fact implicitly.

The following technical lemma gives rough bounds on lattice parameters in terms of representation size.

Lemma 2.1. *Let $\mathcal{L} \subset \mathbb{Q}^n$ be a lattice with basis $\mathbf{B}$. Let ℓ be the bit length of $\mathbf{B}$ in a standard binary representation. Then, $\mu(\mathcal{L}) \leq 2^{O(\ell)}$ and $1/\lambda_1(\mathcal{L}) \leq 2^{O(\ell)}$.*

Proof. Let $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Then clearly

$$\mu(\mathcal{L}) \leq \sum_i \|\mathbf{b}_i\| \leq 2^{O(\ell)} .$$

Similarly, if $\mathbf{b}_i = (p_{i,1}/q_{i,1}, \ldots, p_{i,n}/q_{i,n})$, then any integer linear combination of the $\mathbf{b}_i$ must be expressible as p/q where $q = \prod q_{i,j} \leq 2^{O(\ell)}$. Therefore, $1/\lambda_1(\mathcal{L}) \leq q \leq 2^{O(\ell)}$. □

Given a basis, $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$, we define its Gram-Schmidt orthogonalization $(\widetilde{\mathbf{b}}_1, \ldots, \widetilde{\mathbf{b}}_n)$ by

$$\widetilde{\mathbf{b}}_i = \pi_{\{b_1, \ldots, b_{i-1}\}^\perp}(\mathbf{b}_i) ,$$

and the Gram-Schmidt coefficients $\mu_{i,j}$ by

$$\mu_{i,j} = \frac{\langle \mathbf{b}_i, \widetilde{\mathbf{b}}_j \rangle}{\|\widetilde{\mathbf{b}}_j\|^2} .$$

Here, π_A is the orthogonal projection on the subspace A and $\{b_1, \ldots, b_{i-1}\}^\perp$ denotes the subspace orthogonal to $b_1, \ldots, b_{i-1}$.

Definition 2.2. *A basis $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of $\mathcal{L}$ is a Hermite-Korkin-Zolotarev (HKZ) basis if*

1. $\|\mathbf{b}_1\| = \lambda_1(\mathcal{L})$;
2. *the Gram-Schmidt coefficients of $\mathbf{B}$ satisfy $|\mu_{i,j}| \leq \frac{1}{2}$ for all $j < i$; and*
3. $\pi_{\{\mathbf{b}_1\}^\perp}(\mathbf{b}_2), \ldots, \pi_{\{\mathbf{b}_1\}^\perp}(\mathbf{b}_n)$ *is an HKZ basis of $\pi_{\{\mathbf{b}_1\}^\perp}(\mathcal{L})$.*

2.2 Lattice Problems

Definition 2.3. *For any approximation parameter $\gamma = \gamma(n) \geq 1$, the search problem γ-SVP (Shortest Vector Problem) is defined as follows: The input is a basis $\mathbf{B}$ for a lattice $\mathcal{L} \subset \mathbb{R}^n$. The goal is to output a vector $\mathbf{y} \in \mathcal{L}$ satisfying $\|\mathbf{y}\| \leq \gamma \cdot \lambda_1(\mathcal{L})$.*

Definition 2.4. *For any approximation parameter $\gamma = \gamma(n) \geq 1$, the search problem γ-CVP (Closest Vector Problem) is defined as follows: The input is a basis $\mathbf{B}$ for a lattice $\mathcal{L} \subset \mathbb{R}^n$ and a vector $\mathbf{t} \in \mathbb{R}^n$, the target. The goal is to output a vector $\mathbf{y} \in \mathcal{L}$ satisfying $\|\mathbf{t} - \mathbf{y}\| \leq \gamma \cdot \mathrm{dist}(\mathbf{t}, \mathcal{L})$.*

We often ignore the basis and simply refer to $\mathcal{L}$ and $\mathbf{t}$ as the input.

Definition 2.5. *The decision problems γ-GapCVP is the decision analogue of γ-CVP, defined as follows: The input is a basis $\mathbf{B}$ of a lattice $\mathcal{L} \subset \mathbb{R}^n$ and a target vector $\mathbf{t} \in \mathbb{R}^n$. It is a YES instance if $\mathrm{dist}(\mathbf{t}, \mathcal{L}) \leq 1$. It is a NO instance if $\mathrm{dist}(\mathbf{t}, \mathcal{L}) > \gamma$.*

Dinur et al. [DKRS03] showed the current best known hardness result for γ-GapCVP, which of course immediately implies a hardness result for γ-CVP.

Theorem 2.6 ([DKRS03])**.** *There is some constant $c > 0$ such that γ-GapCVP (and therefore γ-CVP) is NP-hard for $\gamma = n^{c/\log\log n}$.*

Definition 2.7. *Let ϕ be a positive-valued function on lattices and $\gamma(n) \geq 1$. Then, γ-CVP$^\phi$ is the problem of solving γ-CVP when the input lattice $\mathcal{L}$ and target point $\mathbf{t}$ satisfy $\mathrm{dist}(\mathbf{t}, \mathcal{L}) < \phi(\mathcal{L})$. If the target point is outside of this range, any output is acceptable.*

We note that the standard reduction from γ-SVP to γ-CVP (see, for example, [MG02]) is actually a reduction from γ-SVP to γ-CVP$^\phi$ where $\phi(\mathcal{L}) = \lambda_1(\mathcal{L})$.

Theorem 2.8. *There is a polynomial-time reduction from γ-SVP to γ-CVP$^\phi$ where $\phi(\mathcal{L}) = \lambda_1(\mathcal{L})$ for any lattice $\mathcal{L}$ and $\gamma = \gamma(n) \geq 1$.*

Definition 2.9. *An algorithm with preprocessing consists of two phases. The first phase, called the preprocessing algorithm, takes input P and outputs an advice string A. The second phase, called the query algorithm, takes input A and Q, the query, and outputs a solution S. We say that such an algorithm runs in polynomial time if the advice A is polynomial in the length of P and the query algorithm runs in time polynomial in the lengths of P and Q. The preprocessing algorithm may take arbitrary time.*

Definition 2.10. *The search problems γ-CVPP and γ-CVPP$^{\phi}$ (Closest Vector Problem with Preprocessing) are the preprocessing analogues of γ-CVP and γ-CVP$^{\phi}$ respectively, defined as follows: The input to preprocessing is a basis $\mathbf{B}$ of a lattice $\mathcal{L} \subset \mathbb{R}^n$. The input to the query phase is a vector $\mathbf{t} \in \mathbb{R}^n$. The goal is to return a valid solution to γ-CVP or γ-CVP$^{\phi}$ respectively.*

Definition 2.11. *For any approximation parameter $\alpha = \alpha(n)$, the search problem with preprocessing α-BDDP (Bounded Distance Decoding) is simply 1-CVPP$^{\phi}$ where $\phi(\mathcal{L}) = \alpha \cdot \lambda_1(\mathcal{L})$ for any lattice $\mathcal{L}$.*

2.3 The Discrete Gaussian and the Smoothing Parameter

For any $s > 0$, we define the function $\rho_s : \mathbb{R}^n \to \mathbb{R}$ as $\rho_s(\mathbf{t}) = \exp(-\pi\|\mathbf{t}\|^2/s^2)$. When $s = 1$, we simply write $\rho(\mathbf{t})$. For a set A we define $\rho_s(A) = \sum_{\mathbf{x}\in A} \rho_s(\mathbf{x})$.

Definition 2.12. *For a lattice $\mathcal{L} \subset \mathbb{R}^n$ and a vector $\mathbf{t} \in \mathbb{R}^n$, let $D_{\mathcal{L}+\mathbf{t},s}$ be the probability distribution over $\mathcal{L} + \mathbf{t}$ such that the probability of drawing $\mathbf{x} \in \mathcal{L} + \mathbf{t}$ is proportional to $\rho_s(\mathbf{x})$. We call this the discrete Gaussian distribution over $\mathcal{L} + \mathbf{t}$ with parameter s.*

For any lattice $\mathcal{L} \subset \mathbb{R}^n$ and $\mathbf{t} \in \mathbb{R}^n$, let

$$f(\mathbf{t}) = f_{\mathcal{L}}(\mathbf{t}) = \rho(\mathcal{L} + \mathbf{t})/\rho(\mathcal{L}) \,. \tag{5}$$

Banaszczyk proved the following two lemmas in [Ban93]. We include proofs for completeness.

Lemma 2.13. *For an n-dimensional lattice $\mathcal{L}$, shift $\mathbf{c} \in \mathbb{R}^n$, and any $t \geq 1$,*

$$\Pr_{\mathbf{y}\sim D_{\mathcal{L}+\mathbf{c}}}\left[\|\mathbf{y}\| \geq t\sqrt{\frac{n}{2\pi}}\right] \leq \frac{\rho(\mathcal{L})}{\rho(\mathcal{L}+\mathbf{c})} e^{-\frac{n}{2}(t^2 - 2\log t - 1)} \leq \frac{\rho(\mathcal{L})}{\rho(\mathcal{L}+\mathbf{c})} e^{-\frac{n}{2}(t-1)^2} \,.$$

Proof. For any $0 < \alpha < 1$, we have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathbf{y}\sim D_{\mathcal{L}+\mathbf{c}}}[e^{\pi\alpha\|\mathbf{y}\|^2}] &= \frac{\rho(\mathcal{L})}{\rho(\mathcal{L}+\mathbf{c})} \frac{\rho_{1/(\sqrt{1-\alpha})}(\mathcal{L}+\mathbf{c})}{\rho(\mathcal{L})} \\
&= \frac{\rho(\mathcal{L})}{\rho(\mathcal{L}+\mathbf{c})} \Big(\frac{1}{\sqrt{1-\alpha}}\Big)^n \frac{\sum_{\mathbf{y}\in\mathcal{L}^*} e^{2\pi i\langle \mathbf{y},\mathbf{c}\rangle} \rho_{\sqrt{1-\alpha}}(\mathbf{y})}{\rho(\mathcal{L}^*)} \quad \text{(Poisson summation formula)} \\
&\leq \frac{\rho(\mathcal{L})}{\rho(\mathcal{L}+\mathbf{c})} \Big(\frac{1}{\sqrt{1-\alpha}}\Big)^n \frac{\rho_{\sqrt{1-\alpha}}(\mathcal{L}^*)}{\rho(\mathcal{L}^*)} \\
&\leq \frac{\rho(\mathcal{L})}{\rho(\mathcal{L}+\mathbf{c})} \Big(\frac{1}{\sqrt{1-\alpha}}\Big)^n \,.
\end{aligned}$$

Using the above and Markov's inequality, we have that

$$\begin{aligned}
\Pr_{\mathbf{y}\sim D_{\mathcal{L}+\mathbf{c}}}\left[\|\mathbf{y}\| \geq t\sqrt{\frac{n}{2\pi}}\right] &= \Pr\left[e^{\pi\alpha\|\mathbf{y}\|^2} \geq e^{\alpha n t^2/2}\right] \\
&\leq \frac{\rho(\mathcal{L})}{\rho(\mathcal{L}+\mathbf{c})} \frac{(1/\sqrt{1-\alpha})^n}{e^{\alpha n t^2/2}} \\
&= \frac{\rho(\mathcal{L})}{\rho(\mathcal{L}+\mathbf{c})} e^{-\frac{n}{2}(\alpha t^2 + \log(1-\alpha))} \,.
\end{aligned}$$

The first bound now follows by setting $\alpha = 1 - 1/t^2$.

For the simplified bound, using the fact that $0 \le \log t \le t - 1$, for $t \ge 1$, we get that

$$e^{-\frac{n}{2}(t^2 - 2\log t - 1)} \le e^{-\frac{n}{2}(t^2 - 2(t-1) - 1)} = e^{-\frac{n}{2}(t-1)^2},$$

as needed. □

Lemma 2.14. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice of rank n. Then, for all $\mathbf{t} \in \mathbb{R}^n$, $f(\mathbf{t}) \ge \rho(\mathbf{t})$.*

Proof.

$$\rho(\mathcal{L} + \mathbf{t}) = \rho(\mathbf{t}) \sum_{\mathbf{y} \in \mathcal{L}} \cosh(2\pi \langle \mathbf{y}, \mathbf{t} \rangle) \rho(\mathbf{y}) \ge \rho(\mathbf{t}) \rho(\mathcal{L}) \; .$$ □

Definition 2.15. *For $\varepsilon > 0$ and $\mathcal{L} \subset \mathbb{R}^n$ a lattice, we define the smoothing parameter $\eta_\varepsilon(\mathcal{L})$ as the unique value satisfying $\rho_{1/\eta_\varepsilon(\mathcal{L})}(\mathcal{L}^* \setminus \{\mathbf{0}\}) = \varepsilon$.*

The name smoothing parameter comes from the fact that, for $s \ge \eta_\varepsilon(\mathcal{L})$, $\rho_s(\mathcal{L} + \mathbf{t})$ varies by at most a multiplicative factor of $(1 \pm \varepsilon)$ [Reg09].

2.4 Behavior of $\sqrt{\log(1/\varepsilon)}/\eta_\varepsilon(\mathcal{L}^*)$

The function $g(\varepsilon) = \sqrt{\log(1/\varepsilon)}/\eta_\varepsilon(\mathcal{L}^*)$ is quite important for our BDDP algorithm, so we analyze its behavior here. Our first lemma shows that $g(\varepsilon)$ is strictly monotonically decreasing as ε increases. (This is not obvious since both the numerator and the denominator are monotonically decreasing.) It is a simple modification of [CDLP13, Lemma 2.4].

Lemma 2.16. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice of rank at least one. Let $g(\varepsilon) = \sqrt{\log(1/\varepsilon)}/\eta_\varepsilon(\mathcal{L}^*)$ for any $\varepsilon \in (0,1)$. Then, $g(\varepsilon)$ is strictly monotonically decreasing.*

Proof. Our goal is to prove that for any $\varepsilon \in (0,1)$, $r > 1$, $g(\varepsilon/r) > g(\varepsilon)$, or equivalently, that $\eta_{\varepsilon/r}(\mathcal{L}^*) < \eta_\varepsilon(\mathcal{L}^*) \cdot t$ where

$$t = \frac{\sqrt{\log(r/\varepsilon)}}{\sqrt{\log(1/\varepsilon)}} > 1 \; .$$

This follows from

$$\sum_{\mathbf{y} \in \mathcal{L} \setminus \{\mathbf{0}\}} (e^{-\pi \eta_\varepsilon(\mathcal{L}^*)^2 \|\mathbf{y}\|^2})^{t^2} < \Big(\sum_{\mathbf{y} \in \mathcal{L} \setminus \{\mathbf{0}\}} e^{-\pi \eta_\varepsilon(\mathcal{L}^*)^2 \|\mathbf{y}\|^2} \Big)^{t^2} = \varepsilon^{t^2} = \varepsilon/r \; .$$

□

The next lemma and its corollary show the relationship between $g(\varepsilon)$ and $\lambda_1(\mathcal{L})$. Similar analysis appears in [MR07].

Lemma 2.17. *Let $\mathcal{L} \subseteq \mathbb{R}^n$ be an n-dimensional lattice. Then, for $\varepsilon \in (0,1)$,*

$$\frac{\sqrt{\log(2/\varepsilon)/\pi}}{\lambda_1(\mathcal{L})} \le \eta_\varepsilon(\mathcal{L}^*) \le \frac{\sqrt{\log((1+\varepsilon)/\varepsilon)/\pi} + \sqrt{n/(2\pi)}}{\lambda_1(\mathcal{L})} \; .$$

Proof. For the lower bound, we note that for $s \leq \frac{\sqrt{\log(2/\varepsilon)/\pi}}{\lambda_1(\mathcal{L})}$, we have that

$$\rho_{1/s}(\mathcal{L} \setminus \{\mathbf{0}\}) > 2e^{-\pi(s\lambda_1(\mathcal{L}))^2} \geq \varepsilon,$$

as needed. For the upper bound, we note that $\eta_\varepsilon(\mathcal{L}^*) \leq s$ if and only if $\Pr_{\mathbf{y}\sim D_{\mathcal{L},1/s}}[\mathbf{y} \neq \mathbf{0}] \leq \frac{\varepsilon}{1+\varepsilon}$. By Lemma 2.13, letting $s = t\sqrt{n/(2\pi)}/\lambda_1(\mathcal{L})$, for $t \geq 1$, we have that

$$\Pr_{\mathbf{x}\sim D_{\mathcal{L},1/s}}[\mathbf{y} \neq \mathbf{0}] = \Pr_{\mathbf{y}\sim D_{\mathcal{L},1/s}}[\|\mathbf{x}\| \geq \lambda_1(\mathcal{L})] = \Pr_{\mathbf{y}\sim D_{\mathcal{L}}}[\|\mathbf{y}\| \geq t \cdot \sqrt{n/2\pi}] \leq e^{-\frac{n}{2}(t-1)^2}.$$

Setting $t = \sqrt{2\log((1+\varepsilon)/\varepsilon)/n} + 1$, we get that $\Pr_{\mathbf{y}\sim D_{\mathcal{L},1/s}}[\mathbf{y} \neq \mathbf{0}] \leq \frac{\varepsilon}{1+\varepsilon}$. Therefore

$$\eta_\varepsilon(\mathcal{L}^*) \leq t\frac{\sqrt{n/(2\pi)}}{\lambda_1(\mathcal{L})} = \frac{\sqrt{\log((1+\varepsilon)/\varepsilon)/\pi} + \sqrt{n/(2\pi)}}{\lambda_1(\mathcal{L})},$$

as needed. □

Corollary 2.18. *Let $\mathcal{L} \subseteq \mathbb{R}^n$ be an n-dimensional lattice. Then, for $\varepsilon \in (0,1)$,*

$$\frac{\sqrt{\log(2/\varepsilon)/\pi}}{\eta_\varepsilon(\mathcal{L}^*)} \leq \lambda_1(\mathcal{L}) \leq \frac{\sqrt{\log(2/\varepsilon)/\pi}}{\eta_\varepsilon(\mathcal{L}^*)}\left(1 + \frac{\sqrt{n/2}}{\sqrt{\log(2/\varepsilon)}}\right).$$

2.5 Tail Bounds

We next introduce subgaussian and subexponential random variables, and in particular, the subgaussianity of $D_{\mathcal{L},s}$.

