Lossy Codes and a New Variant of the Learning-With-Errors Problem

Nico Döttling and Jörn Müller-Quade

Karlsruhe Institute of Technology
{doettling,mueller-quade}@kit.edu

Abstract

The hardness of the Learning-with-errors (LWE) Problem has become one of the most useful assumptions in cryptography. Many applications, like public key cryptography, hierarchical identity-based encryption, and fully homomorphic encryption can be based on this assumption. Furthermore, the LWE Problem exhibits a worst-to-average-case reduction making the LWE assumption very plausible.

This worst-to-average-case reduction is based on a Fourier argument and therefore the error distribution for current applications of LWE must be chosen with a gaussian distribution. Sampling from a gaussian distribution is cumbersome and it is hence highly desirable to prove worst-to-average-case reductions for other distributions, in particular for the uniform distribution which can be sampled very efficiently.

In this work we present the first worst-to-average case reduction for an LWE problem with non-gaussian noise, namely for uniformly distributed errors. This new worst-to-average-case connection comes with a slight drawback and we need to use a bounded variant of the LWE problem, where the number of samples to be learnt from is fixed in advance. To the best of our knowledge all current applications of LWE can be based on the bounded variant.

The proof is based on a new tool called *lossy codes*. These codes are indistinguishable from good error correcting codes, but provably lose information when noisy codewords are decoded. The concept of lossy codes might be of interest on its own, because it could in future be used to transform other lattice/coding-based hardness assumptions.

Keywords: Learning-With-Errors, Worst-Case Reduction, Uniform Interval Error-Distribution

1 Introduction

The Learning-with-Errors (LWE) Problem asks to recover an unknown vector $\mathbf{x} \in \mathbb{Z}_q^n$, given a random matrix $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ and a *noisy-codeword* $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\mathbf{e} \in \mathbb{Z}_q^m$ is chosen from an error-distribution χ^m. This problem has had a significant impact in cryptography since its conception in 2005 [Reg05]. Maybe the most intriguing feature of this problem is its worst-to-average case connection [Reg05, Pei09]. This basically allows to transform an efficient adversary solving LWE on average, into an efficient (quantum) algorithm solving lattice problems in the worst case. Beyond this very strong hardness-guarantee, the problem has unmatched cryptographic versatility.

It allows for IND-CPA and IND-CCA secure encryption [Reg05, GPV08, Pei09], lossy-trapdoor functions [PW08], (hierarchical) identity-based encryption [CHKP10, ABB10], fully homomorphic encryption [BV11, BGV12, Bra12] and many more. The worst-to-average-case reductions [Reg05, Pei09] crucially rely on the Fourier-properties of gaussian error-distributions. This has the consequence that the cryptographic applications also need to use a gaussian error-distribution. For the above-mentioned encryption-schemes, sampling from a gaussian error-distribution is usually the computationally heaviest step (which occurs mostly during key-generation). It would thus be desirable to have a variant of the LWE problem enjoying the same worst-to-average-case connection, but that comes with an easier-to-sample error-distribution. Micciancio and Mol [MM11a] write:

> "Can lattice-based hardness results for search LWE be extended to noise distributions other than Gaussian? Can we show similar lattice-based hardness results if the noise is distributed uniformly at random modulo 2^i? The latter case is very attractive from a practical viewpoint since arithmetic modulo 2 and sampling from uniform distributions can be implemented very efficiently."

1.1 Our Contribution

In this work we present the first instantiation of the LWE problem with worst-to-average case connection where the error-distribution is the uniform distribution on a small interval $[-r, r]$ (call this distribution $\mathcal{U}([-r, r])$). In particular, setting $r = 2^i$, this answers the question of [MM11a]. Rather than proving a new worst-to-average case reduction, we will build ours on top of existing ones. More precisely, the gaussian error-distributions will appear in the hardness-reduction, but not in the LWE instantiation itself. Our main-lever to achieve this is a technique which we call *lossy codes*. Roughly speaking, lossy codes are pseudorandom codes that seem to be good codes. However, encoding messages with a lossy code and adding certain errors *provably* annihilates the message (on average). On the other hand, encoding the same message using a truly random code and adding the same type of error preserves the message, i.e. the message can be recovered *information theoretically* (yet not efficiently). Using a proof-strategy pioneered by Peikert and Waters [PW08], we conclude that recovering the message when encoding with a random code and adding noise must be computationally hard. If this was not the case, lossy codes could be efficiently distinguished from random codes, contradicting the pseudorandomness-property of lossy codes. The main-part of this work is devoted to proving that a quite simple construction of lossy codes for LWE *actually is lossy* for the error-distribution $\mathcal{U}([-r, r])$. The key-insight for this construction is that the standard LWE problem with gaussian error-distribution allows us to implant many very short vectors into a random looking lattice. More precisely, the lattice generated by $(A\|E)$, where A is random and E is chosen from a gaussian distribution contains at least $n/2$ short vectors (the columns of E). However, a simple basis change yields a new generator $(A\|AT + E)$ for the same lattice. This new lattice is however, by the standard LWE assumption, pseudorandom. We then proceed to show that encoding a random message with $(A\|E)$ and adding errors from $\mathcal{U}([-r, r])$ actually destroys information in the part of the message encoded with E. For this, we need to ensure that the lengths of the columns of E (e.g. in the $\|\cdot\|_2$-norm) are significantly smaller than r. As a consequence of this, we get a drawback of the lossy-codes-technique. Our worst-to-average case connection-factor for LWE with error-distribution $\mathcal{U}([-r, r])$ depends on the number of samples provided by LWE (while those for standard LWE [Reg05, Pei09] do not). We will therefore consider an m-bounded LWE problem $\mathsf{LWE}(n, m, q, \mathcal{U}([-r, r]))$, where the number of samples m has a fixed

$\mathsf{poly}(n)$ upper bound (rather than being arbitrary $\mathsf{poly}(n)$ depending on the adversary, like in the standard LWE problem). As lossy codes are basically an information-theoretical technique, this seems unavoidable. However, this drawback is still quite mild compared to the super-polynomial inapproximability assumptions made in other works [GKPV10, BV11, Bra12]. We now state our main-theorem.

Theorem (Main Theorem)**.** *Let n be a security parameter. Let $q = q(n)$ and $m = m(n) = \mathsf{poly}(n)$ be integers. Let $\rho = \rho(n) \in (0, 1)$ be such that $\rho q \geq 2n^2 m$. Assume there exists a PPT-algorithm that solves $\mathsf{LWE}(n, m, q, \mathcal{U}([-\rho q, \rho q]))$ with non-negligible probability. Then there exists an efficient quantum-algorithm that approximates the decision-version of the shortest vector problem (GAPSVP) and the shortest independent vectors problem (SIVP) to within $\tilde{O}(n^{5/2} m/\rho)$ in the worst case.*

Applying the search-to-decision reduction of [MM11b], we can conclude as a corollary that the decisional variant $\mathsf{DLWE}(n, m, q, \mathcal{U}([-\rho q, \rho q]))$ is also hard. Finally, we believe that the notion of lossy codes might also be useful to transform other lattice/coding-based hardness-assumptions.

1.2 Related Work

Recently, there has been a growing interest to instantiate new LWE variants. In [GKPV10] an LWE variant was introduced where the secret $\mathbf{x}$ is chosen by a distribution with a sufficient amount of min-entropy, rather than uniformly at random. Lyubashevsky et. al [LPR10] introduced the Ring-LWE problem and provided a worst-to-average-case reduction from the GAPSVP problem in ideal lattices to Ring-LWE. In [BPR12] a LWE-variant called Learning-With-Rounding (LWR) was introduced. LWR-samples are of the form $(\mathbf{a}, \lfloor \langle \mathbf{a}, \mathbf{x} \rangle \cdot p/q \rceil)$ (for two moduli p and q). Remarkably, the problem inherits the worst-to-average case connection from the corresponding standard LWE problem modulo q, without making use of a gaussian error-distribution by itself. It is, however, not known whether LWR is sufficient to achieve public-key cryptography. Finally, Pietrzak [Pie12] gave an adaptively secure instantiation of LWE called Subspace-LWE, where the adversary is allowed to learn inner products of the secret $\mathbf{x}$ after it has been projected on an adversarially chosen subspace.

