Research Article

Kaveh Bashiri* and Andreas Wiemers

On the independence heuristic in the dual attack

https://doi.org/10.1515/jmc-2024-0028
received April 11, 2024; accepted May 05, 2025

Abstract: Post-quantum cryptography deals with the development and analysis of cryptographic schemes that are assumed to be secure even against attackers with access to a powerful quantum computer. Along the main candidates for quantum-safe solutions are cryptographic schemes, whose security is based on classic lattice problems such as the *bounded-distance decoding (BDD) problem* or the *learning with error problem*. In this work, we contribute to the analysis of an attack category against these problems called *dual attack*. In recent years, a lot of notable progress was achieved in this topic. Our first contribution is to provide theoretical counterarguments against a so-called independence assumption, which was used in earlier works on this attack, and which was shown in a previous work to be contradicting practical experiments. Then, we provide estimates on the success probability and the cost of the dual attack against the decisional version of the BDD problem. These estimates are derived both rigorously and heuristically. Finally, we also provide experimental evidence that confirms these results.

Keywords: dual attack, learning with errors, lattices, cryptanalysis

MSC 2020: 06B99, 94A60

1 Introduction

In recent years, much research has been done on the lattice problems called *decisional bounded-distance decoding (BDD) problem* or the closely related *learning with error (LWE) problem*. This is due to the fact that many post-quantum crypto schemes rely on their security. As a result, many attacking schemes against the BDD and LWE problem have been established. For instance, there are the algebraic attacks [1], the combinatoric attacks [2], or the *primal lattice attacks* [3]. The latter is based on sampling short vectors in the *primal* lattice Λ, i.e., the lattice, where the BDD or LWE problem is defined on. Another notable way to attack this problem is via the so-called *dual attack*, on which we will focus in this work, and which is based on sampling short vectors in the *dual lattice*. The main idea here is as follows. We are given a sample t, from which we know that it is

- either originated from a *BDD-sample*
- or a uniformly distributed *random-sample*.

Then, short vectors from the dual lattice are sampled in order to compute a statistical quantity, the so-called *score function*. Then, depending on the result of the score function, we make a decision, which case is true, i.e., whether t came from a BDD-sample or a random-sample. If the guess is correct, this in turn breaks the decisional version of the BDD (or LWE) problem. A very important ingredient in this attack is to exploit the fact that lattice sieving algorithms output not only the shortest vector but *exponentially many* short vectors.

*** Corresponding author: Kaveh Bashiri,** Bundesamt für Sicherheit in der Informationstechnik (BSI), Bonn, Germany, e-mail: kaveh.bashiri@bsi.bund.de
Andreas Wiemers: Bundesamt für Sicherheit in der Informationstechnik (BSI), Bonn, Germany, e-mail: andreas.wiemers@bsi.bund.de

The main idea of the dual attack originated in coding theory. The first time these ideas appeared in lattice theory seems to be in the classic paper [4]. In recent years, many seminal results on the dual attack have then been established, where this idea is exploited see e.g. [5–9]). However, many recent advances in this direction rely on an *independence assumption*. Ducas and Pulles [7] especially reported on experiments they made comparing the distributions of scores for random-samples and BDD-samples. They discovered that the distribution of scores for BDD-samples deviates from the predictions made under this independence assumption.

1.1 Main results

In this paper we provide the following contributions.

- We first theoretically show that the independence assumption cannot be true.
- Then, under certain assumptions on the distribution of the dual vectors, we provide rigorous estimates for the success probability of the abovementioned strategy. Moreover, as a byproduct, we also obtain a cost estimate in terms of the number of dual vectors that are needed for such a successful distinction. In this part of the paper we make use of *conditional expectations* and rely on techniques that are inspired from the work by Pouly and Shen [9].
- We then additionally provide a more intuitive approach to derive these results by relying on a central limit theorem heuristic. That is, here we again derive estimates for the success probability and the number of needed dual vectors for the dual attack; however, this time using an intuitive, heuristic approach. We believe that both approaches, the rigorous one and the intuitive one, are essential for a full understanding of the subject; the beginning of Section 5 gives a more detailed explanation on why we believe that both approaches are important.
- Finally, we provide experimental evidence that the just mentioned central limit theorem heuristics indeed hold true. In this way, we verify experimentally our heuristically derived results, which in turn verify the rigorously derived results.

1.2 Outlook

This work is organized as follows. In Section 2, we introduce some preliminary notions that we will need. In Section 3, we compute certain covariances that reveal the incompleteness of the independence assumption. In Section 4, we provide rigorous estimates for the success probability and the cost of the dual attack. In Section 5, we show that these estimates also hold true under certain intuitively justified heuristics. Finally, in Section 6, we provide experimental evidence for our results.

2 Preliminaries

This section is organized as follows. First we provide a glimpse into lattice theory in Section 2.1. Then, we introduce the BDD problem in Section 2.2 and describe the main framework of the dual attack in Section 2.3.

2.1 Lattices

2.1.1 Main definitions

A *lattice* $\Lambda \subset \mathbb{R}^n$ is defined by

$$\Lambda = \mathbb{Z}b_1 + \ldots + \mathbb{Z}b_k \subset \mathbb{R}^n,$$

where $b_1, \ldots, b_k \in \mathbb{R}^n$. We say that the lattice Λ has *full rank* if $k = n$ and $b_1, \ldots, b_n$ are linearly independent. The *volume of a lattice* is defined by

$$\det(\Lambda) \coloneqq \mathrm{Vol}(\mathbb{R}^n/\Lambda),$$

and for full-rank lattices, we have that

$$\det(\Lambda) = |\det(B)|,$$

where $B = (b_1|\ldots|b_n) \in \mathbb{R}^{n,n}$. The *dual lattice* is defined by

$$\hat{\Lambda} \coloneqq \{w \in \mathbb{R}^n | \langle w, \Lambda \rangle \subset \mathbb{Z}\}.$$

2.1.2 Shortest vectors of a lattice

A very important object in lattice-based cryptography are its *shortest nonzero vectors*, i.e., elements of the lattice with length

$$\lambda_1(\Lambda) \coloneqq \min_{v \in \Lambda \setminus \{0\}} \|v\|.$$

It is a common strategy in lattice-based cryptography to rely on the so-called *Gaussian Heuristic*, which is given as follows.

Heuristic 1. The length of the shortest vector of a randomly generated lattice $\Lambda \subset \mathbb{R}^n$ is approximately given by

$$\lambda_1(\Lambda) \approx \sqrt{\frac{n}{2\pi \mathrm{e}}} \det(\Lambda)^{\frac{1}{n}}.$$

2.1.3 Discrete Gaussian distribution

The *discrete Gaussian distribution* is an important probability distribution on a lattice Λ. This distribution is defined via the *Gaussian density function* $\rho_s : \mathbb{R}^n \to (0, \infty)$, which, for $s > 0$, is given by

$$\rho_s(x) = \mathrm{e}^{-\pi \frac{\|x\|^2}{s^2}},$$

where we denote by $\|\cdot\|$ the standard Euclidean norm in $\mathbb{R}^n$. For convenience, for a discrete subset $A \subset \mathbb{R}^n$, we write

$$\rho_s(A) \coloneqq \sum_{x \in A} \rho_s(x).$$

Then, the *Gaussian distribution, $D_{\Lambda,s}$, over Λ of width s* is defined by the following probability mass function:

$$D_{\Lambda,s}(v) = \frac{\rho_s(v)}{\rho_s(\Lambda)} \quad \text{for all } v \in \Lambda.$$

2.1.4 Poisson summation formula

A very useful tool to switch from a lattice to its dual lattice is given by the *Poisson summation formula*, which is already used in the classic and seminal papers in lattice-based cryptography by Regev [10].

Lemma 2.1. *Let Λ be a full-rank lattice, $t \in \mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}$ be such that some growth and integrability conditions are fulfilled. Then,*

$$\sum_{v\in\Lambda} f(v + t) = \frac{1}{\det(\Lambda)} \sum_{w\in\hat{\Lambda}} \hat{f}(w)\mathrm{e}^{2\pi i\langle t,w\rangle},$$

where $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ is the Fourier transform of f defined by

$$y \mapsto \hat{f}(y) = \int_{\mathbb{R}^n} \mathrm{e}^{-2\pi i\langle x,y\rangle} f(x)\mathrm{d}x.$$

In this work, we are very interested to apply this result for the function $x \mapsto \rho_s(x)$, where it has the following form.

Corollary 2.2. *Let Λ be a full-rank lattice, $t \in \mathbb{R}^n$, and $s > 0$. Then,*

$$\sum_{v\in\Lambda} \rho_s(v - t) = \frac{s^n}{\det(\Lambda)} \sum_{w\in\hat{\Lambda}} \rho_{1/s}(w)\mathrm{e}^{-2\pi i\langle t,w\rangle}.$$

2.2 BDD problem

The main object of investigation in this paper is the BDD problem on a lattice Λ. It is defined as follows.

Definition 2.3. Let Λ be a full-rank lattice, and χ, the *error distribution*, be a probability measure on $\mathbb{R}^n$ with

$$\mathbb{E}[\chi] = 0 \quad \text{and with small variance} \operatorname{Var}(\chi),$$

where the notion of "small" depends on the context.

- Suppose that we are given

$$t = v + e, \quad \text{where } v \in \Lambda \text{ and } e \leftarrow \chi \text{ (i.e., } e \text{ is sampled according to } \chi).$$

In the *BDD-search-problem*, we are asked to find v.

- Suppose that we are given a sample $t \pmod{\Lambda}$, for which we know that
  - it is either a *BDD-sample*, i.e.,

$$t = v + e, \quad \text{where } v \in \Lambda \text{ and } e \leftarrow \chi,$$

  - or a *random-sample*, i.e.,

$$t \pmod{\Lambda} \quad \text{is distributed according to } \mathcal{U}(\mathbb{R}^n/\Lambda),$$

where $\mathcal{U}(\mathbb{R}^n/\Lambda)$ denotes the uniform distribution on $\mathbb{R}^n/\Lambda$.

In the *BDD-decision-problem*, we are asked to decide which is the case.

The BDD-problem is closely related to the *LWE problem*, which is the notion more commonly used in post-quantum cryptography. For a survey on the LWE-problem and its relation to the BDD problem, we refer to the excellent notes by Peikert [11]. Moreover, we refer to these notes for further references concerning the question on the equivalence of the search- and the decision-version.

2.3 Dual attack

In this paper, out of the attacks we listed in the introduction, we focus on the dual attack. This attack attracted a lot of interest in recent years (refer for instance [5–9]) and is still the object of ongoing research. In this section, we present the main ideas of this approach.

The dual attack is aimed at the BDD-decision-problem so that we are given a sample $t \pmod{\Lambda}$, for which we know that it is either a BDD-sample or a random-sample. The *idea* is based on first finding a Λ-periodic function $f : \mathbb{R}^n \to \mathbb{R}^n$, the *score/distinguisher*, such that
- $f(t)$ is large in the BDD-sample-case,
- $f(t)$ is small in the random-sample-case.

In practice, we then need to find some $\alpha \in \mathbb{R}$ such that
- if $f(t) \geq \alpha$, then with high probability we know that t must be a BDD-sample,
- if $f(t) < \alpha$, then with high probability we know that t must be a random-sample.

A reasonable candidate for such a score is given by

$$f(t) = \sum_{v \in \Lambda} \rho_s(v - t) \tag{1}$$

as, for $s > 0$ and $\mathrm{Var}(\chi)$ small enough, this function becomes large in the BDD-sample-case. However, this function is not easy to compute as it requires the sum over the whole lattice. This is where the important tool of the Poisson summation formula comes into play. Indeed, for a large enough subset $W \subset \hat{\Lambda}$, we obtain by the Poisson summation formula that

$$\begin{aligned} \sum_{v \in \Lambda} \rho_s(v - t) &= \frac{s^n}{\det(\Lambda)} \sum_{w \in \hat{\Lambda}} \rho_{1/s}(w) \mathrm{e}^{-2\pi i \langle t, w \rangle} \\ &= \frac{s^n}{\det(\Lambda)} \sum_{w \in \hat{\Lambda}} \rho_{1/s}(w) \cos(2\pi \langle t, w \rangle) \\ &\approx \frac{s^n}{\det(\Lambda)} \sum_{w \in W} \rho_{1/s}(w) \cos(2\pi \langle t, w \rangle). \end{aligned} \tag{2}$$

The set $W \subset \hat{\Lambda}$ is usually generated through a lattice-sieve algorithm (such as Blockwise-Korkine-Zolotarev [BKZ] lattice reduction algorithm). The generation of these dual vectors is treated as a black box in this paper. Later on we have to make some concrete assumptions on this output.

In this subsection we are, however, interested in an heuristic way to find a reasonable score. In order to do this, we heuristically assume that the output of the black box yields short dual vectors, which are all roughly of similar norm. This implies in particular that $w \mapsto \rho_{1/s}(w)$ is almost a constant function on W. Let c denote this constant, i.e., $\rho_{1/s}(w) \approx c$ for all $w \in W$. Then,

$$\sum_{v \in \Lambda} \rho_s(v - t) \approx \frac{s^n \cdot c}{\det(\Lambda)} \sum_{w \in W} \cos(2\pi \langle t, w \rangle). \tag{3}$$

As the prefactor $\frac{s^n \cdot c}{\det(\Lambda)}$ does not depend on t, it is reasonable to simply discard this factor for the score function.

Combining (1), (2), and (3) intuitively justifies to choose the following function, which depends on a large enough subset $W \subset \hat{\Lambda}$, as the score.

$$f_W(t) = \sum_{w \in W} f_w(t), \quad \text{where } f_w(t) = \cos(2\pi \langle t, w \rangle).$$

To sum up, the *strategy* for the dual attack against the BDD-decision-problem is to compute $f_W(t)$, and to vote for the BDD-sample-case if $f_W(t) \geq \alpha$ and to vote for the random-sample-case otherwise.

The important remaining tasks are now
- to compute the probabilities with which our strategy is successful,
- and to find out how many dual vectors are needed so that the function f_W is indeed a good choice, such that the computed probabilities are high enough.

We target these problems later in this paper under certain assumptions on the abovementioned black box.

3 Computing the covariances

The score f_W can be seen as a sum of random variables, where these random variables are given by $\{\cos(2\pi\langle t, w\rangle)\}_{w\in W}$, where the randomness is coming from both the sample t and the dual vectors in the set W. Whenever one wishes to compute such a sum f_W of random variables, it is very handy to find and exploit any kind of independence between the random variables. Only under these independence assumptions one is allowed to apply standard results from probability theory such as the law of large numbers or the central limit theorem. This is why in many works [5,6] on the dual attack against BDD (or LWE), it is heuristically assumed that $\{\cos(2\pi\langle t, w\rangle)\}_{w\in W}$ is a family of independent random variables. We refer to this assumption in the following as the *independence assumption*.

However, for $w, w' \in W$ with $w \neq w'$, both random variables $\cos(2\pi\langle t, w\rangle)$ and $\cos(2\pi\langle t, w'\rangle)$ depend on the same sample t. Therefore, it is difficult to justify whether these two random variables are independent of each other. Under certain circumstances, it may be possible that $\cos(2\pi\langle t, w\rangle)$ and $\cos(2\pi\langle t, w'\rangle)$ are *uncorrelated*. For instance, this may happen when w and w' are orthogonal, which is indeed a consequence of Proposition 3.3. In any case, however, in order to apply standard results from probability theory such as the law of large numbers or the central limit theorem, we need independent and not only uncorrelated only random variables.

Of course, this lack of independence remains true if we assume that w, w' are drawn independently of each other (which in turn is a reasonable assumption though, which we will make in Sections 4 and 5). These theoretical doubts on the independence assumption are strengthened in the paper by Ducas and Pulles [7], where (among other results that they achieve) they show that this assumption leads also to contradictions in practical experiments.

In this section, we aim to fully refute the independence assumption by explicitly computing the covariance between the random variables and showing that these are not equal to zero.

We proceed as follows. We first provide some preparatory steps such as the precise definitions of the distributions in the two cases. Then we compute the covariances in the random-sample-case in Section 3.2 and in the BDD-sample-case in Section 3.3. Finally in this section, we provide some heuristic estimates, which underline that covariances in the BDD-sample-case are indeed nonzero.

Remark 3.1. As we change the perspective at some point in this paper, it is useful to emphasize, when and where we consider the sample t and the dual vectors in W as random or as fixed.

Formally, we always consider the family $\{\cos(2\pi\langle t, w\rangle)\}_{w\in W}$ as a family of random variables, where the randomness is coming from both the sample t and the dual vectors in W. However, in Sections 4 and 5, we will use the so-called *probability measure* $\mathbb{P}[\cdot|t]$ *conditioned on* t. As a consequence of the properties of $\mathbb{P}[\cdot|t]$, we implicitly see t as already sampled *a priori*. One can now intuitively say that t is consequently seen as fixed and not random anymore. However, in reality it is still a random variable, but its randomness is not visible under $\mathbb{P}[\cdot|t]$.