Definition 2.19. *We say that a random variable $\mathbf{X}$ (or its distribution) over $\mathbb{R}^n$ is subgaussian with parameter $s > 0$ if $\mathbb{E}[\mathbf{X}] = \mathbf{0}$, and for all $t \in \mathbb{R}$ and all unit vectors $\mathbf{v} \in \mathbb{R}^n$,*

$$\Pr[|\langle \mathbf{X}, \mathbf{v}\rangle| \geq t] \leq 2 \cdot e^{-\pi t^2/s^2}.$$

Lemma 2.20 ([MP12, Lemma 2.8])**.** *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice of rank n. Then for any $s > 0$, $D_{\mathcal{L},s}$ is subgaussian with parameter s.*

Definition 2.21. *We say that a random variable X (or its distribution) over $\mathbb{R}$ is subexponential with parameter s if, for any $t > 0$*

$$\Pr[|X| \geq t] \leq e^{1-t/s}.$$

Vershynin proved a basic relationship between subgaussian and subexponential random variables, from which we derive a simple corollary.

Lemma 2.22 ([Ver12, Lemma 5.14])**.** *If $\mathbf{X}$ is a subgaussian random variable over $\mathbb{R}^n$ with parameter s, then for any unit vector $\mathbf{v} \in \mathbb{R}^n$, $\langle \mathbf{X}, \mathbf{v}\rangle^2$ is subexponential with parameter $O(s)$.*

Corollary 2.23. *If $\mathbf{X}$ and $\mathbf{Y}$ are subgaussian random variables over $\mathbb{R}^n$ with parameter s, then for any two unit vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\langle \mathbf{X}, \mathbf{u}\rangle\langle \mathbf{Y}, \mathbf{v}\rangle$ is subexponential with parameter $O(s)$*

Proof. It follows immediately from the definitions that subgaussian random variables with parameter $O(s)$ are closed under addition and multiplication by constants, as are subexponential random variables with parameter $O(s)$. Therefore,

$$\langle \mathbf{X}, \mathbf{u} \rangle \langle \mathbf{Y}, \mathbf{v} \rangle = \frac{1}{2}(\langle \mathbf{X}, \mathbf{u} \rangle + \langle \mathbf{Y}, \mathbf{v} \rangle)^2 - \frac{1}{2}\langle \mathbf{X}, \mathbf{u} \rangle^2 - \frac{1}{2}\langle \mathbf{Y}, \mathbf{v} \rangle^2 .$$

is subexponential with parameter $O(s)$ as claimed. □

Vershynin showed the next useful property of subexponential random variables.

Lemma 2.24 ([Ver12, Proposition 5.16])**.** *Let $X_1, \ldots, X_N$ be independent subexponential random variables over $\mathbb{R}$ with parameter s, and suppose $\mathbb{E}[X_i] = 0$ for all i. Then, for any $t \geq 0$,*

$$\Pr\left[\frac{1}{N}|\sum_i X_i| \geq t\right] \leq 2^{1-\Omega(N \min(t/s, t^2/s^2))} .$$

We will also need the Chernoff-Hoeffding bound [Hoe63].

Lemma 2.25 (Chernoff-Hoeffding bound)**.** *Let $X_1, \ldots, X_N$ be independent and identically distributed random variables with $-a \leq X_i \leq a$. Then, for $s > 0$*

$$\Pr\left[\left|\mathbb{E}[X_i] - \frac{1}{N} \cdot \sum X_i\right| \geq s\right] \leq 2^{1-\Omega(Ns^2/a^2)} .$$

2.6 Babai's Nearest Plane Algorithm

Babai's nearest plane algorithm (denoted BABAI) is an algorithm introduced by Babai [Bab86] for rounding a target vector to a nearby lattice point one coordinate at a time. The input is a basis $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ for a lattice $\mathcal{L}$ and a target $\mathbf{t} \in \mathbb{R}^n$.

We first project $\mathbf{t}$ onto $\mathrm{span}(\mathcal{L})$. We then choose the last coordinate $c_n \in \mathbb{Z}$ of our nearby lattice point by simple rounding, setting

$$c_n = \lfloor \langle \mathbf{t}, \mathbf{b}_n^* \rangle \rceil .$$

Next we call BABAI recursively on $(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1})$ and $\mathbf{t} - c_n\mathbf{b}_n$ and receive the result $\mathbf{y}$. We then return $\mathbf{y} + c_n\mathbf{b}_n$.

Stated more intuitively, BABAI chooses the lattice hyperplane

$$c_n\mathbf{b}_n + \mathrm{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}) = \{\mathbf{x} \in \mathrm{span}(\mathcal{L}) : \langle \mathbf{x}, \mathbf{b}_n^* \rangle = c_n\}$$

with $c_n \in \mathbb{Z}$ that is nearest to the target and recurses on this hyperplane.

Babai proved the following standard fact about his algorithm.

Lemma 2.26 ([Bab86])**.** *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice of rank n. For any basis, $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of $\mathcal{L}$ with Gram-Schmidt orthogonalization $(\widetilde{\mathbf{b}}_1, \ldots, \widetilde{\mathbf{b}}_n)$ and any target vector $\mathbf{t} \in \mathbb{R}^n$, $\mathrm{BABAI}(\mathbf{t}, \mathbf{B})$ outputs $\mathbf{y} \in \mathcal{L}$ satisfying*

$$\|\mathbf{y} - \mathbf{t}\|^2 \leq \frac{1}{4}\sum_{i=1}^{n}\|\widetilde{\mathbf{b}}_i\|^2 \leq \frac{n}{4} \cdot \max_i \|\widetilde{\mathbf{b}}_i\|^2 .$$

2.7 δ-Nets and the Spectral Norm

Definition 2.27. *For any $\delta > 0$, $A \subset \mathbb{R}^n$ is a δ-net of S if $A \subseteq S$, and for each $\mathbf{v} \in S$, there is some $\mathbf{u} \in A$ such that $\|\mathbf{u} - \mathbf{v}\| \leq \delta$.*

We'll be interested in the case when S is a ball, a sphere, or a shell. The next lemma shows that we can do this without many points. The proof is by a standard packing argument. (See Lemma 5.2 of [Ver12], for example.)

Lemma 2.28. *For any $\delta > 0$, there exists a δ-net of the unit ball in $\mathbb{R}^n$ with $(1 + 2/\delta)^n$ points. Nets of the same cardinality exist for spherical shells of outer radius one, and for the unit sphere.*

A δ-net of the unit sphere can be used to accurately approximate the length of any vector.

Lemma 2.29. *Let $\delta \in (0, 1)$, and let A be a δ-net of the unit sphere in $\mathbb{R}^n$. Then, for any $\mathbf{x} \in \mathbb{R}^n$,*

$$\max_{\mathbf{v} \in A} |\langle \mathbf{v}, \mathbf{x} \rangle| \leq \|\mathbf{x}\| \leq \frac{1}{1 - \delta} \cdot \max_{\mathbf{v} \in A} |\langle \mathbf{v}, \mathbf{x} \rangle|.$$

Proof. Without loss of generality, assume $\|\mathbf{x}\| = 1$. The first inequality is trivial. By hypothesis, there is some $\mathbf{v} \in A$ such that $\|\mathbf{v} - \mathbf{x}\| \leq \delta$. Then,

$$\langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{v} - \mathbf{x}, \mathbf{x} \rangle \geq 1 - \delta \, .$$

The result follows. □

Similarly, a δ-net can be used to approximate the spectral norm of a matrix, as defined below.

Definition 2.30. *For a matrix $M \in \mathbb{R}^{n \times n}$, the spectral norm of M is defined as*

$$\|M\| := \sup_{\|\mathbf{x}\| = 1} \|M\mathbf{x}\| \, .$$

For a symmetric matrix M, $\|M\|$ is equivalently the largest absolute value of an eigenvalue of M.

Lemma 2.31 ([Ver12, Lemma 5.4])**.** *For a symmetric matrix $M \in \mathbb{R}^{n \times n}$ and a δ-net of the unit sphere A with $0 < \delta < 1/2$,*

$$\|M\| \leq \frac{1}{1 - 2\delta} \cdot \max_{\mathbf{x} \in A} |\langle M\mathbf{x}, \mathbf{x} \rangle| \, .$$

3 Exact CVPP with a Promise

In this section we prove the following theorem, which gives an efficient solution to CVPP for points within distance essentially $\sqrt{\log(2/\varepsilon)/\pi}/(2\eta_\varepsilon(\mathcal{L}^*))$. By Corollary 2.18, for $\varepsilon = 1/\mathrm{poly}(n)$ this radius is at least as large as the radius $\sqrt{(\log n)/n} \cdot \lambda_1(\mathcal{L})$ achieved by [LLM06], and moreover, as ε goes to zero, it converges to the unique decoding radius $\lambda_1(\mathcal{L})/2$. Also, by Lemma 2.16, this radius is (essentially) increasing as ε decreases, and thus our algorithm solves a harder problem for smaller ε.

Theorem 3.1. *Let $\varepsilon \in (0, 1/200)$ and $\phi(\mathcal{L}) = \delta_{\mathsf{max}} s_\varepsilon / \eta_\varepsilon(\mathcal{L}^*)$ where $s_\varepsilon = \left(\frac{1}{\pi} \log \frac{2(1+\varepsilon)}{\varepsilon}\right)^{1/2}$ and $\delta_{\mathsf{max}} = \frac{1}{2} - \frac{2}{\pi s_\varepsilon^2}$. Then, there exists an algorithm that solves 1-CVPP$^\phi$ using $O(nN(1 + \log n / \log(1/\varepsilon)) + n^\omega)$ arithmetic operations, where $N = O(n \log(1/\varepsilon)/\sqrt{\varepsilon})$ and n^ω is the number of arithmetic operations needed to compute the inverse of an $n \times n$ matrix. Moreover, the preprocessing consists of N vectors sampled from $D_{\mathcal{L}^*, \eta_\varepsilon(\mathcal{L}^*)}$.*

We note that we can achieve a run-time of $O(nN(1 + \log n / \log(1/\varepsilon)))$ arithmetic operations by computing the inverse of a matrix as part of the preprocessing.

Our result will follow easily from a proposition about f_W, whose proof is in Section 4.

Proposition 3.2. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice with $\rho(\mathcal{L}) = 1 + \varepsilon$ with $\varepsilon \in (0, 1/200)$. Let $s_\varepsilon = \left(\frac{1}{\pi} \log \frac{2(1+\varepsilon)}{\varepsilon}\right)^{1/2}$, $\delta_{\mathsf{max}} = \frac{1}{2} - \frac{2}{\pi s_\varepsilon^2}$, and $\delta(\mathbf{t}) = \max\{\frac{1}{8}, \frac{\|\mathbf{t}\|}{s_\varepsilon}\}$. Let $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ be sampled independently from $D_{\mathcal{L}^*}$. If $N = \Omega(n \log(1/\varepsilon)/\sqrt{\varepsilon})$, then with probability at least $1 - 2^{-\Omega(n)}$,*

$$\left\| \frac{\nabla f_W(\mathbf{t})}{2\pi f_W(\mathbf{t})} + \mathbf{t} \right\| \leq \varepsilon^{(1-2\delta(\mathbf{t}))/4} \|\mathbf{t}\| \tag{6}$$

holds simultaneously for all $\mathbf{t} \in \mathbb{R}^n$ with $\|\mathbf{t}\| \leq \delta_{\mathsf{max}} s_\varepsilon$.

We note that for $\varepsilon < 1/200$, $\varepsilon^{(1-2\delta_{\mathsf{max}})/4} = e^{-\log(1/\varepsilon)/\log(2(1+\varepsilon)/\varepsilon)} \leq 1/2$, so the right hand side of (6) is at most $\|\mathbf{t}\|/2$.

Proof of Theorem 3.1. We present an algorithm with probabilistic preprocessing and argue that with positive probability the preprocessing algorithm will output advice that results in a query algorithm that is successful on all relevant inputs. Clearly this implies a deterministic algorithm.

The preprocessing algorithm takes as input a lattice $\mathcal{L} \subset \mathbb{R}^n$ of rank n. It returns as advice a sequence of samples $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ from $D_{\mathcal{L}^*, \eta_\varepsilon(\mathcal{L}^*)}$ where $N = O(n \log(1/\varepsilon)/\sqrt{\varepsilon})$ is large to satisfy Proposition 3.2.

The query algorithm takes a target point $\mathbf{t} \in \mathbb{R}^n$ and advice W from preprocessing. It then iteratively updates $\mathbf{t} \leftarrow \mathbf{t} + \nabla f_W(\mathbf{t})/(2\pi f_W(\mathbf{t}))$ a total of $1 + \lceil 8 \log(\sqrt{n} s_\varepsilon) / \log(1/\varepsilon) \rceil$ times. It then scans W. Let $V^* = (\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*) \subset W$ be the first n linearly independent vectors it finds of length bounded by $\sqrt{n}\eta_\varepsilon(\mathcal{L}^*)$ (it aborts if no such vectors exist). The algorithm computes $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ satisfying $\langle \mathbf{v}_i^*, \mathbf{v}_j \rangle = \delta_{i,j}$ and returns $\sum c_i \mathbf{v}_i$ for $c_i = \lfloor \langle \mathbf{v}_i^*, \mathbf{t} \rangle \rceil$.

By scaling the lattice appropriately, we can assume without loss of generality that $\rho(\mathcal{L}) = 1 + \varepsilon$ so that $\eta_\varepsilon(\mathcal{L}^*) = 1$. Moreover, it suffices to prove correctness for the case when $\mathbf{0}$ is the closest lattice vector to $\mathbf{t}$, and therefore $\|\mathbf{t}\| \leq \delta_{\mathsf{max}} s_\varepsilon$. The reason is that for $\mathbf{y} \in \mathcal{L}$,

$$f_W(\mathbf{t} + \mathbf{y}) = \frac{1}{N} \sum \cos(2\pi \langle \mathbf{w}_i, \mathbf{t} + \mathbf{y} \rangle) = \frac{1}{N} \sum \cos(2\pi \langle \mathbf{w}_i, \mathbf{t} \rangle) = f_W(\mathbf{t}),$$

so $f_W(\mathbf{t})$ is periodic over the lattice, and so is its gradient, and also

$$\sum \lfloor \langle \mathbf{v}_i^*, \mathbf{t} + \mathbf{y} \rangle \rceil \mathbf{v}_i = \sum \lfloor \langle \mathbf{v}_i^*, \mathbf{t} \rangle \rceil \mathbf{v}_i + \sum \langle \mathbf{v}_i^*, \mathbf{y} \rangle \mathbf{v}_i = \mathbf{y} + \sum \lfloor \langle \mathbf{v}_i^*, \mathbf{t} \rangle \rceil \mathbf{v}_i$$

for any $\mathbf{y} \in \mathcal{L}$.

We now argue that with probability $1 - 2^{-\Omega(n)}$ taken over the preprocessing, the query algorithm succeeds in finding the set V^* (and hence also V). Let $W' = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ for $m = O(n)$. By Lemma 2.13, we have that $\Pr[\|\mathbf{w}_i\| > \sqrt{n}] < e^{-\frac{n}{2}(\sqrt{2\pi}-1)^2} \leq e^{-n}$, and hence with probability at

least $1 - me^{-n} = 1 - 2^{-\Omega(n)}$, all vectors in W' are of norm at most $\sqrt{n}$. In order to show that the vectors in W' span $\mathbb{R}^n$ we can, e.g., apply Lemma 4.5 below to W'. We get that for $m = O(n)$ large enough, the Hessian of $f_{W'}$ satisfies

$$\|Hf_{W'}(\mathbf{0}) + 2\pi I_n\| \leq \frac{4\pi\varepsilon}{1+\varepsilon}\Big(\log\frac{2(1+\varepsilon)}{\varepsilon} + 1\Big) + 1 < 2\pi$$

with probability $1 - 2^{-\Omega(n)}$, where we used $\varepsilon < 1/200$. In particular, the matrix $Hf_{W'}(\mathbf{0}) = -4\pi^2/m \sum_{i=1}^m \mathbf{w}_i\mathbf{w}_i^T$ (see Eq. (3)) is invertible, and hence W' spans $\mathbb{R}^n$.

Now assume that W contains such a subset V^* and satisfies the property in Proposition 3.2. By the union bound this happens with probability at least $1 - 2^{-\Omega(n)}$ over the preprocessing. Then using the remark below Proposition 3.2, for any target $\mathbf{t}$ satisfying $\|\mathbf{t}\| \leq \delta_{\mathsf{max}} s_\varepsilon$, the length of $\mathbf{t}$ shrinks by a factor of at least 2 in the first iteration. In each subsequent iteration, $\|\mathbf{t}\| \leq \delta_{\mathsf{max}} s_\varepsilon/2 < s_\varepsilon/4$, and hence the target shrinks by a factor of at least $\varepsilon^{(1-2(1/4))/4} = \varepsilon^{1/8}$. Therefore, after $1 + \lceil 8\log(\sqrt{n}s_\varepsilon)/\log(1/\varepsilon)\rceil$ total iterations, we have $\|\mathbf{t}\| < 1/(2\sqrt{n})$. So, by Cauchy-Schwarz, $|\langle \mathbf{t}, \mathbf{v}_i^*\rangle| < 1/2$ and $\lfloor \langle \mathbf{t}, \mathbf{v}_i^*\rangle \rceil = 0$ for all i. Therefore, $\sum_{i=1}^n \mathbf{v}_i \lfloor \langle \mathbf{v}_i^*, \mathbf{t}\rangle \rceil = \mathbf{0}$, and correctness follows.

The running time consists of $1 + \lceil 8\log(\sqrt{n}s_\varepsilon)/\log(1/\varepsilon)\rceil = 2 + O(\log(n)/\log(1/\varepsilon))$ iterations, each dominated by the computation of $O(N)$ dot products, followed by a matrix inversion. Each dot product takes $O(n)$ arithmetic operations, and the matrix inversion takes n^ω. So, the total running time is $O(nN(1 + \log(n)/\log(1/\varepsilon)) + n^\omega)$ arithmetic operations as claimed. □

We remark that for small enough $\varepsilon < 1/\mathrm{poly}(n)$ ($\varepsilon < n^{-5}$ suffices), the number of iterations of gradient ascent used by the algorithm is only $1 + \lceil 8\frac{\log(\sqrt{n}s_\varepsilon)}{\log(1/\varepsilon)}\rceil = 2$.

Corollary 3.3. *For $\Omega(1/\sqrt{n}) < \alpha < 1/2$, there exists an algorithm that solves α-BDDP with preprocessing consisting of*

$$N = O\Big(\frac{\alpha^2 n^2}{(1-2\alpha)^2} \cdot \exp\Big(\frac{\alpha^2 n}{(1-2\alpha)^2} + \frac{4}{1-2\alpha}\Big)\Big)$$

vectors using $O(nN(1 + \frac{(1-2\alpha)^2\log n}{\alpha^2 n})) = O(nN(1 + \frac{\log n}{\alpha^2 n}))$ arithmetic operations.