2 Preliminaries

We will use the notation $(\mathbf{A} \| \mathbf{B})$ for the horizontal concatenation of two matrices $\mathbf{A}$ and $\mathbf{B}$ and $(\mathbf{x}, \mathbf{y})$ for the vertical concatenation of two vectors $\mathbf{x}$ and $\mathbf{y}$. Let $\mathsf{sgn}(x)$ be the signum-function, i.e. $\mathsf{sgn}(x) = 1$ if $x > 0$, $\mathsf{sgn}(x) = -1$ if $x < 0$ and $\mathsf{sgn}(x) = 0$ if $x = 0$. We denote computational indistinguishability of two distributions $\mathcal{X}$ and $\mathcal{Y}$ by $\mathcal{X} \approx_c \mathcal{Y}$.

2.1 Norms

We will use the $\| \cdot \|_2$- and the $\| \cdot \|_\infty$-norm in this work. The $\| \cdot \|_2$-norm on $\mathbb{R}^n$ is defined by $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$, the $\| \cdot \|_\infty$-norm on $\mathbb{R}^n$ is defined by $\|\mathbf{x}\|_\infty = \max_{i=1,\ldots,n} |x_i|$. Norms $\| \cdot \|$ are multiplicative and obey the triangle-inequality, i.e. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ it holds that $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ and $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. The set $C = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x}\|_\infty \leq r\}$ forms a hypercube of dimension n, i.e. $C = [-r, r]^n$.

2.2 Min-Entropy

Let χ be a probability distribution with finite support and let X be distributed according to χ. The *min-entropy* $H_\infty(X)$ is defined as $H_\infty(X) = -\log(\max_\xi(\Pr[X = \xi]))$. Let Y be random-variable (possibly correlated with X) and let $\tilde{y}$ be a measurement or outcome of Y. The *conditional min-entropy* $H_\infty(X|Y = \tilde{y})$ is defined as $H_\infty(X|Y = \tilde{y}) = -\log(\max_\xi(\Pr[X = \xi|Y = \tilde{y}]))$. Instead of using the *conditional average min-entropy* [DORS08], we will directly derive laws of the form $\Pr_{\tilde{y}}[H_\infty(X|Y = \tilde{y}) \geq \delta] \geq 1 - \epsilon$, i.e. $H_\infty(X|Y = \tilde{y})$ is at least δ, except with probability ϵ over the choice of the measurement $\tilde{y}$. This will enable a more fine-grained analysis of the lossiness of our constructions (the average conditional min-entropy $\tilde{H}_\infty(X|Y)$ would be lower-bounded by $\tilde{H}_\infty(X|Y) \geq -\log(2^{-\delta} + \epsilon)$, i.e. it compresses δ and ϵ into one scalar).

2.3 Binomial Distributions

Let $X_i \in \{0, 1\}$ for $i = 1, \ldots, n$ be iid. random variables with $\Pr[X_i = 1] = p$. Then $X = \sum_{i=1}^{n} X_i$ is *binomially* distributed with $\Pr[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$. The binomial-distribution assumes its maximum at an index $k_{max} = \lfloor p(n+1) \rfloor$. The tails of a binomial-distribution can be bounded by the Chernoff-bound, which states that $\Pr[X \leq (1-\delta)\mathsf{E}[X]] \leq e^{-\delta^2 \mathsf{E}[X]/2}$, where the expectation is $\mathsf{E}[X] = p \cdot n$

2.4 Gaussian Distributions

We denote Φ_s the normal-distribution with variance $s^2/(2\pi)$, i.e. if X is distributed according to Φ_s, then X has the probability-density function $p_X(x) = e^{-\pi x^2/s^2}/s$. A standard tail-bound for Gaussians is $\Pr[|X| > t \cdot s] < e^{-\pi t^2}$. Following [Reg05], we denote by $\bar{\Psi}_\alpha$ the discretized gaussian distribution over $\mathbb{Z}$ (or $\mathbb{Z}_q$) with variance $(\alpha q)^2/(2\pi)$. More precisely, $\bar{\Psi}_\alpha$ is sampled by taking a sample from $\Phi_{\alpha q}$ and rounding it to the nearest integer. Let Y be distributed according to $\bar{\Psi}_\alpha$, i.e. let $Y = \lceil X \rfloor$ with X distributed by $\Phi_{\alpha q}$. Here, we always assume that q is implicitly given if the distribution is defined over $\mathbb{Z}$. If $t\alpha q \geq 2$, we can derive the tail-bound $\Pr[|Y| > t\alpha q] \leq \Pr[|X| > t\alpha q - 1] \Pr[|X| > t\alpha q/2] \leq e^{-\pi/4 \cdot t^2}$. We will need the following lemma which states that matrices $\mathbf{E}$ chosen from $\bar{\Psi}_\alpha^{m \times n}$ (where $m \geq 2n$) have, with overwhelming probability full rank.

Lemma 1. *Let $\mathbf{E}$ be chosen from $\bar{\Psi}_\alpha^{m \times n}$, where $m \geq 2n$ and $\alpha q \geq 1/\sqrt{\pi}$. Then $\mathbf{E}$ has full rank, except with probability $n \cdot e^{(1/2 - \ln(2))\frac{n}{2}}$ (which is negligible in n).*

The proof of Lemma 1 can be found in Appendix A.

2.5 Lattices

Let $\mathbf{B} \in \mathbb{Z}^{m \times n}$ be a full rank-matrix. The *lattice* $\Lambda(\mathbf{B})$ is defined as $\Lambda(\mathbf{B}) = \{\mathbf{Bx} \in \mathbb{Z}^m : \mathbf{x} \in \mathbb{Z}^n\}$, i.e. the lattice $\Lambda(\mathbf{B})$ is the set of all integer-linear combination of columns of $\mathbf{B}$. Let $q \geq 2$ be an integer. The q-ary lattice $\Lambda_q(\mathbf{B})$ is defined as $\Lambda_q(\mathbf{B}) = \{\mathbf{y} \in \mathbb{Z}^m : \exists \mathbf{x} \in \mathbb{Z}^n : \mathbf{y} \equiv \mathbf{Bx} \mod q\}$. Observe that the lattice $\Lambda_q(\mathbf{B})$ contains $\mathbb{Z}^m$ as a sublattice, therefore $\Lambda_q(\mathbf{B})$ is always full-rank. Moreover, it holds that $\Lambda(\mathbf{B}) \subseteq \Lambda_q(\mathbf{B})$, as $\mathbf{x} \in \Lambda_q(\mathbf{B})$, for each $\mathbf{x} \in \mathsf{columns}(\mathbf{B})$.

We will generally assume that elements of $\mathbb{Z}_q$ are given in the central residue-class representation, i.e. if $x' \in \mathbb{Z}_q$, we will identify $x' = x \mod q$ with an integer x in $\{-\lfloor q/2 \rfloor, \ldots, \lceil q/2 \rceil - 1\}$. We can thus generically lift x' from $\mathbb{Z}_q$ to $\mathbb{Z}$. Moreover, with this we can define a meaningful "norm"

on $\mathbb{Z}_q$ by $\|\mathbf{x} \mod q\|_\infty = \|\mathbf{x}\|_\infty$. In one of our proofs, we will count the number of $\mathbf{x} \in \mathbb{Z}_q^m$ with $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_\infty \leq \rho q$, where $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$, $\mathbf{y} \in \mathbb{Z}_q^m$. A simple calculation shows that if $\rho < 1/2$, this corresponds exactly to the number of $\mathbf{z} \in \Lambda_q(\mathbf{A})$ with $\|\mathbf{y} - \mathbf{z}\|_\infty \leq \rho q$ (where $\mathbf{y}$ is lifted to $\mathbb{Z}^m$). Thus, in this case we can do the *counting* in $\mathbb{Z}^m$.

