

# A Time-Distance Trade-Off for GDD with Preprocessing – Instantiating the DLW Heuristic

Noah Stephens-Davidowitz

Massachusetts Institute of Technology, Cambridge, MA, USA

<http://www.noahsd.com>

noahsd@gmail.com

---

## Abstract

For  $0 \leq \alpha \leq 1/2$ , we show an algorithm that does the following. Given appropriate preprocessing  $P(\mathcal{L})$  consisting of  $N_\alpha := 2^{O(n^{1-2\alpha} + \log n)}$  vectors in some lattice  $\mathcal{L} \subset \mathbb{R}^n$  and a target vector  $\mathbf{t} \in \mathbb{R}^n$ , the algorithm finds  $\mathbf{y} \in \mathcal{L}$  such that  $\|\mathbf{y} - \mathbf{t}\| \leq n^{1/2+\alpha}\eta(\mathcal{L})$  in time  $\text{poly}(n) \cdot N_\alpha$ , where  $\eta(\mathcal{L})$  is the smoothing parameter of the lattice.

The algorithm itself is very simple and was originally studied by Doulgerakis, Laarhoven, and de Weger (to appear in PQCrypto, 2019), who proved its correctness under certain reasonable heuristic assumptions on the preprocessing  $P(\mathcal{L})$  and target  $\mathbf{t}$ . Our primary contribution is a choice of preprocessing that allows us to prove correctness without any heuristic assumptions.

Our main motivation for studying this is the recent breakthrough algorithm for IdealSVP due to Hanrot, Pellet–Mary, and Stehlé (to appear in Eurocrypt, 2019), which uses the DLW algorithm as a key subprocedure. In particular, our result implies that the HPS IdealSVP algorithm can be made to work with fewer heuristic assumptions.

Our only technical tool is the discrete Gaussian distribution over  $\mathcal{L}$ , and in particular, a lemma showing that the one-dimensional projections of this distribution behave very similarly to the continuous Gaussian. This lemma might be of independent interest.

**2012 ACM Subject Classification** Theory of computation → Design and analysis of algorithms

**Keywords and phrases** Lattices, guaranteed distance decoding, GDD, GDDP

**Digital Object Identifier** 10.4230/LIPIcs.CCC.2019.11

**Related Version** A full version of the paper is available at <http://arxiv.org/abs/1902.08340>. (Please read the full version!)

**Acknowledgements** I thank Guillaume Hanrot, Thijs Laarhoven, Alice Pellet–Mary, Oded Regev, and Damien Stehlé for helpful discussions. I also thank Alice Pellet–Mary, Guillaume Hanrot, and Damien Stehlé for sharing early versions of their work with me. I am also grateful to the CCC 2019 reviewers for their very helpful comments, and Daniel Dadush for showing me how to obtain the stronger results to be written up in the full version.

## 1 Introduction

A lattice  $\mathcal{L} \subset \mathbb{R}^n$  is the set of all integer linear combinations

$$\mathcal{L} := \{z_1 \mathbf{b}_1 + \cdots + z_n \mathbf{b}_n : z_i \in \mathbb{Z}\}$$

of linearly independent basis vectors  $\mathbf{b}_1, \dots, \mathbf{b}_n \in \mathbb{R}^n$ . For a lattice  $\mathcal{L} \subset \mathbb{R}^n$  and target vector  $\mathbf{t} \in \mathbb{R}^n$ , the  $d$ -Guaranteed Distance Decoding problem ( $d$ -GDD, or just GDD) asks us to find  $\mathbf{y} \in \mathcal{L}$  such that  $\|\mathbf{y} - \mathbf{t}\| \leq d$  for some distance  $d := d(\mathcal{L})$  that depends only on  $\mathcal{L}$ . In particular, we must have  $d \geq \mu(\mathcal{L})$ , where  $\mu(\mathcal{L}) := \max \text{dist}(\mathbf{t}, \mathcal{L})$  is the covering radius of the lattice.

GDD with preprocessing (GDDP) is the variant of this problem in which we are allowed to perform arbitrary preprocessing on the lattice (but not on  $\mathbf{t}$ ). I.e., formally an “algorithm” for GDDP is really a *pair* of algorithms, a preprocessing algorithm, which takes as input (a



© Noah Stephens-Davidowitz;  
licensed under Creative Commons License CC-BY  
34th Computational Complexity Conference (CCC 2019).