Proof. Let ε be given by

$$1/\varepsilon = \frac{1}{2} \cdot \exp\Big(\frac{2\alpha^2 n}{(1-2\alpha)^2} + \frac{8}{1-2\alpha}\Big) - 1 > 200\,,$$

and notice that

$$\pi s_\varepsilon^2 = \log\Big(2 \cdot \frac{1+\varepsilon}{\varepsilon}\Big) = \frac{2\alpha^2 n + 8(1-2\alpha)}{(1-2\alpha)^2}\,.$$

Using Lemma 2.17, the decoding radius given by Theorem 3.1 satisfies

$$\begin{aligned}
\delta_{\mathsf{max}} \cdot \frac{s_\varepsilon}{\eta_\varepsilon(\mathcal{L}^*)} &\geq \frac{\delta_{\mathsf{max}} s_\varepsilon}{s_\varepsilon + \sqrt{n/(2\pi)}} \cdot \lambda_1(\mathcal{L}) \\
&= \frac{\pi s_\varepsilon^2 - 4}{2\pi s_\varepsilon^2 + \sqrt{2\pi n s_\varepsilon^2}} \cdot \lambda_1(\mathcal{L}) \\
&= \frac{2\alpha^2 n + 4 - 16\alpha^2}{4\alpha^2 n + 16(1-2\alpha) + 2(1-2\alpha)\sqrt{\alpha^2 n^2 + 4n(1-2\alpha)}} \cdot \lambda_1(\mathcal{L}) \\
&\geq \frac{2\alpha^2 n + 4 - 16\alpha^2}{4\alpha^2 n + 16(1-2\alpha) + 2(1-2\alpha)(\alpha n + 2(1-2\alpha)/\alpha)} \cdot \lambda_1(\mathcal{L}) \\
&= \alpha \cdot \lambda_1(\mathcal{L}) \,,
\end{aligned}$$

where we have used the inequality $\sqrt{x+y} \leq \sqrt{x} + y/(2\sqrt{x})$ for $x, y > 0$. □

We remark that one can strengthen the bound in Lemma 2.17 using the first bound in Lemma 2.13, and as a result get improved dependence on α in Corollary 3.3 especially for large α. Since the resulting expressions have no nice closed form, we leave the straightforward calculation to the interested reader.

4 Proof of Proposition 3.2

Our goal is to show that $\nabla f_W(\mathbf{t})/(2\pi f_W(\mathbf{t}))$ is close to $-\mathbf{t}$ when $\|\mathbf{t}\|$ is small. We start by showing in Section 4.1 that this is satisfied by the *exact* function f, i.e., that $\nabla f(\mathbf{t})/(2\pi f(\mathbf{t}))$ is close to $-\mathbf{t}$. We also prove several other bounds on f. We then complete the proof in Section 4.2 by arguing that f_W and f are sufficiently close and so are their gradients.

4.1 Three Bounds on the Periodic Gaussian

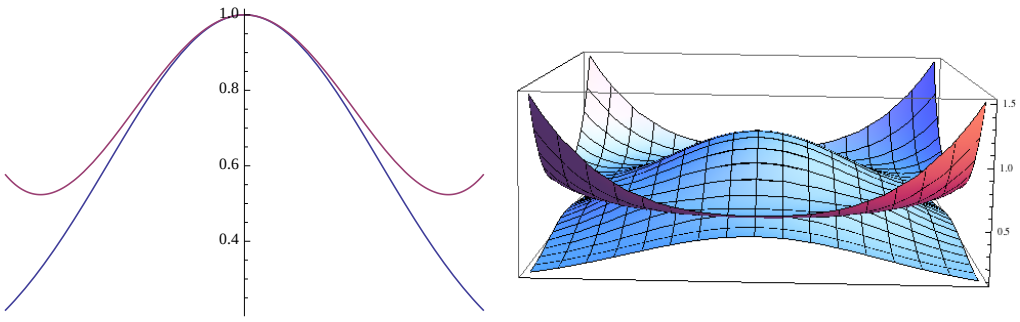
We first give in Lemma 4.1 a general bound (illustrated in Figure 2) on $f(\mathbf{t})$ itself. This will not be used in the sequel and is included here as a warmup and for future reference. We then use a similar idea in Lemma 4.2 to show that $-\nabla f(\mathbf{t})/(2\pi f(\mathbf{t}))$ is close to $\mathbf{t}$, and in Corollary 4.3 bring this bound to a more convenient form. Finally, in Lemma 4.4 we similarly bound the Hessian $Hf(\mathbf{t})$.

Lemma 4.1. *Let $\varepsilon > 0$ and $\mathcal{L} \subset \mathbb{R}^n$ a lattice with $\rho(\mathcal{L}) = 1 + \varepsilon$. Then, for any $\mathbf{t} \in \mathbb{R}^n$,*

$$f(\mathbf{t}) \leq \rho(\mathbf{t})\Big(\frac{1}{1+\varepsilon} + \frac{\varepsilon}{1+\varepsilon} \cdot \cosh(2\pi s_\varepsilon \|\mathbf{t}\|)\Big) + 2\pi\|\mathbf{t}\| \int_{s_\varepsilon - \|\mathbf{t}\|}^{s_\varepsilon + \|\mathbf{t}\|} e^{-\pi z^2} \mathrm{d}z$$

where $s_\varepsilon = \big(\frac{1}{\pi}\log\frac{2(1+\varepsilon)}{\varepsilon}\big)^{1/2}$.

Contrast this with the easy lower bound $f(\mathbf{t}) \geq \rho(\mathbf{t})$ from Lemma 2.14 valid for all lattices and all $\mathbf{t}$.



(a) (b)

Figure 2: $f(\mathbf{t})$ and our bound for $\|\mathbf{t}\| \lesssim s_\varepsilon$ for an example lattice.

Proof. We can write

$$f(\mathbf{t}) = \frac{\rho(\mathcal{L}+\mathbf{t})}{\rho(\mathcal{L})} = \frac{\rho(\mathbf{t})}{\rho(\mathcal{L})} \cdot \sum_{\mathbf{y}\in\mathcal{L}} e^{-2\pi\langle \mathbf{y},\mathbf{t}\rangle}\rho(\mathbf{y}) = \rho(\mathbf{t}) \mathop{\mathbb{E}}_{\mathbf{y}\sim D_{\mathcal{L}}}[\cosh(2\pi\langle \mathbf{y},\mathbf{t}\rangle)]. \tag{7}$$

We now use the fact that for any real-valued random variable X and (sufficiently nice) even function $g : \mathbb{R} \to \mathbb{R}$,

$$\mathop{\mathbb{E}}_{X}[g(X)] = \mathop{\mathbb{E}}_{X}[g(|X|)] = g(0) + \int_0^\infty g'(s) \Pr_X[|X| > s]\mathrm{d}s\ .$$

Therefore, the expectation in Eq. (7) is given by

$$1 + 2\pi\|\mathbf{t}\| \int_{s=0}^{\infty} \Pr\left[|\langle \mathbf{y},\mathbf{t}\rangle| > s\|\mathbf{t}\|\right] \sinh(2\pi s\|\mathbf{t}\|)\mathrm{d}s. \tag{8}$$

We can upper bound the probability using Lemma 2.20 (and noticing that $\mathbf{y}$ is nonzero with probability $\varepsilon/(1+\varepsilon)$) by

$$\Pr[|\langle \mathbf{y},\mathbf{t}\rangle| > s\|\mathbf{t}\|] \le \min\left(\frac{\varepsilon}{1+\varepsilon}, 2e^{-\pi s^2}\right).$$

The minimum is determined by the second term for $s > s_\varepsilon$. We can therefore bound the integral in Eq. (8) from above by the sum of two integrals, the first being

$$\frac{\varepsilon}{1+\varepsilon}\int_{s=0}^{s_\varepsilon} \sinh(2\pi s\|\mathbf{t}\|)\mathrm{d}s = \frac{\varepsilon}{1+\varepsilon}\cdot\frac{\cosh(2\pi s_\varepsilon\|\mathbf{t}\|)-1}{2\pi\|\mathbf{t}\|},$$

and the second being

$$2\int_{s=s_\varepsilon}^{\infty} e^{-\pi s^2}\sinh(2\pi s\|\mathbf{t}\|)\mathrm{d}s = \frac{1}{\rho(\mathbf{t})}\int_{s_\varepsilon-\|\mathbf{t}\|}^{s_\varepsilon+\|\mathbf{t}\|} e^{-\pi z^2}\mathrm{d}z.$$

Putting it all together, we obtain the desired bound

$$\frac{\rho(\mathcal{L}+\mathbf{t})}{\rho(\mathcal{L})} \le \rho(\mathbf{t})\Big(\frac{1}{1+\varepsilon} + \frac{\varepsilon}{1+\varepsilon}\cdot\cosh(2\pi s_\varepsilon\|\mathbf{t}\|) + 2\pi\|\mathbf{t}\|\frac{1}{\rho(\mathbf{t})}\int_{s_\varepsilon-\|\mathbf{t}\|}^{s_\varepsilon+\|\mathbf{t}\|} e^{-\pi z^2}\mathrm{d}z\Big). \qquad \square$$

Lemma 4.2. *Let $\varepsilon > 0$ and $\mathcal{L} \subset \mathbb{R}^n$ a lattice with $\rho(\mathcal{L}) = 1 + \varepsilon$. Then, for any $\mathbf{t} \in \mathbb{R}^n$,*

$$\left\| \frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})} + \mathbf{t} \right\| \leq \frac{\varepsilon}{1+\varepsilon} \cdot (s_\varepsilon \sinh(2\pi s_\varepsilon \|\mathbf{t}\|) + \|\mathbf{t}\| \cosh(2\pi s_\varepsilon \|\mathbf{t}\|)) + \frac{1}{\rho(\mathbf{t})} \cdot (1 + 2\pi \|\mathbf{t}\|^2) \int_{s_\varepsilon - \|\mathbf{t}\|}^{s_\varepsilon + \|\mathbf{t}\|} e^{-\pi z^2} \mathrm{d}z$$

where $s_\varepsilon = \left(\frac{1}{\pi} \log \frac{2(1+\varepsilon)}{\varepsilon}\right)^{1/2}$.

Proof. Using Eq. (7) to compute $\nabla f(\mathbf{t})$ and recalling that $f(\mathbf{t}) = \rho(\mathcal{L} + \mathbf{t})/(1+\varepsilon)$,

$$\left\| \frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})} + \mathbf{t} \right\| = \frac{(1+\varepsilon)\rho(\mathbf{t})}{\rho(\mathcal{L} + \mathbf{t})} \max_{\|\mathbf{v}\|=1} \mathop{\mathbb{E}}_{\mathbf{y} \sim D_\mathcal{L}} [\sinh(2\pi \langle \mathbf{y}, \mathbf{t} \rangle) \langle \mathbf{y}, \mathbf{v} \rangle] \, .$$

Fix a unit vector $\mathbf{v}$. For any $\mathbf{y}$, let $P_{r,s}(\mathbf{y})$ be the indicator that $|\langle \mathbf{y}, \mathbf{v} \rangle| > s$, $|\langle \mathbf{y}, \mathbf{t} \rangle| > r\|\mathbf{t}\|$, and $\langle \mathbf{y}, \mathbf{t} \rangle \langle \mathbf{y}, \mathbf{v} \rangle > 0$. Then,

$$\begin{aligned}
&\sinh(2\pi \langle \mathbf{y}, \mathbf{t} \rangle) \langle \mathbf{y}, \mathbf{v} \rangle \\
&\quad = 2\pi \|\mathbf{t}\| \int_0^\infty \int_0^\infty \cosh(2\pi \|\mathbf{t}\| r) (1_{|\langle \mathbf{y}, \mathbf{v} \rangle| > s} 1_{|\langle \mathbf{y}, \mathbf{t} \rangle| > r\|\mathbf{t}\|} \operatorname{sign}(\langle \mathbf{y}, \mathbf{t} \rangle) \operatorname{sign}(\langle \mathbf{y}, \mathbf{v} \rangle)) \mathrm{d}s \mathrm{d}r \\
&\quad \leq 2\pi \|\mathbf{t}\| \int_0^\infty \int_0^\infty \cosh(2\pi \|\mathbf{t}\| r) P_{r,s}(\mathbf{y}) \mathrm{d}s \mathrm{d}r \, .
\end{aligned}$$

Taking expectations on both sides, we get

$$\mathbb{E}[\sinh(2\pi \langle \mathbf{y}, \mathbf{t} \rangle) \langle \mathbf{y}, \mathbf{v} \rangle] \leq 2\pi \|\mathbf{t}\| \int_0^\infty \int_0^\infty \cosh(2\pi \|\mathbf{t}\| r) \, \mathbb{E}[P_{r,s}(\mathbf{y})] \mathrm{d}s \mathrm{d}r \, .$$

As in the previous proof, note that

$$\mathbb{E}[P_{r,s}(\mathbf{y})] \leq \min\left(\frac{\varepsilon}{1+\varepsilon}, 2e^{-\pi s^2}, 2e^{-\pi r^2}\right)$$

by Lemma 2.20. So, we partition the positive quadrant of the (r,s)-plane into three regions and bound the integral separately in each region.

1. When $s \leq s_\varepsilon$ and $r \leq s_\varepsilon$, $\mathbb{E}[P_{r,s}(\mathbf{y})]$ is at most $\varepsilon/(1+\varepsilon)$, and the integral in this region is bounded by

$$\frac{\varepsilon}{1+\varepsilon} \cdot \int_0^{s_\varepsilon} \int_0^{s_\varepsilon} \cosh(2\pi r \|\mathbf{t}\|) \mathrm{d}s \mathrm{d}r = \frac{\varepsilon}{1+\varepsilon} \cdot \frac{s_\varepsilon}{2\pi \|\mathbf{t}\|} \sinh(2\pi s_\varepsilon \|\mathbf{t}\|) \, .$$

2. When $s \leq r$ and $r > s_\varepsilon$, $\mathbb{E}[P_{r,s}(\mathbf{y})]$ is at most $2e^{-\pi r^2}$, and the integral in this region is bounded by

$$\begin{aligned}
2 \int_{s_\varepsilon}^\infty \int_0^r \cosh(2\pi \|\mathbf{t}\| r) e^{-\pi r^2} \mathrm{d}s \mathrm{d}r &= \frac{1}{\rho(\mathbf{t})} \int_{s_\varepsilon}^\infty (r e^{-\pi(r - \|\mathbf{t}\|)^2} + r e^{-\pi(r + \|\mathbf{t}\|)^2}) \mathrm{d}r \\
&= \frac{1}{2\pi \rho(\mathbf{t})} \left(e^{-\pi(s_\varepsilon - \|\mathbf{t}\|)^2} + e^{-\pi(s_\varepsilon + \|\mathbf{t}\|)^2}\right) + \frac{\|\mathbf{t}\|}{\rho(\mathbf{t})} \int_{s_\varepsilon - \|\mathbf{t}\|}^{s_\varepsilon + \|\mathbf{t}\|} e^{-\pi z^2} \mathrm{d}z \\
&= \frac{1}{2\pi} \frac{\varepsilon}{1+\varepsilon} \cdot \cosh(2\pi s_\varepsilon \|\mathbf{t}\|) + \frac{\|\mathbf{t}\|}{\rho(\mathbf{t})} \int_{s_\varepsilon - \|\mathbf{t}\|}^{s_\varepsilon + \|\mathbf{t}\|} e^{-\pi z^2} \mathrm{d}z \, .
\end{aligned}$$

3. When $s > r$ and $s > s_\varepsilon$, $\mathbb{E}[P_{r,s}(\mathbf{y})]$ is at most $2e^{-\pi s^2}$. So, the integral in this region is bounded by

$$\begin{aligned} 2\int_{s_\varepsilon}^{\infty}\int_0^s \cosh(2\pi\|\mathbf{t}\|r)e^{-\pi s^2}\mathrm{d}r\mathrm{d}s &= \frac{1}{\pi\|\mathbf{t}\|}\int_{s_\varepsilon}^{\infty}\sinh(2\pi\|\mathbf{t}\|s)e^{-\pi s^2}\mathrm{d}r\mathrm{d}s \\ &= \frac{1}{2\pi\|\mathbf{t}\|\rho(\mathbf{t})}\int_{s_\varepsilon-\|\mathbf{t}\|}^{s_\varepsilon+\|\mathbf{t}\|} e^{-\pi z^2}\mathrm{d}z\ . \end{aligned}$$

Combining everything together, and applying Lemma 2.14,

$$\begin{aligned} &\Big\|\frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}+\mathbf{t}\Big\| \\ &\le \frac{\rho(\mathbf{t})}{\rho(\mathcal{L}+\mathbf{t})}\Big(\varepsilon\cdot(s_\varepsilon\sinh(2\pi s_\varepsilon\|\mathbf{t}\|)+\|\mathbf{t}\|\cosh(2\pi s_\varepsilon\|\mathbf{t}\|))+\frac{1+\varepsilon}{\rho(\mathbf{t})}(1+2\pi\|\mathbf{t}\|^2)\int_{s_\varepsilon-\|\mathbf{t}\|}^{s_\varepsilon+\|\mathbf{t}\|}e^{-\pi z^2}\mathrm{d}z\Big) \\ &\le \frac{\varepsilon}{1+\varepsilon}\cdot(s_\varepsilon\sinh(2\pi s_\varepsilon\|\mathbf{t}\|)+\|\mathbf{t}\|\cosh(2\pi s_\varepsilon\|\mathbf{t}\|))+\frac{1}{\rho(\mathbf{t})}\cdot(1+2\pi\|\mathbf{t}\|^2)\int_{s_\varepsilon-\|\mathbf{t}\|}^{s_\varepsilon+\|\mathbf{t}\|}e^{-\pi z^2}\mathrm{d}z\ . \end{aligned}$$

□

Corollary 4.3. *Let $\varepsilon \in (0, 1/200)$ and $\mathcal{L} \subset \mathbb{R}^n$ a lattice with $\rho(\mathcal{L}) = 1+\varepsilon$. Let $s_\varepsilon = \big(\frac{1}{\pi}\log\frac{2(1+\varepsilon)}{\varepsilon}\big)^{1/2}$. Then for all $\mathbf{t}\in\mathbb{R}^n$ satisfying $\|\mathbf{t}\| < s_\varepsilon/2$,*

$$\Big\|\frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}+\mathbf{t}\Big\| \le 12\cdot(\varepsilon/2)^{1-2\delta(\mathbf{t})}\cdot\|\mathbf{t}\|\ ,$$

where $\delta(\mathbf{t}) = \max(1/8, \|\mathbf{t}\|/s_\varepsilon)$. In particular, for $\delta(\mathbf{t}) \le \delta_{\mathsf{max}} = \frac{1}{2} - \frac{2}{\pi s_\varepsilon^2}$,

$$\Big\|\frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}+\mathbf{t}\Big\| \le \frac{\|\mathbf{t}\|}{4}\ .$$

