Smoothing Codes and Lattices: Systematic Study and New Bounds

Thomas Debris-Alazard, Léo Ducas, Nicolas Resch, and Jean-Pierre Tillich

Abstract—In this article we revisit smoothing bounds in parallel between lattices and codes. Initially introduced by Micciancio and Regev, these bounds were instantiated with Gaussian distributions and were crucial for arguing the security of many lattice-based cryptosystems. Unencumbered by direct application concerns, we provide a systematic study of how these bounds are obtained for both lattices and codes, transferring techniques between both areas. We also consider multiple choices of spherically symmetric noise distributions. We found that the best strategy for a worst-case bound combines Parseval’s Identity, the Cauchy-Schwarz inequality, and the second linear programming bound, and this holds for both codes and lattices and all noise distributions at hand. For an average-case analysis, the linear programming bound can be replaced by an expected value computation. This alone gives optimal results for spherically uniform noise over random codes and random lattices. This also improves prior Gaussian smoothing bounds for worst-case lattices, but surprisingly this provides even better results with uniform ball noise than for Gaussian (or Bernoulli noise for codes). This counterintuitive situation can be resolved by adequate decomposition and truncation of Gaussian and noise for codes). This counterintuitive situation can be resolved by adequate decomposition and truncation of Gaussian and noise for codes). This counterintuitive situation can be resolved by adequate decomposition and truncation of Gaussian and noise for codes). This counterintuitive situation can be resolved by adequate decomposition and truncation of Gaussian and noise for codes).

I. INTRODUCTION

A. Smoothing Bounds

In either a code or a lattice, smoothing refers to fact that, as an error distribution grows wider and wider, the associated syndrome distribution tends towards a uniform distribution. In other words, the error distribution, reduced modulo the code or the lattice, becomes essentially flat. This phenomenon is pivotal in arguing security of cryptosystems [14], [16], [28]. In information-theoretic literature, it is also sometimes referred to as flatness [24]. Informally, by a “smoothing bound” we are referring to a result which lower bounds the amount of noise which needs to be added so that the smoothed distribution “looks” flat.

To be more concrete, by a “flat distribution”, we are referring to a uniform distribution over the ambient space modulo the group of interest. For a (linear) code $C \subseteq F_2^n$, this quotient space is $F_2^n/C$; for a lattice $\Lambda \subseteq \mathbb{R}^n$, it is $\mathbb{R}^n/\Lambda$. We then consider some “noise” vector $e$ distributed over the ambient space $F_2^n$ (resp. $\mathbb{R}^n$), and attempt to prove that $e \mod C$ (resp. $e \mod \Lambda$) is “close” to the uniform distribution over the quotient space $F_2^n/C$ (resp. $\mathbb{R}^n/\Lambda$). To quantify “closeness” between distributions, we will use the standard choice of statistical distance.

An important question to be addressed is the choice of distribution for the noise vector $e$. In lattice-based cryptography (where such smoothing bounds originated [28]), the literature ubiquitously uses Gaussian distributions for errors, and smoothness is guaranteed for an error growing as the inverse of the minimum distance of the dual lattice. The original chain [28] of argument goes as follows:

- Apply the Poisson summation formula (PSF);
- Bound variations via the triangle inequality (TI) over all non-zero dual lattice points;
- Bound the absolute sum above via the Banaszczyk tail bound [5] for discrete Gaussian (BT).

An intermediate quantity called the smoothing parameter introduced by [28] before the last step is also often used in the lattice-based cryptographic literature. Each bounding step is potentially non-tight, and indeed more recent works have replaced the last step by the following [3]:

- Bound the number of lattice points in balls of a given radius via the Linear Programming bound [21] (LP) and “sum over all radii” (with care).

With this LP strategy, it is in principle possible to also compute a smoothing bound for spherically symmetric distributions of errors other than the Gaussian; however, we are not aware of prior work doing this explicitly. A very natural choice would be uniform distributions over Euclidean balls.

For codes, there are also two natural distributions of errors: Bernoulli noise, i.e. flip each bit independently with some probability $p$ (a.k.a. the binary symmetric $BS_{p}$ channel), and a uniform noise over a Hamming sphere of a fixed radius. The latter is typically preferred for the design of concrete and practical cryptosystems [4], [14], [27], [30], while the former...
appears more convenient in theoretical works.1 Cryptographic interest for code smoothing has recently arisen [8], [37], but results are so far limited to codes with extreme parameters and specific “balancedness” constraints. However we note that the question is not entirely new in the coding literature (see for instance [19]). In particular, an understanding of the smoothing properties of Bernoulli noise is intimately connected to the undetected error probability of a code transmitted through the BSCp.

In this light, it is interesting to revisit and systematize our understanding of smoothing bounds, unencumbered by direct application concerns. We find it enlightening to do this exploration in parallel between codes and lattices, transferring techniques back and forth between both areas whenever possible.

Furthermore, we keep our arguments agnostic to the specific choice of error distribution, allowing us to apply them with different error distributions and compare the results. To compare different (symmetric) distributions, we advocate parametrizing them by the expected weight/norm of a vector. That is, we quantify the magnitude of a noise vector e by t = E(|e|) (where | · | denotes either the Hamming weight or the Euclidean norm of the vector). Our smoothing bounds will depend on this parameter, and we consider a smoothing bound to be more effective if for the smoothed distribution to be close to uniform we require a smaller lower-bound on t.

B. Contributions

In this work, we collect the techniques that have been used for smoothing, both in the code and lattice contexts. We view individual steps as modular components of arguments, and consider all permissible combinations of steps, thereby determining the most effective arguments. In the following, we outline our systematization efforts, describing the various proof frameworks that we tried before settling on the most effective argument.

Code Smoothing Bounds: Given the relative dearth of results concerning code smoothing, it seems natural to start by adapting the first argument (PSF+TI+BT) to codes following the proof techniques of [5] and [28]. And indeed, the whole strategy translates flawlessly, with only one caveat: it leads to a very poor result, barely better than the trivial bound. Namely, smoothness is established only for Bernoulli errors with parameter very close to p = 1/2.

The adaptation of Banaszczyk tail bound [5] to codes (together with replacing the Gaussian by a Bernoulli distribution) is rather naïve, and it is therefore not very surprising that it leads to a disappointing result. Instead, we can also follow the improved strategy for lattices from [3], and resort to linear programming bounds for codes [1], [6], [29]. Briefly, by an LP bound we are referring to a result that bounds the number of codewords (resp. lattice vectors) of a certain weight (resp. norm) in terms of the dual distance (resp. shortest dual vector) of the code (resp. lattice). In both cases, the results are obtained by considering a certain LP relaxation of the combinatorial quantities one wishes to bound, hence the name. Even more, the bounds for codes and lattices are obtained via essentially the same arguments [9], [11], [29]. We therefore find it natural to apply LP bounds in our effort to develop proof techniques which apply to both code- and lattice-smoothing.

The strategy (PSF+TI+LP) turns out to give a significantly better result, but it nevertheless still appears to be far from optimal. We believe that the application of the triangle inequality in the second step to bound the sum of Fourier coefficients given by the Poisson summation formula leads to the unsatisfactory bound. Indeed, a common heuristic when dealing with sums of Fourier coefficients is that, unless there is a good reason otherwise, the sum should have magnitude roughly the square-root of the order of the group (as is the case for random signs): the triangle inequality is far too crude to notice this.

Instead, we turn to another common upper-bound on a sum, namely, the Cauchy-Schwarz (CS) inequality. It is natural to subsequently apply Parseval’s Identity (PI). It turns out that this strategy yields very promising results, upon which we now elucidate. The upper-bound is described in terms of the weight distribution of a code, i.e. the number of codewords of weight w for each w = 1, . . . , n. Unfortunately, it is quite difficult to understand the weight distribution of arbitrary codes, and the bounds that we do have are quite technical.

Random Codes: For this reason, we first apply our proof template to random codes, as it is quite simple to compute the (expected) weight distribution of a random code. Quite satisfyingly, the simple two steps arguments (PI+CS) already yields optimal results for this case, but when the error is sampled uniformly at random from a sphere! That is, we can show that the support size of the error distribution matches the obvious lower bound that applies to any distribution that successfully smooths a code: namely, for a code C the support size must be at least 2(ε2/n)/ε]. Using coding-theoretic terminology, the weight of the error vector that we need to smooth is given by the ubiquitous Gilbert-Varshamov bound

$$\omega_{GV}(R) = h^{-1}(1 - R)$$

which characterizes the trade-off between a random code’s rate R and its minimum distance. Here, h−1 is the inverse of the binary entropy function.

Moreover, as the argument is versatile enough to apply to essentially all spherical error distributions, we also tried applying it to the Bernoulli distribution, and the random walk distribution of [8]. Comparing them, we were rather surprised that our argument provided better bounds for the uniform distribution over a Hamming sphere than the other two distributions for the same average Hamming weight.

However, while the (PI+CS) sequence of arguments is more effective when the noise is sampled uniformly on the sphere, we can exploit the fact that the Hamming weight of a Bernoulli-distributed vector is tightly concentrated to recover the same smoothing bound for this distribution. In more detail, we use a “truncation” argument. First, we decompose the Bernoulli distribution into a convex combination of uniform sphere distributions. But, by Chernoff’s bound, a Bernoulli distribution is concentrated on vectors whose weight lies in a width εn interval around its expected weight. Therefore,
outside of this interval, the contribution of the Bernoulli on the statistical distance is negligible. Then apply the (PI+CS) sequence of arguments to each constituent distribution close to the expected weight. In this way, we are able to demonstrate that Bernoulli distributions also optimally smooth random codes.

Arbitrary Codes: Next, we turn our attention to smoothing worst-case codes. Motivated by our success in smoothing random codes, we again follow the (PI+CS) sequence of arguments and combine this with LP bounds to derive smoothing bounds when the dual distance of the code is sufficiently large. Again, the sequence of arguments is most effective when the error is distributed uniformly over the sphere, with one caveat: we are also required to assume that the dual code is balanced in the sense that it also does not contain any vectors of too large weight. While this assumption has appeared in other works [8], [37], we find it somewhat unsatisfactory.

Fortunately, this condition is not required if the error is sampled according to the Bernoulli distribution. But then we run into the same issue that we had earlier with random codes: the (PI+CS) argument, followed by LP bounds, natively yields a lesser result when instantiated with Bernoulli noise. Fortunately, we have already seen how to resolve this issue: we pass to the truncated Bernoulli distribution and decompose it into uniform sphere distributions. This yields a best-of-both-worlds result: we obtain the strongest smoothing bound we can in terms of the noise magnitude, while requiring the weakest assumption on the code.

And Back to Lattices: Having now uncovered this better strategy for codes, we can return to lattices and apply our new proof template. Indeed, as we outline in Section II-C, the (PI+CS) sequence of arguments can be applied in a very broad context; see, in particular, Lemma 2.4.

Random Lattices: First, just as we set our expectations for code-smoothing by first studying the random case, we analogously start here by considering random lattices. However, defining a random lattice is a non-trivial task. We actually consider two distributions. The first, which is based on the deep Minkowski-Hlawka-Siegel (MHS) Theorem, we only abstractly describe. Thanks to the MHS Theorem, we can very easily compute the (expected value) of our upper-bound.

For the MHS distribution of lattices, we consider two natural error distributions: the Gaussian distribution (which is used ubiquitously in the literature), as well as the uniform distribution over the Euclidean ball. And again, perhaps surprisingly (although less so now thanks to our experience with the code case), we obtain a better result with the uniform distribution over the Euclidean ball. And moreover, the Euclidean ball result is optimal in the same sense that we had for codes: the support volume of the error distribution is exactly equal to the covolume of the lattice. We view the value \( w \) such that the volume of the \( n \)-ball of radius \( w \) is equal to the covolume of a lattice (which is half the quantity that appears in the Minkowski bound) as being the lattice-theoretic analogue of the Gilbert-Varshamov quantity:

\[
 w_{\text{M}/2} \overset{\text{def}}{=} \sqrt{|\Lambda|} \left( \frac{n/2 + 1}{\sqrt{n}} \right).
\]

However, as Gaussian vectors satisfy many pleasing properties that are often exploited in lattice-theoretic literature, we would like to obtain the same smoothing bound for this error distribution. Fortunately, our experience with codes also tells us how to recover the result for Gaussian noise from the Euclidean ball noise smoothing bound: we decompose the Gaussian distribution appropriately into a convex combination of Euclidean ball distributions. Together with a basic tail bound, we recover the same smoothing bound for Gaussian noise that we had for the uniform ball noise.

We also study random \( q \)-ary lattices, which are more concretely defined: following the traditional lattice-theoretic terminology, they are obtained by applying Construction A to a random code. This does lead to a slight increase in the technicality of the argument – in particular, we need to apply a certain “summing over annuli” trick – but the computations are still relatively elementary. Again, we find that the argument naturally works better when the errors are distributed uniformly over a ball, but we can still transfer the bound to the Gaussian noise.

Interestingly, the same optimal bound has been recovered in a concurrent work [23, Theorem 1] for Gaussian distributions. Their arguments are quite unlike ours: [23] uses the Kullback-Leibler divergence in combination with other information-theoretic arguments. However, contrary to our bounds obtained via the (PI + CS) sequence of arguments, [23, Theorem 1] only holds for random \( q \)-ary lattices.

Arbitrary Lattices: Next, we address the challenge of smoothing arbitrary lattices. And again, we follow the (PI+CS) sequence of arguments, and subsequently use the Kabatiansky and Levenshtein bound [18] to obtain a smoothing bound in terms of the minimum distance of the dual lattice. The Kabatiansky and Levenshtein bound is the lattice analogue of the second LP bound from coding theory. We can directly apply the arguments with both of our error distributions of interest, and again, the uniform ball distribution wins. But the decomposition and tail-bound trick again applies to yield the same result for the Gaussian distribution that we had for the uniform ball distribution.

Comparison: We summarize how our work improves on the state of the art in Table I for lattices, and in Table II and Figure 1 for codes. For this discussion, we let \( U(\mathbb{R}^n/\Lambda) \) (resp. \( U(\mathbb{F}_q^n/\mathcal{C}) \)) denote the uniform distribution over \( \mathbb{R}^n/\Lambda \) (resp. \( \mathbb{F}_q^n/\mathcal{C} \)), and let \( \Delta \) denote the statistical distance.

In the case of lattices (Table I), we fix the smoothing bound target to exponentially small, that is we state the minimal value of \( F > 0 \) such that the bound over the statistical distance implies \( \Delta(\epsilon \mod \Lambda, U(\mathbb{R}^n/\Lambda)) \leq 2^{-\Omega(n)} \) when the error follows the prescribed distribution and of an average Euclidean length of \( \mathbb{E}(|\epsilon|_2) = F \sqrt{n}/\lambda_1(\Lambda)^3 \).

In the case of codes we also fix the smoothing bound target to negligible, but we compare two cases: smoothing bounds

\(^2\text{That is, for a lattice } \Lambda, \text{ the volume of the torus } \mathbb{R}^n/\Lambda. \text{ We will denote this quantity by } |\Lambda| \text{ from now on.} \)

\(^3\text{In fact, the values in this table guarantee exponentially small statistical distance from the uniform distribution.} \)

\(^4\text{Again, it is the same if we insist the statistical distance to uniform is exponentially small.} \)
for random codes (in average) and for a fixed code (worst case). In Figure 1 we compare the minimal value $F > 0$ such that $E_\mathcal{C} (\Delta(e \mod \mathcal{C}^*, U(F_2^n/\mathcal{C}))) \leq 2^{-\Omega(n)}$ when the error $e$ follows the prescribed distribution and with an expectation that is taken over codes of rate $R$. In Table II we make the same comparison but to reach $\Delta(e \mod \mathcal{C}^*, U(F_2^n/\mathcal{C})) \leq 2^{-\Omega(n)}$ for a fixed code $\mathcal{C}^*$ such that the minimum distance of its dual $\mathcal{C}^*$ is known.

### C. Uses for Smoothing Bounds

While the purpose of our paper is to study smoothing bounds unencumbered by direct application concerns, we briefly outline a few situations in which smoothing bounds arise to help the reader put our work in context.

#### Worst-Case to Average-Case Reductions: Smoothing bounds play an important role in worst-case to average-case reductions in cryptography [8], [28], [37]. In such a scenario, given a worst-case instance we must simulate an average-case instance, which means that the simulated instance should be close to uniformly random. Smoothing bounds quantify the performance of this simulation.