Editor: Amir Shpilka; Article No. 11; pp. 11:1–11:8

Leibniz International Proceedings in Informatics



Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



basis for) a lattice  $\mathcal{L} \subset \mathbb{R}^n$  and outputs some preprocessing  $P(\mathcal{L})$ , and a query algorithm which takes as input  $P(\mathcal{L})$  and a target  $\mathbf{t}$  and outputs a valid solution to the GDD instance  $(\mathcal{L}, \mathbf{t})$ . The complexity measure that interests us for such algorithms is the running time of the query algorithm.

In [7], Doulgerakis, Laarhoven, and de Weger (DLW) gave an elegant algorithm for GDDP whose correctness relies on certain heuristic assumptions. (Our presentation here differs quite a bit from DLW's. See Section 1.2.) In fact, [7] gave a family of algorithms parameterized by  $0 \leq \alpha \leq 1/2$  whose preprocessing consists of  $N_\alpha \approx 2^{n^{1-2\alpha}}$  lattice vectors in  $\mathcal{L}$  whose length is roughly  $r$ . Given a target  $\mathbf{t}$ , the query algorithm starts by setting  $\mathbf{t}' = \mathbf{t}$ . The algorithm then simply searches for a vector  $\mathbf{y}$  in the preprocessing list and an integer  $k$  such that  $\|k\mathbf{y} - \mathbf{t}'\| < \|\mathbf{t}'\|$ . If it finds one, it replaces  $\mathbf{t}'$  by  $\mathbf{t}' - k\mathbf{y}$  and repeats this procedure. Finally, it outputs  $\mathbf{y}' := \mathbf{t} - \mathbf{t}' \in \mathcal{L}$ . Under certain heuristic assumptions that in particular imply that the preprocessing is nicely distributed, [7] showed that this algorithm terminates with  $\|\mathbf{y}' - \mathbf{t}\| = \|\mathbf{t}'\| \lesssim n^\alpha \cdot r$  in time  $\text{poly}(n) \cdot N_\alpha$ .

DLW's algorithm is the first to provide a smooth trade-off between the running time and the distance  $d$ . (Such trade-offs are known for other lattice problems. E.g., without preprocessing, block reduction [19, 8] algorithms accomplish this for many lattice problems, and with preprocessing, such trade-offs are known for Bounded Distance Decoding and the Closest Vector Problem [12, 6].) This recently found an exciting application discovered by Pellet–Mary, Hanrot, and Stehlé [18]. [18] showed the best known time-approximation-factor trade-off for the very important problem of finding short non-zero vectors in ideal lattices (given suitable preprocessing on the underlying number field). Their algorithm uses the DLW algorithm as a key subprocedure. However, since DLW's algorithm relies on certain heuristic assumptions, their application crucially relies on the (reasonable but unproven) assumption that these heuristics apply in their particular use case.

## 1.1 Removing the heuristic in DLW's GDDP algorithm

We show how to instantiate DLW's heuristic algorithm in a provably correct way. In particular, we show an explicit distribution over the lattice such that, when the preprocessing consists of independent samples from this distribution, the above algorithm provably succeeds with high probability. Indeed, there is a very natural choice for this distribution: the discrete Gaussian over the lattice,  $D_{\mathcal{L},s}$ . This is the distribution that assigns probability to each lattice vector  $\mathbf{y} \in \mathcal{L}$  proportional to its Gaussian mass  $\exp(-\pi\|\mathbf{y}\|^2/s^2)$ , and it is a ubiquitous tool in lattice algorithms and the study of lattices more generally. (See, e.g., [20].) When the width parameter  $s > 0$  is at least as large as the *smoothing parameter*  $\eta(\mathcal{L})$ , the discrete Gaussian distribution  $D_{\mathcal{L},s}$  provably behaves quite similarly to the continuous Gaussian in many ways [15]. (E.g., its moments are close to those of a continuous Gaussian.) So, one might expect that it will be distributed nicely enough to work for DLW's use case.

We show that for  $s = \eta(\mathcal{L})$ , the discrete Gaussian  $D_{\mathcal{L},s}$  does in fact suffice to provably instantiate DLW's heuristic algorithm with  $r \approx \sqrt{n} \cdot \eta(\mathcal{L})$ . (This is essentially the same value of  $r$  used in [7]. See Section 1.2 for more discussion.) I.e., we prove the following theorem.