This is in contrast to this section. In this section, we consider (implicitly) the *probability measure conditioned on* W. That is, we consider informally the dual vectors in W as fixed. What we will obtain in Proposition 3.3 is an expression for the covariances in terms of the dual vectors in W. This expression shows that the covariance between $\cos(2\pi\langle t, w\rangle)$ and $\cos(2\pi\langle t, w'\rangle)$ is clearly nonzero in the usual case, when w and w' are not orthogonal. In order to obtain a more comprehensible expression for this covariance, we perform some heuristic computations with it in Appendix A.

3.1 Some preparation

We adopt the notation given in the paper by Ducas and Pulles [7] and repeat the approach described by them [7, Section 2.3]. Recall the definition of W, f_W, and f_w from above. In the following, we abbreviate $m_0 = |W|$, which is an important quantity we need to estimate (which in turn yields an indication on the complexity of the dual attack).

3.1.1 BDD-sample-case

In the BDD-sample-case, we are given a sample $t = v + e_0$ with $v \in \Lambda$ and $e_0 \leftarrow \chi$. Here we explicitly assume that χ is the n-dimensional, continuous Gaussian distribution with covariance matrix $\sigma_0^2 \cdot 1_n$ for some $\sigma_0 > 0$. Note that for any dual vector $w \in \hat{\Lambda}$, one has

$$\langle t, w\rangle \equiv \langle e_0, w\rangle \bmod 1.$$

3.1.2 Random-sample-case

We now provide a more precise definition of a random-sample, which is taken from the paper by Laarhoven and Walter [5, Definition 1]. Let Λ be a full-rank n-dimensional lattice. Let B be a basis of Λ. The random-sample distribution for Λ corresponds to the distribution obtained by sampling target vectors t uniformly at random from the fundamental parallelepiped generated by the basis B. We can write t as $t = B\psi$, where the components of ψ are uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$.

For two fixed dual vectors $w, \tilde{w} \in W$, $w \neq \tilde{w}$, we write explicitly

$$w = (B^{-1})^T \mu, \quad \tilde{w} = (B^{-1})^T \tilde{\mu},$$

where the components of μ and $\tilde{\mu}$ are integers.

3.1.3 Covariances

In the paper by Ducas and Pulles [7, Lemma 4], approximations of the expectation values and variances of a *single* $f_w(t)$ are given for the two cases "random-sample vs BDD-sample." In general, we have for the variance of the score

$$V(f_W(t)) = \sum_{w \in W} V(f_w(t)) + \sum_{w, \tilde{w} \in W, w \neq \tilde{w}} \mathrm{Cov}(f_w(t), f_{\tilde{w}}(t)).$$

If the independence assumption (cf. [7, Heuristic 2]) is valid, the second sum over the single covariances is equal to 0. However, in the following, we want to derive approximations of this second sum in the BDD-sample-case, which contradicts this claim. In the end, this might explain that in the experiments in the paper by Ducas and Pulles [7, Table 1], the measured variance is much larger than predicted.

3.2 Computing the covariances for random-samples

We begin with the random-sample-case, where we will see that (at least pairwise) independence assumption holds. We obtain the following result.

Proposition 3.2. *Let t be random-sample and $w, \tilde{w} \in W$ be such that w and $\tilde{w}$ are linearly independent. Let $s, \tilde{s} \in [-1/2, 1/2]$. Then,*

$$\mathbb{P}(\langle t, w\rangle \bmod 1 \leq s, \langle t, \tilde{w}\rangle \bmod 1 \leq \tilde{s}) = \left(s + \frac{1}{2}\right)\left(\tilde{s} + \frac{1}{2}\right).$$

In particular, we have that

$$\mathrm{Cov}(f_w(t), f_{\tilde{w}}(t)) = 0,$$

Proof. Recall the definition of ψ, μ, and $\tilde{\mu}$ from above. In particular recall that the components of μ and $\tilde{\mu}$ are integers. We consider the two-dimensional distribution of

$$\begin{pmatrix} \langle t, w\rangle \\ \langle t, \tilde{w}\rangle \end{pmatrix} = \begin{pmatrix} \langle \psi, \mu\rangle \\ \langle \psi, \tilde{\mu}\rangle \end{pmatrix}$$

and its reduction

$$\begin{pmatrix} \langle \psi, \mu\rangle \bmod 1 \\ \langle \psi, \tilde{\mu}\rangle \bmod 1 \end{pmatrix}$$

as a random variable in ψ. We want to compute the probabilities for $-1/2 \le s, \tilde{s} \le 1/2$.

$$\mathbb{P}(\langle \psi, \mu\rangle \bmod 1 \le s, \langle \psi, \tilde{\mu}\rangle \bmod 1 \le \tilde{s}) = \operatorname{Vol}(\langle \psi, \mu\rangle \bmod 1 \le s, \langle \psi, \tilde{\mu}\rangle \bmod 1 \le \tilde{s}).$$

We can compute this volume as sub-volume of the n-dimensional cube by counting over the points $(\frac{k_1}{p}, \dots, \frac{k_n}{p})$, k_j integers with $-p/2 \le k_j \le p/2$, for very large prime p and going to the limit. As approximation we obtain the sum

$$\sum_{\substack{r,\tilde{r}, \text{ with} \\ r/p \le s, \tilde{r}/p \le \tilde{s}}} \left[\sum_{\substack{k_j, \text{ with} \\ \sum_j \mu_j k_j/p \bmod 1 = r/p, \\ \sum_j \tilde{\mu}_j k_j/p \bmod 1 = \tilde{r}/p}} \frac{1}{p^n} \right] = \sum_{\substack{r,\tilde{r}, \text{with} \\ r \le sp, \tilde{r} \le \tilde{s}p}} \left[\sum_{\substack{k_j, \text{ with} \\ \sum_j \mu_j k_j \bmod p = r, \\ \sum_j \tilde{\mu}_j k_j \bmod p = \tilde{r}}} \frac{1}{p^n} \right],$$

where $r, \tilde{r}$ are integers in $[-p/2, p/2]$. Since μ and $\tilde{\mu}$ are linearly independent (over the rational numbers or the real numbers), the second sum has exactly p^{n-2} solutions. In the end, we derive

$$\sum_{\substack{r,\tilde{r}, \text{ with} \\ r \le sp, \tilde{r} \le \tilde{s}p}} \frac{1}{p^2} = \frac{1}{p^2} \#\{[-p/2, sp] \cap \mathbb{Z}\} \cdot \#\{[-p/2, \tilde{s}p] \cap \mathbb{Z}\} \overset{p\to\infty}{\longrightarrow} \left(s + \frac{1}{2}\right)\left(\tilde{s} + \frac{1}{2}\right).$$

This shows that the random variables $\langle \psi, \mu\rangle \bmod 1$ and $\langle \psi, \tilde{\mu}\rangle \bmod 1$ are independent and the covariances vanish. □

3.3 Computing the covariances for BDD-samples

We now assume that t is chosen as a BDD-sample. Recall that e_0 is sampled from an n-dimensional, continuous Gaussian distribution with covariance matrix $\sigma_0^2 \cdot 1_n$. We obtain the following result.

Proposition 3.3. *Let t be a BDD-sample and $w, \tilde{w} \in W$ be such that w and $\tilde{w}$ are linearly independent. Then,*

$$\operatorname{Cov}(f_w(t), f_{\tilde{w}}(t)) = \frac{1}{2}\Delta_a + \frac{1}{2}\Delta_b - \Delta_c \cdot \Delta_d \tag{4}$$

where we set

$$\Delta_a = \mathrm{e}^{-2\pi^2 \|w+\tilde{w}\|^2 \sigma_0^2}, \quad \Delta_b = \mathrm{e}^{-2\pi^2 \|w-\tilde{w}\|^2 \sigma_0^2},$$
$$\Delta_c = \mathrm{e}^{-2\pi^2 \|w\|^2 \sigma_0^2}, \quad \Delta_d = \mathrm{e}^{-2\pi^2 \|\tilde{w}\|^2 \sigma_0^2}.$$

Proof. We fix two dual vectors $w, \tilde{w} \in W$, $w \ne \tilde{w}$ and consider the two-dimensional distribution of

$$\begin{pmatrix} \langle e_0, w\rangle \\ \langle e_0, \tilde{w}\rangle \end{pmatrix}$$

as a random variable in e_0. This random variable is again Gaussian distributed with covariance matrix

$$\Sigma = \sigma_0^2 \begin{pmatrix} \|w\|^2 & \langle w, \tilde{w}\rangle \\ \langle w, \tilde{w}\rangle & \|\tilde{w}\|^2 \end{pmatrix}.$$

Since we assume that w and $\tilde{w}$ are linear independent and hence, define a two-dimensional positive definite subspace of $\mathbb{R}^n$ and Σ is invertible. We set

$$\tilde{\mathbb{P}}(z) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \mathrm{e}^{-\frac{1}{2}z^T\Sigma^{-1}z}.$$

The distribution of the reduced random variable

$$c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \langle e_0, w\rangle \bmod 1 \\ \langle e_0, \tilde{w}\rangle \bmod 1 \end{pmatrix}$$

is equal to

$$\mathbb{P}(c) = \sum_{\mu\in\mathbb{Z}^2} \tilde{\mathbb{P}}(c+\mu).$$

We use the Poisson summation formula and obtain

$$\mathbb{P}(c) = \sum_{\mu\in\mathbb{Z}^2} \tilde{\mathbb{P}}(c+\mu) = \sum_{v\in\mathbb{Z}^2} \mathrm{e}^{-2\pi i\langle v,c\rangle}\mathrm{e}^{-2\pi^2 v^t\Sigma v}.$$

We now start the computation by

$$\begin{aligned}\mathbb{E}(f_w(t)\cdot f_{\tilde{w}}(t)) &= \int_{c_1,c_2} \cos(2\pi c_1)\cdot\cos(2\pi c_2)\mathbb{P}(c_1,c_2)\,\mathrm{d}c_1\mathrm{d}c_2 \\ &= \sum_{v\in\mathbb{Z}^2} \mathrm{e}^{-2\pi^2 v^t\Sigma v}\int_{c_1}\cos(2\pi c_1)\mathrm{e}^{-2\pi i v_1 c_1}\,\mathrm{d}c_1\cdot\int_{c_2}\cos(2\pi c_2)\mathrm{e}^{-2\pi i v_2 c_2}\,\mathrm{d}c_2.\end{aligned}$$

It is easily seen that each univariate integral (in c_1 or c_2, respectively) vanishes for all v_1, except for $v_1 = \pm 1$ and for $v_2 = \pm 1$, respectively. Namely, we have

$$2\int_0^1 \cos(2\pi nt)\mathrm{e}^{-2\pi imt}\mathrm{d}t = \int_0^1 \mathrm{e}^{2\pi i(n-m)t}\mathrm{d}t + \int_0^1 \mathrm{e}^{2\pi i(-n-m)t}\mathrm{d}t.$$

The first integral on the right-hand side vanishes except for $n = m$ and the second integral vanishes except for $n = -m$. Both integrals are equal to 1 if they do not vanish and the claim follows.

Therefore, we obtain

$$\mathbb{E}(f_w(t)\cdot f_{\tilde{w}}(t)) = \frac{1}{4}\sum_{v_1=\pm1, v_2\pm1} \mathrm{e}^{-2\pi^2 v^t\Sigma v} = \frac{1}{2}\Delta_a + \frac{1}{2}\Delta_b$$

Lemma 4 in the paper by Ducas and Pulles [7] states the equality for the expectation value

$$\mathbb{E}(f_w(t)) = \mathrm{e}^{-2\pi^2\sigma_0^2\|w\|^2}.$$

In the end, we derive for the covariance

$$\mathrm{Cov}(f_w(t), f_{\tilde{w}}(t)) = \frac{1}{2}\Delta_a + \frac{1}{2}\Delta_b - \Delta_c\cdot\Delta_d.$$ □

Remark 3.4. Of course, the just computed covariances seem to be non-vanishing (when w and $\tilde{w}$ are not orthogonal). However, as the formulas are quite hard to be intuitively comprehensible, we provide in Appendix A some heuristic estimates, which underline that the covariances are indeed nonzero and imply the bias of the variance, which was observed experimentally in in the paper by Ducas and Pulles [7].

4 Dual attack

In this section, we estimate the success probability of the above described dual attack and provide a quantification of the number of dual vectors that are needed in theory for a successful attack. The strategy is based on the application of the so-called *Hoeffding's inequality* for $f_W(t)$ with respect to the *probability measure conditioned on t*.

In order to do this, we proceed as follows. We first introduce some technical ingredients that we need in the proof. We then recall some known result on the (conditional) expectation of $f_W(t)$ that we will need. Afterwards, we state and prove the main results given by an estimate of the success probability and the cost of the attack.

4.1 Technical ingredients

4.1.1 Distribution of the dual vectors

In order to be able to quantify the distribution of the score, we need to make assumptions on the output of the lattice sieve algorithm, from which we obtain the dual vectors.

Heuristic 2. We make the following assumptions on the distribution of the family $\{w\}_{w\in W}$ of random dual vectors.
- $\{w\}_{w\in W}$ is a family of independent random variables.
- $\{w\}_{w\in W}$ is independent of the sample t.
- For all $w \in W$, we have that w is distributed according to $D_{\hat{\Lambda},\tau_0\sqrt{2\pi}}$, where $\tau_0 > 0$.

4.1.1.1 Justification of Heuristic 2

The first two bullet points are intuitively justified as the dual vectors are usually generated, via lattice reduction algorithms such as BKZ, independent of each other and also independently from the sample t. However, the justification of the last bullet point, where we make an assumption on the distribution of the dual vectors, is not obvious. While it is clear that some assumption on this distribution has to be made for a rigorous analysis, additional arguments are required to justify why we assume that the dual vectors are distributed according to a discrete Gaussian distribution.

We note that our assumption in Heuristic 2 aligns with the paper by Pouly and Shen [9], whereas in Ducas and Pulles [8] made the assumption that the dual vectors are uniformly distributed on a centered ball $B_r^n \subset \mathbb{R}^n$ of a certain radius r. We believe that the most realistic assumption, which resembles the behavior of the output of the lattice reduction algorithms most suitably, would be a combination of both assumptions, where the dual vectors are uniformly distributed on $(B_r^n \cap \hat{\Lambda})/B_{r'}^n$ for some $r > r' > 0$. However, under this assumption, the rigorous analysis becomes much more complicated as such an analysis would involve hypergeometric and Bessel functions (instead of the Gaussian kernel as in our case). Therefore, these both ways, i.e., the assumption from Heuristic 2 and the paper by Pouly and Shen [9] and the assumption from the paper by Ducas and Pulles [8], provide a first, meaningful step toward this most realistic assumption, the latter being left for future research.

4.1.2 Conditional expectation and conditional probability measure

In our proofs, we use the widely used concept in probability theory of *conditional expectation*. For an introduction into this topic we refer to any standard textbook in probability theory [12, Sections 33 and 34].

Here we only provide the intuitive description that the conditional expectation given t, denoted by $\mathbb{E}[\,\cdot\,|t]$, is the best prediction under the condition that we know t. In particular, it is again a random variable.

Intuitively, this object suits our setting as we are also given a sample t and afterwards we try to estimate the score dependent on this t.

We recall some important facts of the conditional expectation that we will use:

- Many properties of the conditional expectation and the associated conditional probability measure are harvested from the usual expectation and the probability measure. Examples are the linearity and monotonicity of the expectation; refer [12, Sections 33 and 34] for more information.
- The conditional expectation is only defined *almost surely*, i.e., everywhere except of a null set, i.e., a set of measure 0 for the underlying probability space. In our case, this probability space is given as the underlying probability space, on which the random variables t and $\{w\}_{w\in W}$ are formally defined. In this paper, we abuse notation and do not mention the fact that some of the equalities only hold almost surely. This has no effect on the final results as, for the practical relevance of these results, it suffices that the results hold almost everywhere (i.e., everywhere except of a nullset).
- The *conditional probability measure* is, as for the usual probability measure, defined as the conditional expectation of the indicator function $x \mapsto \mathbb{1}_A(x)$ with respect to a Borel subset A. That is, in our case, we have that $\mathbb{P}[A|t] = \mathbb{E}[\mathbb{1}_A|t]$.
- In the special case that a random variable X is independent of t, we have that (for any function G that satisfies some integrability condition)

$$\mathbb{E}[G(X, t)|t](\cdot) = \mathbb{E}_X[G(X, t(\cdot))], \tag{5}$$

  where the notation indicates that the left-hand side is a random variable (as per definition of the conditional expectation) and the right-hand side is an expectation over X but by taking $t(\cdot)$ as a given input parameter for the expectation.