To be more concrete, consider the worst-case decoding problem, where we are given $(G, y = xG + t) \in F_2^{k \times n} \times F_2^n$ and the Hamming norm of $t$ is small, and our task is to recover $x$. We wish to reduce this to the learning parity with noise (LPN) problem, where we are given access to an oracle $\mathcal{O}_x$ and we must recover $x \in F_2^n$, where $\mathcal{O}_x$ behaves as followed: when queried, it outputs $(a, \langle a, x \rangle + \eta)$, where $a \in F_2^k$ is uniformly random and $\eta \in F_2$ is a Bernoulli random variable. Following [8], we use the following strategy. Run the algorithm for LPN as a subroutine, and when it requests a sample from its oracle, we sample a random vector $e \in F_2^n$ according to a smoothing distribution, and then return $(eG^\top, \langle e, xG + t \rangle)$.

#### Connection to Mixing in Markov Chains:

Informally, a time-invariant Markov chain consists of a collection of random variables over a state space such that the distribution of the random variable at a given time $t$ only depends on where it was previously measured; see, e.g., [25] for a modern treatment. A basic question is whether the Markov chain mixes: is there a limiting distribution of the random variables? And, if yes, what is the mixing time, i.e. when do the random variables converge?

We remark that smoothing bounds for codes with respect to many distributions correspond to mixing-time bounds for natural Markov chains. For example, consider the random walk distribution: if we wish to understand the smoothing of a dimension $k$ code $\mathcal{C} \subseteq F_2^n$ with parity-check matrix $H$, this corresponds to a time-invariant discrete Markov chain over the state space $F_2^{n-k}$ where in each time-step we move from $x$ to $x + e_jH^\top$ for uniformly chosen $j \in \{1, \ldots, n\}$, where $e_1, \ldots, e_n$ denote the standard basis vectors. This Markov chain has been studied previously [36]: in fact, analogous Markov chains over broader classes of groups is a well-studied topic [33]. Similarly, Beroulli noise corresponds to a continuous-time Markov chain over the state-space $F_2^{n-k}$, while Gaussian noise over a lattice $\Lambda$ could be viewed as a continuous-time Markov chain over the torus $\mathbb{R}^n/\Lambda$.

### Coding for Wiretap Channels:

In a wiretap channel, a message is encoded and sent to a receiver, who must decode the message; however, an eavesdropper also receives a corrupted version of the codeword, and security requires that the eavesdropper learn essentially nothing about the message. For a lattice $\Lambda$, the mod-$\Lambda$ Gaussian wiretap channel adds Gaussian noise modulo the lattice $\Lambda$ to the codewords for both the receiver and the eavesdropper.

A natural means to argue security is to claim that the distribution of the Gaussian modulo the lattice $\Lambda$ is essentially flat. In other words, one must show that the distribution satisfies a Gaussian smoothing property, which means that the distribution of the Gaussian modulo the lattice $\Lambda$ for any lattice $\Lambda$ is essentially flat. In other words, one must show that the distribution satisfies a Gaussian smoothing property, which means that the distribution of the Gaussian modulo the lattice $\Lambda$ for any lattice $\Lambda$ is essentially flat.

#### Table I: Comparison of Smoothing Bounds for Various Proof Strategies and Error Distributions

| Distribution | Proof strategy | Smoothing factor $\gamma$ | General statement
|--------------|----------------|--------------------------|--------------------
| Gaussian     | PS+TT+BT       | $1/(2\pi) \approx 0.1592\pi$ | [MR07, Lemma 3.2] |
| Gaussian     | PS+TT+LP       | $\gamma_{\text{KL}}(\log \mathcal{C}) \approx 0.19746$ | [ADRS15, Lemma 6.1] |
| Gaussian     | PL+CS+LP       | $\gamma_{\text{KL}}(\log \mathcal{C}) \approx 0.03912$ | Theorem 4.20 |
| Uncl. Encl.  | PS+CS+LP       | $\gamma_{\text{KL}}(\log \mathcal{C}) \approx 0.08731$ | Theorem 4.19 |
| Gaussian     | Uncl. + True   | $\gamma_{\text{KL}}(\log \mathcal{C}) \approx 0.07731$ | Theorem 4.21 |

Another thing we should replace $x$ by $x + x'$ where $x'$ is uniformly random to re-randomize the secret.
Similarly, for two continuous probability density functions \( f \) and \( g \) over a same measure space \( \mathcal{E} \), the statistical distance is defined as:
\[
\Delta(f, g) \defeq \frac{1}{2} \int_{\mathcal{E}} |f - g|.
\]

B. Codes and Lattices

We give here some basic definitions and notation about linear codes and lattices.

Linear Codes: In the whole paper, we will deal exclusively with binary linear codes, namely subspaces of \( \mathbb{F}_2^n \) for some positive integer \( n \). The space \( \mathbb{F}_2^n \) will be embedded with the Hamming weight \( |\cdot| \), namely
\[
\forall x \in \mathbb{F}_2^n, \quad |x| \defeq \sharp \{ i \in [1, n] : x_i \neq 0 \}.
\]

We will denote by \( \mathcal{S}_0 \) the sphere with center 0 and radius \( w \); its size is given by \( \binom{n}{w} \) and we have \( \frac{1}{n} \log_2 \binom{n}{w} = h(w/n) + o(1) \) where \( h \) denotes the binary-entropy, namely \( h(x) \defeq -x \log_2(x) - (1 - x) \log_2(1 - x) \).

An \([n, k]\)-code \( \mathcal{C} \) is defined as a dimension \( k \) subspace of \( \mathbb{F}_2^n \). The rate of \( \mathcal{C} \) is \( \frac{k}{n} \). Its minimal distance is given by
\[
d_{\min}(\mathcal{C}) \defeq \min \{ |c - c'| : c, c' \in \mathcal{C} \text{ and } c \neq c' \}.
\]

The number of codewords of \( \mathcal{C} \) of weight \( t \) will be denoted by \( N_t(\mathcal{C}) \), namely
\[
N_t(\mathcal{C}) \defeq \sharp \{ c \in \mathcal{C} : |c| = t \}.
\]

The dual of a code \( \mathcal{C} \) is defined as
\[
\mathcal{C}^* \defeq \{ c^* \in \mathbb{F}_2^n : \forall c \in \mathcal{C}, \ c \cdot c^* = 0 \}
\]

where \( \cdot \) denotes the standard inner product on \( \mathbb{F}_2^n \).

Lattices: We will consider lattices of \( \mathbb{R}^n \) which is embedded with the Euclidean norm \(|\cdot|_2\), namely
\[
\forall x \in \mathbb{R}^n, \quad |x|_2 \defeq \sqrt{n \sum_{i=1}^n x_i^2}.
\]

We will denote by \( B_w \) the ball with center 0 and radius \( w \); its volume is given by
\[
V_n(w) \defeq \frac{\pi^{n/2} w^n}{\Gamma(n/2 + 1)}.
\]

An \( n \)-dimensional lattice \( \Lambda \) is defined as a discrete subgroup of \( \mathbb{R}^n \). The covolume \( |\Lambda| \defeq \text{vol}(\mathbb{R}^n/\Lambda) \) of \( \Lambda \) is the volume of any fundamental parallelootope. The minimal distance of \( \Lambda \) is given by \( \lambda_1(\Lambda) \defeq \min \{|x|_2 : x \in \Lambda \text{ and } x \neq 0\} \). The number of lattice points of \( \Lambda \) of weight \( \leq t \) will be denoted by \( N_{\leq t}(\Lambda) \), namely
\[
N_{\leq t}(\Lambda) \defeq \sharp \{ x \in \Lambda : |x|_2 \leq t \}.
\]

The dual of a lattice \( \Lambda \) is defined as
\[
\Lambda^* \defeq \{ x^* \in \mathbb{R}^n : \forall x \in \Lambda, \ x \cdot x^* \in \mathbb{Z} \}
\]

where \( \cdot \) denotes the standard inner product on \( \mathbb{R}^n \).

C. Fourier Analysis

We give here a brief introduction to Fourier analysis over arbitrary locally compact Abelian groups. Our general treatment will allow us to apply directly some basic results in a code and lattice context, obviating the need in each case to introduce essentially the same definitions and to provide the same proofs.

Lemma 2.4 at the end of this subsection is the starting point of our smoothing bounds: all of our results are obtained by using different facts to bound the right hand side of the inequality.
Groups and Their Duals: In what follows $G$ will denote a locally compact Abelian group. Such a group admits a Haar measure $\mu$. For instance $G = \mathbb{R}$ with $\mu$ the Lebesgue measure $\lambda$, or $G = \mathbb{F}_2^n$ with $\mu$ the counting measure $\sharp$.

The dual group $\widehat{G}$ is given by the continuous group homomorphisms $\chi$ from $G$ into the multiplicative group of complex numbers of absolute value 1, and it is again a locally compact Abelian group. In Figure 2 we give groups, their duals as well as their associated Haar measures that will be considered in this work.

It is important to note that if $H \subseteq G$ is a closed subgroup, then $G/H$ and $H$ are also locally compact groups. Furthermore, $G/H$ has a dual group that satisfies the following isomorphism

$$\widehat{G/H} \simeq H^\perp \overset{\text{def}}{=} \{ \chi \in \widehat{G} : \forall h \in H, \chi(h) = 1 \}.$$

Norms and Fourier Transforms: For any $p \in [1, \infty]$, $L_p(G)$ will denote the space of measurable functions $f : G \rightarrow \mathbb{C}$ (up to functions which agree almost everywhere) with finite norm $\|f\|_p$ which is defined as

$$\|f\|_p \overset{\text{def}}{=} \sqrt[p]{\int_G |f|^p d\mu}.$$

The Fourier transform of $f \in L_1(G)$ is defined as

$$\hat{f} : \chi \in \widehat{G} \mapsto \int_G f \overline{\chi} d\mu.$$

We omitted here the dependence on $G$. It will be clear from the context.

Theorem 2.1 (Parseval’s Identity): Let $f \in L_1(G) \cap L_2(G)$, then with appropriate normalization of the Haar measure

$$\|f\|_2 = \|\hat{f}\|_2.$$

Poisson Formula: Given $H \subseteq G$ and any function $f : G \rightarrow \mathbb{C}$, its restriction over $H$ is defined as $f|_H : h \in H \mapsto f(h) \in \mathbb{C}$. We define its periodization as follows.

Definition 2.2 (Periodization): Let $H$ be a closed subgroup of $G$ and $f \in L_1(G)$. We define the $H$-periodization of $f$ as

$$f|_H : (g + H) \in G/H \mapsto \int_H f(g + h) d\mu_H(h) \in \mathbb{C},$$

where $\mu_H$ denotes any choice of the Haar measure for $H$.

There always exists a Haar measure $\mu_{G/H}$ such that for any continuous function with compact support $f : G \rightarrow \mathbb{C}$ the quotient integral formula holds

$$\int_{G/H} \left( \int_H f(g + h) d\mu_H(h) \right) d\mu_{G/H}(g + H) = \int_{G/H} f(g) d\mu(g). \quad (1)$$

Theorem 2.3 (Poisson Formula): Let $H \subseteq G$ be a closed subgroup and $f \in L^1(G)$, then with appropriate normalization of the Haar measures,

$$\overline{\left( \hat{f} \right)}_{G/H} = \left( \chi_0 \right)_{G/H}.$$

The following lemma is a simple consequence of the Cauchy-Schwarz inequality, Parseval’s Identity and the Poisson Formula. Our results on smoothing bounds are all based on this lemma.

Lemma 2.4: Let $H$ be a closed subgroup of $G$. Let $a : x \in G/H \mapsto 1$ and $f \in L^1(G)$ such that $\int_{G/H} f d\mu = \mu_{G/H}(G/H)$. Then with appropriate normalization of the Haar measure,\(^6\)

$$\|a - f|_H\|_1 \leq \sqrt{\mu_{G/H}(G/H)} \sqrt{\int_{G/H \setminus \{\chi_0\}} |\hat{f}|^2 d\mu_{G/H}} \quad (2)$$

where $\chi_0$ denotes the identity element of $G/H$.

Proof: We have

$$\|a - f|_H\|_1 = \int_{G/H} |a - f|_H d\mu_{G/H}$$

$$\begin{align*}
&\leq \sqrt{\mu_{G/H}(G/H)} \int_{G/H} |a - f|_H d\mu_{G/H} \\
&= \sqrt{\mu_{G/H}(G/H)} \int_{G/H \setminus \{\chi_0\}} |\hat{f}|^2 d\mu_{G/H} \\
&= \sqrt{\mu_{G/H}(G/H)} \int_{G/H \setminus \{\chi_0\}} |\hat{f}|^2 d\mu_{G/H} \quad (2)
\end{align*}$$

where in the last line we used Poisson and in Equation (2) we used the following equalities:

$$\overline{\hat{f}}(\chi_0) = \int_{G/H} \hat{f} d\mu_{G/H}$$

$$= \int_{G/H} \left( \int_H f(g + h) d\mu_H(h) \right) d\mu_{G/H}(g + H)$$

$$= \int_{G/H} f \quad \text{(By Equation (1))}$$

$$\mu_{G/H}(G/H) \quad \text{(By assumption on f)}$$

and

$$\hat{\overline{a}}(\chi_0) = \int_{G/H} \overline{\hat{a}} d\mu_{G/H} = \mu_{G/H}(G/H)$$

$$\forall \chi \in G/H \setminus \{\chi_0\}, \quad \hat{\overline{a}}(\chi) = \int_{G/H} \overline{\hat{a}} d\mu_{G/H} = 0$$

which concludes the proof. \(\blacksquare\)

\(^6\)We choose the Haar measures $\mu_G$, $\mu_H$, $\mu_{G/H}$ and $\mu_{G/H}$ for which both the Poisson formula and Parseval’s Identity hold.
In this work we will consider $G = \mathbb{R}^n$ and $H = \Lambda$ or $G = \mathbb{F}_2^n$ and $H = \mathcal{C}$. Haar measures associated to $G$, $G/H$ and $G/H$ for which the lemma holds are given in Figure 2. Furthermore, we will use Fourier transforms over $G$ and $G/H$. We describe in Figure 3 the dual groups that we will consider.

**III. SMOOTHING BOUNDS: CODE CASE**

Given a binary linear code $\mathcal{C}$ of length $n$, the aim of a smoothing bound is to quantitatively describe conditions on a noise distribution $f$ to guarantee that $c + e$ is statistically close to the uniform distribution over $\mathbb{F}_2^n$ when $e$ is uniformly drawn from $\mathcal{C}$ and $c$ is sampled according to $f$. Equivalently, we want to understand when $(e \pmod{\mathcal{C}}) \in \mathbb{F}_2^n/\mathcal{C}$ is close to the uniform distribution. We will focus on the case where the distribution of $e$ is radial, meaning that it only depends on the Hamming weight of $e$.

**Notation 3.1:** We will use throughout this section the following notation.

- The uniform probability distribution over the quotient space $\mathbb{F}_2^n/\mathcal{C}$ will frequently recur and for this reason we just denote it by $u$. The uniform distribution over the whole space $\mathbb{F}_2^n$ is denoted by $u_{\text{full}}$ and the uniform distribution over the codewords of $\mathcal{C}$ is denoted by $u_{\mathcal{C}}$.
- We also use the uniform distribution over the sphere $\mathcal{S}$ which we denote by $u_{\mathcal{S}}$.
- For two probability distributions $f$ and $g$ over $\mathbb{F}_2^n$ we denote by $f * g$ the convolution over $\mathbb{F}_2^n$:

$$f * g(x) \overset{\Delta}{=} \sum_{y \in \mathbb{F}_2^n} f(x - y)g(y).$$

It will be more convenient to work in the quotient space and for this we use the following proposition.

**Proposition 3.2:** Let $f$ be a probability distribution over $\mathbb{F}_2^n$ and $\mathcal{C}$ be an $[n, k]$-code. We have

$$\Delta(u_{\text{full}}, u_{\mathcal{C}} * f) = \Delta(u, f^{\mathcal{C}})$$

where

$$f^{\mathcal{C}}(x) \overset{\Delta}{=} 2^k f^{\mathcal{C}}(x) = \sum_{c \in \mathcal{C}} f(x - c).$$

**Proof:** Let $c$ and $e$ be distributed according to $u_{\mathcal{C}}$ and $f$. We have the following computation:

$$\Delta(u_{\text{full}}, u_{\mathcal{C}} * f) = \frac{1}{2} \sum_{x \in \mathbb{F}_2^n} \left| \frac{1}{2^n} - \frac{1}{2^n} \sum_{c_0 \in \mathcal{C}} f(x - c_0) \right| = \frac{1}{2} \sum_{x \in \mathbb{F}_2^n} \left| \frac{1}{2^n} - \frac{1}{2^n} \sum_{c_0 \in \mathcal{C}} f(x - c_0) \right|$$

where in Equation (3) we used that each term of the sum is constant on $x + \mathcal{C}$.