► **Theorem 1.** *For any  $0 \leq \alpha \leq 1/2$ , there is an algorithm that solves  $d$ -GDDP in time  $2^{O(n^{1-2\alpha} + \log n)}$  where  $d(\mathcal{L}) := n^{1/2+\alpha} \cdot \eta(\mathcal{L})$ .*

Theorem 1 is primarily interesting for  $\alpha$  strictly between zero and  $1/2$ . For  $\alpha = 0$ , Theorem 1 is outperformed by existing  $2^{O(n)}$ -time algorithms for CVP [16, 2]. These algorithms do not require preprocessing and are actually guaranteed to find a *closest* vector to the target  $\mathbf{t}$ , so our algorithm is beaten in many respects by the competition in this regime.

Similarly, for  $\alpha = 1/2$ , Babai’s celebrated polynomial-time algorithm [3] always matches or outperforms Theorem 1 when instantiated with an appropriate basis as preprocessing. However, we will show in the full version [21] how to essentially combine ideas from Babai’s algorithm to correct this and to beat all prior algorithms for all (constants)  $0 < \alpha < 1/2$ .

However, even without these improvements, this result is already quite strong for the “typical” lattices that interest us – e.g., that that arise in cryptography and those that satisfy the heuristics in [7] – which in particular satisfy  $\eta(\mathcal{L}) \approx \mu(\mathcal{L})/\sqrt{n}$ . In particular, Theorem 1 is already enough to remove [18]’s reliance on certain heuristic assumptions. ([18] also requires additional unrelated heuristic assumptions, which our result does not remove. We refer the reader to [18] for more information.)

Behind this result is a geometric lemma concerning the discrete Gaussian distribution that, to the author’s knowledge, is novel. The lemma shows that one-dimensional projections of the discrete Gaussian look very much like a continuous Gaussian for parameters above smoothing. (See Theorem 7.)

## 1.2 Relation to DLW

Our presentation here is quite different from the presentation in [7]. (See also an earlier version of the same paper [10] and a closely related paper [11].) We attempt to clarify some of the differences here to avoid confusion.

First of all, DLW described their algorithm as a solution to the *Closest Vector Problem* (CVP), in which the goal is to output a vector  $\mathbf{y} \in \mathcal{L}$  with  $\|\mathbf{y} - \mathbf{t}\| \leq \gamma \cdot \text{dist}(\mathbf{t}, \mathcal{L})$  for some approximation factor  $\gamma \geq 1$ . In contrast, we call the same algorithm a GDD(P) algorithm. This discrepancy arises when one moves from heuristic algorithms to provably correct algorithms. Since  $\text{dist}(\mathbf{t}, \mathcal{L})$  is nearly maximal for “most”  $\mathbf{t}$  [9], DLW’s heuristics quite reasonably imply that  $\text{dist}(\mathbf{t}, \mathcal{L})$  is nearly maximal, i.e.,  $\text{dist}(\mathbf{t}, \mathcal{L}) \approx \mu(\mathcal{L})$ . With this assumption,  $\gamma$ -CVP is essentially equivalent to  $(\gamma\mu(\mathcal{L}))$ -GDD. However, without such a heuristic, the two problems seem to be quite different, so that the distinction is unfortunately necessary here.

Second, since [7] describe their results in terms of CVP and do not mention the smoothing parameter  $\eta(\mathcal{L})$ , their results are formally incomparable with Theorem 1. However, we note that the heuristics in [7] imply that  $\eta(\mathcal{L}) \approx \lambda_1(\mathcal{L})/\sqrt{n} \approx \mu(\mathcal{L})/\sqrt{n}$ , and the DLW algorithm finds vectors within distance roughly  $n^\alpha \lambda_1(\mathcal{L})$  of the target. Since we obtain vectors within distance  $n^{1/2+\alpha}\eta(\mathcal{L})$ , our result essentially matches theirs when their heuristics apply.

Third, while we match DLW’s algorithm asymptotically, we do not claim to match the constants. Indeed, in the language of this paper, much of [7] is devoted to finding vectors within distance  $c_1\sqrt{n} \cdot \eta(\mathcal{L})$  in time  $2^{c_2n+o(n)}$  for small constants  $0 < c_1, c_2 < 1$ . In contrast, we are mostly interested in what appears as a secondary result in that paper: the time-distance trade-off achievable for distance  $n^{1/2+\alpha}\eta(\mathcal{L})$  and time  $2^{O(n^{1-2\alpha} + \log n)}$  for  $0 < \alpha < 1/2$ . And, we make very little effort to optimize the constants. For example, [7] uses nearest neighbor data structures to let the query algorithm avoid reading the entire preprocessing, which we do not attempt to replicate here. Similarly, while [7] proposed specific techniques for computing the preprocessing in  $2^{cn+o(n)}$  time, we ignore this. (We do note, however, that [1] shows how to sample the preprocessing in time  $2^{n+o(n)}$ .)