Let $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n) \in \mathbb{Z}^{m \times n}$ have full rank and let $t > 0$ be an integer. Let S be the convex hull of $t\mathbf{a}_1, \ldots, t\mathbf{a}_n$, i.e. $S = \{\sum_{i=1}^n \alpha_i t\mathbf{a}_i | \sum_{i=1}^n \alpha_i = 1 \text{ and } \forall i : \alpha_i \geq 0\}$. We want to bound the number of points in $S \cap \Lambda(\mathbf{A})$. Let $\mathbf{y} \in S \cap \Lambda(\mathbf{A})$. Then it holds that $\mathbf{y} = \sum_{i=1}^n \alpha_i t\mathbf{a}_i$, $\sum_{i=1}^n \alpha_i = 1$ and for all i $\alpha_i \geq 0$. Since $y \in \Lambda(\mathbf{A})$ and $\mathbf{A}$ has full rank, it must hold that for all i that $\alpha_i t \in \mathbb{Z}$. Thus let $\beta_i = \alpha_i t \in \mathbb{Z}$. The above implies for all i $\beta_i \geq 0$ and $\sum_{i=1}^n \beta_i = t$. Thus, to count the number of points in $S \cap \Lambda(\mathbf{A})$ it suffices to count the number of tuples $(\beta_1, \ldots, \beta_n) \in \mathbb{Z}_{\geq 0}^n$ with $\sum_{i=1}^n \beta_i = t$. By a simple combinatorial argument (see for instance [vLW01] Section 13, Theorem 13.1), this is exactly $\binom{n+t}{t}$. We conclude the following lemma.

Lemma 2. *Let $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n) \in \mathbb{Z}^{m \times n}$ have full rank and let $t > 0$ be an integer. Let S be the convex hull of $t\mathbf{a}_1, \ldots, t\mathbf{a}_n$. Then $|S \cap \Lambda(\mathbf{A})| = \binom{n+t}{n}$.*

2.6 Learning-With-Errors

As mentioned above, we will consider an m-bounded LWE-problem, where the adversary is given $m(n) = \mathsf{poly}(n)$ samples (which we can write conveniently in matrix-form).

Problem 1. **m-bounded LWE Search-Problem, Average-Case Version.** Let n be a security parameter, let $m = m(n) = \mathsf{poly}(n)$ and $q = q(n) \geq 2$ be integers and χ be a distribution on $\mathbb{Z}_q$. Let $\mathbf{x} \in \mathbb{Z}_q^n$ be chosen uniformly at random, let $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ be chosen uniformly at random an let $\mathbf{e}$ be chosen according to χ^m. The goal of the $\mathsf{LWE}(n, m, q, \chi)$ problem is, given $(\mathbf{A}, \mathbf{A}\mathbf{x} + \mathbf{e})$, to find $\mathbf{x}$.

We remark that virtually all cryptographic applications of the LWE problem (at least to which the authors are familiar with) require only an a-priori fixed number of samples. For those applications, the formulation of Problem 1 poses no restriction. Regev [Reg05] and Peikert [Pei09] established worst-to-average-case connections between worst-case lattice problems and Problem 1 for suitable parameter-choices. For our construction, we will rely on the theorem of Regev [Reg05].

Theorem 1 (Worst-to-Average Case Reduction [Reg05])**.** *Let n be a security parameter and $q = q(n)$ be an integer, let $\alpha = \alpha(n) \in (0, 1)$ be such that $\alpha q > 2\sqrt{n}$. If there exists a PPT-algorithm solving $\mathsf{LWE}(n, m, q, \bar{\Psi}_\alpha)$ with non-negligible probability, then there exists an efficient quantum-algorithm that approximates the decision-version of the shortest vector problem (GAPSVP) and the shortest indpendent vectors problem (SIVP) to within $\tilde{O}(n/\alpha)$ in the worst case.*

The LWE distinguishing-problem DLWE asks to distinguish the distribution of Problem 1 from uniform random. Thus, the hardness of the DLWE problem states that the LWE-distribution is pseudorandom.

Problem 2. **m-bounded LWE Distinguishing-Problem** Let n be a security parameter, let $m = m(n) = \mathsf{poly}(n)$ and $q = q(n) \geq 2$ be integers and χ be a distribution on $\mathbb{Z}_q$. Let $\mathbf{x} \in \mathbb{Z}_q^n$ be chosen uniformly at random, let $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ be chosen uniformly at random an let $\mathbf{e}$ be chosen according to χ^m. The goal of the $\mathsf{DLWE}(n, m, q, \chi)$ problem is, given $(\mathbf{A}, \mathbf{y})$, to decide whether $\mathbf{y}$ is distributed by $\mathbf{A}\mathbf{x} + \mathbf{e}$ or chosen uniformly at random from $\mathbb{Z}_q^m$.

There are several search-to-decision reductions basing the hardness of Problem 2 on the hardness of Problem 1 [Reg05, Pei09, MP12, MM11b]. The one most suitable for our instantiation is due to Micciancio and Mol [MM11a, MM11b]. Their search-to-decision reduction works for any error-distribution χ and is *sample-preserving* (i.e. the distinguisher requires the same amount of samples as the search-adversary).

Theorem 2 (Search-to-Decision [MM11b])**.** *Let $q = q(n) = \mathsf{poly}(n)$ be a prime modulus and let χ be* any *distribution over $\mathbb{Z}_q$. Assume there exists a PPT-distinguisher $\mathcal{D}$ that distinguishes $\mathsf{DLWE}(n, m, q, \chi)$ with non-negligible advantage, then there exists a PPT-adversary $\mathcal{A}$ that inverts $\mathsf{LWE}(n, m, q, \chi)$ with non-negligible success-probability.*

Finally, we need a *matrix-version* of Problem 2. The hardness of the matrix-version can be easily established using a hybrid-argument (see e.g. [ACPS09]).

Lemma 3. *Let $m(n), k(n) = \mathsf{poly}(n)$. Assume that $\mathsf{DLWE}(n, m, q, \chi)$ is pseudorandom. Then the distribution $(\mathbf{A}, \mathbf{AX} + \mathbf{E})$ is also pseudorandom, where $\mathbf{A} \in \mathbb{Z}_q^{m\times n}$ and $\mathbf{X} \in \mathbb{Z}_q^{n\times k}$ are chosen uniformly at random and $\mathbf{E}$ is chosen according to $\bar{\Psi}_\alpha^{m\times k}$.*

3 Lossy Codes

In this section, we introduce the main technical tool in this work, lossy codes, and show that the existence of lossy codes implies that the associated decoding problems for random codes are hard. To avoid ambiguities, we will use the variable c to denote error-terms.

Definition 1. Let n be a security parameter, let R_n be a finite Ring and $m(n), k(n) = \mathsf{poly}(n)$ be integers. Let $\mathcal{C}$ be a distribution on $R_n^{m\times k}$ and let χ be a distribution on R_n^m. Finally, let $\mathcal{U}$ be the uniform distribution on $R_n^{m\times k}$. We say that $\mathcal{C}$ is γ-lossy for the error-distribution χ^m, if the following 2 properties hold.

1. $\mathcal{C}$ **is pseudorandom:** It holds that $\mathcal{C} \approx_c \mathcal{U}$.
2. $\mathcal{C}$ **is lossy:** Let $\mathbf{A}, \tilde{\mathbf{A}}$ be distributed according to $\mathcal{C}$, let $\mathbf{x}, \tilde{\mathbf{x}}$ be chosen uniformly from R_n^k and let $\mathbf{c}, \tilde{\mathbf{c}}$ be chosen according to χ^m. Then it holds that $\Pr_{(\tilde{\mathbf{A}},\tilde{\mathbf{x}},\tilde{\mathbf{c}})}[H_\infty(\mathbf{x}|(\mathbf{A}, \mathbf{Ax} + \mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{A}}\tilde{\mathbf{x}} + \tilde{\mathbf{c}})) \geq \gamma] \geq 1 - \mathsf{negl}(n)$.

Our main motivation for defining lossy codes is proving that the decoding-problem of recovering $\mathbf{x}$ from $(\mathbf{A}, \mathbf{Ax}+\mathbf{c})$, where $\mathbf{A}$ and $\mathbf{x}$ are chosen uniformly and $\mathbf{c}$ is chosen from χ^m, is computationally hard, even though $\mathbf{x}$ is information-theoretically (with overwhelming probability) uniquely defined.