Proof. Recall from Lemma 4.2 that

$$\Big\|\frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}+\mathbf{t}\Big\| \le \frac{\varepsilon}{1+\varepsilon}\cdot(s_\varepsilon\sinh(2\pi s_\varepsilon\|\mathbf{t}\|)+\|\mathbf{t}\|\cosh(2\pi s_\varepsilon\|\mathbf{t}\|))+(1+2\pi\|\mathbf{t}\|^2)\cdot e^{\pi\|\mathbf{t}\|^2}\int_{s_\varepsilon-\|\mathbf{t}\|}^{s_\varepsilon+\|\mathbf{t}\|}e^{-\pi z^2}\mathrm{d}z\ .$$

Because sinh is convex on $\mathbb{R}^+$, $\sinh(0) = 0$, and $\|\mathbf{t}\| \le \delta(\mathbf{t})s_\varepsilon$,

$$\sinh(2\pi s_\varepsilon\|\mathbf{t}\|) \le (1-\|\mathbf{t}\|/(\delta(\mathbf{t})s_\varepsilon))\sinh(0)+\frac{\|\mathbf{t}\|}{\delta(\mathbf{t})s_\varepsilon}\cdot\sinh(2\pi\delta(\mathbf{t})s_\varepsilon^2) \le \frac{\|\mathbf{t}\|}{2\delta(\mathbf{t})s_\varepsilon}\cdot e^{2\pi\delta(\mathbf{t})s_\varepsilon^2}\ .$$

Using the above,

$$\begin{aligned} \frac{\varepsilon}{1+\varepsilon}\cdot(s_\varepsilon\sinh(2\pi s_\varepsilon\|\mathbf{t}\|)+\|\mathbf{t}\|\cosh(2\pi s_\varepsilon\|\mathbf{t}\|)) &\le \|\mathbf{t}\|\cdot\frac{\varepsilon}{1+\varepsilon}\cdot\Big(\frac{e^{2\pi\delta(\mathbf{t})s_\varepsilon^2}}{2\delta(\mathbf{t})}+\cosh(2\pi\delta(\mathbf{t})s_\varepsilon^2)\Big) \\ &\le \|\mathbf{t}\|\cdot\frac{\varepsilon}{1+\varepsilon}\cdot\Big(\frac{1}{2\delta(\mathbf{t})}+1\Big)\cdot e^{2\pi\delta(\mathbf{t})s_\varepsilon^2} \\ &= \|\mathbf{t}\|\cdot\Big(\frac{1}{\delta(\mathbf{t})}+2\Big)\cdot e^{-\pi(1-2\delta(\mathbf{t}))s_\varepsilon^2}\ . \end{aligned}$$

Turning to the integral and using the above bound on $\sinh(2\pi s_\varepsilon \|\mathbf{t}\|)$ again,

$$\begin{aligned}
e^{\pi\|\mathbf{t}\|^2}\int_{s_\varepsilon-\|\mathbf{t}\|}^{s_\varepsilon+\|\mathbf{t}\|} e^{-\pi z^2}\mathrm{d}z &\le e^{\pi\|\mathbf{t}\|^2}\int_{s_\varepsilon-\|\mathbf{t}\|}^{s_\varepsilon+\|\mathbf{t}\|} \frac{z}{s_\varepsilon-\|\mathbf{t}\|}e^{-\pi z^2}\mathrm{d}z \\
&= \frac{1}{2\pi(s_\varepsilon-\|\mathbf{t}\|)}e^{\pi\|\mathbf{t}\|^2}(e^{-\pi(s_\varepsilon-\|\mathbf{t}\|)^2}-e^{-\pi(s_\varepsilon+\|\mathbf{t}\|)^2}) \\
&= \frac{1}{\pi(s_\varepsilon-\|\mathbf{t}\|)}e^{-\pi s_\varepsilon^2}\sinh(2\pi s_\varepsilon\|\mathbf{t}\|) \\
&\le \frac{\|\mathbf{t}\|}{(1-\delta(\mathbf{t}))\delta(\mathbf{t})s_\varepsilon}\cdot\frac{1}{2\pi s_\varepsilon}\cdot e^{-\pi s_\varepsilon^2}e^{2\pi\delta(\mathbf{t})s_\varepsilon^2} \\
&\le \|\mathbf{t}\|\cdot\frac{1}{\pi\delta(\mathbf{t})s_\varepsilon^2}\cdot e^{-\pi(1-2\delta(\mathbf{t}))s_\varepsilon^2}\,.
\end{aligned}$$

Combining everything together,

$$\begin{aligned}
\Big\|\frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}+\mathbf{t}\Big\| &\le \|\mathbf{t}\|\Big(\frac{1}{\delta(\mathbf{t})}+2+\frac{1}{\pi\delta(\mathbf{t})s_\varepsilon^2}+2\delta(\mathbf{t})\Big)\cdot e^{-\pi(1-2\delta(\mathbf{t}))s_\varepsilon^2} \\
&\le \|\mathbf{t}\|\Big(\frac{1}{\delta(\mathbf{t})}+2+\frac{1}{\pi\delta(\mathbf{t})s_\varepsilon^2}+2\delta(\mathbf{t})\Big)\cdot(\varepsilon/2)^{1-2\delta(\mathbf{t})}\,.
\end{aligned}$$

The first result follows by noting that $\frac{1}{\delta(\mathbf{t})}+2+\frac{1}{\pi\delta(\mathbf{t})s_\varepsilon^2}+2\delta(\mathbf{t}) < 12$ for $\varepsilon < 1/200$ and $\delta(\mathbf{t}) \in (1/8, 1/2)$. The second result follows by noting that $12(\varepsilon/2)^{1-2\delta(\mathbf{t})} < 1/4$ for $\delta(\mathbf{t}) \le \frac{1}{2}-\frac{2}{\pi s_\varepsilon^2}$. □

Lemma 4.4. *Let $\varepsilon > 0$ and $\mathcal{L} \subset \mathbb{R}^n$ a lattice with $\rho(\mathcal{L}) = 1+\varepsilon$. Then,*

1. $\|Hf(\mathbf{t})\| \le \|Hf(\mathbf{0})\| \le 2\pi$ *for all* $\mathbf{t} \in \mathbb{R}^n$.
2. $\|Hf(\mathbf{0}) + 2\pi I_n\| \le \frac{4\pi\varepsilon}{1+\varepsilon}\Big(\log\frac{2(1+\varepsilon)}{\varepsilon}+1\Big)$.

Proof. From Eq. (2), we have that for any $\mathbf{t} \in \mathbb{R}^n$

$$\begin{aligned}
\|Hf(\mathbf{t})\| &= 4\pi^2\Big\|\mathop{\mathbb{E}}_{\mathbf{w}\sim D_{\mathcal{L}^*}}[\mathbf{w}\mathbf{w}^T\cos(2\pi\langle\mathbf{w},\mathbf{t}\rangle)]\Big\| \\
&\le 4\pi^2\Big\|\mathop{\mathbb{E}}_{\mathbf{w}\sim D_{\mathcal{L}^*}}[\mathbf{w}\mathbf{w}^T]\Big\| \\
&= \|Hf(\mathbf{0})\|\,.
\end{aligned}$$

From Eqs. (2) and (7), we have a representation of $Hf(\mathbf{0})$ in both the primal and the dual,

$$-\frac{1}{2\pi}Hf(\mathbf{0}) = I_n - 2\pi\mathop{\mathbb{E}}_{\mathbf{y}\sim D_{\mathcal{L}}}[\mathbf{y}\mathbf{y}^T] = 2\pi\mathop{\mathbb{E}}_{\mathbf{w}\sim D_{\mathcal{L}^*}}[\mathbf{w}\mathbf{w}^T]\,.$$

Noting that both expectations are positive semidefinite, it follows that $\|Hf(\mathbf{t})\| \le \|Hf(\mathbf{0})\| \le 2\pi$.

For the second bound, following the technique used in the proofs of Lemmas 4.1 and 4.2,

$$\begin{aligned}
\|Hf(\mathbf{0}) + 2\pi I_n\| &= 4\pi^2 \cdot \max_{\|\mathbf{v}\|=1} \mathop{\mathbb{E}}_{\mathbf{y}\sim D_{\mathcal{L}}} [\langle \mathbf{y}, \mathbf{v}\rangle^2] \\
&= 8\pi^2 \cdot \max_{\|\mathbf{v}\|=1} \int_0^\infty r \Pr_{\mathbf{y}\sim D_{\mathcal{L}}} [|\langle \mathbf{y}, \mathbf{v}\rangle| \geq r] \mathrm{d}r \\
&\leq 8\pi^2 \int_0^\infty r \min(\varepsilon/(1+\varepsilon), 2e^{-\pi r^2}) \mathrm{d}r \qquad \text{(Lemma 2.20)} \\
&= \frac{4\pi\varepsilon}{1+\varepsilon}\Big(\log\frac{2(1+\varepsilon)}{\varepsilon} + 1\Big) .
\end{aligned}$$

□

4.2 Completing the Proof

In this section we complete the proof of Proposition 3.2. The basic plan of the proof is straightforward: after having shown in Corollary 4.3 the analogous property for the exact function f, it suffices to show that $\nabla f_W(\mathbf{t})/f_W(\mathbf{t})$ is close to $\nabla f(\mathbf{t})/f(\mathbf{t})$ for all relevant $\mathbf{t}$. It is obviously enough to argue separately that ∇f_W is close to ∇f and that f_W is close to f (with appropriate notions of closeness; see the technical Claim 4.8 for the precise statement). The former will be shown to hold with high probability for any fixed $\mathbf{t}$ in Lemma 4.9 and then to hold with high probability simultaneously for all relevant $\mathbf{t}$ in Lemma 4.10. Similarly, the latter will be shown to hold with high probability for any fixed $\mathbf{t}$ in Lemma 4.11 and then to hold with high probability simultaneously for all relevant $\mathbf{t}$ in Lemma 4.12. In both cases, showing that the result holds simultaneously for all $\mathbf{t}$ is done by taking a union bound over an appropriately chosen net and showing that the functions do not vary much. One minor complication in the proof is that in the former case (closeness of ∇f_W) the net has to become denser as we get closer to the origin. In order to keep the net finite, Lemma 4.10 actually does not handle tiny vectors $\mathbf{t}$. Instead we include Lemma 4.7 which proves Proposition 3.2 directly for the case of tiny vectors. Finally, many of our proofs require quantitative statements about the smoothness f_W and ∇f_W, which are shown in Lemma 4.5 and Lemma 4.6.

Lemma 4.5. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice with $\rho(\mathcal{L}) = 1+\varepsilon$ for some $\varepsilon > 0$, and let $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ be sampled independently from $D_{\mathcal{L}^*}$. Then, for $s \geq 0$, $N\min(s, s^2) \geq \Omega(n)$, and $\Delta_\varepsilon = \frac{4\pi\varepsilon}{1+\varepsilon}(\log\frac{2(1+\varepsilon)}{\varepsilon} + 1)$, we have*

1. $\Pr[\|Hf_W(\mathbf{0}) + 2\pi I_n\| > \Delta_\varepsilon + s] \leq 2^{-\Omega(N\min(s,s^2))}$.
2. $\Pr[\exists \mathbf{t} \in \mathbb{R}^n : \|Hf_W(\mathbf{t}) + 2\pi I_n\| > \Delta_\varepsilon + s + (500n\|\mathbf{t}\|)^2] \leq 2^{-\Omega(n)}$.
3. $\Pr[\exists \mathbf{t} \in \mathbb{R}^n : \|Hf_W(\mathbf{t})\| > 2\pi + s] \leq 2^{-\Omega(N\min(s,s^2))}$.

Proof. For bound (1), using the triangle inequality and Lemma 4.4, we have that

$$\|Hf_W(\mathbf{0}) + 2\pi I_n\| \leq \|Hf(\mathbf{0}) + 2\pi I_n\| + \|Hf_W(\mathbf{0}) - Hf(\mathbf{0})\| \leq \Delta_\varepsilon + \|Hf_W(\mathbf{0}) - Hf(\mathbf{0})\|.$$

It now suffices to bound the probability that $\|Hf_W(\mathbf{0}) - Hf(\mathbf{0})\| > s$. For this, note that

$$\|Hf_W(\mathbf{0})\| = \sup_{\|\mathbf{v}\|=1} |\langle Hf_W(\mathbf{0})\mathbf{v}, \mathbf{v}\rangle| = \frac{4\pi^2}{N} \sup_{\|\mathbf{v}\|=1} \sum_i \langle \mathbf{v}, \mathbf{w}_i\rangle^2 .$$

By Lemma 2.20, $\mathbf{w}_i$ are subgaussian random variables with parameter 1. It follows from Lemma 2.22 that $\langle \mathbf{w}_i, \mathbf{v} \rangle^2$ is subexponential with parameter $O(1)$. Then, applying Lemma 2.24,

$$\Pr[|\langle (Hf_W(\mathbf{0}) - Hf(\mathbf{0}))\mathbf{v}, \mathbf{v} \rangle| > s/2] \leq 2^{1-\Omega(N \min(s,s^2))},$$

for any $s \geq 0$. By Lemma 2.28, there is a $\frac{1}{4}$-net of the unit sphere A with $|A| = 2^{O(n)}$. Taking union bound over A and applying Lemma 2.31 gives

$$\Pr[\|Hf_W(\mathbf{0}) - Hf(\mathbf{0})\| > s] \leq 2^{-\Omega(N \min(s,s^2))+O(n)} = 2^{-\Omega(N \min(s,s^2))}, \tag{9}$$

by our assumption that $N \min(s, s^2) = \Omega(n)$.

For bound (2), using the triangle inequality as above, we have that

$$\begin{aligned} \|Hf_W(\mathbf{t}) + 2\pi I_n\| &\leq \|Hf(\mathbf{0}) + 2\pi I_n\| + \|Hf_W(\mathbf{0}) - Hf(\mathbf{0})\| + \|Hf_W(\mathbf{t}) - Hf_W(\mathbf{0})\| \\ &\leq \Delta_\varepsilon + \|Hf_W(\mathbf{0}) - Hf(\mathbf{0})\| + \|Hf_W(\mathbf{t}) - Hf_W(\mathbf{0})\|. \end{aligned}$$

By equation (9), we know that $\|Hf_W(\mathbf{0}) - Hf(\mathbf{0})\| > s$ with probability at most $2^{-\Omega(N \min(s,s^2))} = 2^{-\Omega(n)}$. Hence it suffices to prove that $\|Hf_W(\mathbf{t}) - Hf_W(\mathbf{0})\| > (500n\|\mathbf{t}\|)^2$, for some $\mathbf{t} \in \mathbb{R}^n$, with probability at most $2^{-\Omega(n)}$.

Using the inequality $1 - \theta^2/2 \leq \cos(\theta) \leq 1$ and Cauchy-Schwarz, we have that

$$\begin{aligned} \|Hf_W(\mathbf{t}) - Hf_W(\mathbf{0})\| &= \frac{4\pi^2}{N} \Big\| \sum_{i=1}^N (\cos(2\pi \langle \mathbf{w}_i, \mathbf{t} \rangle) - 1) \mathbf{w}_i \mathbf{w}_i^T \Big\| \\ &\leq \frac{4\pi^2}{N} \sum_{i=1}^N |\cos(2\pi \langle \mathbf{w}_i, \mathbf{t} \rangle) - 1| \|\mathbf{w}_i \mathbf{w}_i^T\| \\ &\leq \frac{8\pi^4}{N} \sum_{i=1}^N \langle \mathbf{w}_i, \mathbf{t} \rangle^2 \big\|\mathbf{w}_i\big\|^2 \\ &\leq (8\pi^4 n^2 \|\mathbf{t}\|^2) \frac{1}{N} \sum_{i=1}^N \Big\| \frac{\mathbf{w}_i}{\sqrt{n}} \Big\|^4. \end{aligned}$$

It now suffices to bound the sum in the last expression with probability $1 - 2^{-\Omega(n)}$. Let $S_j = \{i \in [N] : \|\mathbf{w}_i\| \geq e^j \sqrt{n}\}$, for $j \geq 0$. Using Lemma 2.13, we have that

$$\mathbb{E}[|S_j|] = N \Pr[\|\mathbf{w}_i\| \geq e^j \sqrt{n}] \leq N e^{-\frac{n}{2}(\sqrt{2\pi} e^j - 1)^2} \leq N e^{-n e^{2j}}.$$

By Markov's inequality, $\Pr[|S_j| \geq N e^{-n e^{2j} + n(j+1)}] \leq e^{-n(j+1)}$. By the union bound, the event $|S_j| \leq N e^{-n e^{2j} + n(j+1)}$, $\forall j \geq 0$, occurs with probability at least $1 - \sum_{j=0}^\infty e^{-n(j+1)} \geq 1 - 2e^{-n}$. Conditioning on this event, we will show the desired bound.

For all $i \in [N]$, we have that $\|\frac{\mathbf{w}_i}{\sqrt{n}}\|^4 \leq 1 + e^4 \sum_{j=0}^\infty e^{4j} 1_{i \in S_j}$. Using this, we get that

$$\frac{1}{N} \sum_{i=1}^N \Big\| \frac{\mathbf{w}_i}{\sqrt{n}} \Big\|^4 \leq 1 + \frac{e^4}{N} \sum_{j=0}^\infty e^{4j} |S_j| \leq 1 + e^4 \sum_{j=0}^\infty e^{-n e^{2j} + n(j+1) + 4j} \leq 1 + e^4 \sum_{j=0}^\infty e^{-j} \leq 2e^4 .$$

Plugging in gives $\|Hf_W(\mathbf{t}) - Hf_W(\mathbf{0})\| \leq (4\pi^2 e^2 n \|\mathbf{t}\|)^2 \leq (500n\|\mathbf{t}\|)^2$.

For bound (3), we simply note that

$$\|Hf_W(\mathbf{t})\| = \frac{4\pi^2}{N}\|\sum \mathbf{w}_i\mathbf{w}_i^T \cos(2\pi\langle \mathbf{w}_i\mathbf{t}\rangle)\| \leq \frac{4\pi^2}{N}\|\sum \mathbf{w}_i\mathbf{w}_i^T\| = \|Hf_W(\mathbf{0})\| .$$

The bound now follows from equation (9), the fact that $\|Hf(\mathbf{0})\| \leq 2\pi$ (Lemma 4.4), and the triangle inequality. □

The following lemma establishes strong continuity properties for f, ∇f, and their respective approximations.