## 2 Preliminaries

Throughout this work, we adopt the common convention of expressing the running times of lattice algorithms in terms of the dimension  $n$  only, ignoring any dependence on the bit length of the input  $B$ . Formally, we should specify a particular input format for the (basis of the)

lattice (e.g., by restricting our attention to rational numbers and using the natural binary representation of a rational matrix to represent a basis for the lattice), and our running time should of course have some dependence on  $B$ . Consideration of the bit length would simply add a  $\text{poly}(B)$  factor to the running time for the algorithm(s) considered in this paper, provided that the input format allows for efficient arithmetic operations.

## 2.1 The discrete Gaussian

For a vector  $\mathbf{x} \in \mathbb{R}^n$  and parameter  $s > 0$ , we write  $\rho_s(\mathbf{x}) := \exp(-\pi\|\mathbf{x}\|^2/s^2)$  for the Gaussian mass of  $\mathbf{x}$  with parameter  $s$ . For a lattice  $\mathcal{L} \subset \mathbb{R}^n$  and shift vector  $\mathbf{t} \in \mathbb{R}^n$ , we write

$$\rho_s(\mathcal{L} - \mathbf{t}) := \sum_{\mathbf{y} \in \mathcal{L}} \rho_s(\mathbf{y} - \mathbf{t})$$

for the Gaussian mass of  $\mathcal{L} - \mathbf{t}$  with parameter  $s$ . We write  $D_{\mathcal{L},s}$  for the probability distribution over  $\mathcal{L}$  defined by

$$\Pr_{\mathbf{X} \sim D_{\mathcal{L},s}} [\mathbf{X} = \mathbf{y}] = \frac{\rho_s(\mathbf{y})}{\rho_s(\mathcal{L})}$$

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

The dual lattice  $\mathcal{L}^* \subset \mathbb{R}^n$  is the set of vectors that have integer inner product with all lattice vectors,

$$\mathcal{L}^* := \{\mathbf{w} \in \mathbb{R}^n : \forall \mathbf{y} \in \mathcal{L}, \langle \mathbf{w}, \mathbf{y} \rangle \in \mathbb{Z}\}.$$

Micciancio and Regev defined the *smoothing parameter*  $\eta(\mathcal{L})$  as the unique parameter  $s$  such that  $\rho_{1/s}(\mathcal{L}^*) = 3/2$  [15].<sup>1</sup> The following claim justifies the name “smoothing parameter,” and it is the only fact about the smoothing parameter that we will need.

▷ **Claim 2.** For any lattice  $\mathcal{L} \subset \mathbb{R}^n$ , parameter  $s \geq \eta(\mathcal{L})$ , and shift  $\mathbf{t} \in \mathbb{R}^n$ ,

$$\frac{1}{3} \leq \frac{\rho_s(\mathcal{L} - \mathbf{t})}{\rho_s(\mathcal{L})} \leq 1.$$

We will also need a simplified version of Banaszczyk’s celebrated tail bound for the discrete Gaussian [4].

► **Theorem 3.** For any lattice  $\mathcal{L} \subset \mathbb{R}^n$  and parameter  $s > 0$ ,

$$\Pr_{\mathbf{X} \sim D_{\mathcal{L},s}} [\|\mathbf{X}\| \geq \sqrt{ns}] \leq 2^{-n}.$$

Finally, we will need the following rather weak consequence of Babai’s algorithm [3].

► **Theorem 4.** There is a polynomial-time algorithm for  $(2^n \eta(\mathcal{L}))$ -GDD.

## 2.2 $\varepsilon$ -nets

For  $\varepsilon > 0$ , we say that a set  $\{\mathbf{v}_1, \dots, \mathbf{v}_M\} \subset \mathbb{R}^n$  of unit vectors with  $\|\mathbf{v}_i\| = 1$  is an  $\varepsilon$ -net of the unit sphere if for any  $\mathbf{t} \in \mathbb{R}^n$  with  $\|\mathbf{t}\| = 1$ , there exists  $\mathbf{v}_i$  such that  $\|\mathbf{v}_i - \mathbf{t}\| \leq \varepsilon$ . We will use a simple bound on the size of such a net, which can be proven via a simple packing argument. See [22, Lemma 5.2], for example.