As a rewriting of Lemma 2.4 we get the following proposition that upper-bounds $\Delta(u, f^{\mathcal{C}})$, namely:

**Proposition 3.3:** Let $\mathcal{C}$ be an $[n, k]$-code and $f$ be a radial distribution on $\mathbb{F}_2^n$. We have

$$\Delta(u, f^{\mathcal{C}}) \leq 2^n \sqrt{\sum_{l = d_{\text{min}}(\mathcal{C}^*)}^n N_l(\mathcal{C}^*) |\hat{f}(t)|^2}$$

where by abuse of notation we denote by $\hat{f}(t)$ the common value of $f$ on vectors of weight $t$.

**Proof:** We have that $\mathcal{C}$ is a closed subgroup of $\mathbb{F}_2^n$ with associated Haar measures:

$$\mu_{\mathbb{F}_2^n} = \frac{1}{2^n} 2^n \quad \text{and} \quad \mu_{\mathbb{F}_2^n/\mathcal{C}} = \frac{2^k}{2^n} 2^n$$

for which we can apply Lemma 2.4. Let $a \overset{\Delta}{=} 2^{n-k}u$ and $b \overset{\Delta}{=} 2^n f$. First, it is clear that $a : x \in \mathbb{F}_2^n/\mathcal{C} \mapsto 1$ and that

$$\int_{\mathbb{F}_2^n} b \, d\mu_{\mathbb{F}_2^n} = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} 2^n f(x) = 1 = \mu_{\mathbb{F}_2^n/\mathcal{C}}(\mathbb{F}_2^n/\mathcal{C})$$

where we used that $f$ is a distribution. Therefore we can apply Lemma 2.4 with functions $a$ and $b$. Furthermore, $b^{\mathcal{C}} = 2^n f^{\mathcal{C}} = 2^{n-k} f^{\mathcal{C}}$ by definition of $f^{\mathcal{C}}$. We get the following computation:

$$||a - b^{\mathcal{C}}||_1 = ||a - 2^{n-k} f^{\mathcal{C}}||_1$$

$$= \sum_{x \in \mathbb{F}_2^n/\mathcal{C}} |1 - 2^{n-k} f^{\mathcal{C}}(x)| \frac{1}{2^{n-k}}$$

$$= \sum_{x \in \mathbb{F}_2^n/\mathcal{C}} \frac{1}{2^{n-k}} - f^{\mathcal{C}}(x)$$

To conclude the proof it remains to apply Lemma 2.4 with Equation (4) and then use that $f$ is radial and therefore $\hat{f}$ is also radial.

The upper-bound in Proposition 3.3 involves the weight distribution of the code $\mathcal{C}^*$, namely $(N_l(\mathcal{C}^*))_{l \geq d_{\text{min}}(\mathcal{C}^*)}$. To understand how our bound behaves for a given distribution $f$, we will start in the following subsection with the case of random codes. The expected value for $N_l$ is well known in this case. This will lead us to estimate our bound on almost all codes and gives us some hints about the best distribution to choose for our smoothing bound in the worst case (which is the case that we treat in Subsection III-B).
A. Smoothing Random Codes

The probabilistic model $C_{n,k}$ that we use for our random code of length $n$ is defined by sampling uniformly at random a generator matrix $G \in \mathbb{F}_2^{k \times n}$ for it, i.e.,

$$\mathcal{C} = \{mG; \, m \in \mathbb{F}_2^k\}.$$ 

The expected number of codewords of weight $t$ in the dual $\mathcal{C}^\ast$ is given by:

**Fact 3.4:** For $\mathcal{C}$ chosen according to $C_{n,k}$

$$\mathbb{E}_\mathcal{C}(N_t(\mathcal{C}^\ast)) = \binom{n}{t} 2^k.$$

As justification note that $x \in \mathcal{C}^\ast$ if and only if $Gx^T = 0$ and we have $\mathbb{P}_G(Gx^T = 0) = 1/2^k$ for $x \neq 0$. As there are $\binom{n}{t}$ vectors of weight $t$ in $\mathbb{F}_2^n$, the fact follows.

This estimation combined with Proposition 3.3 enables us to upper-bound $\mathbb{E}_\mathcal{C}(\Delta(u,f^\mathcal{C}))$.

**Proposition 3.5:** We have:

$$\mathbb{E}_\mathcal{C}(\Delta(u,f^\mathcal{C})) \leq 2^n \sum_{t=0}^n \binom{n}{t} 2^k |\hat{f}(t)|^2. \quad (5)$$

**Proof:** By using Proposition 3.3, we obtain:

$$\mathbb{E}_\mathcal{C}(\Delta(u,f^\mathcal{C})) \leq \mathbb{E}_\mathcal{C}\left(2^n \sqrt{\sum_{t=d_{\min}(\mathcal{C}^\ast)}^n N_t(\mathcal{C}^\ast)|\hat{f}(t)|^2}\right) \leq 2^n \mathbb{E}_\mathcal{C}\left(\sum_{t=d_{\min}(\mathcal{C}^\ast)}^n N_t(\mathcal{C}^\ast)|\hat{f}(t)|^2\right) = 2^n \sum_{t=0}^n \binom{n}{t} 2^k |\hat{f}(t)|^2$$

where we used Jensen’s inequality and in the last line we used the linearity of the expectation and Fact 3.4.

It remains now to choose the distribution $f$. A natural choice in code-based cryptography is the uniform distribution $u_w$ over the sphere $\mathcal{S}_w$ of radius $w$ centered around 0.

**Uniform Distribution Over a Sphere:** The Fourier transform of $u_w$ is intimately connected to Krawtchouk polynomials. The Krawtchouk polynomial of order $n$ and degree $w \in \{0,\ldots,n\}$ is defined as

$$K_w(X;n) \defeq \sum_{j=0}^w (-1)^j \binom{X}{j} \binom{n-X}{w-j}.$$

To simplify notation, since $n$ is clear here from context, we will drop the dependency on $n$ and simply write $K_w(X)$. The following fact allows us to relate $K_w$ to $\bar{u}_w$ (see for instance [34, Lem. 3.5.1, §3.5])

**Fact 3.6:** For any $y \in \mathcal{S}_w$

$$\sum_{e \in \mathcal{S}_w} (-1)^y^e = K_w(t). \quad (6)$$

This leads us to

$$\bar{u}_w(x) = \frac{1}{2^n} K_w(|x|) \binom{n}{w}. \quad (7)$$

By plugging this in Equation (5) of Proposition 3.5 we obtain

$$\mathbb{E}_\mathcal{C}(\Delta(u,f^\mathcal{C})) \leq \sum_{t=0}^n \binom{n}{t} 2^k \left(\frac{K_w(t)}{n} \right)^2. \quad (8)$$

The above sum can be upper-bounded by observing that $\left(\frac{K_w/\sqrt{n}}{w}\right)_{0\leq w\leq n}$ is an orthonormal basis of functions $f : \{0,1,\ldots,n\} \to \mathbb{C}$ for the inner product $(f,g)_{\text{rad}} \defeq \sum_{t=0}^n f(t)g(t)/2^n$. It can be viewed as the standard inner product between radial functions over $\mathbb{F}_2^n$. In particular, $\sum_{t=0}^n K_w(t)^2/2 = 1$ [22, Corollary 2.3]. Therefore, for random codes we obtain the following proposition.

**Proposition 3.7:** We have for any $\mathcal{C}$ chosen according to $C_{n,k}$

$$\mathbb{E}_\mathcal{C}(\Delta(u,u_w^\mathcal{C})) \leq \left(\frac{2^n}{n}\right)^{k-1} \left(\frac{n}{w}\right). \quad (9)$$

In other words, if one wants to smooth a random code with target distance $2^{1-O(n)}$ via the uniform distribution over a sphere, one has to choose its radius $w \leq n/2$ such that $\binom{n}{w} = 2^{O(n)} 2^{-n-k}$. It is readily seen that for fixed code rate $R \defeq \frac{k}{n}$, choosing any fixed ratio $\omega \defeq \frac{w}{n}$ such that $\omega > \omega_{\text{GV}}(R)$ is enough, where $\omega_{\text{GV}}(R)$ corresponds to the asymptotic relative Gilbert-Varshamov (GV) bound

$$\omega_{\text{GV}}(R) \defeq h^{-1}(1-R),$$

with $h^{-1} : [0,1] \to [0,1/2]$ being the inverse of the binary entropy function $h(p) = -p \log_2(p) - (1-p) \log_2(1-p)$. The GV bound $\omega_{\text{GV}}(R)$ appears ubiquitously in the coding-theoretic literature: amongst other contexts, it arises as the (expected) relative minimum distance of a random code of dimension $Rn$, or as the maximum relative minimum error weight for which decoding over the binary symmetric channel can be successful with non-vanishing error probability.

This value of radius $n \omega_{\text{GV}}(R)$ is optimal: clearly, the support size of an error distribution smoothing a code $\mathcal{C}$ must exceed $2^p^\mathcal{C}/\mathcal{C}$. Thus, we cannot expect to smooth a code $\mathcal{C}$ with errors in the sphere $\mathcal{S}_w$ if its volume is smaller than $2^{n-k} / 2^w$. Therefore the uniform distribution over a sphere is optimal for random codes. By this, we mean that it leads to the smallest amount of possible noise (when it is concentrated on a ball) to smooth a random code. Notice that we obtained this result after applying the chain of arguments Cauchy-Schwarz, Parseval and Poisson to bound the statistical distance.

**About the Original Chain of Arguments of Micciancio and Regev:** It can be verified that by coming back to the original steps of [3] and [28], namely the Poisson summation formula and then the triangle inequality, we would obtain

$$\Delta(u,f^\mathcal{C}) \leq 2^n \sum_{t=d_{\min}(\mathcal{C}^\ast)}^n N_t(\mathcal{C}^\ast)|\hat{f}(t)|.\quad (10)$$

By using that $a^2 + b^2 \leq (a+b)^2$ (when $a,b \geq 0$) we see that our bound (Proposition 3.3) is sharper. It turns out that our bound is exponentially sharper for random codes (and even in the worst case) when choosing $f$ as the uniform distribution...
over a sphere of radius \( w \), namely \( f = u_w \). In this case the Miecznik-Regev argument yields the following computation

\[
E_{\theta'} \left( \Delta \left( u, u_{w}^{t} \right) \right) = \sum_{t \geq 2w} N_t(\theta') \left( \frac{K_w(t)}{w} \right)
\]

To carefully estimate this upper-bound (and to compare with (8)) we are going to use the following proposition, which gives the asymptotic behaviour of \( K_w \) (see for instance [15], [17]).

**Proposition 3.9:** Let \( n, t \) and \( w \) be three positive integers. We set \( r \equiv \frac{t}{n} \), \( \omega = \frac{w}{n} \) and \( \omega \equiv \frac{1}{2} - \sqrt{\omega(1-\omega)} \). We assume \( w \leq n/2 \). Let

\[
z \equiv \frac{1 - 2r - \sqrt{D}}{2(1-\omega)}
\]

where

\[
D \equiv (1 - 2r)^2 - 4\omega(1-\omega).
\]

Then,

\[
K_w(t) = O \left( 2^{n (a(\tau, \omega) + o(1))} \right).
\]

- In the case \( \tau \in (0, \omega) \),

\[
a(\tau, \omega) \equiv \tau \log_2 (1 - z) + (1 - \tau) \log_2 (1 + z) - \omega \log_2 z.
\]

- In the case \( \tau \in (\omega, 1/2) \), \( D \) is negative, and

\[
a(\tau, \omega) \equiv \frac{1}{2} \left( 1 + h(\omega) - h(\tau) \right).
\]

We let,

\[
\omega_0 = \lim_{n \to \infty} \left\{ \frac{w}{n} \left[ \frac{2n(1-R)}{n} \right] \geq 1 \right\},
\]

\[
\omega_1 = \lim_{n \to \infty} \left\{ \frac{w}{n} \left[ \sum_{t \geq 0} \left( \frac{n}{2n} \right) K_w(t) \right] \geq 1 \right\}.
\]

In Figure 4 we compare the asymptotic values of \( \omega_0 \) and \( \omega_1 \) as functions of \( R \). Notice that we see that \( \omega_1 \) is undefined for a rate \( R < 1/2 \). In other words, it is impossible to show that \( E_{\theta'}(\Delta(u, u_{w}^{t})) \leq 2^{-\Omega(n)} \) with the standard approach of [3] and [28] when \( R < 1/2 \). Furthermore, for larger rates (and sufficiently large \( n \)), \( \omega_0 \) is much smaller than \( \omega_1 \).

**Bernoulli Distribution:** Another natural distribution to consider when dealing with codes is the so-called “Bernoulli” distribution \( f_{ber, p} \), which is defined for \( p \in [0, 1/2] \) as

\[
\forall x \in \mathbb{F}_2^n, \quad f_{ber, p}(x) = p^{|x|}(1 - p)^{n-|x|}.
\]

This choice leads to simpler computations compared to the uniform distribution over a sphere. For instance we have \( f_{ber, p}(x) = \frac{1}{2^n} (1 - 2p)^{|x|} \). By plugging this in Equation (5) of Proposition 3.5 we obtain

\[
E_{\theta'} \left( \Delta(u, f_{ber, p}^{\theta'}) \right) \leq \sum_{t \geq 0} \left( \frac{n}{2^t} \right) (1 - 2p)^{2t} \leq \sqrt{\frac{1}{2^t} (1 + (1 - 2p)^2)^n}.
\]

Thus, if one wants to smooth a random code at target distance \( 2^{-\Omega(n)} \) with the Bernoulli distribution, the above argument says that one has to choose \( p > p_0 \equiv \frac{1}{2} \left( 1 - \sqrt{2n^2 - 1} \right) \) where \( R = \frac{k}{n} \). As \( E_{f_{ber, p}}(|x|) = pn \), it is meaningful to compare \( p_0 \) and \( \omega_0 \). It is readily seen that \( \omega_0 = \omega_{GV}(R) = h^{-1}(1-R) < \frac{1}{2} \left( 1 - \sqrt{2n^2 - 1} \right) = p_0 \). In other words, this time the upper-bound given by Proposition 3.5 does not give what would be optimal, namely the Gilbert-Varshamov relative distance \( \omega_{GV}(R) \), but a quantity which is bigger. However, it is expected that the average amount of noise to smooth a random code is the same in both cases, since a Bernoulli distribution of parameter \( p \) is extremely concentrated over words of Hamming weight \( pn \) and that therefore \( \Delta(u, f_{ber, p}^{\theta'}) \approx \Delta(u, u_{w}^{\theta'}) \). This suggests that Proposition 3.5 is not tight in this case. This is indeed the case: we can prove that we can smooth a random code with the Bernoulli noise as soon as \( p > \omega_{GV}(R) \). This follows from the following proposition.

**Proposition 3.10:** Let \( \varepsilon > 0 \) and \( p \in [0, 1/2] \). Then,

\[
\Delta(u, f_{ber, p}^{\theta'}) \leq \sum_{r=(1+\varepsilon)np} (1+\varepsilon) np \Delta(u, u_{w}^{\theta'}) + 2^{-\Omega(n)}.
\]

**Proof:** See Appendix A.

This proposition shows that if one wants \( \Delta(u, f_{ber, p}^{\theta'}) \leq 2^{-\Omega(n)} \), it is enough to have \( \Delta(u, f_{unif, r}^{\theta'}) \leq 2^{-\Omega(n)} \) for any \( r \in [(1-\varepsilon)np, (1+\varepsilon)np] \). This can be achieved by choosing \( \varepsilon \) and \( p \) such that \( (1 - \varepsilon)p > \omega_{GV}(R) \).

To summarize this subsection we have the following theorem.