A very important consequence of the above is given in the following lemma.

Lemma 4.1. *Recall that $m_0 = |W|$ and*

$$f_w(t) = \cos(2\pi\langle t, w\rangle) \quad \textit{for all } w \in W.$$

Write $W = \{w_1, \ldots, w_{m_0}\}$. Let $\{B_i\}_{i=1,\ldots,m_0}$ be an arbitrary family of Borel subsets of $\mathbb{R}$. Then, under Heuristic 2,

$$\mathbb{P}\left[\bigcap_{i=1}^{m_0}\{f_{w_i}(t) \in B_i\}|t\right] = \prod_{i=1}^{m_0}\mathbb{P}[f_{w_i}(t) \in B_i|t]. \tag{6}$$

In particular, we have that the sequence $\{f_{w_i}(t)\}_{i=1,\ldots,m_0} = \{f_w(t)\}_{w\in W}$ is independent with respect to the conditional probability measure $\mathbb{P}[\cdot|t]$.

Proof. We have that

$$\mathbb{P}\left[\bigcap_{i=1}^{m_0}\{f_{w_i}(t) \in B_i\}|t\right](\cdot) = \mathbb{E}\left[\prod_{i=1}^{m_0}\mathbb{1}_{B_i}(f_{w_i}(t))|t\right](\cdot) = \mathbb{E}[G(t, w_1, \ldots, w_{m_0})|t](\cdot),$$

where we defined the function $G : \mathbb{R}^n \times \mathbb{R}^{m_0} \to \mathbb{R}$ as $G(t, x_1, \ldots, x_{m_0}) = \prod_{i=1}^{m_0}\mathbb{1}_{B_i}(f_{x_i}(t))$. Then, by applying (5),

$$\mathbb{E}[G(t, w_1, \ldots, w_{m_0})|t](\cdot) = \mathbb{E}[G(t(\cdot), w_1, \ldots, w_{m_0})] = \mathbb{E}\left[\prod_{i=1}^{m_0}\mathbb{1}_{B_i}(f_{w_i}(t(\cdot)))\right].$$

Now we use that, inside the last expectation, $t(\cdot)$ is fixed. In particular, it acts as an additional parameter in the expectation and *not* as a running variable for the integration corresponding to the expectation. Hence, we can use the independence assumption from Heuristic 2 to obtain that

$$\mathbb{E}\left[\prod_{i=1}^{m_0}\mathbb{1}_{B_i}(f_{w_i}(t(\cdot)))\right] = \prod_{i=1}^{m_0}\mathbb{E}[\mathbb{1}_{B_i}(f_{w_i}(t(\cdot)))] = \prod_{i=1}^{m_0}\mathbb{E}[\mathbb{1}_{B_i}(f_{w_i}(t))](\cdot).$$

This concludes the proof. □

Remark 4.2. We emphasize here that the independence claim from Lemma 4.1 is with respect to $\mathbb{P}[\cdot|t]$. In particular, this is a different statement than the non-independence mentioned at the beginning of Section 3.

4.1.3 Hoeffding's inequality

Inspired by the impressive work of Pouly and Shen [9], we will make use of the classic Hoeffding's inequality, which is given as follows.

Lemma 4.3. *Let $X_1, \dots, X_m$ be independent random variables such that $a_i \leq X_i \leq b_i$ almost surely. Let $S_m = X_1 + \dots + X_m$. Then, for all $r > 0$,*

$$\mathbb{P}[S_m - \mathbb{E}[S_m] \geq r] \leq \mathrm{e}^{-2r^2(\sum_{i=1}^{m}(b_i-a_i)^2)^{-1}}.$$

However, in this work we apply this inequality not for $\mathbb{P}[\cdot]$ but for the conditional probability measure $\mathbb{P}[\cdot|t]$. It is straightforward to see that this inequality also holds true for the conditional probability measure, since the original proof only relies on elementary properties of the expectation and the probability measure, and these properties also hold true for their conditional versions.

4.1.4 Threshold

Our threshold for the distinguishing between the random-sample- and the BDD-sample-case is given as follows:

$$\alpha = c \cdot \mathrm{e}^{-2\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2 r_0^2} \tag{7}$$

for some arbitrarily chosen $c \in (0,1)$ and $\xi \in (0, 1/2)$.

4.1.5 Concentration of chi-squared-distribution

Another important result that we use is the following concentration inequalities for the chi-squared-distribution.

Lemma 4.4. *Let e be distributed according to an n-dimensional, continuous Gaussian distribution with covariance matrix $\sigma_0^2 \cdot 1_n$ for some $\sigma_0 > 0$. Then, for all $\xi \in (0, 1/2)$,*

$$\mathbb{P}\left[\frac{||e||^2}{\sigma_0^2} \in [n - 2n^{1/2+\xi}, n + 2n^{1/2+\xi} + 2n^{\xi}]\right] \geq 1 - \mathrm{e}^{-n^{\xi}}. \tag{8}$$

Proof. A proof can be found in the paper by Laurent and Massart [13, Lemma 1]. □

4.2 Estimating the conditional expectation of $f_W(t)$

A crucial step to successfully apply Hoeffding's inequality for $f_W(t)$ is to find good estimates on the conditional expectation of $f_W(t)$ given t. This is done in the following. We first provide the estimates for the BDD-sample-case. Then, we consider two different approaches for the random-sample-case.

4.2.1 BDD-sample-case

Lemma 4.5. *Let $t = v + e_0$ with $v \in \Lambda$ and $e_0 \leftarrow \chi$. Then,*

$$\mathbb{E}\left[\frac{1}{m_0} f_W(t) | t\right] \geq \mathrm{e}^{-2\pi^2\tau_0^2\|e_0\|^2}.$$

Proof. Using the linearity of the conditional expectation that all dual vectors are identically distributed as some $w \sim D_{\hat{\Lambda},\tau_0\sqrt{2\pi}}$ and identity (5), we have that

$$\begin{aligned}\mathbb{E}\left[\frac{1}{m_0} f_W(t) | t\right] &= \frac{1}{m_0} \sum_{w\in W} \mathbb{E}[\cos(2\pi\langle t, w\rangle)|t] \\ &= \frac{1}{m_0} \sum_{w\in W} \mathbb{E}_{w\sim D_{\hat{\Lambda},\tau_0\sqrt{2\pi}}}[\cos(2\pi\langle t, w\rangle)] \\ &= \mathbb{E}_{w\sim D_{\hat{\Lambda},\tau_0\sqrt{2\pi}}}[\cos(2\pi\langle t, w\rangle)].\end{aligned} \tag{9}$$

Applying now the Poisson summation formula, we obtain that

$$\begin{aligned}\mathbb{E}_{w\sim D_{\hat{\Lambda},\tau_0\sqrt{2\pi}}}[\cos(2\pi\langle t, w\rangle)] &= \frac{\sum_{w\in\hat{\Lambda}} \cos(2\pi\langle t, w\rangle)\rho_{\tau_0\sqrt{2\pi}}(w)}{\sum_{w\in\hat{\Lambda}}\rho_{\tau_0\sqrt{2\pi}}(w)} \\ &= \frac{\sum_{w\in\hat{\Lambda}} \mathrm{e}^{-2\pi i\langle t,w\rangle}\rho_{\tau_0\sqrt{2\pi}}(w)}{\sum_{w\in\hat{\Lambda}}\rho_{\tau_0\sqrt{2\pi}}(w)} \\ &= \frac{\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - t)}{\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda)}.\end{aligned}$$

Moreover, using that $t = v + e_0$ and that $v \in \Lambda$, we obtain that

$$\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - t) = \rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - (v + e_0)) = \rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - e_0).$$

At this point, we can apply the following standard lower bound from the thesis by Stephens-Davidowitz [14, Lemma 1.3.10]

$$\begin{aligned}\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - e_0) &= \sum_{x\in\Lambda} \rho_{(\tau_0\sqrt{2\pi})^{-1}}(x - e_0) = \frac{1}{2}\sum_{x\in\Lambda} (\rho_{(\tau_0\sqrt{2\pi})^{-1}}(x - e_0) + \rho_{(\tau_0\sqrt{2\pi})^{-1}}(-x - e_0)) \\ &= \rho_{(\tau_0\sqrt{2\pi})^{-1}}(e_0) \sum_{x\in\Lambda} \rho_{(\tau_0\sqrt{2\pi})^{-1}}(x) \cosh(4\pi^2\tau_0^2\langle x, e_0\rangle) \\ &\geq \rho_{(\tau_0\sqrt{2\pi})^{-1}}(e_0) \sum_{x\in\Lambda} \rho_{(\tau_0\sqrt{2\pi})^{-1}}(x) \\ &= \mathrm{e}^{-2\pi^2\tau_0^2\|e_0\|^2}\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda).\end{aligned}$$

We finally obtain that

$$\mathbb{E}_{w\sim D(\hat{\Lambda},\tau_0\sqrt{2\pi})}[\cos(2\pi\langle e_0, w\rangle)] = \frac{\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - e_0)}{\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda)} \geq \mathrm{e}^{-2\pi^2\tau_0^2\|e_0\|^2}.$$

This concludes the proof. □

4.2.2 Random-sample-case – Approach 1

Lemma 4.6. *Let t be a random-sample. Then, whenever* $\operatorname{dist}(t, \Lambda) \geq r \coloneqq (\tau_0 2\pi)^{-1}\sqrt{n}$,

$$\mathbb{E}\left[\frac{1}{m_0} f_W(t)|t\right] \leq \mathrm{e}^{-2\pi^2\tau_0^2(\operatorname{dist}(t,\Lambda)-r)^2}.$$

Proof. Using the same arguments as above, we have that

$$\mathbb{E}\left[\frac{1}{m_0} f_W(t)|t\right] = \mathbb{E}_{w\sim D_{\hat{\Lambda},\tau_0\sqrt{2\pi}}}[\cos(2\pi\langle t, w\rangle)] = \frac{\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - t)}{\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda)}.$$

Now, we can proceed as in the paper by Pouly and Shen [9, Lemma 8] to bound the right-hand side and to conclude the proof. □

Remark 4.7. Note that the paper by Pouly and Shen [9, Lemma 8], which is the core of the proof of Lemma 4.6, relies heavily on the classic estimates from the paper by Banaszczyk [15]. In Section 5, where we heuristically reprove the statements of this section, we will directly apply the estimates from the paper by Banaszczyk [15]. In order to fully understand the parallels between the approach in this section and the one from Section 5, it is important to have this in mind.

4.2.3 Random-sample-case – Approach 2

Lemma 4.8. *Let t be a random-sample. Let $\zeta \geq 1$. Then,*

$$\mathbb{E}\left[\frac{1}{m_0} f_W(t)|t\right] \leq (\sqrt{2(1+\zeta)})^n \mathrm{e}^{-\pi^2\tau_0^2\left(1-\frac{1}{\zeta}\right)\operatorname{dist}(t,\Lambda)^2}. \tag{10}$$

Proof. Using the same arguments as above, we have that

$$\mathbb{E}\left[\frac{1}{m_0} f_W(t)|t\right] = \frac{\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - t)}{\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda)}. \tag{11}$$

We now estimate the numerator on the right-hand side. It consists of a sum of exponentials of the form $\mathrm{e}^{-2\pi^2\tau_0^2\|v-t\|^2}$, where $v \in \Lambda$. We focus on estimating the factor $\|v - t\|^2$ in the exponent. Here we are inspired by the paper by Chen et al. [16, Appendix A].

Let $K(t)$ be a closest lattice vector to t and let $\delta = \frac{1}{2}(1+\zeta)^{-1}$. Then, by the triangle inequality, we have that

$$\begin{aligned}
\|v - t\|^2 &\geq \frac{1}{2}\|v - t\|^2 + \frac{1}{2}\|K(t) - t\|^2 \\
&= \frac{1}{2}\|v - t\|^2 + \frac{1}{2}\|K(t) - t\|^2 - \delta(\|v - t\| + \|K(t) - t\|)^2 + \delta(\|v - t\| + \|K(t) - t\|)^2 \\
&\geq \frac{1}{2}\|v - t\|^2 + \frac{1}{2}\|K(t) - t\|^2 - \delta(\|v - t\| + \|K(t) - t\|)^2 + \delta(\|v - t + t - K(t)\|)^2 \\
&= \left(\frac{1}{2} - \delta\right)\|v - t\|^2 + \left(\frac{1}{2} - \delta\right)\|K(t) - t\|^2 - 2\delta\|v - t\|\|K(t) - t\| + \delta\|v - K(t)\|^2.
\end{aligned}$$

Applying now the Peter–Paul inequality ($2ab \leq \zeta a^2 + \zeta^{-1}b^2$ for $a, b \in \mathbb{R}$), we obtain that

$$\begin{aligned}
\|v - t\|^2 &\geq \left(\frac{1}{2} - \delta\right)\|v - t\|^2 + \left(\frac{1}{2} - \delta\right)\|K(t) - t\|^2 - \delta\zeta\|v - t\|^2 - \delta\frac{1}{\zeta}\|K(t) - t\|^2 + \delta\|v - K(t)\|^2 \\
&= \left(\frac{1}{2} - \delta - \delta\zeta\right)\|v - t\|^2 + \left(\frac{1}{2} - \delta - \delta\frac{1}{\zeta}\right)\|K(t) - t\|^2 + \delta\|v - K(t)\|^2.
\end{aligned}$$

Due to our choice of δ, we have that

$$\left(\frac{1}{2} - \delta - \delta\zeta\right) = 0, \qquad \left(\frac{1}{2} - \delta - \delta\frac{1}{\zeta}\right) = \frac{1}{2}\left(1 - \frac{1}{\zeta}\right).$$

This yields that

$$\|v - t\|^2 \geq \frac{1}{2}\left(1 - \frac{1}{\zeta}\right)\|K(t) - t\|^2 + \delta\|v - K(t)\|^2.$$

Via this estimate, we obtain that

$$\begin{aligned}
\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda - t) &= \sum_{v\in\Lambda} \mathrm{e}^{-2\pi^2\tau_0^2\|v-t\|^2} \\
&\leq \mathrm{e}^{-2\pi^2\tau_0^2\frac{1}{2}\left(1-\frac{1}{\zeta}\right)\|t-K(t)\|^2} \sum_{v\in\Lambda} \mathrm{e}^{-2\pi^2\tau_0^2\delta\|K(t)-v\|^2} \\
&= \mathrm{e}^{-\pi^2\tau_0^2\left(1-\frac{1}{\zeta}\right)\|t-K(t)\|^2} \sum_{v\in\Lambda} \mathrm{e}^{-2\pi^2\tau_0^2\delta\|v\|^2} \\
&= \mathrm{e}^{-\pi^2\tau_0^2\left(1-\frac{1}{\zeta}\right)\|t-K(t)\|^2} \rho_{(\tau_0\sqrt{2\pi\delta})^{-1}}(\Lambda).
\end{aligned} \tag{12}$$

Via the Poisson summation formula and since $\delta \leq 1$, we have that

$$\begin{aligned}
\rho_{(\tau_0\sqrt{2\pi\delta})^{-1}}(\Lambda) &= \frac{(\tau_0\sqrt{2\pi\delta})^{-n}}{\det(\Lambda)} \sum_{w\in\hat{\Lambda}} \rho_{\tau_0\sqrt{2\pi\delta}}(w) \\
&\leq \frac{(\tau_0\sqrt{2\pi\delta})^{-n}}{\det(\Lambda)} \sum_{w\in\hat{\Lambda}} \rho_{\tau_0\sqrt{2\pi}}(w) = \delta^{-n/2}\rho_{(\tau_0\sqrt{2\pi})^{-1}}(\Lambda).
\end{aligned} \tag{13}$$

Combining (13), (12), and (11) yields (10). □

4.3 Estimating the success probabilities

We now have collected everything we need to formulate some of the main results of this paper in the following three theorems. We first consider the BDD-sample-case. Then, we consider two different approaches for the random-sample-case.

4.3.1 BDD-sample-case

Theorem 4.9. *Let $t = v + e_0$ with $v \in \Lambda$ and $e_0 \leftarrow \chi$. Let α, c, ξ be as in (7). Then,*

$$\mathbb{P}[f_W(t) < m_0\alpha] \leq \mathrm{e}^{-\frac{1}{2}m_0(1-c)^2\mathrm{e}^{-4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2}} + \mathrm{e}^{-n^{\xi}}. \tag{14}$$

Proof. We begin with some elementary reformulations of the desired probability.

$$\begin{aligned}
\mathbb{P}[f_W(t) < m_0\alpha] &= \mathbb{P}[-f_W(t) > -m_0\alpha] \\
&= \mathbb{P}[-f_W(t) - \mathbb{E}[-f_W(t)|t] > \mathbb{E}[f_W(t)|t] - m_0\alpha] \\
&\leq \mathbb{P}[-f_W(t) - \mathbb{E}[-f_W(t)|t] > m_0(\mathrm{e}^{-2\pi^2\|e_0\|^2\tau_0^2} - \alpha)],
\end{aligned} \tag{15}$$

where we have used Lemma 4.5 in the last step.