► **Lemma 5.** For any  $\varepsilon > 0$ , there exists an  $\varepsilon$ -net of the unit sphere in  $\mathbb{R}^n$  with  $(1 + 2/\varepsilon)^n$  points.

<sup>1</sup> This is more commonly referred to as  $\eta_{1/2}(\mathcal{L})$ , where  $\eta_\varepsilon(\mathcal{L})$  is the unique parameter  $s$  such that  $\rho_{1/s}(\mathcal{L}^*) = 1 + \varepsilon$ . Since we will always take  $\varepsilon = 1/2$ , we simply omit it. Our results remain essentially unchanged if we take  $\varepsilon$  to be any constant strictly between zero and one.

### 3 The algorithm

We consider the following algorithm for GDDP. For an input lattice  $\mathcal{L} \subset \mathbb{R}^n$  with  $n \geq 40$ , the preprocessing consists of  $N$  lattice vectors  $\mathbf{y}_1, \dots, \mathbf{y}_N \in \mathcal{L}$ . On input  $\mathbf{t} \in \mathbb{R}^n$ , the query algorithm behaves as follows. It first uses Theorem 4 to find  $\mathbf{t}_0 \in \mathcal{L} + \mathbf{t}$  such that  $\|\mathbf{t}_0\| \leq 2^n \eta(\mathcal{L})$  and sets  $j = 0$ . The algorithm then does the following repeatedly. It finds an index  $i$  and integer  $k$  such that  $\|\mathbf{t}_j - k\mathbf{y}_i\|^2 \leq (1 - 1/n^2) \cdot \|\mathbf{t}_j\|^2$ , sets  $\mathbf{t}_{j+1} := \mathbf{t}_j - k\mathbf{y}_i$ , and increments  $j$ . Once the algorithm fails to find such a vector, it outputs  $\mathbf{t}_j - \mathbf{t} \in \mathcal{L}$ .<sup>2</sup>

Our main theorem shows that this algorithm will succeed when the preprocessing is chosen from the right distribution. We emphasize the order of quantifiers: with high probability over the preprocessing, the algorithm works *for all targets*  $\mathbf{t} \in \mathbb{R}^n$ . In particular, there exists fixed preprocessing that works for all targets  $\mathbf{t}$ .

► **Theorem 6.** *For any  $\alpha$  with  $\frac{2}{\log n} \leq \alpha \leq \frac{1}{2}$ , when the preprocessing of the above algorithm consists of  $N_\alpha := n^2 e^{(n^{1/2-\alpha}+4)^2} = 2^{O(n^{1-2\alpha} + \log n)}$  samples from  $D_{\mathcal{L},s}$  for  $s := \eta(\mathcal{L})$ , it yields a solution to  $d$ -GDDP in time  $\text{poly}(n) \cdot N_\alpha$  with high probability, where  $d := n^{1/2+\alpha} \cdot \eta(\mathcal{L})$ .*

**Proof.** By scaling appropriately, we may assume without loss of generality that  $d = 1$ , and therefore  $s = n^{-1/2-\alpha}$ . Let  $\mathbf{y}_1, \dots, \mathbf{y}_{N_\alpha} \sim D_{\mathcal{L},s}$ . To prove correctness, we must show that, with high probability over the  $\mathbf{y}_i$ , for every  $\mathbf{t} \in \mathbb{R}^n$  with  $\|\mathbf{t}\| \geq 1$ , there exists an index  $i$  and integer  $k$  such that  $\|\mathbf{t} - k\mathbf{y}_i\|^2 \leq (1 - 1/n^2) \cdot \|\mathbf{t}\|^2$ . It suffices to prove that for  $\|\mathbf{t}\| = 1$ , there exists an  $i$  with  $\|\mathbf{t} - \mathbf{y}_i\|^2 \leq 1 - 4/n^2$ .<sup>3</sup> Finally, to prove *this*, it suffices to take a  $(1/n^3)$ -net of the unit sphere,  $\mathbf{v}_1, \dots, \mathbf{v}_M$ , and to show that for each  $j$ , there exists an  $i$  such that  $\|\mathbf{v}_j - \mathbf{y}_i\|^2 \leq 1 - 5/n^2$ .