**Theorem 3.10:** Let \( \mathcal{C} \) be a random code chosen according to \( \mathcal{C}_{n,k}, R \equiv \frac{k}{n} \). Let \( u \) (resp. \( u_{[pn]} \)) be the uniform distribution over \( \mathbb{F}_2^n/\mathcal{C} \) (resp. \( \mathcal{U}_w \)) and \( f_{ber, p} \) be the Bernoulli distribution over \( \mathbb{F}_2^n \) of parameter \( p \). We have,

\[
E_{\theta'} \left( \Delta(u, u_{[pn]}^{\theta'}) \right) \leq 2^{\frac{1}{2} (1-R-h(p) + o(1))}
\]

and,

\[
E_{\theta'} \left( \Delta(u, f_{ber, p}^{\theta'}) \right) \leq 2^{\frac{1}{2} (1-R-h(p) + o(1))}.
\]

In particular, for any fixed \( p > \omega_{GV}(R) \),

\[
E_{\theta'} \left( \Delta(u, u_{[pn]}^{\theta'}) \right) \leq 2^{-\Omega(n)}
\]
and 
\[ \mathbb{E}_{\mathcal{E}}(\Delta(u, f_{\text{ber},p})) \leq 2^{-\Omega(n)}. \]

### B. Smoothing a Fixed Code

Our upper-bound on $\Delta(u, f_{\mathcal{E}})$ given in Proposition 3.3 involves the weight distribution of the dual of $\mathcal{E}$, namely the $N_i(\mathcal{E}^*)$'s. To derive smoothing bounds on a fixed code, our strategy will simply consist of using the best known upper bounds on the $N_i(\mathcal{E}^*)$'s. Roughly speaking, these bounds show that $N_i(\mathcal{E}^*) \leq (\frac{n}{i})2^{-Kn}$ for some constant $K$ which is a function of $d_{\min}(\mathcal{E}^*)$.

**Notation:** Let $\delta \in (0, 1/2)$ and $\delta \leq \tau \leq 1$,

\[ b(\delta, \tau) \overset{\text{def}}{=} \lim_{n \to \infty} \max_{\mathcal{E}} \left\{ \frac{1}{n} \log_2 N_{\tau n}(\mathcal{E}) \right\} \tag{12} \]

where the maximum is taken over all codes $\mathcal{E}$ of length $n$ and minimum distance $\geq \delta n$.

We recall (or slightly extend) results taken from [1]:

**Proposition 3.11:** Let $\delta \in (0, 1/2)$ and

\[ \delta^1 \overset{\text{def}}{=} 1/2 - \sqrt{\delta(1-\delta)}. \]

For any $\delta \leq \tau \leq 1$,

\[ b(\delta, \tau) \leq c(\delta, \tau) = \begin{cases} h(\tau) + h(\delta^1) - 1 & \text{if } \tau \in [\delta, 1-\delta] \\ 2(h(\delta^1) - a(\tau, \delta^1)) & \text{otherwise} \end{cases} \tag{13} \]

where $a(\cdot, \cdot)$ is defined in Proposition 3.8.

**Proof:** See Appendix B.

**Proposition 3.12 ([1, Proposition 4]):** Let $\delta \in (0, 1/2)$,

\[ \delta_{\text{JSB}} \overset{\text{def}}{=} \left( 1 - \sqrt{1 - 2\delta} \right) / 2 \]

and

\[ \tau_0 \overset{\text{def}}{=} \arg \min_{\delta_{\text{JSB}} \leq \alpha \leq 1/2} 1 - h(\alpha) + R_1(\alpha, \delta) \]

where

\[ R_1(\tau, \delta) \overset{\text{def}}{=} \frac{1}{2} \left( 1 - \sqrt{1 - \left( 4\tau(1-\tau) - \delta(2-\delta) - \delta \right)^2} \right) \].

For any $\delta \leq \tau \leq 1$,

\[ b(\delta, \tau) \leq d(\delta, \tau) \tag{14} \]

where:

- if $\tau \in (\delta_{\text{JSB}}, 1 - \delta_{\text{JSB}})$ and $\tau_0 \leq \tau$,

\[ d(\delta, \tau) \overset{\text{def}}{=} h(\tau) - h(\tau_0) + R_1(\tau_0, \delta); \]

- if $\tau \in (\delta_{\text{JSB}}, 1 - \delta_{\text{JSB}})$ and $\tau_0 > \tau$,

\[ d(\delta, \tau) \overset{\text{def}}{=} R_1(\tau, \delta); \]

- otherwise

\[ d(\delta, \tau) \overset{\text{def}}{=} 0. \]

Both of these bounds are derived from “linear programming arguments” which were initially used to upper-bound the size of a code given its minimum distance. Proposition 3.11 is an extension of [1, Theorem 3] in the case of linear codes, in particular we give an upper-bound for any $\tau \in [\delta, 1]$ (rather than just for $\tau \in [\delta, 1/2]$). The proof is in the appendix. The second bound is usually called the second linear programming bound. In terms of $\delta$ and $\tau$, Proposition 3.11 and 3.12 are among the best (known) upper-bounds on $b(\delta, \tau)$. In the case where $0 \leq \delta \leq 0.273$, Proposition 3.12 leads to better smoothing bounds than Proposition 3.11.

**Remark 3.13:** There exist many other bounds on $b(\delta, \tau)$, like [2, Theorem 8] (which holds only for linear codes) or [2, Theorem 7]. However for our smoothing bounds, Propositions 3.11 and 3.12 lead to the best results, partly because these are the best bounds on the number of codewords of Hamming weight close to the minimum distance of the code.

We draw in Figures 5 and 6 the bounds of Propositions 3.11 and 3.12 as function of $\tau \in [\delta, 1]$ for a couple values of $\delta$.

Equipped with these bounds we are ready to give our smoothing bounds for codes in the worst case, namely for a fixed code. Our study with random codes gave a hint that the choice of the uniform distribution over a sphere could give better results than the Bernoulli distribution. However, as we will now show, the distribution on a sphere forces us to assume that no codewords of large weight belong to the dual $\mathcal{E}^*$ when...
we want to smooth $\mathcal{C}$. It corresponds to the hypothesis of balanced-codes made in [8] to obtain a worst-to-average case reduction. We would like to avoid making this assumption as nothing forbids large weight vectors from belonging to a fixed code. Fortunately, as we will see later, we can avoid making this hypothesis while still keeping the advantages of the uniform distribution over a sphere.

**Impossibility of Smoothing a Code Whose Dual Is Not Balanced With the Uniform Distribution Over a Sphere:** It is readily seen that in the case where the dual code $\mathcal{C}^*$ is not balanced, meaning that it contains the all-one vector (and therefore that the dual weight distribution is symmetric: $N_w(\mathcal{C}^*) = N_{n-w}(\mathcal{C}^*)$ for any $w \in \{0, \ldots, n\}$ when the code length is $n$), then it is impossible to smooth it with the uniform distribution over a sphere. Indeed, this implies that all codewords of $\mathcal{C}$ have an even Hamming weight (they have to be orthogonal to the all-one vector). The parity of the Hamming weights of vectors in a coset (i.e. in the class of representatives of some element in $\mathbb{F}_2^n/\mathcal{C}$) will be the same. Therefore, half of the cosets cannot be reached when periodizing $u_w$ over $\mathcal{C}$.

**Difficulty of Using Proposition 3.3 for Proving Smoothness of the Uniform Distribution Over a Sphere:** Even in the case where the dual is balanced, $\mathcal{C}^*$ is not balanced, meaning that it contains the all-one vector and therefore that the dual weight distribution is symmetric: $N_w(\mathcal{C}^*) = N_{n-w}(\mathcal{C}^*)$ for any $w \in \{0, \ldots, n\}$ when the code length is $n$, then it is impossible to smooth it with the uniform distribution $u_w$ over a sphere. Fortunately, as we will later show, we can avoid making this assumption as nothing forbids large weight vectors from belonging to a fixed code. We would like to avoid making this assumption while still keeping the advantages of the uniform distribution over a sphere.
Therefore, using Lemma 3.15 in Equation (16) gives for \( p \in (0, 1/2] \),
\[
\Delta(u, f_{\text{ber}, p}^{\delta'}) \leq \sqrt{n - d_{\text{min}}(\mathcal{E}^*)/2} \sum_{t = d_{\text{min}}(\mathcal{E}^*)} N_t(\mathcal{E}^*)(1 - 2p)^{2t} + 2^{-\Omega(n)}. \tag{17}
\]

In other words, large weight dual codewords (if they exist) have only an exponentially small contribution to our smoothing bound with the Bernoulli distribution. In principle, we could plug in Equation (17) bounds on the \( N_t(\mathcal{E}^*) \)'s given in Propositions 3.11 and 3.12. We will improve on the bounds obtained in this way by truncating the Bernoulli distribution, then

(i) prove that by appropriately truncating both distributions have the same smoothness property,
(ii) show that the truncated distribution has the same nice properties with respect to large weights,
(iii) show that we can apply Proposition 3.3 to the truncated distribution and get appropriate smoothness properties.

We obtain in this way:

Theorem 3.16: Let \( \varepsilon > 0 \), \( \mathcal{E} \) be a binary linear code of length \( n \) and \( p \in (0, 1/2] \) such that \( d_{\text{min}}(\mathcal{E}^*) \geq \delta^*_n \) for some \( \delta^*_n \in [0, 1] \). We have asymptotically,
\[
\frac{1}{n} \log_2 \Delta(u, f_{\text{ber}, p}^{\delta'}) \leq \max_{\delta^*_n \leq \tau \leq 1 - \delta^*_n} \left\{ \frac{1}{2} \min \left\{ c(\delta^*, \tau), d(\delta^*, \tau) \right\} + \max_{(1-\varepsilon)p \leq \lambda \leq (1+\varepsilon)p} \left\{ \lambda \log p + (1 - \lambda) \log_2 (1 - p) + a(\lambda, \tau) \right\} \right\} + O \left( \frac{1}{n} \right)
\]
where \( a(\cdot, \cdot) \), \( c(\cdot, \cdot) \) and \( d(\cdot, \cdot) \) are defined respectively in Propositions 3.8, 3.11 and 3.12.

Proof: See Appendix C.

Let \( i \in \{0, 1\} \) and \( p_i \) be the smallest \( p \in (0, 1/2] \) that enables to reach \( \Delta(u, f_{\text{ber}, p}^{\delta'}) \leq 2^{-\Omega(n)} \) with

• Theorem 3.16 when \( i = 0 \),
• Equation (17) and Propositions 3.11, 3.12 when \( i = 1 \).

In Figure 7 we compare \( p_0 \) and \( p_1 \). As we can see Theorem 3.16 leads to significantly better bounds. Furthermore, it turns out that \( p_0 n \) is roughly equal to the smallest radius \( u \) such that \( \Delta(u, f_{\text{ber}, p}^{\delta'}) \leq 2^{-\Omega(n)} \) if we had supposed that no codewords of weight \( n - d_{\text{min}}(\mathcal{E}^*) \) belong to \( \mathcal{E}^* \). In other words, our proof using the tweak of truncating the Bernoulli enables us to obtain a smoothing bound without the hypothesis of no dual codewords of large Hamming weight which is as good as with the uniform distribution over a sphere if we had made this assumption.

IV. SMOOTHING BOUNDS: LATTICE CASE

Given an \( n \)-dimensional lattice \( \Lambda \) the aim of smoothing bounds is to give a non-trivial model of noise \( e \in \mathbb{R}^n \) for \( e \mod \Lambda \) \( \in \mathbb{R}^n/\Lambda \) (namely the reduction of \( e \) modulo \( \Lambda \)) to be uniformly distributed. Following Micciancio and Regev [28], the standard choice of noise is given by the Gaussian distribution, defined via

\[
\forall x \in \mathbb{R}^n, \quad D_s(x) \overset{\text{def}}{=} \frac{1}{s^n} \rho_s(x)
\]

where \( \rho_s(x) \overset{\text{def}}{=} e^{-\pi |x|^2/s^2} \).

The parametrization is chosen such that \( s \sqrt{n/2\pi} \) is the standard deviation of \( D_s \). Micciancio and Regev showed that when \( e \) is distributed according to \( D_s \), choosing \( s \) large enough enables \( e \mod \Lambda \) to be statistically close to the uniform distribution.

However, following the intuition from the case of codes we will first analyze the case where \( e \) is sampled uniformly from a Euclidean ball. Interestingly, just as with codes where our methodology led to stronger bounds when the uniform distribution over a sphere was used to smooth rather than the Bernoulli distribution, we will obtain better results when we work with the uniform distribution over a ball. Fortunately, using concentration of the Gaussian measure one can translate results from the case where \( e \) is uniformly distributed over a ball to the case that it is sampled according to \( D_s \); see Proposition 4.6. This is analogous to the translation from results for the uniform distribution over a sphere to the Bernoulli distribution for codes elucidated in Proposition 3.9.

For either choice of noise, to obtain a smoothing bound we are required to bound the statistical distance between the distribution of \( e \mod \Lambda \) if \( e \) has density \( g \), and the uniform distribution over \( \mathbb{R}^n/\Lambda \). It is readily seen that \( e \mod \Lambda \) has density \( |\Lambda|g^{\Lambda} \) which is defined as (see Definition 2.2 with the choice of Haar measures given in Table II)

\[
g^{\Lambda}(x) = \frac{1}{|\Lambda|} \sum_{y \in \Lambda} g(x + y)
\]

where we recall that \( |\Lambda| \overset{\text{def}}{=} \text{vol} (\mathbb{R}^n/\Lambda) \) is the covolume of the lattice \( \Lambda \).

Notation 4.1: For any \( g : \mathbb{R}^n \to \mathbb{C} \),

\[
g^{\Lambda} \overset{\text{def}}{=} |\Lambda| g^{\Lambda}.
\]

In the following proposition we specialize Lemma 2.4 to the case of lattices.

Proposition 4.2: Let \( \Lambda \) be an \( n \)-dimensional lattice. Let \( g \) be some density function on \( \mathbb{R}^n \) and \( v \) be the density of the uniform distribution over \( \mathbb{R}^n/\Lambda \). We have

\[
\Delta(v, g^{\Lambda}) \leq \frac{1}{2} \sqrt{\sum_{x \in \Lambda^* \setminus \{0\}} |\overline{g}(x)|^2}.
\]

Fig. 7. Smoothing bounds for a code \( \mathcal{E} \) as function of \( \delta^* \leq d_{\text{min}}(\mathcal{E}^*)/n \) via Theorem 3.16 (for \( \varepsilon = 10^{-2} \)) and Equation (17).

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
We will restrict our instantiations to functions $g$ whose Fourier transforms are radial, that is, $\widehat{g}(\mathbf{x})$ depends only on the Euclidean norm of $\mathbf{x}$, namely $|\mathbf{x}|_2$.

### A. Smoothing Random Lattices

As with codes, we begin our investigation of smoothing lattices by considering the random case. However, defining a “random lattice” is much more involved than the analogous notion of random codes. Fortunately for us, we can apply the Siegel version of the Minkowski-Hlawka theorem to conclude that there exists a random lattice model which behaves very nicely from the perspective of “test functions”. We first state the technical theorem that we require.

**Theorem 4.3 (Minkowski-Hlawka-Siegel):** On the set of all the lattices of covolume $M$ in $\mathbb{R}^n$ there exists a probability measure $\mu$ such that, for any Riemann integrable function $g(\mathbf{x})$ which vanishes outside some bounded region, $7$

$$
\mathbb{E}_{\Lambda \sim \mu} \left( \sum_{\mathbf{x} \in \Lambda \setminus \{0\}} g(\mathbf{x}) \right) = \frac{1}{M} \int_{\mathbb{R}^n} g(\mathbf{x}) d\mathbf{x}.
$$

As intuition for the above theorem, consider the case that $g$ is the indicator function for a bounded, measurable subset $S \subseteq \mathbb{R}^n$. Then, Theorem 4.3 promises that the expected number of lattice points (other than the origin)$^8$ in $S$ is equal to the volume of $S$ over the covolume of the lattice.

**Uniform Distribution Over a Ball:** Let

$$
u_{w} = \frac{1}{V_n(w)} \mathbb{1}_{B_w}(\mathbf{x})
$$

be the density of the uniform distribution over the Euclidean ball of radius $w$. Let us recall that $V_n(w)$ denotes the volume of any ball of radius $w$. From Theorem 4.3, we may obtain the following proposition. This should be compared with Proposition 3.7.