We then rewrite the right-hand side of (15) by using basic rules of probability theory to see that

$$\begin{aligned}
&\mathbb{P}[-f_W(t) - \mathbb{E}[-f_W(t)|t] > m_0(\mathrm{e}^{-2\pi^2\|e_0\|^2\tau_0^2} - \alpha)] \\
&= \mathbb{P}\left[-f_W(t) - \mathbb{E}[-f_W(t)|t] > m_0(\mathrm{e}^{-2\pi^2\|e_0\|^2\tau_0^2} - \alpha), \frac{\|e_0\|^2}{\sigma_0^2} < n + 2n^{1/2+\xi} + 2n^{\xi}\right] \\
&\quad + \mathbb{P}\left[-f_W(t) - \mathbb{E}[-f_W(t)|t] > m_0(\mathrm{e}^{-2\pi^2\|e_0\|^2\tau_0^2} - \alpha), \frac{\|e_0\|^2}{\sigma_0^2} \geq n + 2n^{1/2+\xi} + 2n^{\xi}\right] \\
&\leq \mathbb{P}[-f_W(t) - \mathbb{E}[-f_W(t)|t] > m_0(\mathrm{e}^{-2\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2} - \alpha)] + \mathbb{P}\left[\frac{\|e_0\|^2}{\sigma_0^2} \geq n + 2n^{1/2+\xi} + 2n^{\xi}\right] \\
&= \mathbb{E}[\mathbb{P}[-f_W(t) - \mathbb{E}[-f_W(t)|t] > m_0(1-c)\mathrm{e}^{-2\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2}|t]] + \mathbb{P}\left[\frac{\|e_0\|^2}{\sigma_0^2} \geq n + 2n^{1/2+\xi} + 2n^{\xi}\right] \\
&\leq \mathrm{e}^{-\frac{1}{2}m_0(1-c)^2\mathrm{e}^{-4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2}} + \mathrm{e}^{-n^{\xi}},
\end{aligned}$$

where we applied Hoeffding's inequality for conditional probability measures and the concentration inequality (8) in the last step. This concludes the proof. □

4.3.2 Random-sample-case – Approach 1

Theorem 4.10. *Let t be a random-sample and α, c, ξ be as in* (7)*. Let $\ell \in \mathbb{N}$ be some arbitrary positive integer. Suppose that σ_0, τ_0, and n are such that*

$$\frac{1}{\tau_0 2\pi} + \sigma_0\sqrt{1 - \frac{\log(c/2)}{2\pi^2\tau_0^2\sigma_0^2 n} + 2n^{-(1/2-\xi)} + n^{-(1-\xi)}} \leq \frac{1}{\sqrt{2\pi\mathrm{e}}}\left(\frac{\det(\Lambda)}{\ell}\right)^{\frac{1}{n}}. \tag{16}$$

Then,

$$\mathbb{P}[f_W(t) > m_0\alpha] \leq \mathrm{e}^{-\frac{1}{2}m_0\frac{c^2}{4}\mathrm{e}^{-4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2}} + \frac{1}{\ell}. \tag{17}$$

Proof. Let

$$\delta_n := r + \sqrt{-\log(c/2)(2\pi^2\tau_0^2)^{-1} + (n + 2n^{1/2+\xi} + n^{\xi})\sigma_0^2},$$

where r is defined in Lemma 4.6. Via some elementary reformulations, we obtain that

$$\begin{aligned}
\mathbb{P}[f_W(t) > m_0\alpha] &= \mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - \mathbb{E}[f_W(t)|t]] \\
&= \mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - \mathbb{E}[f_W(t)|t], \mathrm{dist}(t,\Lambda) \leq \delta_n] \\
&\quad + \mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - \mathbb{E}[f_W(t)|t], \mathrm{dist}(t,\Lambda) > \delta_n] \\
&\leq \mathbb{P}[\mathrm{dist}(t,\Lambda) \leq \delta_n] + \mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - m_0\mathrm{e}^{-2\pi^2(\mathrm{dist}(t,\Lambda)-r)^2\tau_0^2}, \mathrm{dist}(t,\Lambda) > \delta_n],
\end{aligned}$$

where we have used Lemma 4.6 in the last step, which is applicable since $\mathrm{dist}(t,\Lambda) > \delta_n > r$.

Estimation of the first term. In order to estimate the first term, note that

$$\mathbb{P}[\mathrm{dist}(t,\Lambda) \leq \delta_n] = \frac{1}{\det(\Lambda)}\mathrm{Vol}(\mathcal{V}(\Lambda) \cap B_{\delta_n}^n) \leq \frac{1}{\det(\Lambda)}\mathrm{Vol}(B_{\delta_n}^n), \tag{18}$$

where $\mathcal{V}(\Lambda)$ is the Voronoi cell of Λ and $B_{\delta_n}^n$ is the n-dimensional ball of radius δ_n around the origin. We show in Appendix B that

$$\mathrm{Vol}(B_{\delta_n}^n) \leq \left(\sqrt{\frac{2\pi\mathrm{e}}{n}}\,\delta_n\right)^n. \tag{19}$$

Moreover, by assumption (16),

$$\delta_n = \frac{1}{\tau_0 2\pi}\sqrt{n} + \sigma_0\sqrt{n}\sqrt{1 - \frac{\log(c/2)}{2\pi^2\tau_0^2\sigma_0^2 n} + 2n^{-(1/2-\xi)} + n^{-(1-\xi)}} \leq \sqrt{n}\frac{1}{\sqrt{2\pi \mathrm{e}}}\left(\frac{\det(\Lambda)}{\ell}\right)^{\frac{1}{n}}. \tag{20}$$

Combining (18), (19), and (20) yields that

$$\mathbb{P}[\mathrm{dist}(t,\Lambda) \leq \delta_n] \leq \frac{1}{\ell}.$$

Estimation of the second term. For the second term, we observe that

$$\begin{aligned}
&\mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - m_0\mathrm{e}^{-2\pi^2(\mathrm{dist}(t,\Lambda)-r)^2\tau_0^2}, \mathrm{dist}(t,\Lambda) > \delta_n] \\
&\quad\leq \mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - m_0\mathrm{e}^{-2\pi^2(\delta_n-r)^2\tau_0^2}] \\
&\quad= \mathbb{P}\left[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - m_0\frac{c}{2}\mathrm{e}^{-2\pi^2(n+2n^{1/2+\xi}+n^{\xi})\sigma_0^2\tau_0^2}\right] \\
&\quad= \mathbb{P}\left[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\frac{\alpha}{2}\right].
\end{aligned}$$

As above, we can now apply Hoeffding's inequality by invoking the probability measure conditioned on t. This concludes the proof. □

4.3.3 Random-sample-case – Approach 2

Theorem 4.11. *Let t be a random-sample and α, c, ξ be as in (7) and ζ as in Lemma 4.8. Let $\ell \in \mathbb{N}$ be some arbitrary positive integer. Suppose that σ_0, τ_0, and n are such that*

$$\sqrt{\left(\pi^2\tau_0^2\left(1 - \frac{1}{\zeta}\right)\right)^{-1}}\sqrt{\frac{1}{2}\log(2(1+\zeta)) - \log(c/2)n^{-1} + 2\pi^2\tau_0^2\sigma_0^2(1 + 2n^{-1/2+\xi} + 2n^{-1+\xi})} \leq \frac{1}{\sqrt{2\pi\mathrm{e}}}\left(\frac{\det(\Lambda)}{\ell}\right)^{\frac{1}{n}}. \tag{21}$$

Then,

$$\mathbb{P}[f_W(t) > m_0\alpha] \leq \mathrm{e}^{-\frac{1}{2}m_0\frac{c^2}{4}\mathrm{e}^{-4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2}} + \frac{1}{\ell}. \tag{22}$$

Proof. Let

$$\delta_n^2 \coloneqq \left(\pi^2\tau_0^2\left(1 - \frac{1}{\zeta}\right)\right)^{-1}(\log((2(1+\zeta))^{n/2}) - \log(\alpha/2)).$$

Then, we know that $\delta_n = O(\sqrt{n})$, since by the definition of α (see (7)),

$$\begin{aligned}
\delta_n^2 &= \left(\pi^2\tau_0^2\left(1 - \frac{1}{\zeta}\right)\right)^{-1}\left(\frac{n}{2}\log(2(1+\zeta)) - \log(c/2) + 2\pi^2\tau_0^2\sigma_0^2(n + 2n^{1/2+\xi} + 2n^{\xi})\right) \\
&= n\left(\pi^2\tau_0^2\left(1 - \frac{1}{\zeta}\right)\right)^{-1}\left(\frac{1}{2}\log(2(1+\zeta)) - \log(c/2)n^{-1} + 2\pi^2\tau_0^2\sigma_0^2(1 + 2n^{-1/2+\xi} + 2n^{-1+\xi})\right).
\end{aligned}$$

We now use the same first step as in the proof of Theorem 4.10 but apply Lemma 4.8 instead of Lemma 4.6. Then, we obtain that

$$\begin{aligned}
\mathbb{P}[f_W(t) > m_0\alpha] = &\mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - \mathbb{E}[f_W(t)|t], \mathrm{dist}(t,\Lambda) \leq \delta_n] \\
&+ \mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - \mathbb{E}[f_W(t)|t], \mathrm{dist}(t,\Lambda) > \delta_n]
\end{aligned}$$

$$\begin{aligned}
&\leq \mathbb{P}[\operatorname{dist}(t,\Lambda) \leq \delta_n] + \mathbb{P}\Big[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - m_0(\sqrt{2(1+\zeta)})^n \mathrm{e}^{-\pi^2\tau_0^2\left(1-\frac{1}{\zeta}\right)\operatorname{dist}(t,\Lambda)^2},\\
&\quad \operatorname{dist}(t,\Lambda) > \delta_n\Big]\\
&\leq \mathbb{P}[\operatorname{dist}(t,\Lambda) \leq \delta_n] + \mathbb{P}\left[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha - m_0(2(1+\zeta))^{n/2}\mathrm{e}^{-\pi^2\tau_0^2\left(1-\frac{1}{\zeta}\right)\delta_n^2}\right]\\
&= \mathbb{P}[\operatorname{dist}(t,\Lambda) \leq \delta_n] + \mathbb{P}[f_W(t) - \mathbb{E}[f_W(t)|t] > m_0\alpha/2].
\end{aligned}$$

For the second term we can, as above, apply Hoeffding's inequality. In order to estimate the first term, we again proceed similar to the proof of Theorem 4.10. That is, we have that

$$\mathbb{P}[\operatorname{dist}(t,\Lambda) \leq \delta_n] \leq \frac{1}{\det(\Lambda)}\operatorname{Vol}(B_{\delta_n}^n) \leq \frac{1}{\det(\Lambda)}\left(\sqrt{\frac{2\pi\mathrm{e}}{n}}\,\delta_n\right)^n. \tag{23}$$

Moreover, by assumption (21),

$$\begin{aligned}
\delta_n &= \sqrt{n}\sqrt{\left(\pi^2\tau_0^2\left(1-\frac{1}{\zeta}\right)\right)^{-1}}\sqrt{\frac{1}{2}\log(2(1+\zeta)) - \log(c/2)n^{-1} + 2\pi^2\tau_0^2\sigma_0^2(1+2n^{-1/2+\xi}+2n^{-1+\xi})}\\
&\leq \sqrt{n}\frac{1}{\sqrt{2\pi\mathrm{e}}}\left(\frac{\det(\Lambda)}{\ell}\right)^{\frac{1}{n}}.
\end{aligned} \tag{24}$$

Combining (23) and (24) yields that

$$\mathbb{P}[\operatorname{dist}(t,\Lambda) \leq \delta_n] \leq \frac{1}{\ell}.$$

This concludes the proof. □

4.4 Conclusion

In order for our strategy described in Section 2.3 to be successful we need that the probabilities on the left-hand side in (14) and (17) (or equivalently (22)) become small. This in turn is given if their respective upper bound, i.e., the respective right-hand sides in (14) and (17) (or equivalently (22)), become small.

In order to achieve this, we choose n and ℓ large enough so that the second terms on the right-hand sides in (14) and (17) (or equivalently (22)), respectively, become smaller than, say, 1/4. It remains to investigate the first terms on the right-hand sides in (14) and (17) (or equivalently (22)).

More precisely, we now need to show that, for some suitable $\ell' \geq 4$, the following two conditions hold true.

$$\mathrm{e}^{-\frac{1}{2}m_0\frac{c^2}{4}}\mathrm{e}^{-4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2} \leq \frac{1}{\ell'} \quad \text{and} \quad \mathrm{e}^{-\frac{1}{2}m_0(1-c)^2}\mathrm{e}^{-4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2} \leq \frac{1}{\ell'}.$$

For that to be true, we derive the condition that

$$\frac{1}{2}m_0 \geq \log(\ell') \cdot \max((1-c)^{-2}, 4c^{-2}) \cdot \mathrm{e}^{4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2}. \tag{25}$$

Obviously, the maximum in (25) is realized for $c = 2/3$, i.e., $\max((1-c)^{-2}, 4c^{-2}) = 9$. Hence, we have the condition that

$$m_0 \geq \log(\ell') \cdot 18 \cdot \mathrm{e}^{4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2}. \tag{26}$$

This yields a lower bound on the necessary number of dual vectors for a successful dual attack.

Remark 4.12. Note that we can choose ℓ and ℓ' such that the probabilities computed in Theorems 4.9–4.11 vanish as $n \to \infty$. It even turns out that choosing ℓ and ℓ' in such a way only has a mild effect on the main results of this paper.

Indeed, as for ℓ', its influence in (26) is of logarithmic order. Therefore, choosing ℓ' increasing in n (e.g., as $\ell' = n$) only has a mild effect on the total complexity of the attack determined by (26).

A similar observation is also true for ℓ. As shown in Section 4.5, its influence on the main results of this paper is given in the parameter constrains and is determined by a factor of the form $\ell^{\frac{1}{n}}$. Now, due to the fact that this term is $1 + o(1)$ in both cases, if we choose $\ell = O(1)$ or if we, e.g., choose $\ell = n$, we observe that choosing ℓ mildly increasing in n only has a mild effect on the parameter constraints of this paper; refer Section 4.5 for more details.

We omit the details concerning these aspects as they are dependent on the concrete scenario, where the results of this paper are used and can be adjusted easily to this scenario.

4.5 Parameter selection

It is important to ask, for which sets of parameters $n, \sigma_0, \tau_0, \xi, \ell$ the results of this section hold true (recall that we already chose $c = 2/3$). We first note that the BDD-sample-result (Theorem 4.9) does not impose any further restrictions on the parameters. This is not the case for the results of Theorems 4.10 and 4.11 in the random-sample-case, where we need to consider conditions (16) and (21), respectively.

In this section, we investigate when these restrictions hold true. In doing so, we restrict our attention to the parameters τ_0 and σ_0 in the following. The reason is that the parameters c, ξ, ℓ do not have a significant influence, since (as n becomes large) they only appear in lower-order terms in conditions (16) and (21) as well as in the complexity estimate (26). Therefore, it is easy to find suitable parameters ξ, ℓ by mildly adjusting the dimension n.

Finally in this section, we investigate if and how the lower bound on m_0 given in (26) has implications on how realistic Heuristic 2 resembles the behavior of lattice reduction algorithms.

4.5.1 Intuition for the parameter τ_0

A very prominent role in this investigation will be played by the parameter τ_0. Recall that this parameter, according to Heuristic 2, determines the distribution of the dual vectors. In particular, due to standard properties of the discrete Gaussian distribution, which are comparable to Lemma 4.4, this yields to the heuristic that, on average,

$$\|w\| \approx \tau_0 \sqrt{n}\,, \tag{27}$$

for all $w \in W$. This intuition suggests to assume that, for some $\vartheta_0 > 1$, τ_0 is to be chosen as

$$\tau_0 \sqrt{n} = \vartheta_0 \lambda_1(\hat{\Lambda}) \approx \vartheta_0 \frac{1}{\sqrt{2\pi e}} \det(\hat{\Lambda})^{1/n} \sqrt{n}\,, \tag{28}$$

since we expect the output of the lattice reduction algorithms to have the length of a multiple of the shortest vector. This translates the search for suitable τ_0 to the search for appropriate ϑ_0.