By Lemma 5, there exists such a net of cardinality  $M = (3n)^{3n}$ . For each  $\mathbf{v}_j$  in this net and each index  $i$ , we have by Corollary 8 (proven below) that

$$\begin{aligned} \Pr[\|\mathbf{v}_j - \mathbf{y}_i\|^2 \leq 1 - 5/n^2] &\geq \exp(-\pi(5/(n^2s) + ns + 4)^2/4) - 2^{-n} \\ &\geq \exp(-(n^{1/2-\alpha} + 4)^2) \\ &= n^2/N_\alpha. \end{aligned}$$

Since the  $\mathbf{y}_i$  are sampled independently, the probability that no such  $i$  exists is at most  $(1 - n^2/N_\alpha)^{N_\alpha} < 2^{-n}/M$ . The result then follows by taking a union bound over the  $\mathbf{v}_j$ . ◀

#### 3.1 One-dimensional projections of the discrete Gaussian

We are interested in the lower bound in the following lemma (whose proof uses a very nice idea from [13]). The upper bound (i.e., the subgaussianity of the discrete Gaussian) applies for all parameters  $s > 0$  and is well known. (It appears in slightly different forms in [5, 17]. See also [14, Lemma 2.8].) We only include the upper bound for comparison, and we make no effort to optimize the lower-order term.

<sup>2</sup> The author made no effort to optimize these parameters. Notice that we can find such an  $i$  and  $k$  (if they exist) in time essentially  $\text{poly}(n) \cdot N$ . To guarantee a total running time of  $\text{poly}(n) \cdot N$ , we can also assume that the algorithm halts and outputs  $\mathbf{t}_j - \mathbf{t} \in \mathcal{L}$  if  $j$  reaches, say,  $100n^3$ . This is not strictly necessary, since we will have  $\|\mathbf{y}_i\| \approx \sqrt{n} \cdot \eta(\mathcal{L})$  with very high probability.

<sup>3</sup> Indeed, suppose that  $\|\mathbf{t} - \mathbf{y}_i\|^2 \leq 1 - 4/n^2$  and  $\|\mathbf{t}\| = 1$ , so that in particular  $\langle \mathbf{y}_i, \mathbf{t} \rangle \geq 0$ . Then for any  $\beta/2 \leq k \leq \beta$ ,

$$\frac{\|\beta\mathbf{t} - k\mathbf{y}_i\|^2}{\|\beta\mathbf{t}\|^2} = 1 - \frac{k^2}{\beta^2} \cdot (1 - \|\mathbf{t} - \mathbf{y}_i\|^2) - \left(\frac{2k}{\beta} - \frac{2k^2}{\beta^2}\right) \cdot \langle \mathbf{y}_i, \mathbf{t} \rangle \leq 1 - \frac{k^2}{\beta^2} \cdot \frac{4}{n^2} \leq 1 - \frac{1}{n^2}.$$

► **Theorem 7.** For any lattice  $\mathcal{L} \subset \mathbb{R}^n$ , parameter  $s \geq \eta(\mathcal{L})$ , unit vector  $\mathbf{v} \in \mathbb{R}^n$  with  $\|\mathbf{v}\| = 1$ , and  $r_0 > 0$ , we have

$$\exp(-\pi(r_0/s + 2)^2) < \Pr_{\mathbf{X} \sim D_{\mathcal{L},s}} [\langle \mathbf{X}, \mathbf{v} \rangle \geq r_0] \leq \exp(-\pi r_0^2/s^2). \quad (1)$$

Before presenting the proof, we provide some of the intuition behind it. The idea is to control the moment-generating function  $g(\beta) := \mathbb{E}[\exp(2\pi\beta\langle \mathbf{X}, \mathbf{v} \rangle)]$  for  $\beta > 0$ . For a continuous Gaussian, this is  $\exp(\pi\beta^2)$ , and the discrete Gaussian behaves similarly, with  $g(\beta) \approx \exp(\pi\beta^2)$  (see Eq. (2)). For a fixed value of  $\beta$ , knowledge of  $g(\beta)$  is insufficient to prove something like Eq. (1). So, we take a weighted combination  $\alpha_1 g(\beta_1) - \alpha_2 g(\beta_2) - \alpha_3 g(\beta_3)$  for appropriately chosen weights to essentially approximate

$$\mathbb{E}[1_{\langle \mathbf{X}, \mathbf{v} \rangle \gtrsim r_0}] \approx \Pr_{\mathbf{X} \sim D_{\mathcal{L},s}} [\langle \mathbf{X}, \mathbf{v} \rangle \geq r_0].$$

More specifically, we define the function  $f(r)$  as in Eq. (3), which satisfies  $f(r) < 0$  unless  $r \approx r_0$  and  $f(r) \approx \exp(2\pi\beta r_0)$  for  $r \approx r_0$ . We then use our bounds on the moment-generating function to show that  $\mathbb{E}[f(\langle \mathbf{X}, \mathbf{v} \rangle)]$  is not too small, which implies the result.