**Proposition 4.4:** On the set of all lattices of covolume $M$ in $\mathbb{R}^n$ there exists a probability measure $\nu$ such that, for any $w > 0$

$$
\mathbb{E}_{\Lambda \sim \nu} \left( \Delta(u, u_{w,\mathbb{B}}^\Lambda) \right) \leq \frac{1}{2} \sqrt{\frac{M}{V_n(w)}}.
$$

In particular, defining

$$
w_0 \overset{\text{def}}{=} \sqrt{n/2\pi e} M^{1/n},
$$

if $w > w_0$ we have

$$
\mathbb{E}_{\Lambda \sim \nu} \left( \Delta(u, u_{w,\mathbb{B}}^\Lambda) \right) \leq O(1) \left( \frac{w_0}{w} \right)^{n/2}.
$$

**Proof:** We define $\nu$ to be the procedure that samples a lattice according to $\mu$ of covolume $M^{-1}$, then outputs its dual. In the following chain, we first apply Proposition 4.2; then, Jensen’s inequality; then, the Minkowski-Hlawka-Siegel (MHS) Theorem (Theorem 4.3) to the function $|u_{w,\mathbb{B}}^\Lambda|^2$; and, lastly, Parseval’s Identity (Theorem 2.1)

$$
\begin{align*}
\mathbb{E}_{\Lambda \sim \mu} \left( 2\Delta(u, u_{w,\mathbb{B}}^\Lambda) \right) &\leq \frac{n}{2} \mathbb{E}_{\Lambda \sim \mu} \left( \sum_{\mathbf{x} \in \Lambda \setminus \{0\}} |u_{w,\mathbb{B}}^\Lambda(\mathbf{x})|^2 \right) \\
&\leq \frac{n}{2} \mathbb{E}_{\Lambda \sim \mu} \left( \sum_{\mathbf{x} \in \Lambda \setminus \{0\}} |u_{w,\mathbb{B}}^\Lambda(\mathbf{x})|^2 \right) \\
&= \frac{1}{M} \int_{\mathbb{R}^n} |u_{w,\mathbb{B}}^\Lambda(\mathbf{x})|^2 d\mathbf{x} \\
&= \frac{M}{V_n(w)} \\
&= O(1) \left( \frac{w_0}{w} \right)^{n/2}.
\end{align*}
$$

For the “in particular” part of the proposition, we use Stirling’s estimate to derive

$$
\begin{align*}
V_n(w) &= \pi^{n/2} w^n \Gamma(n/2 + 1) = \frac{\pi^{n/2} w^n}{(2\pi e)^{n/2}} \left( 1 + o(1) \right)^n \\
\end{align*}
$$

from which it follows that if

$$
w > w_0 = \sqrt{n/2\pi e} M^{1/n},
$$

we have

$$
\sqrt{\frac{M}{V_n(w)}} \leq O(1) \left( \frac{w_0}{w} \right)^{n/2}
$$

which concludes the proof.

It is easily verified that the value of $w_0$ defined in Proposition 4.5 corresponds to the so-called Gaussian heuristic. We view this condition on $w > w_0$ as the equivalent of the Gilbert-Varshamov bound for codes as we discussed just below Proposition 3.7. In particular, as we need the support of the noise to have volume at least $M$ if we hope to smooth a lattice of covolume $M$, we see that the uniform distribution over a ball is optimal for smoothing random lattices, just as the uniform distribution over a sphere was optimal for smoothing random codes.

**Gaussian Noise:** We now turn to the case of Gaussian noise. Following the proof of Proposition 4.4 to the point where we apply Parseval’s identity, but replacing $u_{w,\mathbb{B}}$ by $D_s$, we obtain that

$$
\mathbb{E} \left( \Delta(u, D_s^\Lambda) \right) \leq \sqrt{M} \int_{\mathbb{R}^n} |D_s(\mathbf{x})|^2 d\mathbf{x}.
$$

To conclude, one uses the following routine computation

$$
\int_{\mathbb{R}^n} |D_s(\mathbf{x})|^2 d\mathbf{x} = \frac{1}{s^{2n}} \int_{\mathbb{R}^n} e^{-2\pi s^2 \|\mathbf{x}\|^2} d\mathbf{x}
$$

$$
= \frac{1}{s^{2n}} \int_{\mathbb{R}^n} \rho_{\sqrt{s} / \sqrt{\pi}}(\mathbf{x}) d\mathbf{x} = \left( \frac{1}{s\sqrt{2}} \right)^n.
$$

Thus, we obtain:

**Proposition 4.5:** On the set of all the lattices of covolume $M$ in $\mathbb{R}^n$ there exists a probability measure $\nu$ such that, for

---

$^7$This statement holds for a larger class of functions. In particular it holds for our instantiation with the Gaussian distribution.

$^8$Note that as $0 \in \Lambda$ with certainty, there is really no “randomness” for this event.
any $s > 0$, 
\[ \mathbb{E}_{\Lambda \sim \nu} \left( \Delta(u, D^A_s) \right) \leq \frac{1}{2} \sqrt{\frac{M}{(s/\sqrt{2})^n}}. \]

In particular, if $s > s_0 \equiv M^{1/n}/\sqrt{2}$, we have 
\[ \mathbb{E}_{\Lambda \sim \nu} \left( \Delta(u, D^A_s) \right) \leq \left( \frac{s_0}{s} \right)^{n/2}. \]

To compare Propositions 4.4 and 4.5, we note that a random vector sampled according to $D_s$ has an expected Euclidean norm given by $s \frac{\Gamma \left( \frac{n+1}{2} \right) }{\sqrt{\pi \Gamma(\frac{n}{2})}} \sim s \sqrt{\frac{2}{n}}$. So, it is fair to compare the effectiveness of smoothing with a parameter $s$ Gaussian distribution and the uniform distribution over a ball of radius $s \sqrt{\frac{2}{n}}$. We note that, if $s_0$ is as in Proposition 4.5 and $w_0$ is the radius of the so-called Gaussian heuristic, then 
\[ s_0 \sqrt{\frac{n}{2\pi}} = \frac{M^{1/n}}{\sqrt{2}} \sqrt{\frac{n}{2\pi}} = w_0 \sqrt{e/2}. \]

Thus, we conclude that the parameter $s_0$ from Proposition 4.5 is larger than what we could hope by a factor $\sqrt{e/2}$.

**B. Connecting Uniform Ball Distribution to Gaussian**

However, recall that in the code-case we argued that, as the Hamming weight of a vector sampled according to the Bernoulli distribution is tightly concentrated, we could obtain the same smoothing bound for the Bernoulli distribution as we did for the uniform sphere distribution, essentially by showing that we can approximate a Bernoulli distribution by a convex combination of uniform sphere distributions. Similarly, we can relate the Gaussian distribution to the uniform distribution over a ball, and thereby remove this additional $\sqrt{e/2}$ factor.

We state a general proposition that allows us to translate smoothing bounds for the uniform ball distribution to the Gaussian distribution. It guarantees that if the uniform ball distribution smooths whenever $w > w_0$, the Gaussian distribution smooths whenever $s > s_0 \sqrt{\frac{2}{n}}$. While the intuition for the argument is the same as that which we used in the code-case, the argument is itself a bit more sophisticated.

**Proposition 4.6:** Let $\Lambda$ be a random lattice of covolume $M$ and let $u = u_{2^n/\Lambda}$ be the uniform distribution over its cosets. Suppose that for all $w > w_0$ there is a function $f(n)$ such that 
\[ \mathbb{E}_{\Lambda} \left( \Delta(u, u_{w,\Lambda}) \right) \leq f(n) \left( \frac{w_0}{w} \right)^{n/2}. \]

Let $s_0 \equiv w_0 \sqrt{\frac{2}{n}}$. Then, for all $s > s_0$, defining $\eta \equiv 1 - \frac{w_0}{s} \in (0, 1)$, we have 
\[ \mathbb{E}_{\Lambda} \left( \Delta(u, D^A_s) \right) \leq \exp\left( -\frac{\eta^2}{8} s \right) + f(n) \left( \frac{s_0}{s} \right)^{n/4}. \]

**Proof:** See Appendix D.

Combining the above proposition with Theorem 4.3, setting $f(n) = O(1)$, we obtain the following theorem.

**Theorem 4.7:** Let $\Lambda$ be a random lattice of covolume $M$ sampled according to $\nu$, let $u \equiv u_{2^n/\Lambda}$ be the uniform distribution over its cosets, and let $s_0 \equiv M^{1/n}/\sqrt{e}$.

Then, for any $s > s_0$, setting $\eta \equiv 1 - \frac{w_0}{s} \in (0, 1)$, we have 
\[ \mathbb{E}_{\Lambda} \left( \Delta(u, D^A_s) \right) \leq \exp\left( -\frac{\eta^2}{8} s \right) + O(1) \left( \frac{s_0}{s} \right)^{n/4}. \]

**Remark 4.8:** In analyzing the secrecy rate of the mod-$\Lambda$ Gaussian wiretap channel (cf. Subsection I-C) Mirghesemi and Belfiore [26] essentially showed that for some lattice $\Lambda$ of covolume 1, a Gaussian of parameter $s > 1/\sqrt{2}$ will smooth $\Lambda$ (with statistical distance from uniform bounded by $\exp(-O(n))$). Note that this is a factor $\sqrt{\frac{2}{n}}$ larger than our best bound for random lattices, i.e. Theorem 4.7.

This can be explained as follows: they obtain their bound by proceeding directly with the Gaussian distribution as we outline prior to Proposition 4.5, rather than by relating it to the uniform ball distribution as we do in Theorem 4.7.

**C. Smoothing Random $q$-Ary Lattices**

While the method of sampling lattices promised by the Minkowski-Hlawka-Siegel Theorem (Theorem 4.3) is indeed very convenient for computations, it does not tell us much about how to explicitly sample from the distribution. Furthermore it is not very relevant if one is interested in the random lattices that are used in cryptography.

For a more concrete sampling procedure that is relevant to cryptography, we can consider the randomized Construction $\Lambda$ (or, more precisely, its dual), which gives a very popular random model of lattices which are easily constructed from random codes. Specifically, for a prime $q$ and a linear code $C \subseteq (\mathbb{Z}/q\mathbb{Z})^n$ we obtain a lattice as follows. First, we “lift” the codewords $c \in C$ to vectors in $\mathbb{R}^n$ in the natural way by identifying $\mathbb{Z}/q\mathbb{Z}$ with the set $\{0, 1, \ldots, q-1\}$; denote the lifted vector as $\bar{c}$. Then, we can define the following lattice 
\[ \Lambda_{\bar{c}} \equiv \{ \bar{c} : c \in C \} + q\mathbb{Z}^n. \]

In other words: $\Lambda_{\bar{c}}$ consists of all vectors in the integer lattice $\mathbb{Z}^n$ whose reductions modulo $q$ give an element of $C$.

Fix integers $1 \leq k \leq n$, a prime $q$ and a desired covolume $M$. We sample a random lattice $\Lambda$ as follows:

- First, sample a random linear code $C \subseteq (\mathbb{Z}/q\mathbb{Z})^n$ of dimension $k$ (recall this means that we sample a random $k \times n$ matrix $G$ and define $C = \{ mG : m \in (\mathbb{Z}/q\mathbb{Z})^k \}$),
- Then, we scale $\Lambda_{\bar{c}}$ by \( \frac{1}{M^{1/n}} \sqrt{\frac{1}{q^{1-k/n}}} \Lambda_{\bar{c}} \).

Notice that the scaling is chosen so that, as long as $\Lambda$ is of full rank, the lattice $\Lambda$ we output has the desired covolume $M$. We denote this procedure of sampling $\Lambda$ by $\nu_{\Lambda}$ (the dependence on $q$, $k$ and $n$ is left implicit).

The important fact is that, up to an error term (which decreases as $q$ increases), the expected number of lattice points from $\Lambda^*$ in a Euclidean ball of radius $r$ is roughly $\frac{V_n(r)}{M}$, as one would hope.

**Proposition 4.9 ([38, Lemma 7.9.2]):** For every $n \geq 2$, $1 \leq k \leq n$ and prime power $q$, for $\Lambda \sim \nu_{\Lambda}$ the expected number of lattice points from $\Lambda^*$ in a Euclidean ball of radius $w \equiv t/\sqrt{n}$ satisfies 
\[ \sqrt{\frac{M \mathbb{E}_{\Lambda}(N_{\leq w}(\Lambda^*))}{V_n(w)}} = 1 \pm \delta/t \] where $\delta \equiv \frac{1}{q^{1-k/n}}$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
We now turn to bounding the expected statistical distance between $u$ and $u_{\mathcal{B}_w'}$, where $\Lambda \sim \nu_A$ and $w > 0$ is the radius of the Euclidean ball from which the noise is uniformly sampled. First, we state an explicit formula for the Fourier transform of $1_{\mathcal{B}_w}$, the indicator function of a Euclidean ball of radius $w$, in terms of Bessel functions.

**Notation 4.10:** For a positive real number $\mu > 0$, we denote by $J_{\mu} : \mathbb{R} \to \mathbb{R}$ the Bessel function of the first kind of order $\mu$.

The important fact concerning Bessel functions that we will use is the following.

**Fact 4.11:** We have

$$1_{\mathcal{B}_w}(y) = \left( \frac{w}{|y|^2} \right)^{\mu/2} J_{\mu/2}(2\pi w |y|_2).$$  \hfill (18)

We will refrain from providing an explicit formula for Bessel functions, and instead use the following upper-bound as a black-box.

**Proposition 4.12 (20):** For any $x \in \mathbb{R}$ we have

$$|J_{\mu/2}(x)| \leq |x|^{-1/3}.$$

Using this proposition, we first prove a technical lemma that will be reused when we discuss smoothing arbitrary codes (in expectation) with the uniform distribution over the Euclidean ball from which the noise is uniformly sampled.

**Lemma 4.14:** For any $w > 0$ and any $0 < a \leq b = 1 + \frac{1}{n}$ we have, for some constant $C > 0$

$$\max_{a \leq t \leq b} g_w(t) \leq \frac{C}{V_n(b) w^{3/2}} \left( \frac{1}{a} \right)^{3/2}.$$

**Proof:** First, we notice that for all $t \in [a, b]$

$$V_n(t) = \left( \frac{t}{b} \right)^n V_n(b) \geq \left( \frac{a}{b} \right)^n V_n(b) \geq \left( 1 + \frac{1}{n} \right)^{-n} V_n(b) \geq \frac{1}{C'} V_n(b)$$

for some constant $C' > 0$. We now use Proposition 4.12 to derive

$$\max_{a \leq t \leq b} g_w(t) \leq \frac{C'}{V_n(b) w^{3/2}} \left( \frac{1}{a} \right)^{3/2}.$$

for an appropriate constant $C > 0$ which concludes the proof.

We now provide the main theorem of this section. It demonstrates that to smooth our ensemble of random $q$-ary codes (in expectation) with the uniform distribution over the ball of radius $w$, it still suffices to choose $w > w_0 \text{def} \sqrt{n/2\pi/e} M^{1/n}$, assuming $q$ is not too small.

**Theorem 4.15:** Let $n > 2$ and $1 \leq k < n$. Let $q$ be a prime and set $\gamma \equiv \frac{\sqrt{n}}{q^{1/k}}$. Let $\Lambda \sim \nu_A$. For some constant $C > 0$, we have

$$E_\Lambda \left( \Delta(u, u_{\mathcal{B}_w}) \right) \leq C \left( \frac{n}{w} \right)^{3/2} e^{\gamma/2} \sqrt{M V_n(w)}.$$

In particular, if $w > w_0 \text{def} \sqrt{n/2\pi/e} M^{1/n}$, we have

$$E_\Lambda \left( \Delta(u, u_{\mathcal{B}_w}) \right) \leq O \left( \left( \frac{n}{w} \right)^{1/3} e^{\gamma/2} \right) \left( \frac{w_0}{w} \right)^{n/2}.$$

**Proof:** Let $t_j \equiv \left( 1 + \frac{1}{n} \right)^j$ for $j \in \mathbb{N}$ and

$$N_j = \# \{ x^* \in \Lambda^*: t_j \leq |x^*|^2 < t_{j+1} \} \quad \varphi_j \equiv \max_{t_j \leq t \leq t_{j+1}} g_w(t).$$

Now, we apply Proposition 4.2 and the above definitions to obtain (we use Jensen’s inequality in the second inequality)

$$E_\Lambda \left( 2\Delta(u, u_{\mathcal{B}_w}) \right) \leq E_\Lambda \left( \sum_{x^* \in \Lambda^\ast \setminus \{0\}} |\hat{u}_{\mathcal{B}_w}(x^*)|^2 \right) \leq \frac{1}{V_n(w)} E_\Lambda \left( \sum_{x^* \in \Lambda^\ast \setminus \{0\}} g_w(x) \right) \leq \frac{1}{V_n(w)} \sum_{j=0}^{\infty} E_\Lambda \left( N_{t_{j+1}} \varphi_j \right) \leq \frac{1}{V_n(w)} \sum_{j=0}^{\infty} E_\Lambda \left( N_{t_{j+1}} \Lambda^\ast \right) \varphi_j.$$

By Proposition 4.9, we may upper-bound

$$M V_n(t_{j+1}) \left( 1 + \left( \frac{\sqrt{n}}{1 + \frac{1}{n}} \right)^m q^{1-k/n} \right) \leq E_\Lambda \left( N_{t_{j+1}} \Lambda^\ast \right) \leq M V_n(t_{j+1}) \left( 1 + \left( \frac{\sqrt{n}}{q^{1-k/n}} \right)^{m} \right)^n.$$

Now, recalling $\gamma = \frac{n^{3/2}}{q^{1-k/n}}$ we have for any $j \geq 0$

$$\left( 1 + \left( \frac{\sqrt{n}}{1 + \frac{1}{n}} \right)^m q^{1-k/n} \right)^n \leq \left( 1 + \left( \frac{\sqrt{n}}{q^{1-k/n}} \right)^m \right)^n \leq e^{n \sqrt{n/\gamma}} = e^{\gamma}.$$

Thus, we conclude

$$E_\Lambda \left( 2\Delta(u, u_{\mathcal{B}_w}) \right) \leq \left[ \frac{e^{\gamma} M}{V_n(w)} \sum_{j=0}^{\infty} V_n(t_{j+1}) \varphi_j \right].$$
Now, by Lemma 4.14 we have \( \varphi_j \leq \frac{C_1}{V_n(t_j+1)^{2/3}} t_j^{1/3} \) for all \( j \geq 0 \). Hence,

\[
\sum_{j=0}^{\infty} V_n(t_{j+1}) \varphi_j \leq \frac{C_1}{V_n(t_j+1)^{2/3}} \sum_{j=0}^{\infty} \frac{V_n(t_{j+1})}{V_n(t_j+1)} t_j^{1/3} = C_1 \frac{1}{V_n(t_j+1)^{2/3}} \sum_{j=0}^{\infty} \left( 1 + \frac{1}{n} \right)^{2j/3} = C_1 \frac{1}{V_n(t_j+1)^{2/3}} \left( 1 - \left( 1 + \frac{1}{n} \right)^{-2/3} \right) \leq C_2 \frac{n^{2/3}}{V_n(t_j+1)^{2/3}},
\]

for an appropriate constant \( C_2 > 0 \). Thus, putting everything together we derive

\[
E_\Lambda \left( \Delta(u, u_{w,\Phi}^A) \right) \leq \sqrt{\frac{n}{V_n(w)}} \frac{C_2 t_o^{2/3}}{M^{1/3}} \sqrt{e/2} V_n(w)
\]

for some constant \( C > 0 \). The “in particular” part of the Theorem follows analogously to the corresponding argumentation (Stirling’s estimate) used in the proof of Proposition 4.4.