From our computations, we will find out that, in order for (16) or (21) to be fulfilled, we need a lower bound on ϑ_0. In fact, one can construct special cases of lattices, where our formulas are not valid for small ϑ_0. For example, if there is just one very small dual vector w_0, the score function does not depend mainly on the length of e_0 but on the projection of e_0 onto $\mathrm{span}(w_0)$. However, we believe that the score function should still allow us to distinguish between the BDD-sample-case and the random-sample-case, even if ϑ_0 is small. It is left for future research to quantitatively investigate these cases.

We interpret our results, where we demand a lower bound on ϑ_0, in the way that, on the one hand, an attacker can (or even must) rely on a more coarse lattice reduction algorithm, where ϑ_0 does not need to be so small that only the shortest dual vectors need to be sampled. While, on the other hand, ϑ_0 should not become too large, as the complexity estimate (26) increases exponentially in the parameter τ_0 (and thus in ϑ_0).

4.5.2 Intuition for the parameter σ_0

Another important task is to appropriately choose the parameter σ_0. Also, here we have, according to Lemma 4.4, a similar heuristic, which states that, on average,

$$\|e_0\| \approx \sigma_0\sqrt{n}\,.$$

Similarly as in (28), this suggests to assume that

$$\sigma_0\sqrt{n} \;=\; \theta_0\lambda_1(\Lambda) \approx \theta_0\frac{1}{\sqrt{2\pi e}}\det(\Lambda)^{1/n}\sqrt{n}\,, \tag{29}$$

for some appropriate choice $\theta_0 > 0$. However, due to the fact that σ_0 determines the variance of the error distribution for a BDD-sample, we need to make the assumption $0 < \theta_0 < 1$ here.

4.5.3 Parameter choices for (16)

We first consider condition (16). We abbreviate

$$\kappa_n = \sqrt{1 - \frac{\log(c/2)}{2\pi^2\tau_0^2\sigma_0^2 n} + 2n^{-(1/2-\xi)} + n^{-(1-\xi)}}.$$

Note that $\kappa_n = 1 + o(1)$ asymptotically in n. This transforms condition (16) into

$$\frac{1}{\tau_0 2\pi} + \sigma_0\kappa_n \le \frac{1}{\sqrt{2\pi\mathrm{e}}}\left(\frac{\det(\Lambda)}{\ell}\right)^{\frac{1}{n}}.$$

With our choices (28) and (29), this becomes (by using that $\det(\hat{\Lambda}) = \det(\Lambda)^{-1}$)

$$\frac{\det(\Lambda)^{\frac{1}{n}}\sqrt{e}}{\vartheta_0\sqrt{2\pi}} + \theta_0\frac{1}{\sqrt{2\pi e}}\det(\Lambda)^{1/n}\kappa_n \le \frac{1}{\sqrt{2\pi\mathrm{e}}}\left(\frac{\det(\Lambda)}{\ell}\right)^{\frac{1}{n}},$$

or equivalently,

$$\frac{e}{\vartheta_0} + \theta_0\kappa_n \le \frac{1}{\ell^{\frac{1}{n}}}. \tag{30}$$

We immediately see that if $\vartheta_0 \le e$, inequality (30) is violated (since $\sigma_0 > 0$). Then, one admissible choice that fulfills (30) (and thus, also (16)) would be

$$\vartheta_0 \ge pe\ell^{\frac{1}{n}} \quad \text{and} \quad \theta_0 \le \frac{1}{q\kappa_n\ell^{\frac{1}{n}}} \tag{31}$$

with $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. For example, we could choose $p = q = 2$. We finally note that asymptotically in n (as also $\ell^{\frac{1}{n}} = 1 + o(1)$),

$$\vartheta_0 \ge pe(1 + o(1)) \quad \text{and} \quad \theta_0 \le \frac{1}{q}\,(1 + o(1)). \tag{32}$$

4.5.4 Parameter choices for (21)

We proceed similarly as for (16). We abbreviate

$$\begin{aligned}\varphi_n &= -\log\left(\frac{c}{2}\right)n^{-1} + 2\pi^2\tau_0^2\sigma_0^2(2n^{-1/2+\xi} + 2n^{-1+\xi}) \\ &= -\log\left(\frac{c}{2}\right)n^{-1} + \frac{\vartheta_0^2\theta_0^2}{2e^2}(2n^{-1/2+\xi} + 2n^{-1+\xi})\end{aligned}$$

where we used (28), (29), and that $\det(\hat{\Lambda}) = \det(\Lambda)^{-1}$ in the last step. Note that $\varphi_n = o(1)$ asymptotically in n. Then, again by using (28), (29), and $\det(\hat{\Lambda}) = \det(\Lambda)^{-1}$, condition (21) becomes

$$\frac{\det(\Lambda)^{\frac{1}{n}}\sqrt{2\pi e}}{\vartheta_0}\sqrt{\left(\pi^2\left(1-\frac{1}{\zeta}\right)\right)^{-1}}\sqrt{\frac{1}{2}\log(2(1+\zeta)) + \frac{\vartheta_0^2\theta_0^2}{2e^2} + \varphi_n} \le \frac{1}{\sqrt{2\pi \mathrm{e}}}\left(\frac{\det(\Lambda)}{\ell}\right)^{\frac{1}{n}},$$

or equivalently,

$$\sqrt{\frac{2e^2}{\vartheta_0^2}\left(1-\frac{1}{\zeta}\right)^{-1}(\log(2(1+\zeta)) + 2\varphi_n) + 2\theta_0^2\left(1-\frac{1}{\zeta}\right)^{-1}} \le \frac{1}{\ell^{\frac{1}{n}}}. \tag{33}$$

Then, one admissible choice that fulfills (33) (and thus also (21)) would be

$$\vartheta_0 \ge \ell^{\frac{1}{n}}\sqrt{p}\sqrt{2e^2\left(1-\frac{1}{\zeta}\right)^{-1}(\log(2(1+\zeta)) + 2\varphi_n)} \quad \text{and} \quad \theta_0 \le \frac{1}{\sqrt{q}\,\ell^{\frac{1}{n}}}\sqrt{\frac{1}{2}\left(1-\frac{1}{\zeta}\right)},$$

with $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Asymptotically in n, we obtain the conditions

$$\begin{aligned}\vartheta_0 &\ge \sqrt{p}\sqrt{2e^2\left(1-\frac{1}{\zeta}\right)^{-1}\log(2(1+\zeta))}\ (1+o(1)) \\ \theta_0 &\le \frac{1}{\sqrt{q}}\sqrt{\frac{1}{2}\left(1-\frac{1}{\zeta}\right)}\ (1+o(1)).\end{aligned} \tag{34}$$

A suitable choice of ζ for the most prominent case $p = q = 2$ would be for example $\zeta = 3$. In this case, we obtain that

$$\begin{aligned}\vartheta_0 &\ge \sqrt{6e^2\log(8)}\ (1+o(1)) \approx 3.532e \\ \theta_0 &\le \frac{1}{\sqrt{6}}\ (1+o(1)) \approx 0.408.\end{aligned}$$

which, for this particular choice of ζ, p, and q, shows that Approach 1 seems to induce better results. However, we decided to keep Approach 2 in this work as it has the potential for improvement by finding suitable replacements for the usage of rough inequalities like the Peter–Paul inequality.

4.5.5 Consequences of (26)

Recall that, following our findings in this section,

- on the one hand, we need to sample at least a certain amount (given by the right-hand side of (26)) of dual vectors, and
- on the other hand, following the heuristic according to (27), Heuristic 2 tacitly implies that the sampled dual vectors belong to B_R^n, the n-dimensional ball of radius R, where $R = \tau_0\sqrt{n}$.

Hence, if $|B_R^n \cap \hat{\Lambda}|$ is too small, it is very likely that, during the sampling process for our attack, some dual vectors are sampled more than once.

From a mathematical point of view, this implies no further issues as we modeled the output of the lattice reduction algorithm (via Heuristic 2) as a sequence of independent random variables and, as such, repetitions of the realizations of these random variables are admissible; see for instance coin tossing.

However, from a practical point of view, we have to make sure that Heuristic 2 indeed resembles the behavior of lattice reduction algorithms. Therefore, we have to take into account that, in practice, repeated sample outputs are usually discarded. This, of course, has the negative effect that in such scenarios the joint independence of the sample gets lost. This in turn implies that the independence assumption in Heuristic 2 does not resemble reality in situations, where many repetitions happen.

This is why we now investigate informally under which parameter selections repetitions can be avoided. In other words, we want to find out, when the condition

$$m_0 \ll |B_R^n \cap \hat{\Lambda}|$$

is fulfilled, while at the same time m_0 satisfies (26). In other words, we need that

$$\log(\ell') \cdot 18 \cdot \mathrm{e}^{4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2} \ll |B_R^n \cap \hat{\Lambda}|. \tag{35}$$

In order to quantify the right-hand side of (35), note that, according to Gaussian Heuristics,

$$|B_R^n \cap \hat{\Lambda}| \approx \frac{\mathrm{Vol}(B_R^n)}{\det(\hat{\Lambda})}.$$

In particular, using (28) and similar arguments as in the proof of Theorem 4.10, this implies that

$$|B_R^n \cap \hat{\Lambda}| \approx \frac{\mathrm{Vol}(B_R^n)}{\det(\hat{\Lambda})} \approx \vartheta_0^n.$$

As before, we now ignore the terms, which are (asymptotically in n) of lower-order, as they can be considered by mildly adjusting n. Consequently, condition (35) becomes

$$\mathrm{e}^{4\pi^2 n\sigma_0^2\tau_0^2} \ll \vartheta_0^n.$$

Using (28) and (29), this becomes

$$\mathrm{e}^{n\theta_0^2\vartheta_0^2/e^2} \ll \vartheta_0^n,$$

which in turn, for large n, is equivalent to the condition

$$\theta_0^2\vartheta_0^2/e^2 < \log(\vartheta_0). \tag{36}$$

This imposes additional conditions on θ_0 and ϑ_0.

4.5.6 Provable regime

To sum up, we found the following two sets of parameters, where our attacks work and are practically justified
(1) Using Approach 1: The set of all θ_0 and ϑ_0 that fulfills (32) and (36).
(2) Using Approach 2: The set of all θ_0 and ϑ_0 that fulfills (34) and (36).
We can simplify these conditions by plugging (32) into (36) (or, respectively, (34) into (36)) and ignoring terms that are of lower order in n. We obtain the following conditions.

(1) For Approach 1: By (32) we choose $\vartheta_0 = p\ e$ and $\theta_0 = \frac{1}{q}$, so that (36) becomes

$$\frac{1}{q^2}p^2 < \log(p) + 1. \tag{37}$$

This inequality holds true for many possible choices for p and q (recall that, in our derivation of (32),

we needed to restrict to those $p, q > 1$ that satisfy $\frac{1}{p} + \frac{1}{q} = 1$). For example, we can choose $p = q = 2$ or even $p = 2.363$, $q = 2.363/1.363$. Note that in the latter case

$$\theta_0 = \frac{1.363}{2.363} > \frac{1}{2},$$

which, to the best of our knowledge, is the first time that the provable regime is extended to the case where errors with norm larger than $\frac{1}{2}\lambda_1(\Lambda)$ are allowed [8, p. 11].

(2) For Approach 2: Now, by (34), we choose

$$\vartheta_0^2 = p2e^2\left(1 - \frac{1}{\zeta}\right)^{-1} \log(2(1+\zeta)) \quad \text{and} \quad \theta_0^2 = \frac{1}{q}\frac{1}{2}\left(1 - \frac{1}{\zeta}\right),$$

so that (36) becomes

$$\frac{p}{q}\log(2(1+\zeta)) < \frac{1}{2}\left(\log(p) + \log(2) + 2 - \log\left(1 - \frac{1}{\zeta}\right) + \log(\log(2(1+\zeta)))\right). \tag{38}$$

Recall that, also here, we restrict to $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. (38) is, for example, fulfilled for $p = q = 2, \zeta = 3$.

5 Dual attack – approach based on the conditional central limit theorem

In Section 4, we provided rigorous results for a quantification of the success of the dual attack. However, such rigorous results are always accompanied with many technicalities. Moreover, in order to establish the rigorous results, conservative assumptions and estimates (such as (16), (21), or (26)) are required, which, in many cases, can be relaxed in practice.

We believe that in order to have a full understanding of the subject, to find new results or to even find interdisciplinary connections, it is necessary to also intuitively and more directly understand the behavior of the investigated mathematical objects. This yields the need for a simpler and more intuitive approach to understand this attack, and to benefit from the fact there is no need to worry about mathematical provability of certain statements that can intuitively be justified but not mathematically. This is the goal of this section.

Hence, in this section, we aim to reprove the results from Section 4 via some intuitively justified heuristics and approximations. The strategy is based on a central limit theorem heuristic, which was also used, e.g., in [8]. We emphasize again that we believe that both approaches, the rigorous one from Section 4 and the intuitive one from this section, are essential for a full understanding of the subject.

We proceed as follows. We first provide some preparatory steps. Then, we estimate an important quantity that we need and then we compute the success probabilities both in the BDD-sample-case and in the random-sample-case. The mentioned heuristics are then backed in Section 6 by experiments.

5.1 Some preparation

5.1.1 Conditional central limit theorem

Our computations in this section depend on the following heuristic, which is justified later.

Heuristic 3. Suppose that we are given a sample t (which is either a BDD-sample or a random-sample). We assume that the scaled score function $\frac{1}{m_0}f_W(t)$ can be treated as a realization of the sum $F(t) + \phi(t)\tilde{X}$, where t and $\tilde{X}$ are independent random variables and

- $F(t) = \mathbb{E}[\cos(2\pi\langle t, w\rangle) \mid t]$ for some $w \in W$ (recall that the family $\{w\}_{w\in W}$ is identically distributed), and

- $\tilde{X}$ is a Gaussian random variable with expectation value 0 and variance 1.
- $\phi(t)^2 = \frac{1}{m_0}\mathrm{Var}[\cos(2\pi\langle t, w\rangle) \mid t]$.

5.1.1.1 Representation of $F(t)$ and the variance of $\tilde{X}$

Before we justify Heuristic 3, we note that we can simplify the expressions in this heuristic as follows. Using property (5), we have that

$$F(t) = \mathbb{E}[\cos(2\pi\langle t, w\rangle) \mid t] = \mathbb{E}_{w\leftarrow D_{\hat{\Lambda},\tau_0\sqrt{2\pi}}}[\cos(2\pi\langle t, w\rangle)]. \tag{39}$$

And for the variance of $\tilde{X}$, we apply the identity $\cos^2(x) = 1/2(1 + \cos(2x))$ (cf. [7, Proof of Lemma 4]) to obtain that

$$\begin{aligned}\frac{1}{m_0}\mathrm{Var}[\cos(2\pi\langle t, w\rangle) \mid t] &= \frac{1}{m_0}\mathrm{E}_w[\cos(2\pi\langle t, w\rangle)^2 \mid t] - \frac{1}{m_0}\mathrm{E}_w[\cos(2\pi\langle t, w\rangle) \mid t]^2 \\ &= \frac{1}{2m_0}(1 + F(2t) - 2F(t)^2).\end{aligned}$$

5.1.1.2 Justification of Heuristic 3

First note that by combining (9) and (39), we obtain that

$$\mathbb{E}\left[\frac{1}{m_0}f_W(t)|t\right] = \mathbb{E}_{w\leftarrow D_{\hat{\Lambda},\tau_0\sqrt{2\pi}}}[\cos(2\pi\langle t, w\rangle)] = F(t).$$

Therefore, for all $x \in \mathbb{R}$

$$\begin{aligned}\mathbb{P}\left[\frac{1}{m_0}f_W(t) \le x\right] &= \mathbb{P}\left[\frac{1}{m_0}f_W(t) - \mathbb{E}\left[\frac{1}{m_0}f_W(t)|t\right] + \mathbb{E}\left[\frac{1}{m_0}f_W(t)|t\right] \le x\right] \\ &= \mathbb{P}\left[\frac{1}{m_0}(f_W(t) - \mathbb{E}[f_W(t)|t]) + F(t) \le x\right] \\ &= \mathbb{E}\left[\mathbb{P}\left[\frac{1}{m_0}(f_W(t) - \mathbb{E}[f_W(t)|t]) + F(t) \le x \mid t\right]\right].\end{aligned}$$

Then, by using the conditional central limit theorem formulated in the paper by Yuan et al. [17, Theorem 3.1], we have that $\frac{1}{\sqrt{m_0}}(f_W(t) - \mathbb{E}[f_W(t)|t])$ converges, as $m_0 \to \infty$, in distribution to a normal distribution. We therefore have that, in distribution and for large enough m_0,

$$\mathbb{P}\left[\frac{1}{m_0}f_W(t) \le x\right] \approx \mathbb{E}[\mathbb{P}[\phi(t)\tilde{X} + F(t) \le x \mid t]], \tag{40}$$

where $\tilde{X}$ is a Gaussian random variable with expectation value 0, variance 1, and

$$\phi(t)^2 = \frac{1}{m_0}\mathrm{Var}[\cos(2\pi\langle t, w\rangle) \mid t].$$

Then, by revoking the conditional probability measure, we have that

$$\mathbb{E}[\mathbb{P}[\phi(t)\tilde{X} + F(t) \le x \mid t]] = \mathbb{P}[\phi(t)\tilde{X} + F(t) \le x]. \tag{41}$$

Combining (40) and (41) yields that $\frac{1}{m_0}f_W(t)$ can be treated as a realization of the sum $F(t) + \phi(t)\tilde{X}$.