Theorem 3. *Let the distribution χ be efficiently samplable. Let $\gamma = \gamma(n) = \omega(\log(n))$. Let $\mathcal{C}$ be γ-lossy for the error-distribution χ. Let $\mathbf{A}$ be chosen uniformly at random from $R_n^{m\times k}$, $\mathbf{x}$ be chosen uniformly at random from R_n^k and $\mathbf{c}$ be distributed according to χ^m. Set $\mathbf{y} = \mathbf{Ax}+\mathbf{c}$. Then it holds for every PPT $\mathcal{A}$ that $\Pr[\mathcal{A}(\mathbf{A}, \mathbf{y}) = \mathbf{x}] \leq \mathsf{negl}(n)$.*

The proof will proceed in two parts. First, we will show that the uniformly random $\mathbf{A}$ can be replaced with a pseudorandom $\mathbf{A}'$ chosen from $\mathcal{C}$, while the success-probability of $\mathcal{A}$ changes at most by a negligible amount. This will be established using the pseudorandomness-property of $\mathcal{C}$. After this modification, even an unbounded $\mathcal{A}$ has only negligible success-probability in recovering x, due to the lossiness-property of $\mathcal{C}$.

Proof. Let $\mathcal{A}$ be a PPT-adversary. Consider the following two experiments.

- **Experiment 1** Choose $\mathbf{A} \in R_n^{m \times k}$ uniformly at random, choose $\mathbf{x} \in R_n^k$ uniformly at random and choose $\mathbf{c}$ from the distribution χ^m. Set $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{c}$. Input $(\mathbf{A}, \mathbf{y})$ to $\mathcal{A}$. Let $\mathbf{x}'$ be the output of $\mathcal{A}$. If $\mathbf{x} = \mathbf{x}'$ output win.
- **Experiment 2** Choose $\mathbf{A}' \in R_n^{m \times k}$ according to $\mathcal{C}$, choose $\mathbf{x} \in R_n^k$ uniformly at random and choose $\mathbf{c}$ from the distribution χ^m. Set $\mathbf{y} = \mathbf{A}'\mathbf{x} + \mathbf{c}$. Input $(\mathbf{A}', \mathbf{y})$ to $\mathcal{A}$. Let $\mathbf{x}'$ be the output of $\mathcal{A}$. If $\mathbf{x} = \mathbf{x}'$ output win.

We will first show that from $\mathcal{A}$'s view, experiment 1 and experiment 2 are indistinguishable, given that $\mathcal{C}$ is pseudorandom. For contradiction, assume there is a PPT $\mathcal{A}$ that distinguishes between experiment 1 and experiment 2 with non-negligible advantage ϵ, i.e. $\Pr[\mathsf{Exp}_1(\mathcal{A}) = \mathsf{win}] > \Pr[\mathsf{Exp}_2(\mathcal{A}) = \mathsf{win}] + \epsilon$. We construct a distinguisher $\mathcal{B}$ that distinguishes uniformly random $\mathbf{A}$ from pseudorandom $\mathbf{A}'$ chosen from $\mathcal{C}$. Let $\mathbf{A}^*$ be $\mathcal{B}$'s input. $\mathcal{B}$ chooses $\mathbf{x}$ uniformly at random from R_n^k, samples a $\mathbf{c}$ from distribution χ^m and sets $\mathbf{y} = \mathbf{A}^*\mathbf{x} + \mathbf{c}$. $\mathcal{B}$ now runs $\mathcal{A}$ on input $(\mathbf{A}^*, \mathbf{A}^*\mathbf{x} + \mathbf{e})$. Let $\mathbf{x}'$ be $\mathcal{A}$'s output. If it holds that $\mathbf{x}' = \mathbf{x}$, $\mathcal{B}$ outputs 1 (for random), otherwise 0 (for pseudorandom). Observe that if $\mathbf{A}^*$ is distributed uniformly, then $\mathcal{A}$'s view in $\mathcal{B}$'s simulation is distributed identically to experiment 1. Likewise, if $\mathbf{A}^*$ is distributed according to $\mathcal{C}$, then $\mathcal{A}$'s view in $\mathcal{B}$'s simulation is distributed as in experiment 2. Therefore, it holds that $|\Pr[\mathcal{B}(\mathbf{A}) = 1] - \Pr[\mathcal{B}(\mathbf{A}') = 1]| = |\Pr[\mathsf{Exp}_1(\mathcal{A}) = \mathsf{win}] - \Pr[\mathsf{Exp}_2(\mathcal{A}) = \mathsf{win}]| > \epsilon$. Thus $\mathcal{B}$ has advantage ϵ distinguishing the uniform distribution on $R_n^{m \times k}$ from $\mathcal{C}$, contradicting the pseudorandomness-property of $\mathcal{C}$

Second, we will show that in experiment 2 $\mathcal{A}$'s success-probability of recovering $\mathbf{x}$ is negligible. Let $(\tilde{\mathbf{A}}, \tilde{\mathbf{y}})$ be $\mathcal{A}$'s input. By the lossiness-property of $\mathcal{C}$, the it holds that $\Pr_{(\tilde{\mathbf{A}}, \tilde{\mathbf{x}}, \tilde{\mathbf{c}})}[H_\infty(\mathbf{x}|(\mathbf{A}, \mathbf{A}\mathbf{x} + \mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{A}}\tilde{\mathbf{x}} + \tilde{\mathbf{c}})) \geq \gamma] \geq 1 - \epsilon(n)$ for a negligible $\epsilon(n)$. Assume now that $H_\infty(\mathbf{x}|(\mathbf{A}, \mathbf{A}\mathbf{x} + \mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{A}}\tilde{\mathbf{x}} + \tilde{\mathbf{c}})) \geq \gamma$. In this case it holds that $\Pr[\mathcal{A}(\tilde{\mathbf{A}}, \tilde{\mathbf{y}}) = \mathbf{x}] \leq 2^{-H_\infty(\mathbf{x}|(\mathbf{A}, \mathbf{A}\mathbf{x}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{A}}\tilde{\mathbf{x}}+\tilde{\mathbf{c}}))} \leq 2^{-\gamma(n)}$. All together, it holds that $\Pr[\mathcal{A}(\tilde{\mathbf{A}}, \tilde{\mathbf{y}}) = \mathbf{x}] \leq \epsilon + 2^{-\gamma(n)}$, which is negligible in n, as $\epsilon = \mathsf{negl}(n)$ and $\gamma(n) = \omega(\log(n))$. □

4 Lossy Codes from LWE

We will first outline the ideas on which our construction of lossy codes from LWE is based. Our starting-point is the observation that the standard LWE problem allows us to construct pseudorandom lattices that contains many vectors that are significantly shorter than one would expect for random lattices. More specifically, let $\mathbf{E} \in \mathbb{Z}_q^{m \times n}$ be component-wise chosen according to a (short) discretized gaussian distribution $\bar{\Psi}_\alpha$. As error-distribution we will set $\chi = \mathcal{U}([-\rho q, \rho q])$, i.e. the uniform distribution on the interval $[-\rho q, \rho q]$, for some parameter ρ. We want to set the parameters α and ρ such that the lattice generated by E is "bad" on average against errors from χ. Put differently, if $\mathbf{y} = \mathbf{E}\mathbf{x} + \mathbf{c}$, where $\mathbf{x}$ is chosen uniformly at random and $\mathbf{c}$ is chosen from χ, we want that there exist many other "admissible" $\mathbf{x}'$ and $\mathbf{c}' \in [-\rho q, \rho q]^m$ such that $\mathbf{y} = \mathbf{E}\mathbf{x}' + \mathbf{c}'$. As $\mathbf{c}$ is distributed uniformly on the volume $[-\rho q, \rho q]^m$, each $\mathbf{x}'$ will have the same posterior-probability given $\mathbf{E}$ and $\mathbf{y}$. If there is a super-polynomial number of such $\mathbf{x}'$, then $\mathbf{y}$ statistically hides $\mathbf{x}$. The main effort in the rest of this section will be to show that such a randomly chosen $\mathbf{E}$ actually is lossy $\mathcal{U}([-\rho q, \rho q])$.

To make our lossy-code pseudorandom, we will *implant* $\Lambda_q(\mathbf{E})$ as a sub-lattice into a bigger lattice $\Lambda_q(\mathbf{A})$. This can be achieved in a pretty standard way. Let $\mathbf{A}' \in \mathbb{Z}_q^{m\times n}$ be chosen uniformly at random. Define $\mathbf{B} = (\mathbf{A}'\|\mathbf{E})$ as the concatenation of $\mathbf{A}'$ and $\mathbf{E}$.