Next, turning to Gaussian noise, we could again prove a smoothing bound “directly,” but this will lose the same factor of \( \sqrt{e/2} \) as we had earlier. Instead, we apply Proposition 4.6 with the function \( f(n) = O \left( \left( \frac{n}{\pi} \right)^{1/3} e^{n/2} \right) \) to conclude the following.

**Theorem 4.16:** Let \( n > 2 \) and \( 1 \leq k < n \). Let \( q \) be a prime and set \( \gamma = \frac{n^{2/3}}{q^{1/3} \gamma} \). Let \( \Lambda \) be a random \( q \)-ary lattice sampled according to \( \nu_A \), let \( u \equiv u_{W,\Phi}^A \) be the uniform distribution over its cosets, and let

\[
s_0 \equiv M^{1/3} \frac{1}{\sqrt{e}}.
\]

Then, for any \( s > s_0 \), setting \( \eta \equiv \frac{2}{n} \in (0, 1) \), we have

\[
E_\Lambda \left( \Delta(u, D_s^A) \right) \leq \exp \left( -\frac{s^2}{8} n \right) + O(1) \frac{(s/s_0)^{n/4} e^{n/2}}{\sqrt{e}}.
\]

**D. Smoothing Arbitrary Lattices**

We now turn our attention to the task of smoothing arbitrary lattices.

Analogously to how we used the minimum distance of the dual code to give our smoothing bound for worst-case codes, we will use the shortest vector of the dual lattice in order to provide our smoothing bound for worst-case lattices. The lemma that we will apply is the following where

\[
C_{KL} \equiv 2^{0.401}.
\]

**Lemma 4.17 ([32, Lemma 3]):** For any \( n \)-dimensional lattice \( \Lambda \),

\[
\forall t \geq \lambda_1(\Lambda), \quad N_{\leq t}(\Lambda) \leq \frac{V_n(t)}{V_n(\lambda_1(\Lambda))} C_{KL}^{o(1)}.
\]

**Remark 4.18:** This lemma is a consequence of the Kabatiansky and Levenshtein’ bound [18] on the size of spherical codes, historically known as the second linear programming bound. We will also occasionally use this terminology.

We begin by considering the effectiveness of smoothing with noise uniformly sampled from the ball. The following theorem is proved using similar techniques to those we used for Theorem 4.15, although instead of using Proposition 4.9 to bound the \( N_{\leq t}(\Lambda^*) \)'s, we use Lemma 4.17.

**Theorem 4.19:** Let \( \Lambda \) be an \( n \)-dimensional lattice and \( u \equiv u_{W,\Phi}^A \) be the uniform distribution over its cosets. Then, it holds that

\[
\Delta(u, u_{w,\Phi}^A) \leq \sqrt{\frac{C_{KL}^{n(1+\omega(1))}}{V_n(\lambda_1(\Lambda^*))}} \sqrt{V_n(w)}.
\]

In particular, setting

\[
w_0 \equiv \frac{1}{C_{KL}^{1+\omega(1/n)}}\frac{V_n(\lambda_1(\Lambda^*))}{2\pi e \lambda_1(\Lambda^*)}
\]

for all \( w > w_0 \), it holds that

\[
\Delta(u, u_{w,\Phi}^A) \leq O(1) (w_0/w)^{n/2}.
\]

**Proof:** Define,

\[
t_0 \equiv \lambda_1(\Lambda^*), \quad t_{j+1} \equiv (1 + \frac{t_j}{n}) t_j \quad \varphi_j \equiv \max \{ g_w(t) \} \quad \text{for } j \geq 0,
\]

where we recall the definition of \( g_w(t) = \frac{1}{V_n(w)} \sum_{x \in \Lambda^* \setminus \{0\}} \Phi_{w,\Phi}(x)^2 \) with \( |x|^2 = t \) (see Notation 4.13). We also define

\[
N_j \equiv \# \{ x^* \in \Lambda^* : t_j \leq |x^*|^2 \leq t_{j+1} \}.
\]

With this notation and Proposition 4.2 we have

\[
2 \Delta(u, u_{w,\Phi}^A) \leq \sqrt{\sum_{x \in \Lambda^* \setminus \{0\}} \frac{|\Phi_{w,\Phi}(x)|^2}{V_n(w)}}
\]

\[
\leq \frac{1}{V_n(w)} \sum_{x \in \Lambda^* \setminus \{0\}} g_w(x)
\]

\[
\leq \frac{1}{V_n(w)} \sum_{j=0}^{\infty} N_j \varphi_j
\]

\[
\leq \frac{1}{V_n(w)} \sum_{j=0}^{\infty} N_{\leq t_{j+1}}(\Lambda^*) \varphi_j.
\]

By Lemma 4.14, for some constant \( C_1 > 0 \), we obtain

\[
\varphi_j \leq \frac{C_1}{V_n(t_{j+1})^{2/3}} \frac{1}{t_j^{2/3}}.
\]

Combining this with the upper-bound on \( N_{\leq t_{j+1}}(\Lambda^*) \) provided by Lemma 4.17 (note that \( t_{j+1} \geq \lambda_1(\Lambda^*) \) for all \( j \geq 0 \),
we find
\[
\sum_{j=0}^{\infty} N_{x_j+1} (c) \varphi_j
\]
\[
\leq \sum_{j=0}^{\infty} \frac{V_n(t_{j+1})}{V_n(c)} C_n^{1(o(1))} \frac{C_1}{V_n(t_{j+1})} w^{2/3} \frac{1}{t_j^{2/3}}
\]
\[
= \frac{C_n^{1(o(1))}}{V_n(c)} \sum_{j=0}^{\infty} \frac{1}{t_j^{2/3}}
\]
\[
= \frac{C_n^{1(o(1))}}{V_n(c)} \left( \frac{n}{w(c)} \right)^{2/3} \lambda(c)^{2/3} (1 + \frac{1}{n})^{2j/3}
\]
In the above, all necessary constants were absorbed into the \( C_n^{o(n)} \) term. Combining this with (20), we obtain the first part of the theorem. The “in particular” part again follows using Stirling’s approximation.

Next, we can consider the effectiveness of smoothing with the Gaussian distribution. As usual, we could follow the steps of the proof of Theorem 4.19 and obtain the same result, but with an additional multiplicative factor of \( \sqrt{\frac{n}{e}} \). That is, we can obtain:

**Theorem 4.20:** Let \( \Lambda \) be an \( n \)-dimensional lattice and \( u \defeq u_{\mathbb{R}^n/\Lambda} \) be the uniform distribution over its cosets. Then, it holds that

\[
\Delta (u, D^A) \leq \sqrt{\frac{C_n^{1(o(1))}}{V_n(c)}} \left[ \frac{1}{w(c)} \right]^{n/2} \frac{n}{w(c)} \lambda(c)^{2/3} (1 + \frac{1}{n})^{2j/3}.
\]

In particular, setting

\[
s_0 \defeq \sqrt{n} \frac{C_n^{1(o(1))}}{2\sqrt{\pi e} \lambda(c)^{2/3}}
\]

it holds for any \( s > s_0 \) that \( \Delta (u, D^A) \leq O(1) (s_0/s)^{n/2} \).

However, as usual it is more effective to combine the bound for the uniform ball distribution and decompose the Gaussian as a convex combination of uniform ball distributions, i.e. to apply Proposition 4.6. In this way, we can obtain the following theorem, improving the smoothing bound \( s_0 \) by another \( \sqrt{\frac{n}{e}} \) factor. In the following theorem, we are setting the \( f(n) \) function of Proposition 4.6 with the \( O(1) \) term in the bound of Theorem 4.19.

**Theorem 4.21:** Let \( \Lambda \) be an \( n \)-dimensional lattice, \( u \defeq u_{\mathbb{R}^n/\Lambda} \) the uniform distribution over its cosets, and

\[
s_0 \defeq \sqrt{n} \frac{C_n^{1(o(1))}}{2\sqrt{\pi e} \lambda(c)^{2/3}}
\]

Then, for any \( s > s_0 \) and letting \( \eta \defeq 1 - \frac{s_0}{s} \in (0, 1) \), it holds that

\[
\Delta (u, D^A) \leq \exp \left( - \frac{\eta^2}{8} n \right) + O(1) \left( \frac{s_0}{s} \right)^{n/4}.
\]

**APPENDIX A**

**PROOF OF PROPOSITION 3.9**

Our aim in this section is to prove the following proposition. **Proposition 3.9:** Let \( \varepsilon > 0 \) and \( p \in [0, 1/2] \). Then,

\[
\Delta(u, f_\varepsilon^{\text{trunc}}) \leq \sum_{r = (1-\varepsilon)p}^{(1+\varepsilon)p} \Delta(u, u_r^{\varepsilon}) + 2^{-\Omega(n)}.
\]

Roughly speaking, this proposition is a consequence of the fact that a Bernoulli distribution concentrates Hamming weights over a small number of slices close to the expected weight \( (1 \pm \varepsilon)pn \) and, on each slice, the Bernoulli distribution is uniform. Let us introduce the truncated Bernoulli distribution over words of Hamming weight \( [(1 - \varepsilon)pn, (1 + \varepsilon)pn] \) for some \( \varepsilon > 0 \), namely:

- if \( |x| \in [(1 - \varepsilon)pn, (1 + \varepsilon)pn] \),

\[
f_{\text{truncBer}, p}(x) \defeq \frac{1}{Z} ; f_{\text{Ber}, p}(x)
\]

otherwise,

\[
f_{\text{truncBer}, p}(x) \defeq 0 ;
\]

where

\[
Z \defeq \sum_{|y| = (1 - \varepsilon)pn} f_{\text{Ber}, p}(y)
\]

is the probability normalizing constant.

**Proposition 3.9** is a consequence of the following lemmas.

**Lemma 4.1:** Let \( \varepsilon > 0 \). We have

\[
\Delta(f_{\text{Ber}, p}, f_{\text{truncBer}, p}) = 2^{-\Omega(n)}.
\]

**Proof:** By Chernoff’s bound

\[
1 - Z = \sum_{|y| \notin [(1 - \varepsilon)pn, (1 + \varepsilon)pn]} f_{\text{Ber}, p}(y) \leq 2e^{-\varepsilon^2 n} = 2^{-\Omega(n)}.
\]

Therefore for any \( |x| \in [(1 - \varepsilon)pn, (1 + \varepsilon)pn] \),

\[
f_{\text{truncBer}, p}(x) = \frac{1}{1 - 2^{-\Omega(n)}} f_{\text{Ber}, p}(x) = \left( 1 + 2^{-\Omega(n)} \right) f_{\text{Ber}, p}(x).
\]

We have now the following computation:

\[
2\Delta(f_{\text{Ber}, p}, f_{\text{truncBer}, p}) = \sum_{x \in \mathbb{F}_2^n} |f_{\text{Ber}, p}(x) - f_{\text{truncBer}, p}(x)|
\]

\[
= \sum_{x \in \mathbb{F}_2^n} \sum_{|x| \notin [(1 - \varepsilon)pn, (1 + \varepsilon)pn]} |f_{\text{Ber}, p}(x) - f_{\text{truncBer}, p}(x)|
\]

\[
+ \sum_{|x| \notin [(1 - \varepsilon)pn, (1 + \varepsilon)pn]} |f_{\text{Ber}, p}(x)|
\]

\[
= 2^{-\Omega(n)} \left( \sum_{|x| \notin [(1 - \varepsilon)pn, (1 + \varepsilon)pn]} |f_{\text{Ber}, p}(x)| + 2^{-\Omega(n)} \right)
\]

where we used Equations (22), (23) and in the last line we used that \( f_{\text{Ber}, p} \) is a probability distribution.


Lemma A.2: We have
\[ \Delta(u, f_{\text{trunc}}^\ell) \leq \Delta(u, f_{\text{trunc}}^\ell) + 2^{-\Omega(n)}. \]

Proof: By the triangle inequality,
\[ \Delta(u, f_{\text{trunc}}^\ell) \leq \Delta(u, f_{\text{trunc}}^\ell) + \Delta(f_{\text{trunc}}^\ell, f_{\text{trunc}}^\ell). \]

Focusing on the second term now
\[
\Delta(f_{\text{trunc}}^\ell, f_{\text{trunc}}^\ell)
= \frac{1}{2} \sum_{y \in \mathbb{F}_2^n \setminus \mathbb{F}} |f_{\text{trunc}}^\ell(y) - f_{\text{trunc}}^\ell(y)|
\]
\[
= \frac{1}{2} \sum_{y \in \mathbb{F}_2^n \setminus \mathbb{F}} \left| \sum_{c \in \mathbb{F}} f_{\text{trunc}}^\ell(c + y) - \sum_{c \in \mathbb{F}} f_{\text{trunc}}^\ell(c + y) \right|
\]
\[
\leq \frac{1}{2} \sum_{y \in \mathbb{F}_2^n \setminus \mathbb{F}} \sum_{c \in \mathbb{F}} |f_{\text{trunc}}^\ell(c + y) - f_{\text{trunc}}^\ell(c + y)|
\]
\[= \Delta(f_{\text{trunc}}^\ell, f_{\text{trunc}}^\ell). \]

which concludes the proof by Lemma A.1.

The following lemma is a basic property of the statistical distance.

Lemma A.3: For any distribution \( f \) and \( (g_i)_{1 \leq i \leq m} \) we have
\[ \Delta\left(f, \sum_{i=1}^{m} \lambda_i g_i\right) \leq \sum_{i=1}^{m} \lambda_i \Delta(f, g_i) \]
where the \( \lambda_i \)'s are positive and sum to one.

We are now ready to prove Proposition 3.9.

Proof of Proposition 3.9: First, by Lemma A.2 we have
\[ \Delta(u, f_{\text{trunc}}^\ell) \leq \Delta(u, f_{\text{trunc}}^\ell) + 2^{-\Omega(n)}. \]

To upper-bound \( \Delta(u, f_{\text{trunc}}^\ell) \) we are going to use Lemma A.3. Notice that
\[ f_{\text{trunc}}^\ell = \sum_{r=0}^{n} \binom{n}{r}^\ell (1-p)^n-r u_r. \]

Therefore it is readily seen that
\[ f_{\text{trunc}}^\ell = \sum_{r=(1-\epsilon)n p}^{(1+\epsilon)n p} \lambda_r u_r \]
where,
\[ \lambda_r \triangleq \frac{1}{2} \binom{n}{r}^\ell (1-p)^n-r. \]

By using Lemma A.3 we obtain:
\[ \Delta(u, f_{\text{trunc}}^\ell) \leq \sum_{r=(1-\epsilon)n p}^{(1+\epsilon)n p} \lambda_r \Delta(u, u_r^\ell) \]
\[ \leq \sum_{r=(1-\epsilon)n p}^{(1+\epsilon)n p} \Delta(u, u_r^\ell) \]
(25)
where in the last line we used that the \( \lambda_r \)'s are smaller than one. To conclude the proof we plug Equation (25) in (24). ■

Appendix B

Proof of Proposition 3.11

Our aim in this section is to prove the following proposition which is an extension of [1, Theorem 3] for \( \tau \in [\delta, 1] \) ([1, Theorem 3] only applied for \( \tau \in [\delta, 1/2] \).