**Proof.** By scaling appropriately, we may assume that  $s = 1$ . Let  $\beta > 0$  to be chosen later. By completing the square in the exponent, we see that

$$\mathbb{E}[\exp(2\pi\beta\langle \mathbf{X}, \mathbf{v} \rangle)] = \exp(\pi\beta^2) \cdot \frac{\rho_1(\mathcal{L} - \beta\mathbf{v})}{\rho_1(\mathcal{L})}.$$

Therefore, by Claim 2,

$$\frac{1}{3} \leq \exp(-\pi\beta^2) \cdot \mathbb{E}[\exp(2\pi\beta\langle \mathbf{X}, \mathbf{v} \rangle)] \leq 1. \quad (2)$$

I.e., we know the moment generating function of  $\langle \mathbf{X}, \mathbf{v} \rangle$  to within a multiplicative constant. The upper bound in Eq. (1) then follows from taking  $\beta = r_0$  and applying Markov's inequality. (This proof of the upper bound is identical to the proof in [14]. See their Lemma 2.8 and their discussion above it.)

Turning to the lower bound, for  $r \in \mathbb{R}$ , let

$$f(r) := \exp(2\pi\beta r) \cdot (1 - \exp(2\pi(r_0 - r)) - \exp(2\pi(r - r_0 - 2))). \quad (3)$$

Notice that  $f(r) < 0$  unless  $r_0 < r < r_0 + 2$ . And,  $f(r) < \exp(2\pi\beta r)$ , which together with the previous two inequalities implies that  $f(r) < \exp(2\pi\beta(r_0 + 2))$  for all  $r$ . Therefore,

$$\mathbb{E}[f(\langle \mathbf{X}, \mathbf{v} \rangle)] < \exp(2\pi\beta(r_0 + 2)) \cdot \Pr[\langle \mathbf{X}, \mathbf{v} \rangle \geq r_0]. \quad (4)$$

By applying Eq. (2) term-wise and taking  $\beta = r_0 + 1$ , we have

$$\begin{aligned} \mathbb{E}[f(\langle \mathbf{X}, \mathbf{v} \rangle)] &= \mathbb{E}[\exp(2\pi\beta\langle \mathbf{X}, \mathbf{v} \rangle)] \\ &\quad - \exp(2\pi r_0) \mathbb{E}[\exp(2\pi(\beta - 1)\langle \mathbf{X}, \mathbf{v} \rangle)] \\ &\quad - \exp(-2\pi(r_0 + 2)) \mathbb{E}[\exp(2\pi(\beta + 1)\langle \mathbf{X}, \mathbf{v} \rangle)] \\ &\geq \frac{1}{3} \cdot \exp(\pi\beta^2) - \exp(\pi(\beta - 1)^2 + 2\pi r_0) - \exp(\pi(\beta + 1)^2 - 2\pi(r_0 + 2)) \\ &= \exp(\pi r_0^2 + 2\pi r_0) \cdot (e^\pi/3 - 2) \\ &> \exp(\pi r_0^2 + 2\pi r_0). \end{aligned} \quad (5)$$

Combining Eqs. (4) and (5) and rearranging, we have

$$\Pr[\langle \mathbf{X}, \mathbf{v} \rangle \geq r_0] > \exp(\pi r_0^2 + 2\pi r_0 - 2\pi\beta(r_0 + 2)) = \exp(-\pi(r_0 + 2)^2),$$

as needed. ◀

► **Corollary 8.** For any  $0 < r < 1$ , lattice  $\mathcal{L} \subset \mathbb{R}^n$ , unit vector  $\mathbf{v} \in \mathbb{R}^n$  with  $\|\mathbf{v}\| = 1$ , and  $s \geq \eta(\mathcal{L})$ , we have

$$\Pr_{\mathbf{X} \sim D_{\mathcal{L},s}} [\|\mathbf{v} - \mathbf{X}\|^2 \leq 1 - r] > \exp(-\pi(r/s + ns + 4)^2/4) - 2^{-n}.$$