Clearly, the lattice $\Lambda_q(\mathbf{A})$ contains the lattice $\Lambda_q(\mathbf{E})$ as a sub-lattice. Thus, $\Lambda_q(\mathbf{A})$ has a *lossy sub-code*. We will show that having a lossy code as a sub-lattice is sufficient to be lossy, thus $\Lambda_q(\mathbf{A})$ is also lossy. We can randomize the generator-matrix $\mathbf{B} = (\mathbf{A}'\|\mathbf{E})$ by applying the transformation

$$\mathbf{T} = \begin{pmatrix} \mathbf{I} & \mathbf{T}' \\ \mathbf{0} & \mathbf{I} \end{pmatrix},$$

for a $\mathbf{T}' \in \mathbb{Z}_q^{n\times n}$ chosen uniformly at random. This yields the randomized generator $\mathbf{A} = \mathbf{BT} = (\mathbf{A}'\|\mathbf{A}'\mathbf{T}' + \mathbf{E})$ for $\Lambda_q(\mathbf{B})$, i.e. $\Lambda_q(\mathbf{A}) = \Lambda_q(\mathbf{B})$. By the LWE-assumption (for specific parameters), the matrix $\mathbf{A}$ is pseudorandom. As $\mathbf{A}$ and $\mathbf{B}$ generate the same lattice, $\Lambda_q(\mathbf{A})$ is lossy. This is summarized in Construction 1.

Construction 1. Let $q = q(n)$ be an integer and $m = m(n) = \mathsf{poly}(n)$ an integer. The distribution $\mathcal{C}$ defined on $\mathbb{Z}_q^{m\times 2n}$ is specified as follows. Choose $\mathbf{A}' \in \mathbb{Z}_q^{m\times n}$ uniformly at random, choose $\mathbf{T}' \in \mathbb{Z}_q^{n\times n}$ uniformly at random and sample $\mathbf{E} \in \mathbb{Z}_q^{m\times n}$ from $\bar{\Psi}_\alpha^{m\times n}$. Output $\mathbf{A} = (\mathbf{A}'\|\mathbf{A}'\mathbf{T}' + \mathbf{E})$.

The pseudorandomness of the distribution $\mathcal{C}$ follows directly from Lemma 3.

We will now prove that Construction 1 actually yields a lossy code. We first define the notion of ambiguous pairs.

Definition 2. We say that a matrix-vector-pair $(\mathbf{A}, \mathbf{y}) \in \mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^m$ is (r, N)-*ambiguous*, if $|\{\mathbf{x} \in \mathbb{Z}_q^n | \|\mathbf{y} - \mathbf{Ax}\|_\infty \ \leq r\}| \geq N$.

Ambiguousity formalizes the idea that a vector $\mathbf{y} = \mathbf{Ax}+\mathbf{c}$ has many different possible *preimages* $\mathbf{x}$, i.e. there exist at least N pairs $(\mathbf{x}, \mathbf{c})$ with $\|\mathbf{c}\|_\infty \leq r$ such that $\mathbf{y} = \mathbf{Ax} + \mathbf{c}$. Since we want to establish lossiness for the distribution $\mathcal{U}([-r, r])$, counting the number of possible preimages is sufficient, as each preimage is equally likely. This is formalized in the following lemma.

Lemma 4. *Let m, n, q, r and N be integers. Let $\mathbf{A} \in \mathbb{Z}_q^{m\times n}$ be distributed according to a distribution $\mathcal{C}$, $\mathbf{x}$ be chosen uniformly at random from $\mathbb{Z}_q^n$ and $\mathbf{c}$ be chosen from $\mathcal{U}([-r, r])^m$. Assume that the pair $(\mathbf{A}, \mathbf{Ax} + \mathbf{c})$ is (r, N)-ambiguous, except with probability ϵ, i.e. $\Pr[(\mathbf{A}, \mathbf{Ax} + \mathbf{c})$ is not (r, N)-ambigous$] < \epsilon$. Then it holds for $(\tilde{\mathbf{A}}, \tilde{\mathbf{y}})$ chosen iid to $(\mathbf{A}, \mathbf{Ax}+\mathbf{c})$ that $H_\infty(\mathbf{x}|(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})) \geq \log(N)$, except with probability ϵ over the choice of $(\tilde{\mathbf{A}}, \tilde{\mathbf{y}})$.*

Proof. First, fix $\tilde{\mathbf{A}}$, $\tilde{\mathbf{x}}$, $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{y}} = \tilde{\mathbf{A}}\tilde{\mathbf{x}} + \tilde{\mathbf{c}}$ such that $(\tilde{\mathbf{A}}, \tilde{\mathbf{y}})$ is (r, N)-ambiguous, i.e. it holds that $|\{\tilde{\mathbf{x}} \in \mathbb{Z}_q^n | \|\tilde{\mathbf{y}} - \tilde{\mathbf{A}}\tilde{\mathbf{x}}\|_\infty \ \leq r\}| \geq N$. We will first compute the probability $\Pr[\mathbf{x} = \tilde{\mathbf{x}}|(\mathbf{A}, \mathbf{Ax} + \mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})]$. First notice that

$$\Pr[(\mathbf{A}, \mathbf{Ax} + \mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})|\mathbf{x} = \tilde{\mathbf{x}}] = \Pr[\mathbf{A} = \tilde{\mathbf{A}} \text{ and } \mathbf{c} = \tilde{\mathbf{y}} - \tilde{\mathbf{A}}\tilde{\mathbf{x}}] = \Pr[\mathbf{A} = \tilde{\mathbf{A}}] \cdot \Pr[\mathbf{c} = \tilde{\mathbf{y}} - \tilde{\mathbf{A}}\tilde{\mathbf{x}}].$$

As $\mathbf{c}$ is distributed according to $\mathcal{U}([-r, r])^m$, it holds that $\Pr[\mathbf{c} = \tilde{\mathbf{y}} - \tilde{\mathbf{A}}\tilde{\mathbf{x}}] = (2r)^{-m}$ for $\|\tilde{\mathbf{y}} - \tilde{\mathbf{A}}\tilde{\mathbf{x}}\|_\infty \leq r$ and $\Pr[\mathbf{c} = \tilde{\mathbf{y}} - \tilde{\mathbf{A}}\tilde{\mathbf{x}}] = 0$ otherwise. As $\mathbf{x}$ is chosen uniformly from $\mathbb{Z}_q^n$, it holds for all $\tilde{\mathbf{z}} \in \mathbb{Z}_q^n$ that $\Pr[\mathbf{x} = \tilde{\mathbf{z}}] = q^{-n}$. Therefore, it holds that

$$\Pr[\tilde{\mathbf{A}}\mathbf{x}+\mathbf{c} = \tilde{\mathbf{y}}] = \sum_{\tilde{\mathbf{z}}\in\mathbb{Z}_q^n} \Pr[\tilde{\mathbf{A}}\mathbf{x}+\mathbf{c} = \tilde{\mathbf{y}}|\mathbf{x} = \tilde{\mathbf{z}}]\cdot\Pr[\mathbf{x} = \tilde{\mathbf{z}}] = q^{-n}\cdot \sum_{\tilde{\mathbf{z}}\in\mathbb{Z}_q^n, \|\tilde{\mathbf{y}}-\tilde{\mathbf{A}}\tilde{\mathbf{z}}\|_\infty \leq r} (2r)^{-m} \geq q^{-n}(2r)^{-m}N.$$

Thus, it holds that

$$\Pr[(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})] = \Pr[\mathbf{A} = \tilde{\mathbf{A}}] \cdot Pr[\tilde{\mathbf{A}}\mathbf{x}+\mathbf{c} = \tilde{\mathbf{y}}] \leq \Pr[\mathbf{A} = \tilde{\mathbf{A}}] \cdot q^{-n}(2r)^{-m}N.$$

Finally, Bayes' rule yields

$$\begin{aligned}\Pr[\mathbf{x} = \tilde{\mathbf{x}}|(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})] &= \Pr[(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})|\mathbf{x} = \tilde{\mathbf{x}}] \cdot \frac{\Pr[\mathbf{x} = \tilde{\mathbf{x}}]}{\Pr[(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})]} \\ &\leq \frac{\Pr[\mathbf{A} = \tilde{\mathbf{A}}](2r)^{-m}q^{-n}}{\Pr[\mathbf{A} = \tilde{\mathbf{A}}](2r)^{-m}q^{-n}N} = \frac{1}{N}.\end{aligned}$$