Proposition 3.11: Let \( \delta \in (0, 1/2) \) and
\[ \delta^\perp \triangleq 1/2 - \sqrt{\delta(1-\delta)}. \]

For any \( \delta \leq \tau \leq 1, \)
\[ b(\delta, \tau) \triangleq c(\delta, \tau) \triangleq \begin{cases} h(\delta) + h(\delta^\perp) - 1 & \text{if } \tau \in [\delta, 1-\delta] \\ 2(h(\delta^\perp) - \alpha(\tau, \delta^\perp)) & \text{otherwise} \end{cases} \]

where \( \alpha(\cdot, \cdot) \) is defined in Proposition 3.8.

Our proof is mainly a rewriting of the proof of [1, Theorem 3] which relies on the following proposition.

Proposition B.1 ([1, Proposition 2 with \( d' = 0 \)]): Let \( \mathcal{C} \) be a binary code of length \( n \) such that \( d_{\text{min}}(\mathcal{C}) = \Omega(n) \). Let \( t \triangleq n/2 - \sqrt{d_{\text{min}}(\mathcal{C})}(n - d_{\text{min}}(\mathcal{C})) \) and \( a \) be such that
\[ x_{1}^{(t+1)} < a < x_{1}^{(t)}: \quad \frac{K_{1}(a)}{K_{t+1}(a)} = -1 \]
where \( x_{1}^{(\mu)} \) denotes the first root of the Krawtchouk polynomial of order \( \mu \), namely \( K_{\mu} \).

When \( 0 \leq w < t \leq n/2 \), we have
\[ \sum_{c \in \mathcal{C} \setminus \{0\}} K_w(|c|)^2 \leq \frac{n}{2a} \left( \frac{n}{t} \right)^2 \left( \binom{n}{t+1} + \binom{n}{t} \right)^2 \]
(26)

The approach is to optimize on the choice of \( w \) in Proposition B.1 to give an upper-bound on \( N_1(\mathcal{C}) \). More precisely we observe that
\[ N_1(\mathcal{C}) \leq \frac{1}{K_w(\ell)^2} \sum_{c \in \mathcal{C} \setminus \{0\}} K_w(|c|)^2 \]
\[ \leq \frac{1}{K_w(\ell)^2} \left( \frac{n}{w} \right)^2 \left( \binom{n}{t+1} + \binom{n}{t} \right)^2 \]
(27)
and then choose \( w \) to minimize \( \frac{1}{K_w(\ell)^2} \).

Proof of Proposition 3.11: It will be helpful to define the following map:
\[ x \in [0, 1] \mapsto x^\perp \triangleq \frac{1}{2} - \sqrt{x(1-x)}. \]

It can be verified that this application is an involution, is symmetric \( (1-x)^\perp = x^\perp \) and decreasing on \( [0, 1/2] \).

Let \( \mathcal{C} \) be a binary code of length \( n \) such that \( d_{\text{min}}(\mathcal{C}) = \delta n \) where \( \delta \in (0, 1/2) \) and \( t \) be defined as in Proposition B.1.

Let \( \omega \triangleq w/n \), \( \lambda \triangleq t/n \) and \( \delta^\perp \triangleq 1/2 - \sqrt{\delta(1-\delta)} \). Then by Proposition B.1 we have (see Equation (27))
\[ \log_2 N_1(\mathcal{C}) \leq h(\omega) + h(\delta^\perp) - 2\log_2 |K_w(\ell)| + o(1). \]
(28)

Case 1: \( \lambda \in [\delta, 1-\delta] \):

It is optimal to choose in this case \( w \) such that \( \omega = \lambda^\perp - \varepsilon \) where \( \varepsilon > 0 \) and \( \varepsilon = o(1) \) as \( n \) tends to infinity. Let us first notice that \( \lambda \in [\delta, 1-\delta] \) implies that \( \lambda^\perp \leq \delta^\perp \) which together with \( \omega < \lambda^\perp \) implies that \( \omega < \delta^\perp \) which in turn is equivalent to the condition \( w < t \) for being able to apply
Proposition B.1. Moreover \( \omega < \lambda^\perp \) also implies \( \lambda < \omega^\perp \) and by using Proposition 3.8 we obtain
\[
\frac{2 \log_2|K_w(\ell)|}{n} \leq h(\omega) + 1 - h(\lambda) + o(1).
\]

Therefore
\[
\frac{\log_2 N_\ell(\ell')}{n} \leq h(\omega) + h(\delta^\perp) - h(\omega) - 1 + h(\lambda) + o(1)
= h(\delta^\perp) + h(\lambda) - 1 + o(1).
\]

Case 2: \( \lambda \in (1-\delta,1] \):
In that case, let \( \omega = \delta^\perp - \varepsilon \) with \( \varepsilon > 0 \) and \( \varepsilon = o(1) \) as \( n \) tends to infinity. Here we can write
\[
\frac{2 \log_2 |K_w(\ell)|}{n} = \frac{\log_2(K_w(\ell)^2)}{n} - \frac{\log_2(K_w(n-\ell)^2)}{n}.
\]
Since \( \lambda > 1-\delta \), we have \( 1-\lambda < \delta \). On the other hand, \( \omega < \delta^\perp \) implies \( \delta < \omega^\perp \). We deduce from these two inequalities that \( 1-\lambda < \omega^\perp \). By using Proposition 3.8 again, we get
\[
\frac{\log_2(K_w(n-\ell)^2)}{n} = 2a(1-\lambda, \delta^\perp) + o(1) = 2a(\lambda, \delta^\perp) + o(1).
\]
By plugging this estimate in (28) we get
\[
\frac{\log_2 N_\ell(\ell')}{n} \leq 2 \cdot h(\delta^\perp) - 2a(\lambda, \delta^\perp).
\]
This concludes the proof.

**Appendix C**

**Proof of Theorem 3.16**

Our aim in this appendix is to prove the following theorem.

**Theorem 3.16:** Let \( \varepsilon > 0 \), \( \ell' \) be a binary linear code of length \( n \) and \( p \in (0, 1/2) \) such that \( d_{\max}(\ell') \geq \delta n \) for some \( \delta^\perp \in [0, 1] \). We have asymptotically,
\[
\frac{1}{n} \log_2 \Delta(u, f_{ber, p}^\ell) \leq \max_{\delta^\perp \leq \delta' \leq 1-\delta^\perp} \left\{ \frac{1}{2} \min \left\{ c(\delta', \tau), d(\delta', \tau) \right\} \right\} + \max_{1-\varepsilon \leq \lambda' \leq 1+\varepsilon} \left\{ \lambda \log_2 p + (1-\lambda) \log_2 (1-p) + a(\lambda, \tau) \right\}
+ O\left( \frac{1}{n} \right)
\]
where \( a(\cdot, \cdot), c(\cdot, \cdot) \) and \( d(\cdot, \cdot) \) are defined respectively in Propositions 3.8, 3.11 and 3.12.

**Sketch of Proof:** We will use the following proof strategy.

1. By Lemma A.2 we know that
\[
\Delta(u, f_{ber, p}^\ell) = \Delta(u, f_{\text{trunc}, p}^\ell) + 2^{-\Omega(n)}.
\]
This is actually a consequence of Chernoff’s bound. This argument can also be used to show that the Fourier transforms are also close to each other pointwise:
\[
\forall x \in \mathbb{F}_2^n, 2^n \left| \hat{f}_{\text{trunc}, p}(x) - \hat{f}_{\text{ber}, p}(x) \right| = 2^{-\Omega(n)}.
\]

2. Equation (30) together with Lemma 3.15 are then used to show that:
\[
\Delta(u, f_{\text{trunc}, p}^\ell) \leq 2^n \sum_{t=\delta_{\min}(\ell')}^{n-\delta_{\max}(\ell')} N_t(\ell'), f_{\text{trunc}, p}(t)^2 + 2^{-\Omega(n)}.
\]

3. We use the two previous points to upper-bound
\[
\Delta(u, f_{ber, p}^\ell) \text{ as in the equation above and conclude by using bounds of Propositions 3.11 and 3.12.}
\]

**Proof of Step 1:** As we explained above (29) is just Lemma A.2. Let us now prove that
\[
\text{Lemma C.1: We have}
\forall x \in \mathbb{F}_2^n, 2^n \left| \hat{f}_{\text{trunc}, p}(x) - \hat{f}_{\text{ber}, p}(x) \right| = 2^{-\Omega(n)}.
\]

**Proof:** Recall that \( Z = \sum_{|y|=1-\varepsilon} f_{ber, p}(y) \) where by Chernoff’s bound, we have
\[
Z = 1 - 2^{-\Omega(n)}.
\]
(32)

Notice now that
\[
f_{ber, p} = \sum_{r=0}^{n} \binom{n}{r} p^r (1-p)^{n-r} u_r
\]
and
\[
f_{\text{trunc}, p} = \frac{1}{Z} \sum_{r=(1-\varepsilon)p}^{(1+\varepsilon)p} \binom{n}{r} p^r (1-p)^{n-r} u_r.
\]

Let \( \mathcal{S} \) be defined as in the equation above and conclude by
\[
2^{-\Omega(n)} \max_r |\hat{u}_r(x)| \leq 2^{-\Omega(n)} \max_r |\hat{u}_r(x)|
\]
where in the last line we used Equation (32). Recall now that by definition of the Fourier transform for functions over \( \mathbb{F}_2^n \) we have:
\[
|u_r(x)| = \frac{1}{2^n} \sum_{y : |y|=r} (-1)^{x \cdot y} \binom{n}{r} \leq \frac{1}{2^n}.
\]
(33)

By plugging this in Equation (33) we get:
\[
\left| f_{\text{trunc}, p}(x) - f_{\text{ber}, p}(x) \right| \leq 2^{-\Omega(n)} \max_{r} \binom{n}{r} p^r (1-p)^{n-r} + 2^{-\Omega(n)} \frac{2^n}{2^n} \leq 2^{-\Omega(n)} \frac{2^n}{2^n}
\]
which concludes the proof. ■

**Proof of Step 2:** This corresponds to proving the following lemma.
\textbf{Lemma C.2:}

\[ \Delta(u, f_{\text{truncBer}, p}) \leq 2^n \sqrt{\sum_{t \in \Delta_u(\mathcal{C})} N_t(\mathcal{C}^*) |f_{\text{truncBer}, p}(t)|^2} \]

\[ + 2^{-\Omega(n)} \] (34)

\textbf{Proof:} By applying Proposition 3.3 to \( f_{\text{truncBer}, p} \) we obtain

\[ \Delta(u, f_{\text{truncBer}, p}) \leq 2^n \sqrt{\sum_{t \in \Delta_u(\mathcal{C})} N_t(\mathcal{C}^*) |f_{\text{truncBer}, p}(t)|^2} \]

where \( f_{\text{truncBer}, p}(t) \) denotes the common value of the radial function \( f_{\text{truncBer}, p} \) on vectors of Hamming weight \( t \). Recall now that \( f_{\text{Ber}, p}(x) = \frac{1}{p} \) \((1 - 2p)|x|\) and by Lemma C.1 that

\[ 2^n \left| f_{\text{truncBer}, p}(x) - f_{\text{Ber}, p}(x) \right| = 2^{-\Omega(n)}. \]

Therefore,

\[ \Delta(u, f_{\text{truncBer}, p}) \leq 2^n \sqrt{\sum_{t \in \Delta_u(\mathcal{C})} N_t(\mathcal{C}^*) |f_{\text{truncBer}, p}(t)|^2} \]

which concludes the proof.

\textbf{Proof of Step 3:} We finish the proof of Theorem 3.16 by noticing that

\[ f_{\text{truncBer}, p} = \frac{1}{Z} \sum_{\ell = 0}^{(1+\varepsilon)n} \binom{n}{\ell} p^\ell (1 - p)^{n-\ell} \hat{u}_\ell \]

where \( Z \overset{\text{def}}{=} \sum_{|y| = (1-\varepsilon)n} f_{\text{Ber}, p}(y) = 1 - 2^{-\Omega(n)} \) by Chernoff’s bound. Therefore,

\[ f_{\text{truncBer}, p} = \left( 1 + 2^{-\Omega(n)} \right) \sum_{\ell = 0}^{(1+\varepsilon)n} \binom{n}{\ell} p^\ell (1 - p)^{n-\ell} \hat{u}_\ell. \]

By plugging this in Equation (35) and using \( \hat{u}_\ell = \frac{1}{\sqrt{\ell}} \frac{K_{\ell}(\varepsilon)}{\ell} \) we obtain

\[ \Delta(u, f_{\text{truncBer}, p}) \leq \left( 1 + 2^{-\Omega(n)} \right) \times \]

\[ 2^n \sum_{t = 0}^{(1+\varepsilon)n} \binom{n}{\ell} p^\ell (1 - p)^{n-\ell} \frac{K_{\ell}(\varepsilon)}{\ell}^2 \]

\[ + 2^{-\Omega(n)} \]

\[ \text{APPENDIX D}

\textbf{PROOF OF PROPOSITION 4.6}

Our aim in this section is to prove the following proposition.

\textbf{Proposition 4.6:} Let \( \Lambda \) be a random lattice of covolume \( M \) and let \( u = u_{w_0/\Lambda} \) be the uniform distribution over its cosets. Suppose that for all \( w > w_0 \) there is a function \( f(n) \) such that

\[ \mathbb{E}_\Lambda \left( \Delta(u, u_{w_0/\Lambda}) \right) \leq f(n) \left( \frac{w_0}{w} \right)^{n/2}. \]

Let \( s_0 \overset{\text{def}}{=} \frac{w_0}{w} \sqrt{\frac{2\pi}{n}} \). Then, for all \( s > s_0 \), defining \( \eta \overset{\text{def}}{=} 1 - \frac{s_0}{s} \in (0, 1) \), we have

\[ \mathbb{E}_\Lambda \left( \Delta(u, D_s^\Lambda) \right) \leq \exp(-\frac{\eta^2}{8} n) + f(n) \left( \frac{s_0}{s} \right)^{n/4}. \]

It will be a consequence of the following lemmas. We begin with the following result decomposing the Gaussian as a convex combination of balls.

\textbf{Lemma D.1:} The Gaussian distribution in dimension \( n \) of parameter \( s \) is the following convex combination of uniform distributions over balls:

\[ D_s = \frac{1}{s} \int_0^\infty G_n(x/s) u_{w_0/\Lambda} \ dw \]

where \( G_n(x) = x^{n+1} V_n(1) 2\pi \ exp(-\pi x^2) \geq 0 \). Furthermore, we have \( \frac{1}{s} \int_0^\infty G_n(x/s) \ dw = 1 \).

\textbf{Proof:} First, let \( g_s(w) \overset{\text{def}}{=} \frac{1}{s} \ exp(-\pi \frac{w^2}{s^2}) \) (i.e. the value the probability density function \( D_s \) takes on vectors of weight \( w \) and denote \( h_s(w) = -g_s'(w) = \frac{2\pi w}{s^2} \ exp(-\pi \frac{w^2}{s^2}) \). For any \( x \in \mathbb{R}^n \), setting \( u = |x|_2 \), as \( \lim_{w \to \infty} g_s(w) = 0 \) we have

\[ D_s(x) = g_s(u) \]

\[ = \int_u^\infty h_s(w) \ dw \]

\[ = \int_0^\infty h_s(w) 1\{u \leq w\} \ dw \]

\[ = \int_0^\infty h_s(w) 1_{\beta_w}(x) \ dw. \]

Above, we denoted by \( 1\{u \leq w\} \) the function which takes value 1 on input \( w \) if \( u \leq w \), and 0 otherwise. To conclude, note that \( \frac{1}{s} \int_0^\infty G_n(x/s) \ dw = h_s(w) V_n(w) \) and recall \( u_{w_0/\Lambda} = \frac{1}{V_n(w_0)} \).