Another, maybe more elementary, way to understand the step from (40) to (41) is to note that the left-hand side of (41) is nothing but a double integral as in the situation of Fubini's theorem, i.e.,

$$\begin{aligned}\mathbb{E}[\mathbb{P}[\phi(t)\tilde{X} + F(t) \le x \mid t]] &= \int\left(\int \mathbb{1}_{\{\phi(s)\tilde{x}+F(s)\le x\}}\rho_{\tilde{X}}(\tilde{x})d\tilde{x}\right)\rho_t(s)ds \\ &= \int \mathbb{1}_{\{\phi(s)\tilde{x}+F(s)\le x\}}\rho_{\tilde{X}}(\tilde{x})\ \rho_t(s)d(\tilde{x}, s) \\ &= \mathbb{P}[\phi(t)\tilde{X} + F(t) \le x],\end{aligned} \tag{42}$$

where $\rho_{\tilde{X}}$ and ρ_t denote the probability density functions of the random variables $\tilde{X}$ and t, respectively.

Moreover, from the representation in (42), it is visible that the right-hand side in (41) is equal to the same expression but by taking random variables $F(t)$ and $\tilde{X}$ that are, in addition, assumed to be independent. We can do this as $\tilde{X}$ does not depend on the random variable t. This justifies Heuristic 3.

Another justification for Heuristic 3 is that the experiments in Section 6 confirm the soundness of this assumption. Moreover, we note that Heuristic 3 is also used in the paper by Ducas and Pulles [8], where it was also justified by experimental evidence.

5.2 Estimating $F(t)$

The expectation on the right-hand side of (39) is given by

$$F(t) = \frac{\sum_{w\in\hat{\Lambda}} \cos(2\pi\langle t, w\rangle)\mathrm{e}^{-\|w\|^2/(2\tau_0^2)}}{\sum_{w\in\hat{\Lambda}}\mathrm{e}^{-\|w\|^2/(2\tau_0^2)}}. \tag{43}$$

The denominator and the numerator in the quotient (43) can be expressed as a sum over the lattice Λ by using the Poisson summation formula. Therefore, we further derive

$$\begin{aligned} F(t) &= \frac{\sum_{z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z-t\|^2}}{\sum_{z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z\|^2}} \\ &= \frac{\mathrm{e}^{-2\pi^2\tau_0^2\|t\|^2} + \sum_{0\neq z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z-t\|^2}}{1 + \sum_{0\neq z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z\|^2}}. \end{aligned} \tag{44}$$

In typical cases, we can expect that the value of $\sum_{0\neq z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z\|^2}$ is very small. The paper by Banaszczyk [15, Lemma (1.5, (i))] gives a concrete bound for this value.

Lemma 5.1. *Let $c = \frac{\sqrt{2\pi}}{\sqrt{n}}\tau_0\lambda_1(\Lambda)$ and assume that $c \geq \frac{1}{\sqrt{2\pi}}$. We set $C = c\sqrt{2\pi e}\cdot e^{-\pi c^2}$. Then, we have the bound*

$$\sum_{0\neq z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z\|^2} \leq C^n/(1-C^n).$$

Proof. The smallest non-trivial vector of the stretched lattice $\sqrt{2\pi\tau_0^2}\Lambda$ is of length $c\sqrt{n}$ and the lemma from the paper by Banaszczyk [15, Lemma (1.5, (i))] applies. □

In the following, we simplify our computations by assuming the following heuristic.

Heuristic 4. We have that

$$F(t) \approx \mathrm{e}^{-2\pi^2\tau_0^2\|t\|^2} + \sum_{0\neq z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z-t\|^2}. \tag{45}$$

That is, $\sum_{0\neq z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z\|^2}$ can be neglected in equation (44) for computing $F(t)$.

5.2.1 Justification of Heuristic 4

Recall the Gaussian heuristic Heuristic 1. And recall the definition of ϑ_0 in (28), which is determined through the equation

$$\tau_0 = \vartheta_0\frac{1}{\sqrt{2\pi e}}\det(\Lambda)^{-1/n}. \tag{46}$$

If $\vartheta_0 \geq e$, we can apply Lemma 5.1 with $c = \frac{\vartheta_0}{e\sqrt{2\pi}} \geq \frac{1}{\sqrt{2\pi}}$. As an example, for $\vartheta_0 \geq 4$, (resp. $\vartheta_0 \geq 5$), we have the concrete bound $C^n/(1 - C^n)$ with $C = 0.82$, (resp. $C = 0.56$). Note that in (44), we just need the approximation $1 + \sum_{0 \neq z \in \Lambda} e^{-2\pi^2\tau_0^2\|z\|^2} \approx 1$.

We continue the estimation of $F(t)$ by splitting into the two cases (the BDD-sample-case and the random-sample-case).

5.2.2 Estimating $F(t)$ in the BDD-sample-case

We now rely on the paper by Banaszczyk [15, Lemma (1.5, (ii))] for a simplification of the numerator of $F(t)$.

Lemma 5.2. *Let $c = \frac{\sqrt{2\pi}}{\sqrt{n}}\tau_0(\lambda_1(\Lambda) - \|t\|)$ and assume that $c \geq \frac{1}{\sqrt{2\pi}}$. We set $C = c\sqrt{2\pi e} \cdot e^{-\pi c^2}$. Then, we have the bound*

$$\sum_{0 \neq z \in \Lambda} e^{-2\pi^2\tau_0^2\|z-t\|^2} \leq 2C^n \sum_{z \in \Lambda} e^{-2\pi^2\tau_0^2\|z\|^2}.$$

Proof. For every $z \in \Lambda$, $z \neq 0$, we certainly have $\|z + t\| \geq \lambda_1(\Lambda) - \|t\|$. We can apply the paper by Banaszczyk [15, Lemma (1.5, (ii))] to the stretched lattice $\sqrt{2\pi\tau_0^2}\Lambda$. □

This lemma leads us to the following heuristic.

Heuristic 5. Suppose that the BDD-sample $t = v + e_0$ is such that $\|e_0\|$ is much smaller than $\lambda_1(\Lambda)$, the length of the shortest vector in Λ. Then, approximately, for large enough n,

$$F(t) \approx e^{-2\pi^2\tau_0^2\|e_0\|^2}.$$

5.2.1.1 Justification of Heuristic 5

From Lemma 4.5, we already know that $F(t)$ is greater or equal to $e^{-2\pi^2\tau_0^2\|e_0\|^2}$. It remains to show the other bound. Recall that $F(t) = F(v + e_0) = F(e_0)$ for any $v \in \Lambda$ (by using the same arguments as in the proof of Lemma 4.5). Then, by using (45),

$$F(t) = F(e_0) \approx e^{-2\pi^2\tau_0^2\|e_0\|^2} + \sum_{0 \neq z \in \Lambda} e^{-2\pi^2\tau_0^2\|z-e_0\|^2}. \tag{47}$$

We now show, by applying Lemma 5.2, that the second term on the right-hand side of (47) is negligible with respect to the first term. Let $c = \frac{\sqrt{2\pi}}{\sqrt{n}}\tau_0(\lambda_1(\Lambda) - \|e_0\|)$. Then, by using the parameters θ_0 and ϑ_0 from Section 4.5 and the same arguments as in Section 4.5 we obtain that

$$c = \frac{\vartheta_0}{e\sqrt{2\pi}} \frac{\lambda_1(\Lambda) - \|e_0\|}{\lambda_1(\Lambda)} \approx \frac{\vartheta_0}{e\sqrt{2\pi}} \frac{\lambda_1(\Lambda) - \theta_0\lambda_1(\Lambda)}{\lambda_1(\Lambda)} = \frac{\vartheta_0}{e\sqrt{2\pi}}(1 - \theta_0).$$

For the corresponding parameter C from Lemma 5.2, we obtain that

$$C = c\sqrt{2\pi e} \cdot e^{-\pi c^2} \approx \frac{\vartheta_0(1 - \theta_0)}{\sqrt{e}} e^{-\vartheta_0^2(1-\theta_0)^2/(2e^2)}.$$

By assuming that $c \geq \frac{1}{\sqrt{2\pi}}$, we conclude (via Lemma 5.2) that the second term on the right-hand side of (47) is bounded from above by

$$2\frac{\vartheta_0^n(1 - \theta_0)^n}{\sqrt{e}^n} e^{-n\vartheta_0^2(1-\theta_0)^2/(2e^2)} \sum_{z \in \Lambda} e^{-2\pi^2\tau_0^2\|z\|^2} \approx 2\frac{\vartheta_0^n(1 - \theta_0)^n}{\sqrt{e}^n} e^{-n\vartheta_0^2(1-\theta_0)^2/(2e^2)}, \tag{48}$$

where we used in the last step that, according to Heuristic 4, the term $\sum_{0 \neq z \in \Lambda} e^{-2\pi^2\tau_0^2\|z\|^2}$ can be neglected. In particular, we implicitly assumed here that $\vartheta \geq e$ in order to apply Heuristic 4.

We need to compare bound (48) to the first term on the right-hand side in (47). Using similar arguments as the ones that led to (48), we can show for the first term on the right-hand side in (47) that

$$\mathrm{e}^{-2\pi^2\tau_0^2\|e_0\|^2} \approx \mathrm{e}^{-n\vartheta_0^2\theta_0^2/(2e^2)}.$$

To sum up, we need the following conditions to hold true in order to justify Heuristic 5:

$$2\frac{\vartheta_0^n(1-\theta_0)^n}{\sqrt{e}^n}\mathrm{e}^{-n\vartheta_0^2(1-\theta_0)^2/(2e^2)} \ll \mathrm{e}^{-n\vartheta_0^2\theta_0^2/(2e^2)} \quad \text{and} \quad \frac{\vartheta_0}{e\sqrt{2\pi}}(1-\theta_0) \geq \frac{1}{\sqrt{2\pi}}. \tag{49}$$

Note that the second condition in (49) is equivalent to

$$\vartheta_0 \geq \frac{e}{(1-\theta_0)}, \tag{50}$$

which implicitly contains the condition that $\vartheta_0 \geq e$. As an example, for $\vartheta_0 = 4$ (resp. $\vartheta_0 = 5$) and $\theta_0 \leq 1/6$, (resp. $\theta_0 = 1/4$), condition (49) is fulfilled for n large enough. This justifies Heuristic 5.

We conclude that $Z(e_0) \coloneqq e^{-2\pi^2\tau_0^2\|e_0\|^2}$ is a valid approximation of $F(e_0)$. The distribution function of Z can be explicitly computed as

$$t \mapsto c_0(-\ln(t))^{n/2-1}\, t^{-1/(2\gamma_0)-1},$$

with a certain real number c_0 and $t \in [0,1]$ and where

$$\gamma_0 = -2\pi^2\sigma_0^2\tau_0^2. \tag{51}$$

If $|\gamma_0|$ is small, then the distribution function looks roughly like a Gaussian function. If $|\gamma_0| > \frac{1}{2}$, the exponent of t is negative and the distribution function looks completely different.

Furthermore, we can compute the expectation value and the variance of $Z(e_0)$. Similar computations as in Appendix A yield that its expectation is given by $(1-2\gamma_0)^{-n/2}$ and its variance by

$$(1-4\gamma_0)^{-n/2} - (1-2\gamma_0)^{-n}.$$

5.2.3 Estimating $F(t)$ in the random-sample-case

In the random-sample-case, it is difficult to find a closed formula for $F(t)$. Therefore, we analyze its properties as a random variable. In particular, we now estimate its expectation and its variance. Note that here we compute the expectations with respect to t.

Lemma 5.3. *Let t be a random-sample. Then,*

$$\mathbb{E}[F(t)] = \frac{1}{\sqrt{2\pi\tau_0^2}^n \det(\Lambda)\sum_{z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z\|^2}}$$

and

$$\mathbb{E}[F(t)^2] = \frac{\sum_{z\in\Lambda}\mathrm{e}^{-\pi^2\tau_0^2\|z\|^2}}{2^n\sqrt{\pi\tau_0^2}^n \det(\Lambda)[\sum_{z\in\Lambda}\mathrm{e}^{-2\pi^2\tau_0^2\|z\|^2}]^2}.$$

Proof. Since $\langle t, w\rangle \bmod 1$ is equally distributed on $[-\frac{1}{2}, \frac{1}{2}]$ for fixed $w \neq 0$ (cf. Section 3.2) we can compute (via elementary properties of the conditional expectation) the expectation value of $F(t)$ as

$$\begin{aligned}\mathbb{E}[F(t)] &= \mathbb{E}[\mathbb{E}[\cos(2\pi\langle t, w\rangle) \mid t]] = \mathbb{E}[\cos(2\pi\langle t, w\rangle)] = \mathbb{E}[\mathbb{E}[\cos(2\pi\langle t, w\rangle) \mid w]] \\ &= \mathbb{E}[\mathbb{E}[\cos(2\pi\langle t, w\rangle) \mid w]\mathbb{1}_{w\neq 0}] + \mathbb{E}[\mathbb{E}[\cos(2\pi\langle t, w\rangle) \mid w]\mathbb{1}_{w=0}] \\ &= \mathbb{E}\left[\mathbb{E}_{u\xleftarrow{\$}[-\frac{1}{2},\frac{1}{2}]}[\cos(2\pi u)]\mathbb{1}_{w\neq 0}\right] + \mathbb{P}[w=0]\end{aligned}$$

$$= \mathbb{E}_{u \xleftarrow{\$} [-\frac{1}{2}, \frac{1}{2}]}[\cos(2\pi u)]\mathbb{P}[w \neq 0] + \frac{1}{\sum_{w \in \hat{\Lambda}} e^{-\|w\|^2/(2\tau_0^2)}}$$
$$= \frac{1}{\sum_{w \in \hat{\Lambda}} e^{-\|w\|^2/(2\tau_0^2)}}$$
$$= \frac{1}{\sqrt{2\pi\tau_0^2}^n \det(\Lambda) \sum_{z \in \Lambda} e^{-2\pi^2\tau_0^2\|z\|^2}},$$

where we used the Poisson summation formula in the last step.

Note that

$$\sum_{w,w' \in \hat{\Lambda}} \cos(2\pi\langle t, w\rangle)\cos(2\pi\langle t, w'\rangle) e^{-\|w\|^2/(2\tau_0^2)} e^{-\|w'\|^2/(2\tau_0^2)}$$
$$= \frac{1}{2} \sum_{w,w' \in \hat{\Lambda}} [\cos(2\pi\langle t, w + w'\rangle) + \cos(2\pi\langle t, w - w'\rangle)] e^{-\|w\|^2/(2\tau_0^2)} e^{-\|w'\|^2/(2\tau_0^2)}.$$

This property allows us to compute the second moment of $F(t)$ as

$$\frac{\sum_{w \in \hat{\Lambda}} e^{-2\|w\|^2/(2\tau_0^2)}}{\left[\sum_{w \in \hat{\Lambda}} e^{-\|w\|^2/(2\tau_0^2)}\right]^2}.$$

Again, we use the Poisson summation formula for the denominator and numerator of this quotient, which gives the claim.

$$\mathbb{E}[F(t)^2] = \frac{\sqrt{\pi\tau_0^2}^n \det(\Lambda) \sum_{z \in \Lambda} e^{-\pi^2\tau_0^2\|z\|^2}}{\left[\sqrt{2\pi\tau_0^2}^n \det(\Lambda) \sum_{z \in \Lambda} e^{-2\pi^2\tau_0^2\|z\|^2}\right]^2} = \frac{\sum_{z \in \Lambda} e^{-\pi^2\tau_0^2\|z\|^2}}{2^n \sqrt{\pi\tau_0^2}^n \det(\Lambda) \left[\sum_{z \in \Lambda} e^{-2\pi^2\tau_0^2\|z\|^2}\right]^2}.$$ □

This lemma leads us to the following heuristic.