Lemma 4.6. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice. Then, for all $\mathbf{t}, \mathbf{t}' \in \mathbb{R}^n$,*

1. $\|\nabla f(\mathbf{t}') - \nabla f(\mathbf{t})\| \leq 2\pi\|\mathbf{t} - \mathbf{t}'\|$.
2. $|f(\mathbf{t}') - f(\mathbf{t})| \leq 2\pi \max(\|\mathbf{t}\|, \|\mathbf{t}'\|)\|\mathbf{t}' - \mathbf{t}\|$.

Let $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ be sampled independently from $D_{\mathcal{L}^*}$. Then for $s > 0$, $N \min(s, s^2) = \Omega(n)$, the following both hold simultaneously for all $\mathbf{t}, \mathbf{t}' \in \mathbb{R}^n$ with probability at least $1 - 2^{-\Omega(N\min(s,s^2))}$.

1. $\|\nabla f_W(\mathbf{t}') - \nabla f_W(\mathbf{t})\| \leq (2\pi + s)\|\mathbf{t}' - \mathbf{t}\|$
2. $|f_W(\mathbf{t}') - f_W(\mathbf{t})| \leq (2\pi + s)\max(\|\mathbf{t}\|, \|\mathbf{t}'\|)\|\mathbf{t}' - \mathbf{t}\|$

Proof. By Lemma 4.4, we have that $\|Hf(\mathbf{x})\| \leq 2\pi$ for all $\mathbf{x} \in \mathbb{R}^n$. From this, we get that

$$\begin{aligned}\|\nabla f(\mathbf{t}') - \nabla f(\mathbf{t})\| &= \Big\| \int_0^1 Hf((1-r)\mathbf{t} + r\mathbf{t}') \cdot (\mathbf{t}' - \mathbf{t}) dr \Big\| \\ &\leq \|\mathbf{t}' - \mathbf{t}\| \int_0^1 \|Hf((1-r)\mathbf{t} + r\mathbf{t}')\| dr \\ &\leq 2\pi\|\mathbf{t}' - \mathbf{t}\| .\end{aligned}$$

Since $\nabla f(\mathbf{0}) = \mathbf{0}$, using the above we get that $\|\nabla f(\mathbf{x})\| = \|\nabla f(\mathbf{x}) - \nabla f(\mathbf{0})\| \leq 2\pi\|\mathbf{x}\|$, for all $\mathbf{x} \in \mathbb{R}^n$. Using this inequality, we get that

$$\begin{aligned}|f(\mathbf{t}') - f(\mathbf{t})| &= \Big| \int_0^1 \langle \nabla f((1-r)\mathbf{t} + r\mathbf{t}'), \mathbf{t}' - \mathbf{t}\rangle dr \Big| \\ &\leq \|\mathbf{t}' - \mathbf{t}\| \int_0^1 \|\nabla f((1-r)\mathbf{t} + r\mathbf{t}')\| dr \\ &\leq 2\pi \max(\|\mathbf{t}\|, \|\mathbf{t}'\|)\|\mathbf{t}' - \mathbf{t}\| .\end{aligned}$$

For the second part, by Lemma 4.5 the event $\|Hf_W(\mathbf{x})\| \leq 2\pi + s$, for all $\mathbf{x} \in \mathbb{R}^n$, holds with probability $1 - 2^{\Omega(N\min(s,s^2))}$. The claim now follows by the same proof as above replacing f by f_W. □

Lemma 4.7. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice with $\rho(\mathcal{L}) = 1 + \varepsilon$ for $\varepsilon \in (0, 1/200)$. Let $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ be sampled independently from $D_{\mathcal{L}^*}$ with $N \geq \Omega(n/\sqrt{\varepsilon})$. Then,*

$$\Pr\left[\exists \mathbf{t}, \|\mathbf{t}\| \leq \varepsilon^{1/8}/(1000n) : \left\| \frac{\nabla f_W(\mathbf{t})}{2\pi f_W(\mathbf{t})} + \mathbf{t} \right\| > \varepsilon^{1/4}\|\mathbf{t}\|\right] \leq 2^{-\Omega(n)} .$$

Proof. Let $\Delta_\varepsilon = \frac{4\pi\varepsilon}{1+\varepsilon}(\log\frac{2(1+\varepsilon)}{\varepsilon}+1)$ as in Lemma 4.5, and note that $\Delta_\varepsilon \le 3\varepsilon^{1/4}$ for $\varepsilon < 1/200$. Then, by Lemma 4.5, setting $s = 3\varepsilon^{1/4}/4$, we have that

$$\|Hf_W(\mathbf{x}) + 2\pi I_n\| < \Delta_\varepsilon + (500n\|\mathbf{x}\|)^2 + 3\varepsilon^{1/4}/4 \le 3\varepsilon^{1/4} + \varepsilon^{1/4}/4 + 3\varepsilon^{1/4}/4 = 4\varepsilon^{1/4}$$

holds simultaneously for all $\mathbf{x}$ with $\|\mathbf{x}\| \le \varepsilon^{1/8}/(1000n)$ with probability at least $1 - 2^{-\Omega(n)}$. Suppose this holds. Noting that $\nabla f_W(\mathbf{0}) = \mathbf{0}$, it follows that for all $\mathbf{x}'$, $\|\mathbf{x}'\| \le \varepsilon^{1/8}/(1000n)$, we have that

$$\begin{aligned}\|\nabla f_W(\mathbf{x}') + 2\pi\mathbf{x}'\| &= \Big\|\int_0^1 Hf_W(r\mathbf{x}')\mathbf{x}'dr + 2\pi\mathbf{x}'\Big\| = \Big\|\int_0^1 (Hf_W(r\mathbf{x}') + 2\pi I_n)\mathbf{x}'dr\Big\| \\ &\le \|\mathbf{x}'\|\int_0^1 \|Hf_W(r\mathbf{x}') + 2\pi I_n\|dr \le 4\varepsilon^{1/4}\cdot\|\mathbf{x}'\| \,.\end{aligned}$$

In particular, $\|\nabla f_W(\mathbf{x}')\| \le (2\pi + 4\varepsilon^{1/4})\|\mathbf{x}'\|$. Since $f_W(\mathbf{0}) = 1$, it follows that for any $\mathbf{t}$ with $\|\mathbf{t}\| \le \varepsilon^{1/8}/(1000n)$, we have

$$1 \ge f_W(\mathbf{t}) \ge 1 - (2\pi + 4\varepsilon^{1/4})\|\mathbf{t}\|^2 \ge 1 - \varepsilon^{1/4}/100 \,.$$

Putting it all together,

$$\begin{aligned}\Big\|\frac{\nabla f_W(\mathbf{t})}{2\pi f_W(\mathbf{t})} + \mathbf{t}\Big\| &\le \Big\|\frac{\nabla f_W(\mathbf{t})}{2\pi} + \mathbf{t}\Big\| + \Big(\frac{1}{f_W(\mathbf{t})} - 1\Big)\Big\|\frac{\nabla f_W(\mathbf{t})}{2\pi}\Big\| \\ &\le \frac{4\varepsilon^{1/4}}{2\pi}\|\mathbf{t}\| + \Big(\frac{1-f_W(\mathbf{t})}{f_W(\mathbf{t})}\Big)\Big(1 + \frac{4\varepsilon^{1/4}}{2\pi}\Big)\|\mathbf{t}\| \\ &\le \frac{2}{3}\varepsilon^{1/4}\|\mathbf{t}\| + \frac{4}{3}\Big(\frac{1-f_W(\mathbf{t})}{f_W(\mathbf{t})}\Big)\|\mathbf{t}\| \\ &\le \varepsilon^{1/4}\|\mathbf{t}\| \,,\end{aligned}$$

as needed. □

Claim 4.8. *Let $\varepsilon \in (0, 1/200)$ and $\mathcal{L} \subset \mathbb{R}^n$ be a lattice with $\rho(\mathcal{L}) = 1+\varepsilon$. Let $s_\varepsilon = \big(\frac{1}{\pi}\log\frac{2(1+\varepsilon)}{\varepsilon}\big)^{1/2}$, $\delta_{\mathsf{max}} = \frac{1}{2} - \frac{2}{\pi s_\varepsilon^2}$, and $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ be vectors in $\mathcal{L}^*$. Suppose that for some $\gamma > 0$ and $\mathbf{t} \in \mathbb{R}^n$ it holds that*

1. $\|\mathbf{t}\| \le \min\{\delta_{\mathsf{max}} s_\varepsilon,\ \sqrt{\log(1/(4\gamma))/\pi}\}$,
2. $\|\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})\| \le \frac{\pi}{2}\gamma\|\mathbf{t}\|$, *and*
3. $|f_W(\mathbf{t}) - f(\mathbf{t})| \le \gamma$.

Then,

$$\Big\|\frac{\nabla f_W(\mathbf{t})}{2\pi f_W(\mathbf{t})} - \frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\Big\| \le \frac{2\gamma}{\rho(\mathbf{t})}\|\mathbf{t}\| \,.$$

Proof. By Lemma 2.14 and the first assumption, we see that $f(\mathbf{t}) \ge \rho(\mathbf{t}) \ge 4\gamma$.

By the triangle inequality

$$\begin{aligned}\left\|\frac{\nabla f_W(\mathbf{t})}{2\pi f_W(\mathbf{t})} - \frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\right\| &= \left\|\frac{\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\,\frac{f(\mathbf{t})}{f_W(\mathbf{t})} + \frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\left(\frac{f(\mathbf{t})}{f_W(\mathbf{t})} - 1\right)\right\| \\ &\le \left\|\frac{\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\right\|\frac{f(\mathbf{t})}{f_W(\mathbf{t})} + \left\|\frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\right\|\left|\frac{f(\mathbf{t})}{f_W(\mathbf{t})} - 1\right| . \qquad (10)\end{aligned}$$

For the first term in (10), by the second and third assumption, we have

$$\left\|\frac{\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\right\|\frac{f(\mathbf{t})}{f_W(\mathbf{t})} \le \frac{\gamma\|\mathbf{t}\|}{4f(\mathbf{t})}\cdot\frac{f(\mathbf{t})}{f(\mathbf{t})-\gamma} = \frac{\gamma}{4(f(\mathbf{t})-\gamma)}\|\mathbf{t}\| . \qquad (11)$$

For the second term in (10), by Corollary 4.3 and the first and third assumption,

$$\left\|\frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\right\|\left|\frac{f(\mathbf{t})}{f_W(\mathbf{t})} - 1\right| \le \frac{5}{4}\|\mathbf{t}\|\left(\frac{f(\mathbf{t})}{f(\mathbf{t})-\gamma} - 1\right) = \frac{5\gamma}{4(f(\mathbf{t})-\gamma)}\|\mathbf{t}\| . \qquad (12)$$

Combining (10), (11), and (12) together, we have

$$\left\|\frac{\nabla f_W(\mathbf{t})}{2\pi f_W(\mathbf{t})} - \frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})}\right\| \le \frac{6}{4}\frac{\gamma}{f(\mathbf{t})-\gamma}\|\mathbf{t}\| \le \frac{2\gamma}{\rho(\mathbf{t})}\|\mathbf{t}\| ,$$

as needed. □

Lemma 4.9. *For $\mathcal{L} \subset \mathbb{R}^n$ a lattice, $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ sampled independently from $D_{\mathcal{L}^*}$, $\mathbf{t} \in \mathbb{R}^n$, and $s \ge 0$,*

$$\Pr[\|\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})\| > s\|\mathbf{t}\|] \le 2^{-\Omega(N\min(s,s^2))+O(n)} .$$

Proof. For any i and any unit vector $\mathbf{v}$,

$$|\langle \nabla f_{\{\mathbf{w}_i\}}(\mathbf{t}), \mathbf{v}\rangle| = 2\pi|\langle \mathbf{w}_i, \mathbf{v}\rangle \sin(2\pi\langle \mathbf{w}_i, \mathbf{t}\rangle)| \le 4\pi^2|\langle \mathbf{w}_i, \mathbf{v}\rangle\langle \mathbf{w}_i, \mathbf{t}\rangle| .$$

It follows from the subgaussianity of the discrete Gaussian and Corollary 2.23 that $\langle \nabla f_{\{\mathbf{w}_i\}}(\mathbf{t}), \mathbf{v}\rangle / \|\mathbf{t}\|$ is subexponential with parameter $O(1)$. Applying Lemma 2.24, we get that

$$\Pr[|\langle \nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t}), \mathbf{v}\rangle| > (s/2)\|\mathbf{t}\|] \le 2^{1-\Omega(N\min(s,s^2))} .$$

By Lemma 2.28, there is a $\frac{1}{2}$-net of the sphere, A with $|A| = 2^{O(n)}$. Taking a union bound over A and applying Lemma 2.29 gives

$$\Pr[\|\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})\| > s\|\mathbf{t}\|] \le 2^{-\Omega(N\min(s,s^2))+O(n)},$$

as needed. □

Lemma 4.10. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice with $\rho(\mathcal{L}) = 1 + \varepsilon$ with $\varepsilon \in (0, 1/200)$. Let $s_\varepsilon = \left(\frac{1}{\pi}\log\frac{2(1+\varepsilon)}{\varepsilon}\right)^{1/2}$. Let $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ be sampled independently from $D_{\mathcal{L}^*}$. Then, for $\varepsilon^2 \le s \le 10$, if $N \ge \Omega(n\log(1/\varepsilon)/s^2)$,*

$$\Pr[\exists \mathbf{t} \in \mathbb{R}^n, \varepsilon^{1/8}/(1000n) \le \|\mathbf{t}\| \le s_\varepsilon : \|\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})\| > s\|\mathbf{t}\|] \le 2^{-\Omega(Ns^2)} .$$

Proof. We wish to find a set a vectors $A = \{\mathbf{t}_j\}$ such that for any $\mathbf{t}$ with $\varepsilon^{1/8}/(1000n) \le \|\mathbf{t}\| \le s_\varepsilon$, there is a $\mathbf{t}_j \in A$ with $\|\mathbf{t} - \mathbf{t}_j\| \le s\|\mathbf{t}_j\|/100$. For $i = -\lceil \log n \rceil - \lceil \log 1/\varepsilon \rceil - 10$ to $\lceil \log s_\varepsilon \rceil$, let A_i be a $(e^i s/100)$-net of the shell of inner radius radius e^i and outer radius e^{i+1}. By Lemma 2.28, we can take $|A_i| = 2^{O(n \log(1/\varepsilon))}$. Let $A = \cup A_i$. There are $O(\log n + \log(1/\varepsilon))$ such nets, so $|A| = 2^{O(n \log(1/\varepsilon))}$.

We show that two bounds hold with high probability.

1. By Lemma 4.6, $\|\nabla f_W(\mathbf{x}) - \nabla f_W(\mathbf{y})\| \le 3\pi\|\mathbf{x} - \mathbf{y}\|$ holds simultaneously for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with probability at least $1 - 2^{-\Omega(N)}$.

2. By Lemma 4.9 and union bound over A, $\|\nabla f_W(\mathbf{t}_j) - \nabla f(\mathbf{t}_j)\| \le s\|\mathbf{t}_j\|/10$ holds simultaneously for all $\mathbf{t}_j$ with probability at least $1 - 2^{-\Omega(Ns^2)+O(n \log(1/\varepsilon))}$.

Suppose that both bounds hold, which happens with probability at least $1 - 2^{-\Omega(Ns^2)}$. For a target vector $\mathbf{t}$ with $\varepsilon^{1/8}/(1000n) \le \|\mathbf{t}\| \le s_\varepsilon$, let $\mathbf{t}_j$ be the closest vector to $\mathbf{t}$ in A. Then, by the first bound, $\|\nabla f_W(\mathbf{t}) - \nabla f_W(\mathbf{t}_j)\| \le 3\pi\|\mathbf{t} - \mathbf{t}_j\| \le s\|\mathbf{t}_j\|/10$. Again, using Lemma 4.6, $\|\nabla f(\mathbf{t}) - \nabla f(\mathbf{t}_j)\| \le s\|\mathbf{t}_j\|/10$. Applying triangle inequality repeatedly and noting that $\|\mathbf{t}_j\| \le e\|\mathbf{t}\|$,

$$\begin{aligned}\|\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})\| &\le s\|\mathbf{t}_j\|/5 + \|\nabla f_W(\mathbf{t}_j) - \nabla f(\mathbf{t}_j)\| \\ &\le s\|\mathbf{t}\| \ .\end{aligned}$$

□

Lemma 4.11 (Implicit in [AR05, Lemma 1.3])**.** *For a lattice $\mathcal{L} \subset \mathbb{R}^n$, $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ sampled independently from $D_{\mathcal{L}^*}$, $\mathbf{t} \in \mathbb{R}^n$, and $s \ge 0$,*

$$\Pr[|f_W(\mathbf{t}) - f(\mathbf{t})| > s] \le 2^{1-\Omega(Ns^2)} \ .$$

Proof. The result follows immediately from Lemma 2.25 (the Chernoff-Hoeffding bound) and the definitions of $f_W(\mathbf{t})$ and $f(\mathbf{t})$ (see Eqs. (2) and (3)). □

Lemma 4.12. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice with $\rho(\mathcal{L}) = 1 + \varepsilon$ for $\varepsilon \in (0, 1/200)$. Let $s_\varepsilon = \left(\frac{1}{\pi} \log \frac{2(1+\varepsilon)}{\varepsilon}\right)^{1/2}$. Let $W = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$ be sampled independently from $D_{\mathcal{L}^*}$. Then, for $\varepsilon^2 \le s \le 10$, if $N \ge \Omega(n \log(1/\varepsilon)/s^2)$,*

$$\Pr[\exists \mathbf{t}, \|\mathbf{t}\| \le s_\varepsilon : |f_W(\mathbf{t}) - f(\mathbf{t})| > s] \le 2^{-\Omega(Ns^2)} \ .$$

Proof. Our proof is quite similar to that of Lemma 4.10. Let A be a $s/(100 s_\varepsilon)$-net of the ball of radius $\delta_{\mathsf{max}} s_\varepsilon$. By Lemma 2.28, and since $s \ge \varepsilon^2$, we can take $|A| \le (1 + 200 s_\varepsilon^2/s)^n = 2^{O(n \log(1/\varepsilon))}$.

The following events hold with high probability.