Thus, it holds that

$$H_\infty(\mathbf{x}|(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})) = -\log(\max_{\tilde{\mathbf{x}}}(\Pr[\mathbf{x} = \tilde{\mathbf{x}}|(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})])) \geq -\log(1/N) = \log(N).$$

Now, let $(\tilde{\mathbf{A}}, \tilde{\mathbf{y}})$ be iid to $(\mathbf{A}, \mathbf{Ax}+\mathbf{c})$. Since it holds that $\Pr[(\tilde{\mathbf{A}}, \tilde{\mathbf{y}})$ is (r, N)-typical$] \geq 1-\epsilon$, we conclude

$$\Pr[H_\infty(\mathbf{x}|(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})) \geq \log(N)] \geq 1-\epsilon.$$

□

In the rest of this section, we will be concerned with finding suitable parameters N and ϵ in order to maximize $H_\infty(\mathbf{x}|(\mathbf{A}, \mathbf{Ax}+\mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}}))$ (i.e. maximize N) and minimize ϵ, while keeping the worst-to-average connection of the underlying lattice-problem (parametrized by α) as tight as possible. We will first show that (r, N)-ambigousity is preserved if additional columns are added to the matrix A.

Lemma 5. *Let $(\mathbf{A}_2, \mathbf{y}_2) \in \mathbb{Z}_q^{m \times n_2} \times \mathbb{Z}_q^m$ be (r, N)-ambiguous, let $\mathbf{A}_1 \in \mathbb{Z}_q^{m \times n_1}$. Set $\mathbf{A} = (\mathbf{A}_1 \| \mathbf{A}_2)$. Then it holds for any $\mathbf{x}_1 \in \mathbb{Z}_q^{n_1}$ that $(\mathbf{A}, \mathbf{A}_1\mathbf{x}_1 + \mathbf{y}_2)$ is also (r, N)-ambiguous.*

Proof. Set $\mathbf{y} = \mathbf{A}_1\mathbf{x}_1 + \mathbf{y}_2$. Define the sets $M = \{\mathbf{x} \in \mathbb{Z}_q^n | \|\mathbf{y}_2 - \mathbf{A}_2\mathbf{x}\|_\infty \leq r\}$ and $M' = \{\mathbf{x} \in \mathbb{Z}_q^n | \|\mathbf{y} - \mathbf{Ax}\|_\infty \leq r\}$. We will show that if $\mathbf{x} \in M$ then it holds for the vertical concatenation $\tilde{\mathbf{x}} = (\mathbf{x}_1, \mathbf{x})$ of $\mathbf{x}_1$ and $\mathbf{x}$ that $\tilde{\mathbf{x}} \in M'$. Since concatenation is one-to-one, it then follows that $N = |M| \leq |M'|$, i.e. $|\{\mathbf{x} \in \mathbb{Z}_q^n | \|\mathbf{y} - \mathbf{Ax}\|_\infty \leq r\}| \geq N$, which means that $(\mathbf{A}, \mathbf{y})$ is (r, N)-ambiguous. Thus, let $\mathbf{x} \in M$, i.e. $\|\mathbf{y}_2 - \mathbf{A}_2\mathbf{x}\|_\infty \leq r$. It holds that $\|\mathbf{y} - \mathbf{A}\tilde{\mathbf{x}}\|_\infty = \|\mathbf{y} - \mathbf{A}_1\mathbf{x}_1 - \mathbf{A}_2\mathbf{x}\|_\infty = \|\mathbf{A}_1\mathbf{x}_1 + \mathbf{y}_2 - \mathbf{A}_1\mathbf{x}_1 - \mathbf{A}_2\mathbf{x}\|_\infty = \|\mathbf{y}_2 - \mathbf{A}_2\mathbf{x}\|_\infty \leq r$, i.e. $\tilde{\mathbf{x}} \in M'$ which concludes the proof. □

Henceforth let $r = \rho q$. Our proof-strategy will continue as follows. We first observe that a $\mathbf{c}$ chosen by $\mathcal{U}([-\rho q, \rho q])$ is, with overwhelming probability *typical* in the following sense. Roughly speaking, $\mathbf{c}$ is typical, if at most a small (more precisely logarithmic) number of components of $\mathbf{c}$ are "too" close to the boundaries of the interval $[-\rho q, \rho q]$. Let $t > 0$ be an integer. If we fix such a typical $\mathbf{c}$ and choose $\mathbf{e}$ according to $\bar{\Psi}_\alpha^m$, then it holds with high probability (over the choice of $\mathbf{e}$), that $\|\mathbf{c} - t\mathbf{e}\|_\infty \leq \rho q$. If $\mathbf{E}$ is distributed according to $\bar{\Psi}_\alpha^{m \times n}$, then it holds for a sufficiently high number of columns $\mathbf{e}_i$ of $\mathbf{E}$ that $\|\mathbf{c} - t\mathbf{e}\|_\infty \leq \rho q$. It thus holds for all convex-combinations $\mathbf{z}$ of these $t \cdot \mathbf{e}_i$ that $\|\mathbf{c} - \mathbf{z}\|_\infty \leq \rho q$. Among those $\mathbf{z}$ are a super-polynomial number of vectors of the form $\mathbf{Ex}'$. We then conclude that the pair $(\mathbf{E}, \mathbf{Ex}+\mathbf{c})$ is $(\rho q, N)$-ambiguous for a super-polynomial N.

We will now define the notion of typical points in a hypercube of dimension m.

Definition 3. We say that a $\mathbf{c} \in [-\rho q, \rho q]^m \subseteq \mathbb{Z}^m$ is (k,δ)-typical, if it holds for at most k components c_i of $\mathbf{c}$ that $|c_i| > (1-\delta)\rho q$.

Lemma 6. *Let m, k and δ be such that $\delta(m+1) \leq k$. It holds that a $\mathbf{c}$ chosen from $\mathcal{U}([-\rho q, \rho q])^m \subseteq \mathbb{Z}^m$ is (k,δ)-typical, except with probability $m \cdot (m\delta)^k$.*

Proof. For $i = 1, \ldots, m$ let Z_i be a random-variable that is 1 if $|c_i| > (1-\delta)\rho q$ and 0 otherwise. Let $p = \Pr[Z_1 = 1] = \cdots = \Pr[Z_m = 1]$. As c_1 is distributed by $\mathcal{U}([-\rho q, \rho q])$, it holds that $p = \Pr[Z_1 = 1] = \Pr[|c_1| > (1-\delta)\rho q] \leq \delta$. Set $Z = \sum_{i=1}^{m} Z_i$. Clearly, Z is binomially distributed and $\mathbf{c}$ is (k,δ)-typical in $[-\rho q, \rho q]^m$ if and only if $Z \leq k$. We can bound the probability $\Pr[Z > k]$ by

$$\Pr[Z > k] = \sum_{i=k+1}^{m} \binom{m}{i} p^i (1-p)^{m-i} \leq m \underbrace{\binom{m}{k}}_{\leq m^k} \underbrace{p^k}_{\leq \delta^k} \underbrace{(1-p)^{m-k}}_{\leq 1} \leq m \cdot (m\delta)^k.$$

The first inequality holds as $\binom{m}{i} p^i (1-p)^{m-i}$ assumes its maximum $i_{max} = \lfloor p(m+1) \rfloor \leq \delta(m+1) \leq k$. □

Lemma 7. *Let $t > 0$ be an integer and let $\mathbf{e}$ be distributed according to $\bar{\Psi}_\alpha^m$. Let $\delta \geq \frac{\sqrt{n} \cdot t\alpha}{\rho}$ and $k + 1 \leq \pi/4 \cdot n \log(e) - \log(m)$. Fix a $\mathbf{c} \in [-\rho q, \rho q]^m$ that is (k,δ)-typical. Then it holds that $\Pr[\|\mathbf{c} - t\mathbf{e}\|_\infty \leq \rho q] \geq 2^{-(k+1)}$.*