**Proof.** Notice that  $\|\mathbf{v} - \mathbf{X}\|^2 = 1 + \|\mathbf{X}\|^2 - 2\langle \mathbf{v}, \mathbf{X} \rangle$ . By Theorem 3, we have that  $\|\mathbf{X}\|^2 \leq ns^2$  except with probability at most  $2^{-n}$ . By Theorem 7, we see that

$$\Pr[\langle \mathbf{v}, \mathbf{X} \rangle \geq (ns^2 + r)/2] > \exp(-\pi(r/s + ns + 4)^2/4).$$

The result follows from union bound. ◀

---

## References

- 1 Divesh Aggarwal, Daniel Dadush, Oded Regev, and Noah Stephens-Davidowitz. Solving the Shortest Vector Problem in  $2^n$  time via Discrete Gaussian Sampling. In *STOC*, 2015. [arXiv:abs/1412.7994](https://arxiv.org/abs/1412.7994).
- 2 Divesh Aggarwal, Daniel Dadush, and Noah Stephens-Davidowitz. Solving the Closest Vector Problem in  $2^n$  time—The discrete Gaussian strikes again! In *FOCS*, 2015. [arXiv:1504.01995](https://arxiv.org/abs/1504.01995).
- 3 L. Babai. On Lovász’ lattice reduction and the nearest lattice point problem. *Combinatorica*, 6(1), 1986.
- 4 Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4), 1993.
- 5 Wojciech Banaszczyk. Inequalities for convex bodies and polar reciprocal lattices in  $\mathbb{R}^n$ . *Discrete & Computational Geometry*, 13, 1995.
- 6 Daniel Dadush, Oded Regev, and Noah Stephens-Davidowitz. On the Closest Vector Problem with a distance guarantee. In *CCC*, 2014. [arXiv:1409.8063](https://arxiv.org/abs/1409.8063).
- 7 Emmanouil Doulgerakis, Thijs Laarhoven, and Benne de de Weger. Finding closest lattice vectors using approximate Voronoi cells. In *PQCrypto*, 2019. (To appear.) <https://eprint.iacr.org/2016/888>.
- 8 Nicolas Gama and Phong Q. Nguyen. Finding Short Lattice Vectors Within Mordell’s Inequality. In *STOC*, 2008.
- 9 Ishay Haviv, Vadim Lyubashevsky, and Oded Regev. A Note on the Distribution of the Distance from a Lattice. *Discrete & Computational Geometry*, 41(1), 2009.
- 10 Thijs Laarhoven. Finding closest lattice vectors using approximate Voronoi cells, 2016. URL: <https://eprint.iacr.org/2016/888/20161219:141310>.
- 11 Thijs Laarhoven. Sieving for Closest Lattice Vectors (with Preprocessing). In *SAC*, 2016.
- 12 Yi-Kai Liu, Vadim Lyubashevsky, and Daniele Micciancio. On Bounded Distance Decoding for general lattices. In *RANDOM*, 2006.
- 13 J. E. Mazo and A. M. Odlyzko. Lattice points in high-dimensional spheres. *Monatshefte für Mathematik*, 110(1), 1990.
- 14 Daniele Micciancio and Chris Peikert. Trapdoors for Lattices: Simpler, Tighter, Faster, Smaller. In *EUROCRYPT*, 2012. URL: <https://eprint.iacr.org/2011/501>.
- 15 Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM Journal of Computing*, 37(1), 2007.
- 16 Daniele Micciancio and Panagiotis Voulgaris. A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations. *SIAM Journal on Computing*, 42(3), 2013.
- 17 Chris Peikert. Limits on the Hardness of Lattice Problems in  $\ell_p$  Norms. *Computational Complexity*, 17(2), 2008.
- 18 Alice Pellet–Mary, Guillaume Hanrot, and Damien Stehlé. Approx-SVP in Ideal Lattices with Pre-processing. In *Eurocrypt*, 2019. (to appear).



## 11:8 GDD with Preprocessing

- 19 Claus-Peter Schnorr. A Hierarchy of Polynomial Time Lattice Basis Reduction Algorithms. *Theor. Comput. Sci.*, 53(23), 1987.
- 20 Noah Stephens-Davidowitz. *On the Gaussian measure over lattices*. Ph.D. thesis, New York University, 2017.
- 21 Noah Stephens-Davidowitz. A time-distance trade-off for GDD with preprocessing – Instantiating the DLW heuristic, 2019. [arXiv:1902.08340](https://arxiv.org/abs/1902.08340).
- 22 Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.