1. By Lemma 4.6, we have that $|f_W(\mathbf{x}) - f_W(\mathbf{y})| \le 3\pi s_\varepsilon \|\mathbf{x} - \mathbf{y}\|$ holds simultaneously for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with $\|\mathbf{x}\|, \|\mathbf{y}\| \le s_\varepsilon$ with probability at least $1 - 2^{-\Omega(N)}$.

2. By Lemma 4.11 and union bound, we have that $|f_W(\mathbf{t}_j) - f(\mathbf{t}_j)| \le s/10$ holds simultaneously for all $\mathbf{t}_j \in A$ with probability at least $1 - 2^{-\Omega(s^2 N)+O(n \log(1/\varepsilon))}$.

Suppose that both bounds hold, which happens with probability at least $1 - 2^{-\Omega(s^2 N)}$. For a target vector $\mathbf{t}$ with $\|\mathbf{t}\| \le s_\varepsilon$, let $\mathbf{t}_j$ be the closest point to $\mathbf{t}$ in A. From the first event, we have that $|f_W(\mathbf{t}) - f_W(\mathbf{t}_j)| \le 3\pi s_\varepsilon \|\mathbf{t} - \mathbf{t}_j\| < s/10$. Similarly, by Lemma 4.6, we have $|f(\mathbf{t}) - f(\mathbf{t}_j)| < s/10$. Then, using the triangle inequality,

$$\begin{aligned} |f_W(\mathbf{t}) - f(\mathbf{t})| &\le |f_W(\mathbf{t}) - f_W(\mathbf{t}_j)| + |f_W(\mathbf{t}_j) - f(\mathbf{t}_j)| + |f(\mathbf{t}_j) - f(\mathbf{t})| \\ &< s/10 + s/10 + s/10 \\ &< s \,. \end{aligned}$$

□

Proof of Proposition 3.2. Lemma 4.7 shows that the proposition is satisfied for all $\mathbf{t}$ with $\|\mathbf{t}\| \le \varepsilon^{1/8}/(1000n)$ with probability at least $1 - 2^{-\Omega(n)}$. So, we consider the case when $\varepsilon^{1/8}/(1000n) \le \|\mathbf{t}\| \le \delta_{\mathsf{max}} s_\varepsilon$. By Lemma 2.14, for such $\mathbf{t}$,

$$f(\mathbf{t}) \ge \rho(\mathbf{t}) \ge e^{-\pi \delta_{\mathsf{max}}^2 s_\varepsilon^2} > \varepsilon^{\delta_{\mathsf{max}}^2}/2 \ge \varepsilon^{1/4}/2 \,. \tag{13}$$

We first show that the estimators $f_W, \nabla f_W$ are close to their expectations.

1. By Lemma 4.10, we have that $\|\nabla f_W(\mathbf{t}) - \nabla f(\mathbf{t})\| \le \varepsilon^{1/4}\|\mathbf{t}\|/100$ holds simultaneously for all $\mathbf{t}$ with $\varepsilon^{1/8}/(1000n) \le \|\mathbf{t}\| \le \delta_{\mathsf{max}} s_\varepsilon$ with probability at least $1 - 2^{-\Omega(\varepsilon^{1/2} N)} = 1 - 2^{-\Omega(n)}$.
2. By Lemma 4.12, we have that $|f_W(\mathbf{t}) - f(\mathbf{t})| \le \varepsilon^{1/4}/100$ holds simultaneously for all relevant $\mathbf{t}$ with probability at least $1 - 2^{-\Omega(\varepsilon^{1/2} N)} = 1 - 2^{\Omega(n)}$.

Suppose that both of these bounds hold, which happens with probability at least $1 - 2^{\Omega(n)}$. Then, applying Claim 4.8 with $\gamma = \varepsilon^{1/4}/100$, we have that for all relevant $\mathbf{t}$,

$$\begin{aligned} \Big\| \frac{\nabla f_W(\mathbf{t})}{2\pi f_W(\mathbf{t})} + \mathbf{t} \Big\| &\le \frac{2\gamma}{\rho(\mathbf{t})}\|\mathbf{t}\| + \Big\| \frac{\nabla f(\mathbf{t})}{2\pi f(\mathbf{t})} + \mathbf{t} \Big\| \\ &\le \frac{\varepsilon^{1/4}}{50} \cdot e^{\pi\|\mathbf{t}\|^2}\|\mathbf{t}\| + 12(\varepsilon/2)^{1-2\delta(\mathbf{t})}\|\mathbf{t}\| \qquad \text{(Corollary 4.3)} \\ &\le \frac{\varepsilon^{1/4}}{50} \cdot \Big(\frac{2(1+\varepsilon)}{\varepsilon}\Big)^{\delta(\mathbf{t})^2}\|\mathbf{t}\| + 12\varepsilon^{1-2\delta(\mathbf{t})}\|\mathbf{t}\| \\ &\le \frac{\varepsilon^{1/4-\delta(\mathbf{t})^2}}{40} \cdot \|\mathbf{t}\| + \frac{9\varepsilon^{(1-2\delta(\mathbf{t}))/4}}{10} \cdot \|\mathbf{t}\| \\ &\le \varepsilon^{(1-2\delta(\mathbf{t}))/4}\|\mathbf{t}\| \,, \end{aligned}$$

as needed. In the next-to-last inequality we used the straightforward inequality $12\varepsilon^{3(1-2\delta_{\mathsf{max}})/4} = 12\exp(-3\log(1/\varepsilon)/(\pi s_\varepsilon^2)) < 9/10$. □

5 Reduction from CVPP to CVPP with a Promise

In this section, we present our Kannan-style reductions from γ'-CVPP to γ-CVPP$^\phi$.

Theorem 5.1. *Let $\gamma(n) \geq 1$, and let $\alpha(n) > 0$ be a non-increasing function. Then, a polynomial-time algorithm that solves γ-CVPP$^\phi$, where $\phi(\mathcal{L}) = \alpha(n) \cdot \lambda_1(\mathcal{L})$ for any lattice $\mathcal{L}$ of rank n, implies a polynomial-time algorithm that solves γ'-CVPP, where*

$$\gamma'(n) := \max_{i \in \{0,\ldots,n\}} \left(\gamma(n-i)^2 + \frac{i}{4\alpha(n)^2}\right)^{1/2}$$

with the convention that $\gamma(0) = 0$. In particular, if $\alpha(n) \leq 1/2$ and $\gamma(n) = \sqrt{n}/(2\alpha(n))$, we have $\gamma'(n) = \gamma(n)$.

The reduction of Theorem 5.1 uses as preprocessing an HKZ basis and the preprocessing of the underlying γ-CVPP$^\phi$ algorithm on n lattices of dimension $O(n)$, so it incurs a blowup of roughly n in the size of the preprocessing. We now present a more elaborate reduction based on similar ideas that incurs almost no blowup in the size of preprocessing for an appropriate setting of parameters.

Theorem 5.2. *Let $0 < \alpha(n) \leq 1/2$ be a non-increasing function and $g(n) \geq 1$ be a non-decreasing function. Let $\gamma(n) = g(n)^{h(n)}/(2\alpha(n))$ where $0 \leq h(n) < n$ is a non-decreasing integer-valued function satisfying $g(n)^{h(n)-1} \leq \sqrt{n}$. Let $\gamma'(n) = g(n)\sqrt{n}/(2\alpha(n))$ and $\phi(\mathcal{L}) = \alpha(n)\lambda_1(\mathcal{L})$ for any lattice $\mathcal{L}$ of rank n. Then, a polynomial-time algorithm that solves γ-CVPP$^\phi$ implies a polynomial-time algorithm that solves γ'-CVPP using as preprocessing only an HKZ basis of the input lattice and the preprocessing of the γ-CVPP$^\phi$ algorithm for a collection of lattices $\{\mathcal{L}_k\}$ with $\sum \dim \mathcal{L}_k \leq n \cdot (h(n)+1)$, where n is the dimension of the input lattice.*

Of particular interest to us is the special case $g(n) = 1$ and $h(n) = 0$ in Theorem 5.2, which we highlight in the following corollary. With these parameters, the reduction achieves $\sum \dim \mathcal{L}_k = n$, which is intuitively optimal.

Corollary 5.3. *Let $0 < \alpha(n) < 1/2$ be a non-increasing function and define $\gamma(n) = \sqrt{n}/(2\alpha(n))$. Then, there is a polynomial-time reduction from γ-CVPP to α-BDDP that uses as preprocessing an HKZ basis of the input lattice and the preprocessing of the α-BDDP algorithm for a collection of lattices $\{\mathcal{L}_k\}$ with $\sum \dim \mathcal{L}_k = n$, where n is the dimension of the input lattice.*

Another interesting special case, obtained by setting $g(n) = 2$ and $h(n) = \lfloor (\log_2 n)/2 \rfloor + 1$, gives a reduction that matches the approximation factor γ achieved by Theorem 5.1 up to a factor of 2 but incurs only a logarithmic blow-up in preprocessing, $\sum \dim \mathcal{L}_k \leq O(n \log n)$ (as opposed to linear). Finally, setting $g(n) = n^{1/(2m)}$ and $h(n) = m + 1$ for any integer $m \geq 1$ gives a reduction with $\gamma'(n) = \gamma(n) = n^{1/2+1/(2m)}/(2\alpha(n))$ that achieves $O(m)$ blow-up, $\sum \dim \mathcal{L}_k \leq (m+2) \cdot n$.

Lastly, we show that similar ideas can be made to work without preprocessing with worse parameters.

Proposition 5.4. *Let $\gamma(n) \geq g(n)\sqrt{n+3}/2$ where $g(n) \geq 1$ is a non-decreasing function. Let $\phi(\mathcal{L}) = \lambda_1(\mathcal{L})$ for any lattice. Then, there is a polynomial-time reduction from γ-CVP to g-CVP$^\phi$.*

5.1 Proof of Theorem 5.1

Proof of Theorem 5.1. Suppose that we have an efficient algorithm that solves γ-CVPP$^\phi$ with preprocessing algorithm P and query algorithm Q. We assume without loss of generality that $\gamma(1) = 1$. We construct an algorithm that solves γ'-CVPP as follows.

On input $\mathcal{L} \subset \mathbb{R}^n$, the preprocessing algorithm first computes an HKZ basis $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of $\mathcal{L}$. For $i = 0, \ldots, n$, let $\pi_i = \pi_{\{\mathbf{b}_1, \ldots, \mathbf{b}_i\}^\perp}$ and $\mathcal{N}_i = \pi_i(\mathcal{L})$. Then, the preprocessing algorithm returns as its advice $\mathbf{B}$ and the advice strings $A_i = \mathrm{P}(\mathcal{N}_i)$ for all i.

On input $\mathbf{t} \in \mathbb{R}^n$, the query algorithm does the following for each $i = 0, \ldots, n$. It computes $\mathbf{x}_i = \mathrm{Q}(A_i, \pi_i(\mathbf{t})) \in \mathcal{N}_i$. Write $\mathbf{x}_i = \sum_{j=i+1}^n a_{i,j} \pi_i(\mathbf{b}_j)$ for some coefficients $a_{i,j} \in \mathbb{Z}$ and let $\mathbf{y}_i = \sum_{j=i+1}^n a_{i,j} \mathbf{b}_j \in \mathcal{L}$ be a "lift" of $\mathbf{x}_i$. Let $\mathcal{M}_i = \mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_i) \subseteq \mathcal{L}$ and

$$\mathbf{z}_i = \textsc{Babai}(\pi_{\mathrm{span}(\mathcal{M}_i)}(\mathbf{t} - \mathbf{y}_i), (\mathbf{b}_1, \ldots, \mathbf{b}_i)) \in \mathcal{L}.$$

The query algorithm then returns the vector nearest to the target $\mathbf{t}$ among the vectors $\mathbf{y}_i + \mathbf{z}_i \in \mathcal{L}$. In other words, for each $i = 0, \ldots, n$, we use Babai to compute a close point to $\mathbf{t}$ in $\mathcal{M}_i + \mathbf{y}_i = \{\mathbf{w} \in \mathcal{L} : \pi_i(\mathbf{w}) = \mathbf{x}_i\} \subseteq \mathcal{L}$, and output the closest.

Clearly, the advice from preprocessing has polynomial length and the query algorithm runs in polynomial time. Let $i \in \{0, \ldots, n-1\}$ be minimal such that $\mathrm{dist}(\pi_i(\mathbf{t}), \mathcal{N}_i) < \phi(\mathcal{N}_i) = \alpha(n - i) \cdot \|\widetilde{\mathbf{b}}_{i+1}\|$, where $(\widetilde{\mathbf{b}}_1, \ldots, \widetilde{\mathbf{b}}_n)$ is the Gram-Schmidt orthogonalization of $\mathbf{B}$. If no such i exists, we take i to be n. We will complete the proof by showing that $\mathbf{y}_i + \mathbf{z}_i$ is close to $\mathbf{t}$. By separating the norm into its projection on the two orthogonal subspaces,

$$\begin{aligned}\|\mathbf{y}_i + \mathbf{z}_i - \mathbf{t}\|^2 &= \|\pi_i(\mathbf{y}_i - \mathbf{t})\|^2 + \|\pi_{\mathrm{span}(\mathcal{M}_i)}(\mathbf{y}_i + \mathbf{z}_i - \mathbf{t})\|^2 \\ &= \|\mathbf{x}_i - \pi_i(\mathbf{t})\|^2 + \|\mathbf{z}_i - \pi_{\mathrm{span}(\mathcal{M}_i)}(\mathbf{t} - \mathbf{y}_i)\|^2 .\end{aligned}$$

For the first term, using the definition of γ-CVPP$^\phi$ and our choice of i, we have that

$$\|\mathbf{x}_i - \pi_i(\mathbf{t})\|^2 \leq \gamma(n-i)^2 \,\mathrm{dist}(\pi_i(\mathbf{t}), \mathcal{N}_i)^2 \leq \gamma(n-i)^2 \,\mathrm{dist}(\mathbf{t}, \mathcal{L})^2 .$$

For the second term, by Lemma 2.26 and again by our choice of i,

$$\begin{aligned}\|\mathbf{z}_i - \pi_{\mathrm{span}(\mathcal{M}_i)}(\mathbf{t} - \mathbf{y}_i)\|^2 &\leq \frac{i}{4} \max_{j<i} \|\widetilde{\mathbf{b}}_{j+1}\|^2 \\ &\leq \frac{i}{4} \max_{j<i} \frac{1}{\alpha(n-j)^2} \,\mathrm{dist}(\pi_j(\mathbf{t}), \mathcal{N}_j)^2 \\ &\leq \frac{i}{4\alpha(n)^2} \cdot \mathrm{dist}(\mathbf{t}, \mathcal{L})^2 .\end{aligned}$$

The theorem follows by combining the two inequalities. □

5.2 Proof of Theorem 5.2

Proof of Theorem 5.2. Suppose that we have an algorithm that solves γ-CVPP$^\phi$ in polynomial time with preprocessing algorithm P and query algorithm Q. We construct an algorithm that solves γ'-CVPP as follows.

On input $\mathcal{L} \subset \mathbb{R}^n$ a lattice of rank n, the preprocessing algorithm first computes an HKZ basis $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of $\mathcal{L}$ with Gram-Schmidt orthogonalization $(\widetilde{\mathbf{b}}_1, \ldots, \widetilde{\mathbf{b}}_n)$. Fix $r = h(n) + 1$ and $c = g(n)$. We define a series of indices $n = i_0 > i_1 > i_2 > \cdots > i_\ell = 0$ in the following recursive way: for each k such that $i_k > 0$, define $0 \leq i_{k+1} < i_k$ to be minimal such that

$$\|\widetilde{\mathbf{b}}_{i_{k+1}+1}\| \geq \max_{j \leq i_k} \|\widetilde{\mathbf{b}}_j\| / c ,$$

or equivalently, the largest such that

$$\max_{j \le i_{k+1}} \|\widetilde{\mathbf{b}}_j\| < \max_{j \le i_k} \|\widetilde{\mathbf{b}}_j\| / c \ . \tag{14}$$

Notice that we have

$$\max_{j \le i_k} \|\widetilde{\mathbf{b}}_j\| / c \le \|\widetilde{\mathbf{b}}_{i_{k+1}+1}\| \le \max_{j \le i_k} \|\widetilde{\mathbf{b}}_j\| \ . \tag{15}$$

Let $\pi_k = \pi_{\{\mathbf{b}_1,\ldots,\mathbf{b}_{i_k}\}^\perp}$ and $\mathcal{L}_k = \pi_k(\mathcal{L}(\mathbf{b}_{i_k+1},\ldots,\mathbf{b}_{i_{\max(k-r,0)}}))$. Then, the preprocessing algorithm returns as its advice $\mathbf{B}$ and the advice strings $A_k = \mathrm{P}(\mathcal{L}_k)$ for all k. Notice that each vector $\mathbf{b}_j$ is included in the definition of $\mathcal{L}_k$ for at most $r = h(n)+1$ different values of k. As a result, $\sum \dim \mathcal{L}_k \le n \cdot (h(n)+1)$ as claimed.

Let $\mathcal{N}_k = \pi_k(\mathcal{L})$. Before describing the query algorithm, we define a key recursive sub-procedure $\mathrm{S}(\mathbf{t},k)$ that will be used to find solutions to γ-CVP$^\phi(\pi_k(\mathbf{t}),\mathcal{N}_k)$. On input $\mathbf{t}$ and k, if $k \le r$, then S simply outputs $\mathrm{Q}(A_k,\pi_k(\mathbf{t}))$. Otherwise, it calls itself recursively, setting $\mathbf{x} = \mathrm{S}(\mathbf{t},k-r) \in \mathcal{N}_{k-r}$. Write $\mathbf{x} = \sum_{j=i_{k-r}+1}^{n} a_j \pi_{k-r}(\mathbf{b}_j)$, and let $\mathbf{y} = \sum_{j=i_{k-r}+1}^{n} a_j \pi_k(\mathbf{b}_j) \in \mathcal{N}_k$ be a "lift" of $\mathbf{x}$. Then S outputs $\mathbf{z} = \mathrm{Q}(A_k, \pi_k(\mathbf{t}) - \mathbf{y}) + \mathbf{y}$. In other words, S uses Q to find a close point to $\pi_k(\mathbf{t})$ in $\mathcal{L}_k + \mathbf{y} = \{\mathbf{w} \in \mathcal{N}_k : \pi_{k-r}(\mathbf{w}) = \mathbf{x}\} \subseteq \mathcal{N}_k$ and outputs it.