Proof. Denote by A the event that $\|\mathbf{c} - t \cdot \mathbf{e}\|_\infty \leq \rho q$ and by B the event that $\|t \cdot \mathbf{e}\|_\infty \leq \delta \rho q$. The probability of event $\neg B$ is bounded by

$$\begin{aligned}\Pr[\neg B] = \Pr[\exists i : |t \cdot e_i| > \delta \rho q] &\leq \sum_{i=1}^{m} \Pr[|e_i| > \frac{\delta \rho}{t\alpha} \alpha q] \leq m \cdot e^{-\pi/4 \cdot \left(\frac{\delta \rho}{t\alpha}\right)^2} \\ &\leq m \cdot e^{-\pi/4 \cdot n} = 2^{-(\pi/4 \cdot n \log(e) - \log(m))} \leq 2^{-(k+1)}.\end{aligned}$$

Thus it holds that

$$\Pr[\|\mathbf{c} - t\mathbf{e}\|_\infty \leq \rho q] = \Pr[A] \geq \Pr[A|B] \underbrace{\Pr[B]}_{\geq 1 - 2^{-(k+1)})} \geq \Pr[A|B] - \underbrace{\Pr[A|B]}_{\leq 1} \cdot 2^{-(k+1)} \geq \Pr[A|B] - 2^{-(k+1)}$$

We will now show that $\Pr[A|B] \geq 2^{-k}$, from which follows with the above that $\Pr[\|\mathbf{c} - t\mathbf{e}\|_\infty \leq \rho q] \geq 2^{-(k+1)}$. As we condition on $\|t\mathbf{e}\|_\infty \leq \delta \rho q$, assume that for all $i = 1, \ldots, m$ $|te_i| \leq \delta \rho q$. As $\mathbf{c} \in [-\rho q, \rho q]^m$ is (k,δ)-typical, there are at most k indices $i_1, \ldots, i_k$ such that $|c_{i_l}| > (1-\delta)\rho q$ (for $l = 1 \ldots k$). Let $i \in \{i_1, \ldots, i_k\}$. If $\mathsf{sgn}(e_i) = \mathsf{sgn}(c_i)$, then it holds that $|c_i - te_i| = |c_i| - |te_i| \leq |c_i| \leq \rho q$. As $\bar{\Psi}_\alpha$ is a symmetrical distribution, it holds that $\Pr[\mathsf{sgn}(e_i) = \mathsf{sgn}(c_i)] \geq \frac{1}{2}$. Therefore, it holds that $\Pr[|c_i - te_i| \leq \rho q || te_i| \leq \delta \rho q] \geq \frac{1}{2}$. For all other indices $j \notin \{i_1, \ldots, i_k\}$ it holds that $|c_j| \leq (1-\delta)\rho q$. The triangle-inequality yields $|c_j - te_j| \leq |c_j| + |te_j| \leq (1-\delta)\rho q + \delta \rho q = \rho q$. Therefore, we have that $\Pr[|c_j - te_j| \leq \rho q || te_j| \leq \delta \rho q] = 1$. Putting it all together, we get that

$$\Pr[A|B] = \Pr[\|\mathbf{c} - t\mathbf{e}\|_\infty \leq \rho q | \|t\mathbf{e}\|_\infty \leq \delta \rho q] = \prod_{i=1}^{m} \Pr[|c_i - te_i| \leq \rho q || te_i| \leq \delta \rho q] \geq 2^{-k},$$

which concludes the proof. □

On one side, we want the probability that a random $\mathbf{c}$ distributed by $\mathcal{U}([-\rho q, \rho q])^m$ is not (k,δ)-typical to be as low as possible. On the other side, we want for a (k,δ)-typical $\mathbf{c}$ that $\Pr[\|\mathbf{c} - t\mathbf{e}\|_\infty \leq \rho q]$ is reasonably high. We therefore set $k = \lceil 1/2 \cdot \log_2 n \rceil$ and $\delta = \frac{1}{m\sqrt{n}}$. First note that $\delta(m+1) = \frac{m+1}{\sqrt{n}m} \leq 1 \leq 1/2 \cdot \log(n) \leq k$ and $k+1 \leq 1/2 \cdot \log(n) + 2 \leq \pi/4 \cdot n \log(e) - \log(m)$ for sufficiently large n, thus the requirements of Lemma 6 are given. Lemma 6 then yields that $\Pr[\mathbf{c} \text{ is not } (k,\delta)\text{-typical}] \leq m \cdot n^{-1/4 \cdot \log_2(n)}$ (which is negligible in n). Moreover, if the constraint $\delta = \frac{1}{m\sqrt{n}} \geq \frac{\sqrt{n} \cdot t\alpha}{\rho}$ is met, 7 yields, that $\Pr[\|\mathbf{c} - t\mathbf{e}\|_\infty \leq \rho q] \geq \frac{1}{2\sqrt{n}}$. Plugging in $\delta = \frac{1}{m\sqrt{n}}$ in the above constraint yields $\alpha \leq \frac{\rho}{n \cdot t \cdot m}$.

Lemma 8. *Let $m \geq 2n$, $\rho < 1/2$ and $\alpha \leq \frac{8\rho}{n^{3/2} \cdot m}$. Let $\mathbf{x}$ be chosen uniformly at random from $\mathbb{Z}_q^n$, let $\mathbf{c}$ be distributed by $\mathcal{U}([-\rho q, \rho q])^m$ and let $\mathbf{E}$ be distributed by $\bar{\Psi}_\alpha^{m \times n}$. Then the pair $(\mathbf{E}, \mathbf{E}\mathbf{x} + \mathbf{c})$ is $(\rho q, 2^{\sqrt{n}/16})$-ambiguous, except with probability $m \cdot n^{-1/4 \cdot \log_2(n)} + e^{-\sqrt{n}/32} + n \cdot e^{(1/2 - \ln(2/))\frac{n}{2}}$.*

Proof. Let $t > 0$ be an integer specified later. Setting $k = \lceil 1/2 \cdot \log_2 n \rceil$ and $\delta = \frac{1}{m\sqrt{n}}$, Lemma 6 yields $\Pr[\mathbf{c} \text{ is not } (\lceil 1/2 \cdot \log_2 n \rceil, \frac{1}{m\sqrt{n}})\text{-typical}] \leq m \cdot n^{-1/4 \cdot \log_2(n)}$ for a $\mathbf{c}$ distributed by $\mathcal{U}([-\rho q, \rho q])^m$. Assume now that $\mathbf{c}$ is $(\lceil 1/2 \cdot \log_2 n \rceil, \frac{1}{m\sqrt{n}})$-typical. Let $\mathbf{E}$ be distributed by $\bar{\Psi}_\alpha^{m \times n}$. By Lemma 7, it holds for each column $\mathbf{e}_i$ of $\mathbf{E}$ that $\Pr[\|\mathbf{c} - t \cdot \mathbf{e}_i\|_\infty \leq r] \geq \frac{1}{4\sqrt{n}}$. Let L be the number of columns $\mathbf{e}_i$ of $\mathbf{E}$ for which $\|\mathbf{c} - t \cdot \mathbf{e}_i\|_\infty \leq r$ holds. As all $\mathbf{e}_i$ are independently identically distributed according to $\bar{\Psi}_\alpha^m$, L is binomially distributed and it holds that $\mathsf{E}[L] = \Pr[\|\mathbf{c} - t\mathbf{e}\|_\infty \leq r] \cdot n \geq \frac{1}{4\sqrt{n}} n = \sqrt{n}/4$. A Chernoff-bound yields $\Pr[L \leq \sqrt{n}/8] \leq \Pr[L \leq \mathsf{E}[L]/2] \leq e^{-\mathsf{E}[L]/4} \leq e^{-\sqrt{n}/32}$. Thus it holds that $L > \sqrt{n}/8$, except with probability $e^{-\sqrt{n}/32}$. Now let $l := \lfloor \sqrt{n}/8 \rfloor$ and assume that $i_1, \ldots, i_l$ are such that for all $j = 1, \ldots, l$ that $\|\mathbf{c} - t \cdot \mathbf{e}_{i_j}\|_\infty \leq r$. By Lemma 1 the lattice $\Lambda(\mathbf{E})$ has full rank, except with probability $n \cdot e^{(1/2 - \ln(2/))\frac{n}{2}}$ (which is negligible in n).