For the “furthermore” part of the lemma, we compute

\[ \frac{1}{s} \int_0^\infty G_n(x/s) \ dw \]

\[ = \frac{1}{s} \int_0^\infty (w/s)^{n+1} V_n(1) 2\pi \ exp(-\pi (w/s)^2) \ dw \]

(36)

We make the substitution \( t = \pi (w/s)^2 \), which means

\[ dw = \frac{s^2}{2\pi w} dt = \frac{s}{2\sqrt{\pi}} dt. \]
Also, we recall $V_n(1) = \frac{n^{5/2}}{\Gamma(n/2 + 1)}$. Thus,

$$
\frac{1}{s} \int_0^\infty G_n(w/s) \, dw
= \frac{1}{s} \frac{\pi^{n/2}}{\Gamma(n/2 + 1)} \int_0^\infty \left( \frac{t}{\pi} \right)^{(n+1)/2} 2\pi \, e^{-t} \, \frac{s}{2\sqrt{1-t}} \, dt
= \frac{1}{\Gamma(n/2 + 1)} \int_0^{\infty} \frac{t^n}{s} \, e^{-t} \, dt = \frac{\Gamma(n/2 + 1)}{\Gamma(n/2 + 1)} = 1
$$

which concludes the proof.

We now quote the following bound, which makes precise the intuition that it is exponentially unlikely that a random Gaussian vector has norm $(1 - \eta)$ factor smaller than its expected norm. This result provides the analogy for the Chernoff bound that we used for the code-case.

Lemma D.2 ([35, Example 2.5]): Let $X$ be a random Gaussian vector of dimension $n$ and parameter 1. Let $0 < \eta < 1$. Then

$$
P \left( |X|^2 \leq (1 - \eta) \frac{n}{2\pi} \right) \leq \exp(-\frac{\eta^2}{8} n).
$$

This lemma allows us to prove the following lemma bounding $\frac{1}{s} \int_0^\infty G_n(w/s) \, dw$ when $w < s \sqrt{n}/(2\pi)$.

Lemma D.3: Let $\eta \in (0, 1)$ and $w = \sqrt{1 - \eta} \frac{s}{\sqrt{n}(2\pi)}$. Then

$$
\frac{1}{s} \int_0^\infty G_n(w/s) \, dw \leq \exp(-\frac{\eta^2}{8} n).
$$

Proof: Let $w \overset{\text{def}}{=} \sqrt{1 - \eta} \frac{s}{\sqrt{n}(2\pi)}$. By Lemma D.2, if $X$ denotes a random Gaussian vector of dimension $n$ and parameter 1, we have

$$
\int_0^{\infty} \exp(-\pi |x|^2) \, dx = P \left( |X|^2 \leq (1 - \eta) \frac{n}{2\pi} \right) \leq \exp(-\frac{\eta^2}{8} n).
$$

To compute this last integral, note that

$$
\int_0^{\infty} \exp(-\pi |x|^2) \, dx = \int_0^{\infty} \int_{|u| \leq 1} e^{-\pi u^2} \, dA \, du
$$

where $u \overset{\text{def}}{=} u^{n-1}$ denotes the Euclidean sphere of radius $u$ and $dA$ is the area element. If $A_{n-1}(u)$ denotes the surface area of $u^{n-1}$, then $A_{n-1}(u) = u^{n-1} A_{n-1}(1)$ and thus

$$
\int_0^{\infty} \int_{|u| \leq 1} e^{-\pi u^2} \, dA \, du = A_{n-1}(1) \int_0^{\infty} u^{n-1} \exp(-\pi u^2) \, du
$$

Further, it is known that $A_{n-1}(1) = \frac{2\pi^{n/2}}{\Gamma(n/2)}$. Therefore, plugging Equations (38) and (39) into (37) leads to

$$
\int_0^{\infty} u^{n-1} \exp(-\pi u^2) \, du \leq \frac{1}{A_{n-1}(1)} \exp(-\frac{\eta^2}{8} n).
$$

Now, we look at the left-hand side of the inequality we wish to prove. We begin by making the substitution $u = w/s$. So then $dw = s \, du$. Moreover, let $w \overset{\text{def}}{=} \sqrt{1 - \eta} \frac{s}{\sqrt{n}(2\pi)}$

and note that when $w = \bar{w}$ we have $u = \bar{w}/s = \sqrt{1 - \eta} \frac{s}{\sqrt{n}(2\pi)} = \bar{w}$.

$$
\frac{1}{s} \int_0^{\bar{w}} G_n(w/s) \, dw = \int_0^{\bar{w}} G_n(u) \, du
= V_n(1) \frac{2\pi}{A_{n-1}(1)} \int_0^{\bar{w}} u^{n-1} \exp(-\pi u^2) \, du
\leq V_n(1) \frac{2\pi \bar{w}^2}{A_{n-1}(1)} \int_0^{\bar{w}} u^{n-1} \exp(-\pi u^2) \, du.
$$

Plugging this last inequality into (40) yields

$$
\frac{1}{s} \int_0^{\bar{w}} G_n(w/s) \, dw \leq \frac{V_n(1) 2\pi \bar{w}^2}{A_{n-1}(1)} \exp(-\frac{\eta^2}{8} n).
$$

To conclude the proof, note that $V_n(1) = \int_0^1 A_{n-1}(u) \, du = \int_0^1 u^{n-1} A_{n-1}(1) \, du = \frac{A_{n-1}(1)}{n}$ and therefore

$$
V_n(1) \frac{2\pi \bar{w}^2}{A_{n-1}(1)} \exp(-\frac{\eta^2}{8} n) = \frac{2\pi (1 - \eta)n}{2\pi n} = 1 - \eta \leq 1.
$$

This completes the proof.

We are now ready to prove Proposition 4.6.

Proof of Proposition 4.6: By Lemma D.1, $D_s$ is a convex combination of uniform distribution over balls, namely $D_s = \frac{1}{s} \int_0^{\infty} G_n(w/s) \, dw$. Therefore (we use here the analogue of Lemma A.3 in the context of the statistical distance between two probability density functions)

$$
E_A \left( \Delta(u, D_s^A) \right) \leq \frac{1}{s} \int_0^{\infty} G_n(w/s) \, E_A \left( \Delta(u, w_{u_{w_{u^A}}}) \right) \, dw.
$$

We split the integral in two parts at radius $\bar{w} = \sqrt{1 - \eta} \frac{s}{\sqrt{n}(2\pi)}$. For the first part $w \leq \bar{w}$, we use the trivial bound $E_A \left( \Delta(u, w_{u_{w_{u^A}}}) \right) \leq 1$ which gives

$$
\frac{1}{s} \int_0^{\bar{w}} G_n(w/s) \, E_A \left( \Delta(u, w_{u_{w_{u^A}}}) \right) \, dw \leq \frac{1}{s} \int_0^{\bar{w}} G_n(w/s) \, dw.
$$

We then apply Lemma D.3, which bounds this part by $\exp(-\frac{\eta^2}{8} n)$.

For the second part $w \geq \bar{w}$, we use the trivial bound

$$
\frac{1}{s} \int_{\bar{w}}^{\infty} G_n(w/s) \, dw \leq 1 \text{ and, noting }
$$

$$
w \geq \bar{w} = \sqrt{1 - \eta} \frac{s}{\sqrt{n}(2\pi)}
\geq \frac{1}{\sqrt{1 - \eta}} \frac{s_0}{\sqrt{n}(2\pi)} = \frac{1}{\sqrt{1 - \eta}} \frac{s_0}{\sqrt{n}(2\pi)} = w_0,
$$

we may apply the assumption of the proposition, yielding

$$
E_A \left( \Delta(u, w_{u_{w_{u^A}}}) \right) \leq f(n) \left( \frac{w_0}{w} \right)^{n/2}
\leq f(n) \left( \frac{w_0}{w} \right)^{1/2}
= f(n) \left( \sqrt{1 - \eta} \right)^{n/2} = f(n) \left( \frac{s_0}{s} \right)^{n/4}.
$$

Adding these bounds yields the proposition.

Acknowledgment

The authors would like to thank Josif Pinelis for help with the proof of Proposition 4.6.
REFERENCES

[1] A. Ashikhmin, A. Barg, and S. Litsyn, “Estimates of the distance distribution of codes and designs,” Electron. Notes Discrete Math., vol. 6, pp. 4–14, Apr. 2001.

[2] A. E. Ashikhmin, G. D. Cohen, M. Krivelevich, and S. N. Litsyn, “Bounds on distance distributions in codes of known size,” IEEE Trans. Inf. Theory, vol. 51, no. 1, pp. 250–258, Jan. 2005.

[3] D. Aggarwal, D. Dadush, O. Regev, and N. Stephens-Davidowitz, “Solving the shortest vector problem in $2^{n}$th time using discrete Gaussian sampling,” in Proc. 47th Annu. ACM Symp. Theory Comput., Jun. 2015, pp. 733–742.

[4] M. Alekhnovich, “More on average case vs approximation complexity,” J. Comput. Complexity, vol. 20, no. 4, pp. 755–786, Dec. 2011.

[5] W. Banaszczyk, “New bounds in some transference theorems in the geometry of numbers,” Mathematische Annalen, vol. 296, no. 1, pp. 625–635, Dec. 1993.

[6] L. Bassalygo, “New upper bounds for codes correcting errors,” Problemy Peredachi Informatsii, vol. 1, pp. 41–44, Jan. 1965.

[7] M. Blomer, A. Couvreur, and T. Debris-Alazard, “On the pseudorandomness of the decoding problem via the Oracle comparison problem,” J. Comput. Syst. Sci., vol. 92, pp. 20–37, Dec. 2018.

[8] Z. Brakerski, V. Lyubashevsky, V. Vaikuntanathan, and D. Wichs, “Worst-case hardness for LPN and cryptographic hashing via code smoothing,” in Proc. Ann. Int. Conf. Theory Appl. Cryptograph. Techn. Cham, Switzerland: Springer, 2019, pp. 619–635.

[9] H. Cohn and N. Elkies, “New upper bounds on sphere packings I,” Ann. Math., vol. 157, no. 2, pp. 689–714, Mar. 2003.

[10] R. K. Chung, Spectral Graph Theory, vol. 92. Providence, RI, USA: Amer. Math. Soc., 1997.

[11] T. Debris-Alazard and N. Resch, “Worst and average case hardness of decoding via smoothing bounds,” Cryptol. ePrint Arch., 2022.

[12] T. Debris-Alazard, N. Sendrier, and J.-P. Tillich, “Wave: A new family of trapdoor one-way preimage sampleable functions based on codes,” in 25th Int. Conf. Theory Appl. Cryptol. Inf. Secur. (ASIACRYPT), S. D. Galbraith and S. Moriai, Eds., Kobe, Japan, Dec. 2019.

[13] T. Debris-Alazard and J.-F. Tillich, “Statistical decoding,” 2017, arXiv:1701.07416.

[14] C. Gentry, C. Peikert, and V. Vaikuntanathan, “Trapdoors for hard lattices and new cryptographic constructions,” in Proc. 14th Annu. ACM Symp. Theory Comput., May 2008, pp. 197–206.

[15] M. E. H. Ismail and P. Simeonov, “Strong asymptotics for Kravtchouk polynomials,” J. Comput. Appl. Math., vol. 100, no. 2, pp. 121–144, Dec. 1998.

[16] G. Kabatiansky and I. V. Levenshtein, “Bounds for packings on a sphere and in space,” Problemy Inf. Transmiss., vol. 14, no. 1, pp. 1–17, 1978.

[17] T. Klove, Codes for Error Detection (Series on Coding Theory and Cryptology). Singapore: World Scientific, 2007.

[18] I. Krasikov, “Uniform bounds for Bessel functions,” J. Appl. Anal., vol. 12, no. 1, pp. 83–91, Jan. 2006.

[19] I. V. Levenshtein, “On bounds for packings in $n$-dimensional Euclidean space,” Doklady Akademii Nauk SSSR, vol. 245, pp. 1299–1303, Jan. 1979.

[20] V. I. Levenshtein, “Kravtchouk polynomials and universal bounds for codes and designs in Hamming spaces,” IEEE Trans. Inf. Theory, vol. 41, no. 5, pp. 1303–1321, Sep. 1995.

[21] L. Luzzi, C. Ling, and R. M. Bloch, “Secret key generation from Gaussian wiretap channels,” IEEE Trans. Inf. Theory, vol. 60, no. 10, pp. 6399–6416, Oct. 2014.

[22] D. A. Levin and Y. Peres, Markov Chains and Mixing Times. Providence, RI, USA: American Mathematical Soc., 2017.

[23] H. Mirghasemi and J.-C. Belfiore, “The semantic secrecy rate of the lattice Gaussian coding for the Gaussian wiretap channel,” in Proc. IEEE Inf. Theory Workshop, Nov. 2014, pp. 112–116.

[24] J. McEliece, “A public-key system based on algebraic coding theory,” Jet Propulsion Lab., Bengkulu, India, DSN Prog. Rep. 44, 1978, pp. 114–116.

[25] D. Micciancio and O. Regev, “Worst-case to average-case reductions based on Gaussian measures,” SIAM J. Comput., vol. 37, no. 1, pp. 267–302, 2007.

[26] R. McEliece, E. Rodemich, H. Rumsey, and L. Welch, “New upper bounds on the rate of a code via the Delsarte–MacWilliams inequalities,” IEEE Trans. Inf. Theory, vol. IT-23, no. 2, pp. 157–166, Mar. 1977.

[27] R. Misoczki, J.-P. Tillich, N. Sendrier, and P. S. L. M. Barreto, “MDPC-McEliece: New McEliece variants from moderate density parity-check codes,” in Proc. IEEE Int. Symp. Inf. Theory, Jul. 2013, pp. 2069–2073.

[28] P. Q. Nguyen and O. Regev, “Learning a parallelepiped: Cryptanalysis of GGH and NTRU signatures,” in Proc. 24th Annu. Int. Conf. Theory Appl. Cryptograph. Techn. Cham, Switzerland: Springer, 2006, pp. 271–288.

[29] X. Pujol and J. Stehle, “Solving the shortest lattice vector problem in time $2^{0.265n}$,” Cryptol. ePrint Arch., vol. 2009, p. 605, Jan. 2009.

[30] L. Saloff-Coste, “Random walks on finite groups,” in Probability on Discrete Structures, H. Kesten, Ed. Berlin, Germany: Springer, 2004, pp. 263–346.

[31] J. H. V. Lint, “Introduction to coding theory,” in Graduate Texts in Mathematics, 3rd ed. Cham, Switzerland: Springer, 1999.

[32] M. Wainwright, High-Dimensional Statistics: A Nonasympotic Approach. Cambridge, UK: Cambridge Univ. Press, 2019.

[33] D. B. Wilson, “Random walks on $Z_{d}^{n}$,” Probab. Theory Rel. Fields, vol. 110, no. 4, pp. 441–457, 1997.

[34] Y. Yu and J. Zhang, “Smoothing out binary linear codes and worst-case sub-exponential hardness for LPN,” in Proc. 41st Annu. Int. Conf. Cryptol., in Lecture Notes in Computer Science, vol. 12827. Cham, Switzerland: Springer, 2021, pp. 473–501.

[35] Z. Wang, Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation and Multiuser Information Theory. Cambridge, UK: Cambridge Univ. Press, 2014.

Thomas Debris-Alazard received the Ph.D. degree in computer science from the INRIA, Sorbonne Université, Paris, France, in 2019. He was with the INRIA as a Junior Researcher (chargé de recherche) in 2020. His research interests include code and lattice-based cryptography.

Léo Duc received the Ph.D. degree in lattice-based cryptography from ENS Paris. He was a Post-Doctoral Researcher with UCSD. Then, he joined Centrum Wiskunde & Informatica (CWI) in 2015. He received the VENI Grant in 2016. He held a tenure-track position at CWI in 2017. He has been a Professor with Leiden University since 2021. His research interests include the algorithmic aspects of lattices, in particular cryptography and cryptanalysis.

Nicolas Resch received the B.Sc. degree from McGill University in 2015 and the Ph.D. degree from Carnegie Mellon University in 2020. During his Ph.D. degree, he supported by the NSERC PGS-D Fellowship. He was a Post-Doctoral Researcher with the Cryptology Group, Centrum Wiskunde & Informatica (CWI), from 2020 to 2022. He is currently an Assistant Professor with the Theoretical Computer Science Group, Informatics Institute, University of Amsterdam. His research interests include coding theory, pseudorandomness, and crypotology, with a particular emphasis on questions that bridge these fields.

Jean-Pierre Tillich received the Ph.D. degree from École Nationale Supérieure des Télécommunications (ENST), Paris, in 1994. From 1997 to 2003, he was an Assistant Professor with University Paris XI. He is currently a Senior Researcher with the Institut de Recherche en Informatique et Automatique (INRIA), Paris, France. His research interests include classical and quantum coding theory, cryptography, and graph theory. From 2009 to 2012, he was an Associate Editor for Coding Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.