On input $\mathbf{t} \in \mathbb{R}^n$, the query algorithm does the following for each k. It first computes $\mathbf{x}_k = \mathrm{S}(\mathbf{t},k) \in \mathcal{N}_k$. Let $\mathbf{y}_k \in \mathcal{L}$ be a "lift" of $\mathbf{x}_k$. Let $\mathcal{M}_k = \mathcal{L}(\mathbf{b}_1,\ldots,\mathbf{b}_{i_k}) \subseteq \mathcal{L}$ and

$$\mathbf{z}_k = \textsc{Babai}(\pi_{\mathrm{span}(\mathcal{M}_k)}(\mathbf{t}-\mathbf{y}_k), \mathcal{M}_k) \in \mathcal{L}.$$

The query algorithm then returns the vector nearest to the target $\mathbf{t}$ among the vectors $\mathbf{y}_k + \mathbf{z}_k \in \mathcal{L}$. In other words, for each k, we use BABAI to compute a close point to $\mathbf{t}$ in $\mathcal{M}_k + \mathbf{y}_k = \{\mathbf{w} \in \mathcal{L} : \pi_k(\mathbf{w}) = \mathbf{x}_k\} \subseteq \mathcal{L}$, and output the closest. It is clear that the algorithm runs in polynomial time.

First, assume that $\mathrm{S}(\mathbf{t},k)$ returns a valid solution to γ-CVP$^\phi(\pi_k(\mathbf{t}),\mathcal{N}_k)$. Then, the proof of correctness proceeds nearly identically to that of Theorem 5.1. In particular, let $k > 0$ be maximal such that $\mathrm{dist}(\mathbf{t},\mathcal{L}) < \alpha(n)\|\widetilde{\mathbf{b}}_{i_k+1}\|$. If no such k exists, we take $k = 0$. As in the previous proof,

$$\|\mathbf{y}_k + \mathbf{z}_k - \mathbf{t}\|^2 = \|\pi_k(\mathbf{y}_k - \mathbf{t})\|^2 + \|\mathbf{z}_k - \pi_{\mathrm{span}(\mathcal{M}_k)}(\mathbf{t}-\mathbf{y}_k)\|^2 \ .$$

For the first term, since $\mathrm{dist}(\pi_k(\mathbf{t}),\mathcal{N}_k) \le \mathrm{dist}(\mathbf{t},\mathcal{L}) < \alpha(n)\|\widetilde{\mathbf{b}}_{i_k+1}\| \le \phi(\mathcal{N}_k)$, we have

$$\begin{aligned}\|\mathbf{x}_k - \pi_k(\mathbf{t})\|^2 &\le \gamma(n-i_k)^2 \,\mathrm{dist}(\pi_k(\mathbf{t}),\mathcal{N}_k)^2 \\ &\le \gamma'(n-i_k)^2 \,\mathrm{dist}(\pi_k(\mathbf{t}),\mathcal{N}_k)^2 \le c^2 \frac{n-i_k}{4\alpha(n)^2} \,\mathrm{dist}(\mathbf{t},\mathcal{L})^2 \ .\end{aligned}$$

For the second term, by Lemma 2.26, Eq. (15), and our choice of k,

$$\|\mathbf{z}_k - \pi_{\mathrm{span}(\mathcal{M}_k)}(\mathbf{t}-\mathbf{y}_k)\|^2 \le \frac{i_k}{4} \max_{j \le i_k} \|\widetilde{\mathbf{b}}_j\|^2 \le c^2 \frac{i_k}{4} \|\widetilde{\mathbf{b}}_{i_{k+1}+1}\|^2 \le c^2 \frac{i_k}{4\alpha(n)^2} \,\mathrm{dist}(\mathbf{t},\mathcal{L})^2 \ .$$

Combining the two inequalities, we get $\|\mathbf{y}_k + \mathbf{z}_k - \mathbf{t}\| \le \gamma'(n)\,\mathrm{dist}(\mathbf{t},\mathcal{L})$.

It remains to show that the sub-procedure $\mathrm{S}(\mathbf{t},k)$ returns a valid solution to γ-CVP$^\phi(\pi_k(\mathbf{t}),\mathcal{N}_k)$. We prove this by induction. If $k \le r$, the claim follows immediately from the fact that $\mathcal{L}_k = \mathcal{N}_k$. Otherwise, we claim that $\mathcal{L}_k + \mathbf{y}$ contains the closest vector to $\pi_k(\mathbf{t})$ in $\mathcal{N}_k$. This claim immediately implies the correctness of S using the correctness of Q and the fact that $\gamma(\dim \mathcal{L}_k) \le \gamma(\dim \mathcal{N}_k)$

and $\phi(\mathcal{L}_k) \geq \phi(\mathcal{N}_k)$. To prove the claim, first notice from Eqs. (14) and (15) that $\|\widetilde{\mathbf{b}}_{i_k+1}\| \leq \|\widetilde{\mathbf{b}}_{i_{k-r}+1}\|/c^{r-1}$, and so

$$\operatorname{dist}(\pi_k(\mathbf{t}), \mathcal{N}_k) < \phi(\mathcal{N}_k) = \alpha(n-i_k)\|\widetilde{\mathbf{b}}_{i_k+1}\| \leq \frac{\alpha(n-i_k)}{c^{r-1}} \cdot \|\widetilde{\mathbf{b}}_{i_{k-r}+1}\| = \frac{\alpha(n-i_k)}{c^{r-1}} \cdot \lambda_1(\mathcal{N}_{k-r}) \,. \quad (16)$$

As a result, $\operatorname{dist}(\pi_{k-r}(\mathbf{t}), \mathcal{N}_{k-r}) \leq \operatorname{dist}(\pi_k(\mathbf{t}), \mathcal{N}_k) < \phi(\mathcal{N}_{k-r})$, and so by the induction hypothesis and Eq. (16),

$$\begin{aligned}
\|\mathbf{x} - \pi_{k-r}(\mathbf{t})\| &\leq \gamma(n-i_{k-r}) \operatorname{dist}(\pi_{k-r}(\mathbf{t}), \mathcal{N}_{k-r}) \\
&\leq \frac{c^{r-1}}{2\alpha(n-i_k)} \operatorname{dist}(\pi_{k-r}(\mathbf{t}), \mathcal{N}_{k-r}) \\
&\leq \frac{c^{r-1}}{2\alpha(n-i_k)} \operatorname{dist}(\pi_k(\mathbf{t}), \mathcal{N}_k) \\
&< \frac{\lambda_1(\mathcal{N}_{k-r})}{2} \,.
\end{aligned}$$

So, $\mathbf{x}$ is the unique closest vector in $\mathcal{N}_{k-r}$ to $\pi_{k-r}(\mathbf{t})$. Finally, by Eq. (16), $\operatorname{dist}(\pi_k(\mathbf{t}), \mathcal{N}_k) < \lambda_1(\mathcal{N}_{k-r})/2$, yet all vectors $\mathbf{y}' \in \mathcal{N}_k \setminus (\mathcal{L}_k + \mathbf{y})$ must be at distance at least

$$\|\pi_{k-r}(\mathbf{y}') - \pi_{k-r}(\mathbf{t})\| \geq \lambda_1(\mathcal{N}_{k-r}) - \operatorname{dist}(\pi_{k-r}(\mathbf{t}), \mathcal{N}_{k-r}) > \frac{\lambda_1(\mathcal{N}_{k-r})}{2}$$

from $\pi_k(\mathbf{t})$ and hence cannot be closest to $\pi_k(\mathbf{t})$ in $\mathcal{N}_k$. □

5.3 Proof of Proposition 5.4

Proof of Proposition 5.4. Let A be an algorithm solving g-CVP$^{\phi}$. We say that a basis $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of $\mathcal{L}$ is a g-HKZ basis if $\|\mathbf{b}_1\| \leq g(n)\lambda_1(\mathcal{L})$ and $(\pi_{\{\mathbf{b}_1\}^\perp}(\mathbf{b}_2), \ldots, \pi_{\{\mathbf{b}_1\}^\perp}(\mathbf{b}_n))$ is a g-HKZ basis. Note that Theorem 2.8 immediately implies that A can be used to compute a g-HKZ basis in polynomial time.

On input $\mathcal{L}$ and target vector $\mathbf{t}$, first use A to compute a g-HKZ basis, $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of $\mathcal{L}$. Then, as in the proof of Theorem 5.1, for $i = 0, \ldots, n$, let $\pi_i = \pi_{\{\mathbf{b}_1, \ldots, \mathbf{b}_i\}^\perp}$ and $\mathcal{N}_i = \pi_i(\mathcal{L})$. Compute $\mathbf{x}_i = \mathrm{A}(\pi_i(\mathbf{t}), \mathcal{N}_i) \in \mathcal{N}_i$ and lift it to a vector $\mathbf{y}_i \in \mathcal{L}$. Similarly, let $\mathcal{M}_i = \mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_i) \subseteq \mathcal{L}$ and

$$\mathbf{z}_i = \textsc{Babai}(\pi_{\operatorname{span}(\mathcal{M}_i)}(\mathbf{t} - \mathbf{y}_i), (\mathbf{b}_1, \ldots, \mathbf{b}_i)) \in \mathcal{L}.$$

Finally, return the vector nearest to the target $\mathbf{t}$ among the vectors $\mathbf{y}_i + \mathbf{z}_i \in \mathcal{L}$.

Let $i \in \{0, \ldots, n-1\}$ be minimal such that $\operatorname{dist}(\pi_i(\mathbf{t}), \mathcal{N}_i) < \|\widetilde{\mathbf{b}}_{i+1}\|/g(n-i)$. If no such i exists, we take $i = n$. As in the proof of Theorem 5.1,

$$\|\mathbf{y}_i + \mathbf{z}_i - \mathbf{t}\|^2 = \|\mathbf{x}_i - \pi_i(\mathbf{t})\|^2 + \|\mathbf{z}_i - \pi_{\operatorname{span}(\mathcal{M}_i)}(\mathbf{t} - \mathbf{y}_i)\|^2 \,.$$

By our choice of i and the definition of a g-HKZ basis, $\operatorname{dist}(\pi_i(\mathbf{t}), \mathcal{N}_i) < \lambda_1(\mathcal{N}_i) = \phi(\mathcal{N}_i)$, so A is guaranteed to output $\mathbf{x}_i$ satisfying

$$\|\mathbf{x}_i - \pi_i(\mathbf{t})\|^2 \leq g(n-i)^2 \operatorname{dist}(\pi_i(\mathbf{t}), \mathcal{N}_i)^2 \leq g(n-i)^2 \operatorname{dist}(\mathbf{t}, \mathcal{L})^2 \,,$$

where we define $g(0) = 0$. By Lemma 2.26,

$$\|\mathbf{z}_i - \pi_{\mathrm{span}(\mathcal{M}_i)}(\mathbf{t} - \mathbf{y}_i)\|^2 \le \frac{i}{4} \max_{j<i} \|\widetilde{\mathbf{b}}_{j+1}\|^2 \le \frac{i}{4} \max_{j<i} g(n-j)^2 \operatorname{dist}(\pi_j(\mathbf{t}), \mathcal{N}_j)^2 \le \frac{i}{4} g(n)^2 \operatorname{dist}(\mathbf{t}, \mathcal{L})^2 .$$

Combining the two inequalities gives

$$\begin{aligned} \|\mathbf{y}_i + \mathbf{z}_i - \mathbf{t}\|^2 &\le g(n-i)^2 \operatorname{dist}(\mathbf{t}, \mathcal{L})^2 + \frac{i}{4} g(n)^2 \operatorname{dist}(\mathbf{t}, \mathcal{L})^2 \\ &\le \frac{n+3}{4} \cdot g(n)^2 \operatorname{dist}(\mathbf{t}, \mathcal{L})^2 \\ &= \gamma(n)^2 \operatorname{dist}(\mathbf{t}, \mathcal{L})^2 \end{aligned}$$

as claimed. □

6 Reduction to bounded distance using sparsification

In this section we prove Theorem 6.1, our second reduction to the bounded distance case.

Theorem 6.1. *For any $\tau = \tau(n) > 0$ and $\gamma = \gamma(n) \ge 1$, there is a randomized polynomial-time reduction from $\gamma \cdot \sqrt{1+\tau^2}$-CVP to γ-CVP$^{\phi}$ where $\phi(\mathcal{L}) = \sqrt{1+\tau^{-2}} \cdot \lambda_1(\mathcal{L})$.*

Note that for $\tau \ge \sqrt{n-1}/2$, Proposition 5.4 provides a strictly stronger reduction. The above theorem and Theorem 2.6 (the NP-hardness of $n^{c/\log\log n}$-CVP) immediately imply a hardness result for γ-CVP$^{\phi}$.

Corollary 6.2. *There exists a constant $c > 0$ such that γ-CVP$^{\phi}$ is NP-hard for $\phi(\mathcal{L}) = (1 + n^{-c/\log\log n}) \cdot \lambda_1(\mathcal{L})$ and $\gamma = n^{c/\log\log n}$.*

We follow the sparsification idea of [DK13]. Basically, given a target $\mathbf{t}$, we try to find a sublattice $\mathcal{L}'$ of $\mathcal{L}$, such that $\mathcal{L}'$ has minimum distance proportional to $\operatorname{dist}(\mathbf{t}, \mathcal{L}')$ with $\operatorname{dist}(\mathbf{t}, \mathcal{L}')$ not much larger than $\operatorname{dist}(\mathbf{t}, \mathcal{L})$. Notice that the first condition is needed to ensure that a distance-bounded CVP solver will succeed on $\mathcal{L}'$ and $\mathbf{t}$, and the second condition allows us to bound the loss in approximation when passing from $\mathcal{L}$ to $\mathcal{L}'$. Implicit in the work of [DK13] is the fact that a random sublattice $\mathcal{L}'$ of $\mathcal{L}$ of index p (for an appropriate p) will work. To obtain the approximation factor stated in the theorem, we actually work with a random coset of $\mathcal{L}'$, and we also do a slightly more careful analysis in order to avoid the loss incurred by a triangle inequality.

For a full rank lattice $\mathcal{L}$ with basis $\mathbf{B}$, a prime p, a vector $\mathbf{z} \in \mathbb{Z}_p^n$, and $c \in \mathbb{Z}_p$, we define

$$\mathcal{L}_{p,c}(\mathbf{B}, \mathbf{z}) = \{\mathbf{y} \in \mathcal{L} : \langle \mathbf{z}, \mathbf{B}^{-1}\mathbf{y} \rangle = c \pmod p\} ,$$

and $\mathcal{L}_p(\mathbf{B}, \mathbf{z}) = \mathcal{L}_{p,0}(\mathbf{B}, \mathbf{z})$. Note that $\mathcal{L}_p(\mathbf{B}, \mathbf{z})$ is a sublattice of $\mathcal{L}$ and $\mathcal{L}_{p,c}(\mathbf{B}, \mathbf{z})$ is a coset of $\mathcal{L}_p(\mathbf{B}, \mathbf{z})$. We wish to argue that, for any $\mathbf{t}$ and appropriate p, if $\mathbf{z}$ and c are chosen uniformly at random, then with constant positive probability, $\lambda_1(\mathcal{L}_p(\mathbf{B}, \mathbf{z}))$ will be relatively large but $\operatorname{dist}(\mathbf{t}, \mathcal{L}_{p,c}(\mathbf{B}, \mathbf{z}))$ will be relatively close to $\operatorname{dist}(\mathbf{t}, \mathcal{L})$. The next lemma is a modification of [DK13, Lemma 4.3] more suited to our purposes and is the key to the reduction.

Lemma 6.3. *Let $r > 0$, $\mathcal{L} \subset \mathbb{R}^n$ a full rank lattice with basis $\mathbf{B}$, $N = |\mathcal{L} \cap rB_2^n|$, and $p > N$ a prime. Let $\mathbf{z}$ be sampled uniformly from $\mathbb{Z}_p^n$, and define*

$$C = \{c \in \mathbb{Z}_p : |\mathcal{L}_{p,c}(\mathbf{B}, \mathbf{z}) \cap rB_2^n| > 0\} .$$

Then,

- $\Pr_{\mathbf{z}}[\lambda_1(\mathcal{L}_p(\mathbf{B}, \mathbf{z})) \leq r] \leq \dfrac{N}{p}$, *and*
- $\Pr_{\mathbf{z}}\left[|C| \leq \varepsilon \cdot \dfrac{N}{p+N-1} \cdot p\right] \leq \varepsilon$ *for any* $\varepsilon \in (0,1)$.

Proof. First, we wish to show that for any $\mathbf{x} \in (\mathcal{L} \cap 2rB_2^n) \setminus \{\mathbf{0}\}$, $\langle \mathbf{z}, \mathbf{B}^{-1}\mathbf{x} \rangle$ is uniformly distributed mod p over the choice of $\mathbf{z}$. Let $\mathbf{x} = \sum y_i \mathbf{b}_i$ and $\mathbf{z} = (z_1, \ldots, z_n)$. Then, $\langle \mathbf{z}, \mathbf{B}^{-1}\mathbf{x} \rangle = \sum z_i y_i$. So, it suffices to show that at least one y_i is not 0 mod p, or equivalently that $\mathbf{x} \notin p\mathcal{L}$. Suppose $\mathbf{x} \in p\mathcal{L}$. Then there is some $\mathbf{x}' \in \mathcal{L}$ such that $p\mathbf{x}' = \mathbf{x}$, so clearly the vectors $\lfloor -p/2 \rfloor \mathbf{x}', \lfloor -p/2 + 1 \rfloor \mathbf{x}', \ldots, \mathbf{0}, \ldots, \lfloor p/2 \rfloor \mathbf{x}'$ are all in $(\mathcal{L} \cap rB_2^n)$. This contradicts the fact that there are exactly $N < p$ vectors in $(\mathcal{L} \cap rB_2^n)$. It follows that $\langle \mathbf{z}, \mathbf{B}^{-1}\mathbf{x} \rangle$ is uniformly distributed mod p over the choice of $\mathbf{z}$.

Now, to prove the first result, let $\mathbf{x} \in (\mathcal{L} \cap rB_2^n) \setminus \{\mathbf{0}\}$. Since $\langle \mathbf{z}, \mathbf{B}^{-1}\mathbf{x} \rangle$ is uniformly distributed mod p, $\Pr_{\mathbf{z}}[\mathbf{x} \in \mathcal{L}_p(\mathbf{B}, \mathbf{z})] = 1/p$. We simply apply union bound and recall the definition of N to get $\Pr_{\mathbf{z}}[\lambda_1(\mathcal{L}_p(\mathbf{B}, \mathbf{z})) \leq r] \leq N/p$ as claimed.