Heuristic 6. Let t be a random-sample. In typical cases, we can expect that the distribution of $F(t)$ is extremely close to 0 since both the expectation value and the standard deviation are negligible. Recall the definition of ϑ_0 in (46). Then, approximately,

$$\mathbb{E}[F(t)] \approx \left[\frac{\vartheta_0}{\sqrt{e}}\right]^{-n} \quad \text{and} \quad \mathbb{E}[F(t)^2] \approx \left[\frac{\vartheta_0\sqrt{2}}{\sqrt{e}}\right]^{-n}.$$

5.2.2.1 Justification of Heuristic 6

We assume as in Heuristic 4 that we can neglect terms $z \neq 0$ in the series in Lemma 5.3, which gives

$$\mathbb{E}[F(t)] \approx \frac{1}{\sqrt{2\pi\tau_0^2}^n \det(\Lambda)}$$

and

$$\mathbb{E}[F(t)^2] = \frac{1}{2^n \sqrt{\pi\tau_0^2}^n \det(\Lambda)}.$$

Again as above, we want to assume that the Gaussian Heuristic 1 is valid for both lattices Λ, $\hat{\Lambda}$ and that $\tau_0\sqrt{n}$ is not too small a multiple ϑ_0 of the shortest vector in $\hat{\Lambda}$ as in (46). Then, the expectation value of $F(t)$ is of size

$$\frac{1}{\sqrt{2\pi\tau_0^2}^n \det(\Lambda)} = \left[\frac{\vartheta_0}{\sqrt{e}}\right]^{-n},$$

and the second moment is of size

$$\frac{1}{\sqrt{4\pi\tau_0^2}^n \det(\Lambda)} = \left[\frac{\vartheta_0\sqrt{2}}{\sqrt{e}}\right]^{-n}.$$

5.3 Parameter selection for assuming the Heuristics

We want to assume that the previous heuristics are valid, i.e., Heuristic 3, Heuristic 4, and Heuristic 5. Recall the definition of ϑ_0 in (46). Moreover, following the reasoning of Section 4.5.5, we will also assume an upper bound for m_0. That is, we assume that the size m_0 of W is much smaller compared to the number of all lattice vectors of length $\leq\tau_0\sqrt{n}$. Here with respect to the conditions derived in Section 4.5.5, we even make the stronger assumption that

$$2m_0 \leq (\vartheta_0/2)^n. \tag{52}$$

This stronger condition implies in particular that in the random sample case (cf. Heuristic 4),

$$\left[\frac{\vartheta_0\sqrt{2}}{\sqrt{e}}\right]^{-n} \ll \left[\frac{\vartheta_0}{2}\right]^{-n} \leq \frac{1}{2m_0}.$$

Therefore, in the random-sample case, we can approximate $\frac{1}{m_0}f_W(t)$ by

$$F(t) + \phi(t)\tilde{X} \approx \phi(t)\tilde{X} \approx \frac{1}{\sqrt{2m_0}}\tilde{X}.$$

In the BDD-sample case, we expect (cf. Heuristic 5)

$$F(t) \approx \mathrm{e}^{-2\pi^2\tau_0^2\|e_0\|^2} \approx e^{-n\theta_0^2\vartheta_0^2/(2e^2)}.$$

As an example, for $\theta_0 = 1/6$ and $\vartheta_0 = 4$, (resp. $\theta_0 = 1/4$ and $\vartheta_0 = 5$), we have $e^{-\theta_0^2\vartheta_0^2/(2e^2)} \approx 0.97$, (resp. $e^{-\theta_0^2\vartheta_0^2/(2e^2)} \approx 0.90$). For $n \geq 100$, we therefore expect that $F(t)$ is very small compared to 1, so that in Heuristic 3 $\phi(t)^2$ can be approximated by $\frac{1}{2m_0}$.

5.4 Success probabilities

We want to distinguish the cases "random-samples vs BDD-samples" with good probability by checking if the score is higher or lower than a certain value α. If we have good approximations for the distribution of $F(e_0)$, we can compute numerically a condition on m_0 based on the heuristics above. In the following, we want to derive a simple formula that gives us a plausible condition on m_0. To this end, we assume that the approximations in the previous sections are valid. As argued in Section 5.3, we further approximate $\phi(t)^2$ by $\frac{1}{2m_0}$ in Heuristic 3. We set $X = \frac{1}{\sqrt{2m_0}}\tilde{X}$.

Following the arguments of Section 5.3, if t is chosen uniformly in $\mathbb{R}^n/\Lambda$, we assume that $F(t)$ is extremely small compared to X. Therefore, we approximate

$$\begin{aligned}&\mathbb{P}\left(\frac{1}{m_0}\sum_{w\in W}\cos(2\pi\langle t, w\rangle) \leq \alpha \,|\, \text{case "random-samples"}\right)\\ &\quad = \mathbb{P}(F(t) + X \leq \alpha \,|\, \text{case "random-samples"}) \approx \mathbb{P}(X \leq \alpha).\end{aligned} \tag{53}$$

We therefore choose for simplicity

$$\alpha = \frac{\mu_0}{\sqrt{2m_0}},$$

with a certain number $\mu_0 \geq 1$, say $\mu_0 \in \{2, 3, 4\}$.

On the other hand, in the BDD-sample-case, we want that

$$\mathbb{P}\left(\frac{1}{m_0}\sum_{w\in W}\cos(2\pi\langle e_0, w\rangle) \geq \alpha \,|\, \text{case “BDD-samples”}\right)$$

is notably larger than 0.5 (note that we tacitly used that $\langle t, w\rangle = \langle v + e_0, w\rangle = \langle e_0, w\rangle$, since $v \in \Lambda$ and w is a dual vector). We consider

$$\begin{aligned}&\mathbb{P}\left(\frac{1}{m_0}\sum_{w\in W}\cos(2\pi\langle e_0, w\rangle) \geq \alpha \,|\, \text{case “BDD-samples”}\right)\\ &= \mathbb{P}(F(e_0) + X \geq \alpha \,|\, \text{case “BDD-samples”})\\ &\approx \mathbb{P}(Z(e_0) + X \geq \alpha).\end{aligned}$$

We know the distribution function of Z and X, so that the probability $\mathbb{P}(Z + X \geq \alpha)$ can be computed numerically by a two-dimensional integral, since Heuristic 3 gives us independence of $F(t)$ and X. However, here we want to derive a rough estimate on m_0 that gives us a simple formula. Since the standard deviation of X is equal to $\frac{1}{\sqrt{2m_0}} = \frac{\alpha}{\mu_0}$, we can approximate $\mathbb{P}(Z + X \geq \alpha)$ by $\mathbb{P}(Z \geq \alpha)$ for moderate μ_0. We further compute

$$\mathbb{P}(Z + X \geq \alpha) \approx \mathbb{P}(Z \geq \alpha) = \mathbb{P}(e^{\gamma_0\|e_0\|^2/\sigma_0^2} \geq \alpha) = \mathbb{P}\left(\|e_0\|^2/\sigma_0^2 \leq \frac{\ln(\alpha)}{\gamma_0}\right),$$

$\|e_0\|^2/\sigma_0^2$ is χ^2-distributed. If we choose

$$n + \sqrt{2n} \leq \frac{\ln(\alpha)}{\gamma_0} \iff \alpha \leq e^{\gamma_0(n+\sqrt{2n})}$$

we obtain a “good” probability for $\mathbb{P}(Z \geq \alpha)$. (For $n \geq 50$, this probability of the χ^2-distribution is very near to 0.84.) In the end, we derive the condition

$$e^{\gamma_0(n+\sqrt{2n})} \geq \frac{\mu_0}{\sqrt{2m_0}} \iff 2m_0 \geq \mu_0^2 e^{-2\gamma_0(n+\sqrt{2n})} = \mu_0^2 e^{4\pi^2\sigma_0^2\tau_0^2(n+\sqrt{2n})}. \tag{54}$$

Remark 5.4.

- Note the resemblance of condition (54) with condition (26) identified in the rigorous analysis. More precisely, the leading-order term on the right-hand side of (54) is given by

  $$e^{-2\gamma_0(n+\sqrt{2n})} = e^{4\pi^2\sigma_0^2\tau_0^2(n+\sqrt{2n})},$$

  while the leading-order term on the right-hand side of (26) is given by

  $$\mathrm{e}^{4\pi^2(n+2n^{1/2+\xi}+2n^{\xi})\sigma_0^2\tau_0^2},$$

  which only differs in the lower-order terms in the exponent. These differences are due to the fact that, in our rigorous approach, we needed to apply quite coarse concentration inequalities.
- The conditions (26) or (54) are also very similar to the usual condition on m_0 computed under the independence heuristic. The main difference, ignoring $\mu_0 \geq 1$, is a slightly higher number for the number of vectors m_0 due to the new term $n + \sqrt{2n}$ instead of n, which results in an additional factor of the form $\mathrm{e}^{-2\gamma_0\sqrt{2n}}$.
- Recall the condition formulated in (52), so that we have to consider

  $$(\vartheta_0/2)^n \geq 2m_0 \geq \mu_0^2 e^{4\pi^2\sigma_0^2\tau_0^2(n+\sqrt{2n})} \approx \mu_0^2 e^{2(n+\sqrt{2n})\theta_0^2\vartheta_0^2/(2e^2)}.$$

For typical parameter sets as $\theta_0 = 1/6, \vartheta_0 = 4$, (resp. $\theta_0 = 1/4, \vartheta_0 = 5$), this condition can be fulfilled. For instance, in the case $\vartheta_0 = 4$, the upper bound on m_0 is given by 2^{n-1}, which for typical values for n becomes very large and hence, is not restrictive in practice.

6 Experiments

In order to check the reasonability of our heuristics we conducted a few experiments. In our experiments we chose the lattice Λ as generated by the columns of the $n \times n$-matrix B with

$$B = \begin{pmatrix} \mathbb{1}_{n/2} & 0 \\ C & p\mathbb{1}_{n/2} \end{pmatrix},$$

where p is the prime number $2^{16} + 1$, C is a $n/2 \times n/2$-matrix with integer entries chosen uniformly between 0 and $p - 1$. The dual lattice $\hat{\Lambda}$ is generated by the columns of the matrix

$$\frac{1}{p}\begin{pmatrix} p\mathbb{1}_{n/2} & -C^T \\ 0 & \mathbb{1}_{n/2} \end{pmatrix}.$$

In our experiments, we did not consider concrete lattice reduction algorithms for generating the subset $W \subset \hat{\Lambda}$. Instead, we used a setting such that the distribution in W should be near to a Gaussian distribution.

We emphasize that we do not aim at an efficient implementation of our attacks here. Instead, we rather aim for a simple implementation, which resembles the theoretic setting of our paper, with the goal to practically verify the soundness of our heuristics.

6.1 Experiments for small n

For small n we chose the distribution in W by the following random process over m_0 repetitions:

(1) Choose z_1 as realization of a Gaussian variable with expectation value 0 and covariance matrix $\tau_1\mathbb{1}_n$.

(2) Compute w as the closest vector in $\hat{\Lambda}$ to z_1.

Note that we had to restrict ourselves to small n for being able to compute step 2 in reasonable time. Using the above notations, we further chose

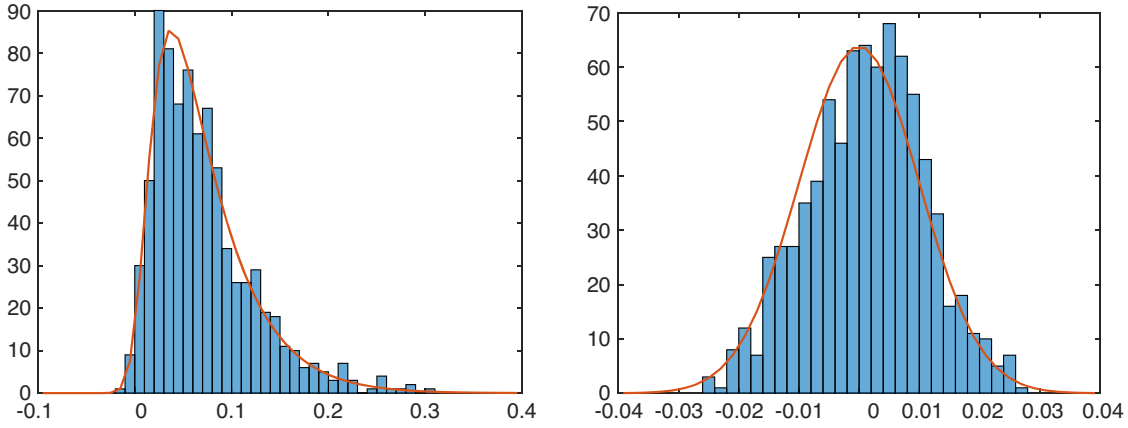
$$n = 30, \quad m_0 = 5{,}000, \quad \sigma_0 = 18, \quad \vartheta_0 = 4$$
$$\tau_1 = \frac{\vartheta_0}{\sqrt{2\pi e}} \det(\Lambda)^{-1/n} = \frac{\vartheta_0}{\sqrt{2\pi e}\sqrt{p}} = 0.0038.$$

We can expect that the average of the length w_j generated by this process is slightly larger than $\tau_1\sqrt{n}$ and we obtained in our experiments

$$\frac{1}{m_0}\sum_{w\in W} \|w\|^2 = \tau_0^2 n, \quad \text{where } \tau_0 = 0.0039.$$

In Figure 1 (a), we computed the score function $\frac{1}{m_0}\sum_{w\in W}\cos(2\pi\langle e_0, w\rangle)$ for 800 repetitions in e_0, where e_0 is the realization of a Gaussian random variable with expectation value 0 and covariance matrix $\sigma_0 1_n$. For the sake of comparison, we also include in Figure 1 (a) a red colored graph, which represents the distribution of the sum $F(e_0) + X$ with two independent random variables e_0 and X as in Heuristic 3 (and where X was introduced in Section 5.4).

In Figure 1 (b), we compute the score function $\frac{1}{m_0}\sum_{w\in W}\cos(2\pi\langle t, w\rangle)$ for 800 repetitions in t, where t is chosen as a random-sample. As expected, this distribution is close to a Gaussian distribution with standard deviation 0.01, which is represented by the red colored graph in Figure 1 (b).



(a) Score function, e_0 variable (b) Score function, random-sample

Figure 1: Experiments for small n. (a) Score function, e_0 variable and (b) score function, random-sample.

6.2 Experiments for moderate n

For moderate n, we generated a set $\tilde{W}$ of many vectors in $\hat{\Lambda}$ that have length near ϑ_0-times the shortest vector expected by the Gaussian heuristic. Then, we chose the distribution in $\tilde{W}$ by the following ad hoc random process over m_0 repetitions:

(1) Choose z_1 as realization of a Gaussian variable with expectation value 0 and covariance matrix $\tau_0 \mathbb{1}_n$.

(2) Compute w as the closest vector from $\tilde{W}$ to z_1.

Let W be the collection of all these outputted w.

We were able to conduct experiments in this way for $n = 70$. Using the above notations, we further chose

$$n = 70, \quad \#\tilde{W} = 100 \cdot m_0, \quad m_0 = 500, \quad \sigma_0 = 10, \quad \vartheta_0 = 4,$$
$$\tau_0 = \frac{\vartheta_0}{\sqrt{2\pi e}} \det(\Lambda)^{-1/n} = \frac{\vartheta_0}{\sqrt{2\pi e}\sqrt{p}} = 0.0038.$$

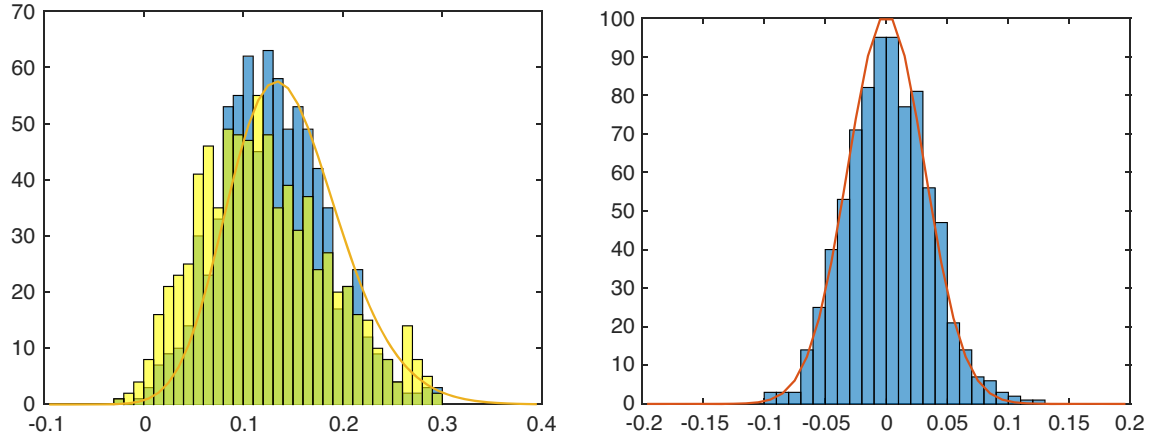
In our experiments, we observed that the average of the length w_j generated by this process was very close to $\tau_0\sqrt{n}$.

We computed the score function $\frac{1}{m_0}\sum_{w\in W}\cos(2\pi\langle e_0, w\rangle)$ for 800 repetitions in e_0, where e_0 is the realization of a Gaussian random variable with expectation value 0 and covariance matrix $\sigma_0 1_n$. Furthermore, we compared this with the score function $\frac{1}{m_0}\sum_{w\in \hat{W}}\cos(2\pi\langle e_0, w\rangle)$ for 800 repetitions in e_0, where

- e_0 is again the realization of a Gaussian random variable with expectation value 0 and covariance matrix $\sigma_0 1_n$, and
- we chose $w \in \hat{W}$ not in the dual lattice, but as a realization of a continuous Gaussian random variable with expectation value 0 and covariance matrix $\tau_0 1_n$.

The reason why we also computed this additional comparison is due to the observation that in our theoretic results, we never used the geometry of the lattice. Consequently, intuition tells us that the results should be similar, and we would like to see, whether this intuition is sound.

Figure 2 (a) shows both score functions in one figure. The distributions are similar, but we note a slight difference in the probabilities for larger $F(e_0)$. As above, we also include in Figure 2 (a), a red colored graph, which represents the distribution of the sum $F(e_0) + X$. As expected, the distributions in Figure 2 (a) are close to this theoretical distribution.



(a) Score function, e_0 variable, yellow corresponds to W, blue to $\hat{W}$

(b) Score function, random-sample

Figure 2: Experiments for moderate n. (a) Score function, e_0 variable, yellow corresponds to W and blue to $\hat{W}$ and (b) score function, random-sample.

In Figure 2(b), we compute the score function $\frac{1}{m_0}\sum_{w\in W}\cos(2\pi\langle t, w\rangle)$ for 800 repetitions in t, where t is chosen as a random-sample. As expected, this distribution is close to a Gaussian distribution with standard deviation $\frac{1}{2m_0} \approx 0.03$, which is represented by the red colored graph in this figure.

7 Discussion

7.1 Comparison with other recent works

As pointed out in Section 1 of this paper, there is a lot of recent progress and research on the dual attack for BDD or LWE. Most notably with respect to our results are the two interesting and impressive works by Pouly and Shen [9] and Ducas and Pulles [8]. We now briefly compare their results with our results in this work.

7.1.1 Results from the paper by Pouly and Shen [9]

The first obvious difference between our work and the work of [9] is that they studied LWE, while we studied BDD. However, it is possible to translate the results into the other respective framework by using the arguments from the paper by Pouly and Shen [9, Section 6]. The obvious similarity to our work is that they also assumed that the dual vectors are distributed according to the discrete Gaussian distribution.

Our approach in Section 4, which is based on a direct application of Hoeffding's inequality, is inspired by the paper by Pouly and Shen [9]. However, we use their technique for the *conditional probability measure*, while they use it for the *original probability measure*. This is why they apply this strategy only in their "basic dual attack" where dual vectors are sampled afresh for many different samples (in order to obtain independence, which is required to apply Hoeffding's inequality), while we use it for the case when there is only one sample and one family of dual vectors (i.e., the usual dual attack).

The main result in the paper by Pouly and Shen [9] is their "modern dual attack". First note that via this attack, the authors solve the search version of LWE (where the secret is to be extracted directly) and not the

decision version (as in our case). The main difference to our approach is that they do not choose the secret according to a single evaluation of the score function and checking whether it is above or below a threshold. What they do is that they evaluate many score functions and choose the candidate with the maximum score as the secret. We believe that this approach is more precise than ours, since in the approach by Pouly and Shen [9], they do not need to choose a threshold or rely on concentration inequalities. However, we did not rigorously translated our results into their setting in order to make a precise comparison. They also relied on geometric properties of the lattice in order to estimate the success probabilities of their attack, which is more involved than the strategy based on Hoeffding's inequality.

7.1.2 Results from the paper by Ducas and Pulles [8]

In the other independent work by Ducas and Pulles [8], similar results as in the present paper were obtained. The work by Ducas and Pulles [8] is a follow-up of their previous paper [7] and builds upon the observations made there. Also, our work can be seen as being initiated by the observations from the paper by Ducas and Pulles [7], our setting and notation is very similar to that in the paper by Ducas and Pulles [8]. However, a key difference between their work and the results of this paper is the assumption on the output of the dual lattice sieve algorithm. While we assume that the obtained dual vectors are distributed according to a discrete Gaussian distribution on the dual lattice, Ducas and Pulles [8] assume that the dual vectors are uniformly distributed on a ball of a certain radius.

The strategy in the paper by Ducas and Pulles [8] is, as in Section 5, based on a central limit theorem-type argument. In this way, they derive estimates on the cumulative distribution function of the score function. Besides, in their justification of their central limit heuristic ([8, Heuristic 3]), they also implicitly apply conditional independence in their computations in the paper by Ducas and Pulles [8, Section 3.3] by changing from a spherical error distribution to a radial error distribution.

Their theoretic results are accompanied by impressive experimental results, which underline and substantiate the former.

7.2 Future work

There are a lot of directions for improvements and future research with respect to our results in this paper. In the following, we would like to emphasize some of them.

- What happens in the asymptotic case, when m_0 is extremely large? In (53), we assume that $F(t)$ is extremely small compared to X, if e_0 is chosen uniformly in $\mathbb{R}^n/\Lambda$. This is certainly not true, when m_0 is extremely large. Instead, we consider

$$\begin{aligned}&\mathbb{P}\left(\frac{1}{m_0}\sum_{w\in W}\cos(2\pi\langle t,w\rangle)\le\alpha\,|\,\text{case “random-samples”}\right)\\&\quad=\mathbb{P}(F(t)+X\le\alpha\,|\,\text{case “random-samples”})\approx\mathbb{P}(F(t)\le\alpha\,|\,\text{case “random-samples”}).\end{aligned}$$

  If we assume that the approximations in Section 5 are valid, we have different distributions for $F(t)$ in both cases. This allows us to distinguish the cases “random-samples vs BDD-samples” with certain fixed probabilities that do not depend on m_0.
- A natural question arises: What are good weights in the formula for the total score taking into account the approach above? We therefore consider weights β_w in

$$f_\beta(t)=\sum_{w\in W}\beta_w\cos(2\pi\langle t,w\rangle).$$

We can adapt the formulas from Section 5 if we restrict ourselves by choosing

$$\beta_w = \mathrm{e}^{-\gamma_0 \|w\|^2}.$$

In this way, one can find an optimal γ_0 that gives a certain lower condition on m_0 compared to (5). Also, the results from Section 4 should be easily transferable to this new score function. This is left for future research.

- It is very interesting to see whether Heuristic 2 can be modified or generalized in order to be more suitable for modern lattice sieve algorithms. For instance, as we pointed out after Heuristic 2, we believe that the most realistic assumption, which resembles the behavior of the output of the lattice reduction algorithms most suitably, would be the case, where the dual vectors are uniformly distributed on $(B_r^n \cap \hat{\Lambda})/B_{r'}^n$ for some $r > r' > 0$.
- As outlined in Section 4.5, our theoretic results demand a lower bound on ϑ_0. It would be interesting to also find quantitative results on the success probabilities and the cost estimates for our attacks in the cases, where ϑ_0 is small.
- The main purpose of our experiments from Section 6 is the verification of the soundness of Heuristic 3, which in turn (via the computations from Section 5) verifies the soundness of the results from Section 4. In that sense, the main focus of this work is the theoretical part. Hence, there are many possible ways to further conduct the experimental aspects related to this paper. In particular, it would be interesting
  - to see how good Heuristic 2 resembles the behavior of concrete lattice reduction algorithms, or
  - to check whether the choice, where the dual vectors are uniformly distributed on $(B_r^n \cap \hat{\Lambda})/B_{r'}^n$, is indeed the most suitable case.

Acknowledgment: We gratefully acknowledge helpful discussions with Léo Ducas, Stephan Ehlen, and Ludo N. Pulles. We moreover would like to express our gratitude toward the anonymous referees for numerous helpful suggestions and remarks.

Funding information: Authors state no funding involved.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

References

[1] Arora S, Ge R. New algorithms for learning in presence of errors. In: International colloquium on automata, languages, and programming. Berlin, Heidelberg: Springer; 2011. p. 403–15.

[2] Kirchner P, Fouque PA. An improved BKW algorithm for LWE with applications to cryptography and lattices. In: Advances in Cryptology-CRYPTO 2015: 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part I 35. Berlin, Heidelberg: Springer; 2015. p. 43–62.

[3] Gama N, Nguyen PQ. Predicting lattice reduction. In: Advances in Cryptology-EUROCRYPT 2008: 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13–17, 2008. Proceedings 27. Berlin, Heidelberg: Springer; 2008. p. 31–51.

[4] Aharonov D, Regev O. Lattice problems in NP ∩ coNP. JACM. 2005;52(5):749–65.

[5] Laarhoven T, Walter M. Dual lattice attacks for closest vector problems (with preprocessing). In: Cryptographers Track at the RSA Conference. Cham: Springer; 2021. p. 478–502.

[6] Guo Q, Johansson T. Faster dual lattice attacks for solving LWE with applications to CRYSTALS. In: Advances in Cryptology-ASIACRYPT 2021: 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6–10, 2021, Proceedings, Part IV 27. Cham: Springer; 2021. p. 33–62.

[7] Ducas L, Pulles LN. Does the dual-sieve attack on learning with errors even work? In: Annual International Cryptology Conference. Cham: Springer; 2023. p. 37–69.
[8] Ducas L, Pulles LN. Accurate score prediction for dual-sieve attacks. Cryptology ePrint Archive, Paper 2023/1850. https://eprint.iacr.org/2023/1850.
[9] Pouly A, Shen Y. Provable dual attacks on learning with errors. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer; 2024. p. 256–85.
[10] Regev O. New lattice-based cryptographic constructions. JACM. 2004;51(6):899–942.
[11] Peikert C. A decade of lattice cryptography. Foundations Trends® Theoret Comp Sci. 2016;10(4):283–424.
[12] Billingsley P. Probability and Measure. New York: John Wiley & Sons Inc.; 2012.
[13] Laurent B, Massart P. Adaptive estimation of a quadratic functional by model selection. Ann Statist. 2000;28(5):1302–38.
[14] Stephens-Davidowitz N. On the Gaussian measure over lattices. USA: New York University; 2017.
[15] Banaszczyk W. New bounds in some transference theorems in the geometry of numbers. Mathematische Annalen. 1993;296:625–35.
[16] Chen Y, Hu Z, Liu Q, Luo H, Tu Y. LWE with quantum amplitudes: algorithm, hardness, and oblivious sampling; 2023. Cryptology ePrint Archive, Paper 2023/1498. https://eprint.iacr.org/2023/1498.
[17] Yuan DM, Wei LR, Lei L. Conditional central limit theorems for a sequence of conditional independent random variables. J Korean Math Soc. 2014;51(1):1–15.
[18] Batir N. Inequalities for the gamma function. Archiv der Math. 2008;91(6):554–63.

Appendix

A Some approximations of the covariances in the BDD-sample-case

In this section, we provide some heuristic estimates, which underline that the covariances computed in Section 3.3 are indeed nonzero.

Our goal is to intuitively estimate the sum over all single covariances

$$\sum_{w,\tilde{w}\in W, w\neq\tilde{w}} \mathrm{Cov}(f_w(t), f_{\tilde{w}}(t)).$$

Instead of computing this sum via (4) directly, we aim to find plausible approximations that give simple formulas. Recall that $m_0 = |W|$, and note that

$$\frac{1}{m_0^2} \sum_{w,\tilde{w}\in W, w\neq\tilde{w}} \mathrm{Cov}(f_w(t), f_{\tilde{w}}(t)), \tag{A1}$$

can be interpreted as a computation of a mean value (by using that $m_0^2 - m_0 \approx m_0^2$ is the approximate number of summands in the sum). In order to find good approximations on this mean value, we suppose (similarly as in the Sections 4 and 5) that $w, \tilde{w} \in W, w \neq \tilde{w}$ are random variables. In the simplest approximation, these random variables are Gaussian distributed with covariance matrix $\tau_0^2 1_n$. Note that here we make a more simplified assumption than in Heuristic 2, since our only goal here is to intuitively comprehend why the covariances do not vanish.

By using again the techniques of conditional independence and other simple elementary techniques from probability theory, we can assume that (A1) is a sum of uncorrelated, square-integrable random variables so that we are in the setting of (a generalized form of) the *law of large numbers*. We omit the details here as they are similar to the detailed computations in Sections 4 and 5.

Consequently, we can expect that the expression (A1) is close to the expectation value of the $\mathrm{Cov}(f_w(t), f_{\tilde{w}}(t))$. Recall from Proposition 3.3 that

$$\mathrm{Cov}(f_w(t), f_{\tilde{w}}(t)) = \frac{1}{2}\Delta_a + \frac{1}{2}\Delta_b - \Delta_c \cdot \Delta_d,$$

where

$$\Delta_a = \mathrm{e}^{-2\pi^2\|w+\tilde{w}\|^2\sigma_0^2}, \quad \Delta_b = \mathrm{e}^{-2\pi^2\|w-\tilde{w}\|^2\sigma_0^2}$$
$$\Delta_c = \mathrm{e}^{-2\pi^2\|w\|^2\sigma_0^2}, \quad \Delta_d = \mathrm{e}^{-2\pi^2\|\tilde{w}\|^2\sigma_0^2}.$$

Hence, the expectation value of the $\mathrm{Cov}(f_w(t), f_{\tilde{w}}(t))$ (by considering w and $\tilde{w}$ as the random variables) is a sum of terms of the form

$$\mathbb{E}(\mathrm{e}^{\gamma Y}),$$

where Y is (standard)-χ-square distributed. For $\gamma < 0.5$, this is identical to

$$\mathbb{E}(\mathrm{e}^{\gamma Y}) = (1 - 2\gamma)^{-k/2},$$

where k denotes the degrees of freedom of Y. Δ_a (resp. Δ_b) depends on

$$\|w + \tilde{w}\|^2, \quad \text{resp. } \|w - \tilde{w}\|^2,$$

which has n degrees of freedom, whereas $\Delta_c \cdot \Delta_d$ depends on

$$\|w\|^2 + \|\tilde{w}\|^2,$$

which has $2n$ degrees of freedom. In the end, we derive as an approximation of (A1)

$$(1 - 4\gamma_0)^{-n/2} - (1 - 2\gamma_0)^{-n},$$

where γ_0 was defined in (51). For the total variance, we therefore expect as approximation

$$V(f_W(t)) = \sum_{w \in W} V(f_w(t)) + \sum_{w,\tilde{w} \in W, w \neq \tilde{w}} \mathrm{Cov}(f_w(t), f_{\tilde{w}}(t))$$
$$\approx \frac{m_0}{2} + m_0^2[(1 - 4\gamma_0)^{-n/2} - (1 - 2\gamma_0)^{-n}].$$

This yields heuristically the bias of the variance, which was observed experimentally in [7].

B Bound on the volume of a ball

Lemma B.1. *For all $R > 0$, we have that*

$$\mathrm{Vol}(B_R^n) \leq \left(\sqrt{\frac{2\pi \mathrm{e}}{n}} R\right)^n.$$

Proof. It is known that

$$\mathrm{Vol}(B_R^n) = \frac{(\sqrt{\pi} R)^n}{\Gamma(1 + n2)}, \tag{A2}$$

where Γ is the gamma function. Recall the estimate on the gamma function given by

$$\Gamma\left(1 + \frac{n}{2}\right) > \left(\frac{n}{2\mathrm{e}}\right)^{\frac{n}{2}} \sqrt{\frac{n}{2} + \frac{1}{2}}\, \mathrm{e}^{-1/\left(6\left(\frac{n}{2} + \frac{3}{8}\right)\right)} \sqrt{2}\, \mathrm{e}^{\frac{4}{9}}, \tag{A3}$$

which is proven in e.g. [18, Corollary 1.2]. Note that for all n,

$$\sqrt{\frac{n}{2} + \frac{1}{2}}\, \mathrm{e}^{-1/\left(6\left(\frac{n}{2} + \frac{3}{8}\right)\right)} \sqrt{2}\, \mathrm{e}^{\frac{4}{9}} \geq 1,$$

which, together with (A2) and (A3), concludes the proof. □