To prove the second result, for $c \in C$ let $S_c = \mathcal{L}_{p,c}(\mathbf{B}, \mathbf{z}) \cap rB_2^n$, and let

$$A = \bigcup_{c \in C} S_c^2 ,$$

the set of pairs of short vectors in the same coset. Then, recalling the definition of N and applying Cauchy-Schwarz,

$$N^2 = \Big(\sum_{c \in C} |S_c| \Big)^2 \leq \Big(\sum_{c \in C} 1 \Big) \Big(\sum_{c \in C} |S_c|^2 \Big) = |C| \cdot |A| .$$

Therefore $|C| \geq N^2/|A|$. So, it suffices to bound $\Pr[|A| \geq N \cdot (p+N-1)/(\varepsilon p)]$.

Let $\mathbf{x}, \mathbf{x}' \in (\mathcal{L} \cap rB_2^n)$ be distinct. Since $\mathbf{x} - \mathbf{x}' \in (\mathcal{L} \cap 2rB_2^n) \setminus \{\mathbf{0}\}$, it follows that $\langle \mathbf{z}, \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}') \rangle$ is uniformly distributed mod p over the choice of $\mathbf{z}$. So, $\Pr[\langle \mathbf{z}, \mathbf{B}^{-1}\mathbf{x} \rangle = \langle \mathbf{z}, \mathbf{B}^{-1}\mathbf{x}' \rangle \pmod p] = 1/p$. Therefore,

$$\mathop{\mathbb{E}}_{\mathbf{z}}[|A|] = N + N(N-1)/p = N \cdot \frac{p+N-1}{p} .$$

Applying Markov's inequality,

$$\Pr_{\mathbf{z}}\left[|A| \geq N \cdot \frac{p+N-1}{\varepsilon p}\right] \leq \varepsilon ,$$

and the result follows. □

Proof of Theorem 6.1. Let A be an algorithm that solves γ-CVP$^{\phi}$. Our input is a lattice $\mathcal{L} \subset \mathbb{Q}^n$ with basis $\mathbf{B}$ and target vector $\mathbf{t} \in \mathbb{R}^n$. We assume without loss of generality that $\mathcal{L}$ is full rank. Let $r = \tau \operatorname{dist}(\mathbf{t}, \mathcal{L})$, and $N = |\mathcal{L} \cap rB_2^n| > 0$. Our reduction needs to have a prime number p satisfying $2N \leq p \leq 8N$. Since we do not know N, we simply run the reduction with each of polynomially

many values for p, one of which is guaranteed to be in the right range, and then output the closest of all lattice vectors we find. In more detail, assume $\tau < \sqrt{n}$ since otherwise the reduction already follows from Proposition 5.4. By Lemma 2.1 and a simple packing argument, N is at most $2^{\mathrm{poly}(\ell)}$ for some fixed polynomial in the bit length ℓ of the description of $\mathcal{L}$. So it suffices to try for each $i = 1, \ldots, \mathrm{poly}(\ell)$, a prime p with $2^i < p \leq 2^{i+1}$. We now continue with the description of the reduction assuming we know a prime p satisfying $2N \leq p \leq 8N$.

With this, the reduction is straightforward. It samples $\mathbf{z} \in \mathbb{Z}_p^n$ and $c \in \mathbb{Z}_p$ uniformly at random. It then returns $\mathrm{A}(\mathbf{t} - \mathbf{y}, \mathcal{L}_p(\mathbf{B}, \mathbf{z})) + \mathbf{y}$ where $\mathbf{y}$ is an arbitrary point in $\mathcal{L}_{p,c}(\mathbf{B}, \mathbf{z})$. I.e., we find a close vector to $\mathbf{t}$ in the coset $\mathcal{L}_{p,c}(\mathbf{B}, \mathbf{z})$.

By Lemma 6.3, we have that $\lambda_1(\mathcal{L}_p(\mathbf{B}, \mathbf{z})) > r$ and $|C| \geq p/50$ where

$$C = \{c' \in \mathbb{Z}_p : |\mathcal{L}_{p,c'}(\mathbf{B}, \mathbf{z}) \cap rB_2^n| > 0\} ,$$

with probability at least $1/4$ over the choice of $\mathbf{z}$. Suppose both of these hold.

Let $\mathbf{x} \in \mathcal{L}$ be the closest lattice vector to $\mathbf{t}$, and for each coset c', let $\mathbf{y}_{c'} \in \mathcal{L}_{p,c'}(\mathbf{B}, \mathbf{z})$ be a closest vector in $\mathcal{L}_{p,c'}(\mathbf{B}, \mathbf{z})$ to $\mathbf{x}$. If there are multiple choices for $\mathbf{y}_{c'}$, we take one that maximizes $\langle \mathbf{x} - \mathbf{y}_{c'}, \mathbf{x} - \mathbf{t} \rangle$. We wish to argue that, with positive constant probability over the choice of the random coset c, both (1) $\|\mathbf{x} - \mathbf{y}_c\| \leq r$ and (2) $\langle \mathbf{x} - \mathbf{y}_c, \mathbf{x} - \mathbf{t} \rangle \geq 0$ hold. Since $\mathcal{L}_{p,c}(\mathbf{B}, \mathbf{z}) - \mathbf{x}$ is a uniformly distributed random coset, our assumption on $|C|$ implies that at least $p/50$ cosets satisfy condition (1). Let c^* be such that $\mathbf{x} \in \mathcal{L}_{p,c^*}(\mathbf{B}, \mathbf{z})$. Note that for all c', $2\mathbf{x} - \mathbf{y}_{c'}$ is a closest vector to $\mathbf{x}$ in $\mathcal{L}_{p,2c^*-c'}(\mathbf{B}, \mathbf{z})$. It follows that $\|\mathbf{x} - \mathbf{y}_{2c^*-c'}\| = \|\mathbf{x} - \mathbf{y}_{c'}\|$, and if $\langle \mathbf{x} - \mathbf{y}_{c'}, \mathbf{x} - \mathbf{t} \rangle < 0$, then $\langle \mathbf{x} - \mathbf{y}_{2c^*-c'}, \mathbf{x} - \mathbf{t} \rangle \geq \langle \mathbf{x} - (2\mathbf{x} - \mathbf{y}_{c'}), \mathbf{x} - \mathbf{t} \rangle > 0$. It follows that for each coset c' that satisfies (1) but not (2), $2c^* - c'$ satisfies both (1) and (2). Since the map $c' \mapsto 2c^* - c'$ is a bijection on $\mathbb{Z}_p$, we obtain that with probability $1/100$ over the choice of the coset c, both (1) and (2) hold. When this is the case, by expanding the squared norm as an inner product, we have

$$\begin{aligned} \|\mathbf{y}_c - \mathbf{t}\|^2 &= \|(\mathbf{x} - \mathbf{t}) - (\mathbf{x} - \mathbf{y}_c)\|^2 \\ &\leq \|\mathbf{x} - \mathbf{t}\|^2 + \|\mathbf{x} - \mathbf{y}_c\|^2 \\ &\leq (1 + \tau^2)\,\mathrm{dist}(\mathbf{t}, \mathcal{L})^2 . \end{aligned}$$

Finally, note that $\sqrt{1+\tau^2}\,\mathrm{dist}(\mathbf{t}, \mathcal{L}) = \sqrt{1+\tau^{-2}} \cdot \tau\,\mathrm{dist}(\mathbf{t}, \mathcal{L}) < \phi(\mathcal{L}_p(\mathbf{B}, \mathbf{z}))$. So, by the definition of A, the distance of our output from $\mathbf{t}$ is at most

$$\gamma \cdot \|\mathbf{y}_c - \mathbf{t}\| \leq \gamma \cdot \sqrt{1+\tau^2}\,\mathrm{dist}(\mathbf{t}, \mathcal{L}) .$$

It follows that the reduction succeeds with probability at least $1/400$. □

7 Local Maxima of $f(\mathbf{t})$

Claim 7.1. *For any sufficiently large n there exists a lattice $\mathcal{L} \subset \mathbb{R}^n$ such that the function f has a local maximum that is not a global one. Furthermore, the local maximum is at distance $\lambda_1(\mathcal{L})/\sqrt{2}$ from the lattice, and the value of f at this point is exponentially close to 1 (the value at global maxima).*

Proof. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis of $\mathbb{R}^n$, and let $\mathcal{L} = \{\mathbf{z} \in \mathbb{Z}^n : \sum \langle \mathbf{e}_i, \mathbf{z} \rangle \equiv 0 \mod 2\}$. Note that the shortest non-zero vectors of $\mathcal{L}$ are of the form $\mathbf{e}_i + \mathbf{e}_j$, $i \neq j$, and hence $\lambda_1(\mathcal{L}) = \sqrt{2}$. Then, it is easy to see that $\mathcal{L}^* = \mathbb{Z}^n \cup (\mathbb{Z}^n + \mathbf{u})$, where $\mathbf{u} = \sum_{i=1}^n \mathbf{e}_i/2$.

Let $\mathbf{t}$ be any point in $\mathbb{Z}^n \setminus \mathcal{L}$, say $\mathbf{t} = (1, 0, \ldots, 0)$. Note that $\mathrm{dist}(\mathcal{L}, \mathbf{t}) = 1 = \lambda_1(\mathcal{L})/\sqrt{2}$. Since f is a periodic function and $2\mathbf{t} \in \mathcal{L}$, $\nabla f(\mathbf{t}) = \nabla f(-\mathbf{t})$. On the other hand, ∇f is an odd function, and therefore $\nabla f(\mathbf{t}) = \mathbf{0}$.

We will now show that $f(\mathbf{t})$ approaches $f(\mathbf{0}) = 1$ as n approaches ∞ by exploiting the multiplicative structure of ρ on $\mathbb{Z}^n$ and $\mathbb{Z}^n + \mathbf{u}$. In particular, $\rho(\mathbb{Z}^n) = \rho(\mathbb{Z})^n$ and $\rho(\mathbb{Z}^n + \mathbf{u}) = \rho(\mathbb{Z} + 1/2)^n$. So,

$$\begin{aligned} f(\mathbf{t}) &= \mathop{\mathbb{E}}_{\mathbf{w} \sim D_{\mathcal{L}^*}} [\cos(2\pi \langle \mathbf{w}, \mathbf{t} \rangle)] \\ &= \frac{1}{\rho(\mathcal{L}^*)} \cdot (\rho(\mathbb{Z}^n) - \rho(\mathbb{Z}^n + \mathbf{u})) \\ &= \frac{1}{\rho(\mathcal{L}^*)} \cdot \left(\rho(\mathbb{Z})^n - \rho(\mathbb{Z} + 1/2)^n\right) . \end{aligned}$$

Similarly, we have that

$$f(\mathbf{0}) = \frac{1}{\rho(\mathcal{L}^*)} \cdot \left(\rho(\mathbb{Z})^n + \rho(\mathbb{Z} + 1/2)^n\right) = 1 .$$

Since, $\rho(\mathbb{Z} + 1/2) < \rho(\mathbb{Z})$ the difference between $f(\mathbf{0})$ and $f(\mathbf{t})$ is exponentially small in n.

It remains to show that $Hf(\mathbf{t})$ is negative definite. Note that

$$\begin{aligned} Hf(\mathbf{t}) &= -4\pi^2 \mathop{\mathbb{E}}_{\mathbf{w} \sim D_{\mathcal{L}^*}} [\mathbf{w}\mathbf{w}^T \cos(2\pi \langle \mathbf{w}, \mathbf{t} \rangle)] \\ &= -\frac{4\pi^2}{\rho(\mathcal{L}^*)} \cdot \Big(\sum_{\mathbf{z} \in \mathbb{Z}^n} \mathbf{z}\mathbf{z}^T \rho(\mathbf{z}) - \sum_{\mathbf{z} \in \mathbb{Z}^n + \mathbf{u}} \mathbf{z}\mathbf{z}^T \rho(\mathbf{z}) \Big) . \end{aligned}$$

Again exploiting the multiplicative structure of ρ on $\mathbb{Z}^n$ and $\mathbb{Z}^n + \mathbf{u}$, we have

$$\begin{aligned} \sum_{\mathbf{z} \in \mathbb{Z}^n} \mathbf{z}\mathbf{z}^T \rho(\mathbf{z}) &= I_n \cdot \sum_{\mathbf{z} \in \mathbb{Z}^n} z_1^2 \rho(\mathbf{z}) \\ &= I_n \cdot \rho(\mathbb{Z})^{n-1} \sum_{z \in \mathbb{Z}} z^2 \rho(z) . \end{aligned}$$

A similar calculation shows that

$$\sum_{\mathbf{z} \in \mathbb{Z}^n + \mathbf{u}} \mathbf{z}\mathbf{z}^T \rho(\mathbf{z}) = I_n \cdot \rho(\mathbb{Z} + 1/2)^{n-1} \sum_{z \in \mathbb{Z}} (z + 1/2)^2 \rho(z + 1/2) .$$

The result then follows by again noting that $\rho(\mathbb{Z} + 1/2) < \rho(\mathbb{Z})$, so for sufficiently large n, the $\rho(\mathbb{Z})^{n-1}$ term dominates. (In fact, $n = 7$ suffices.) □

References

[AKKV11] M. Alekhnovich, S. Khot, G. Kindler, and N. K. Vishnoi. Hardness of approximating the closest vector problem with pre-processing. *Computational Complexity*, 20(4):741–753, 2011.

[AKS01] M. Ajtai, R. Kumar, and D. Sivakumar. A sieve algorithm for the shortest lattice vector problem. In *Proc. 33rd ACM Symposium on Theory of Computing*, pages 601–610. 2001.

[AR05] D. Aharonov and O. Regev. Lattice problems in NP intersect coNP. *Journal of the ACM*, 52(5):749–765, 2005. Preliminary version in FOCS'04.

[Bab86] L. Babai. On Lovász' lattice reduction and the nearest lattice point problem. *Combinatorica*, 6(1):1–13, 1986.

[Ban93] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625–635, 1993.

[CDLP13] K.-M. Chung, D. Dadush, F.-H. Liu, and C. Peikert. On the lattice smoothing parameter problem. In *Proc. IEEE Conference on Computational Complexity*. 2013.

[DK13] D. Dadush and G. Kun. Lattice sparsification and the approximate closest vector problem. In *SODA*. 2013.

[DKRS03] I. Dinur, G. Kindler, R. Raz, and S. Safra. Approximating CVP to within almost-polynomial factors is NP-hard. *Combinatorica*, 23(2):205–243, 2003.

[FM04] U. Feige and D. Micciancio. The inapproximability of lattice and coding problems with preprocessing. *Journal of Computer and System Sciences*, 69(1):45–67, 2004. Preliminary version in CCC 2002.

[GG00] O. Goldreich and S. Goldwasser. On the limits of nonapproximability of lattice problems. *Journal of Computer and System Sciences*, 60(3):540–563, 2000.

[Hoe63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963.

[HR12] I. Haviv and O. Regev. Tensor-based hardness of the shortest vector problem to within almost polynomial factors. *Theory of Computing*, 8(23):513–531, 2012.

[Kan87] R. Kannan. Minkowski's convex body theorem and integer programming. *Mathematics of Operations Research*, 12(3):pp. 415–440, 1987.

[Kho04] S. Khot. Hardness of approximating the shortest vector problem in lattices. In *Proc. 45th Annual IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 126–135. IEEE, 2004.

[Kho10] S. Khot. Inapproximability results for computational problems on lattices. In P. Q. Nguyen and B. Valle, editors, *The LLL Algorithm*, Information Security and Cryptography, pages 453–473. Springer Berlin Heidelberg, 2010.

[Kle00] P. Klein. Finding the closest lattice vector when it's unusually close. In *Proc. 11th ACM-SIAM Symposium on Discrete Algorithms*, pages 937–941. 2000.

[KPV14] S. Khot, P. Popat, and N. K. Vishnoi. $2^{\log^{1-\varepsilon} n}$ hardness for closest vector problem with preprocessing. *SIAM Journal on Computing*, 43(3):1184–1205, 2014.

[LLL82] A. K. Lenstra, H. W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261:515–534, 1982.

[LLM06] Y.-K. Liu, V. Lyubashevsky, and D. Micciancio. On bounded distance decoding for general lattices. In *International Workshop on Randomization and Computation - Proceedings of RANDOM 2006*, volume 4110 of *Lecture Notes in Computer Science*, pages 450–461. Springer, Barcellona, Spain, August 2006.

[LLS90] J. C. Lagarias, H. W. Lenstra, Jr., and C.-P. Schnorr. Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice. *Combinatorica*, 10(4):333–348, 1990.

[MG02] D. Micciancio and S. Goldwasser. *Complexity of Lattice Problems: a cryptographic perspective*, volume 671 of *The Kluwer International Series in Engineering and Computer Science*. Kluwer Academic Publishers, Boston, Massachusetts, March 2002.

[Mic01a] D. Micciancio. The hardness of the closest vector problem with preprocessing. *IEEE Transactions on Information Theory*, 47(3):1212–1215, 2001.

[Mic01b] D. Micciancio. The shortest vector problem is NP-hard to approximate to within some constant. *SIAM Journal on Computing*, 30(6):2008–2035, March 2001. Preliminary version in FOCS 1998.

[MP12] D. Micciancio and C. Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In D. Pointcheval and T. Johansson, editors, *Advances in Cryptology—EUROCRYPT 2012*, volume 7237 of *Lecture Notes in Computer Science*, pages 700–718. Springer Berlin Heidelberg, 2012.

[MR07] D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM Journal on Computing*, 37(1):267–302 (electronic), 2007.

[MV13] D. Micciancio and P. Voulgaris. A deterministic single exponential time algorithm for most lattice problems based on voronoi cell computations. *SIAM Journal on Computing*, 42(3):1364–1391, 2013. Preliminary version in STOC'10.

[Reg04] O. Regev. Improved inapproximability of lattice and coding problems with preprocessing. *IEEE Transactions on Information Theory*, 50(9):2031–2037, 2004. Preliminary version in CCC'03.

[Reg09] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM*, 56(6):Art. 34, 40, 2009.

[Reg10] O. Regev. On the complexity of lattice problems with polynomial approximation factors. In P. Q. Nguyen and B. Vallée, editors, *The LLL Algorithm*, Information Security and Cryptography, pages 475–496. Springer Berlin Heidelberg, 2010.

[Sch87] C.-P. Schnorr. A hierarchy of polynomial time lattice basis reduction algorithms. *Theoretical Computer Science*, 53(2-3):201–224, 1987.

[Ver12] R. Vershynin. Introduction to the non-asymptotic analysis of random matrices. In Y. Eldar and G. Kutyniok, editors, *Compressed Sensing: Theory and Applications*, pages 210–268. Cambridge Univ Press, 2012.