Let A denote the event that $\mathbf{c}$ is $(\lceil 1/2 \cdot \log_2 n \rceil, \frac{1}{m\sqrt{n}})$-typical, let B denote the event that at least l rows $\mathbf{e}_i$ of $\mathbf{E}$ suffice $\|\mathbf{c} - t \cdot \mathbf{e}_i\|_\infty \leq r$ and let C denote the event that $\Lambda(\mathbf{E})$ has full rank. It holds that $\Pr[\neg A] \leq m \cdot n^{-1/4 \cdot \log_2(n)}$, $\Pr[\neg B | A] \leq e^{-\sqrt{n}/32}$ and $\Pr[\neg C] \leq n \cdot e^{(1/2 - \ln(2/))\frac{n}{2}}$. Thus, a union-bound yields

$$\begin{aligned}
\Pr[\neg A \vee \neg B \vee \neg C] &\leq \Pr[\neg A \vee \neg B] + \Pr[\neg C] \\
&\leq \underbrace{\Pr[\neg A]}_{m \cdot n^{-1/4 \cdot \log_2(n)}} \cdot \underbrace{\Pr[\neg A \vee \neg B | \neg A]}_{=1} + \underbrace{\Pr[A]}_{\leq 1} \cdot \underbrace{\Pr[\neg A \vee \neg B | A]}_{= \Pr[\neg B | A] \leq e^{-\sqrt{n}/32}} + \underbrace{\Pr[\neg C]}_{\leq n \cdot e^{(1/2 - \ln(2/))\frac{n}{2}}} \\
&\leq m \cdot n^{-1/4 \cdot \log_2(n)} + e^{-\sqrt{n}/32} + n \cdot e^{(1/2 - \ln(2/))\frac{n}{2}},
\end{aligned}$$

which is negligible in n.

Assume henceforth that the vectors $\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_l}$ have full rank and it holds for all j that $\|\mathbf{c} - t \cdot \mathbf{e}_{i_j}\|_\infty \leq \rho q$. As $[-\rho q, \rho q]^m$ is convex, it holds for all points $\mathbf{z}$ in the convex hull S of $t\mathbf{e}_{i_1}, \ldots, t\mathbf{e}_{i_l}$ that $\|\mathbf{c} - \mathbf{z}\|_\infty \leq \rho q$ (i.e. $\mathbf{c} - \mathbf{z} \in [-\rho q, \rho q]^m$). By Lemma 2 it holds that $|S \cap \Lambda(\mathbf{E})| = \binom{l+t}{l}$. Setting $t = l = \lfloor \sqrt{n}/8 \rfloor$ yields $|S \cap \Lambda(\mathbf{E})| = \binom{2l}{l} \geq 2^l / l \geq 4 \cdot 2^{\sqrt{n}/8} / \sqrt{n} \geq 2^{\sqrt{n}/16}$. Thus, there are at least $2^{\sqrt{n}/16}$ $\mathbf{x}' \in \mathbb{Z}^n$ such that $\|\mathbf{c} - \mathbf{E} \cdot \mathbf{x}'\|_\infty \leq \rho q$. This remains true if $\mathbf{c}$ is shifted by any lattice-point $\mathbf{E}\mathbf{x}$. I.e. if $\mathbf{y} = \mathbf{E}\mathbf{x} + \mathbf{c}$, it holds that $\|\mathbf{y} - \mathbf{E} \cdot (\mathbf{x} + \mathbf{x}')\|_\infty = \|\mathbf{E}\mathbf{x} + \mathbf{c} - \mathbf{E} \cdot \mathbf{x} - \mathbf{E}\mathbf{x}'\|_\infty = \|\mathbf{c} - \mathbf{E} \cdot \mathbf{x}'\|_\infty \leq \rho q$. Therefore, there are at least $2^{\sqrt{n}/16}$ $\hat{\mathbf{x}} = \mathbf{x} + \mathbf{x}' \in \mathbb{Z}^n$ such that $\|\mathbf{y} - \mathbf{E} \cdot \hat{\mathbf{x}}\|_\infty \leq \rho q$. As $\rho q < q/2$, this holds also modulo q. This means that $(\mathbf{E}, \mathbf{y})$ is $(\rho q, 2^{\sqrt{n}/16})$-ambiguous. The choice of $t = \lfloor \sqrt{n}/8 \rfloor$ imposes the constraint $\alpha \leq \frac{8\rho}{n^{3/2} m}$, which is fulfilled by $\alpha \leq \frac{\rho}{n^{3/2} m}$ (as in the statement of the Lemma). This concludes the proof. □

We can summarize all the above in the following theorem.

Theorem 4. *Let $\alpha \leq \frac{\rho}{n^{3/2} \cdot m}$. Assuming that $\mathsf{LWE}(n, m, q, \bar{\Psi}_\alpha)$ is hard, the distribution $\mathcal{C}$ given in Construction 1 is $\sqrt{n}/16$-lossy for the distribution $\mathcal{U}([-\rho q, \rho q])$.*

Proof. First, let $m \geq 2n$. The distribution $\mathcal{U}([-\rho q, \rho q])$ is (trivially) efficiently samplable. Assuming $\mathsf{LWE}(n, m, q, \bar{\Psi}_\alpha)$ is hard, the pseudorandomness of the distribution $\mathcal{C}$ follows directly from Lemma 3. We will now establish the lossiness of $\mathcal{C}$. Let $\mathbf{A'}$ be chosen uniformly from $\mathbb{Z}_q^{m \times n}$ and $\mathbf{E}$ be chosen according to $\bar{\Psi}_\alpha^{m \times n}$ and let $\mathbf{T'}$ be chosen uniformly from $\mathbb{Z}_q^{n \times n}$. Set $\mathbf{B} = (\mathbf{A'} \| \mathbf{E})$ and $\mathbf{A} = (\mathbf{A'} \| \mathbf{A'T'} + \mathbf{E})$. Let $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$ be chosen uniformly at random from $\mathbb{Z}_q^{2n}$ and let $\mathbf{c}$ be distributed by $\mathcal{U}([-\rho q, \rho q])^m$. Lemma 8 yields that $(\mathbf{E}, \mathbf{E}\mathbf{x}_2 + \mathbf{c})$ is $(\rho q, 2^{\sqrt{n}/16})$-ambiguous, except with negligible probability. Given that $(\mathbf{E}, \mathbf{E}\mathbf{x}_2 + \mathbf{c})$ is $(\rho q, 2^{\sqrt{n}/16})$-ambiguous, Lemma 5 yields that $((\mathbf{A'} \| \mathbf{E}), \mathbf{A'}\mathbf{x}_1 + \mathbf{E}\mathbf{x}_2 + \mathbf{c}) = (\mathbf{B}, \mathbf{B}\mathbf{x} + \mathbf{c})$ is also $(\rho q, 2^{\sqrt{n}/16})$-ambiguous. This holds also for the transformed generator $\mathbf{A}$, as the mapping $(\mathbf{A'} \| \mathbf{E}) \mapsto (\mathbf{A'} \| \mathbf{A'T'} + \mathbf{E})$ is invertible. Now, this holds also if $m < 2n$, as removing rows from $\mathbf{A}$ will only increase the number N of possible $\mathbf{x}$ with $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_\infty \leq \rho q$. Finally, Lemma 4 establishes that it holds for $(\tilde{\mathbf{A}}, \tilde{\mathbf{y}})$ chosen iid to $(\mathbf{A}, \mathbf{A}\mathbf{x} + \mathbf{c})$ that $H_\infty(\mathbf{x} | (\mathbf{A}, \mathbf{A}\mathbf{x} + \mathbf{c}) = (\tilde{\mathbf{A}}, \tilde{\mathbf{y}})) \geq \log(2^{\sqrt{n}/16}) = \sqrt{n}/16$, except with negligible probability over the choice of $(\tilde{\mathbf{A}}, \tilde{\mathbf{y}})$. This concludes the proof. □

5 LWE with Uniform Interval Error-Distribution

Combining Theorems 1, 2, 3 and 4 yields the main theorem.

Theorem 5 (Main Theorem)**.** *Let n be a security parameter. Let $q = q(n)$ and $m = m(n) = \mathsf{poly}(n)$ be integers. Let $\rho = \rho(n) \in (0, 1)$ be such that $\rho q \geq 2n^2 m$. Assume there exists a PPT-algorithm that solves $\mathsf{LWE}(n, m, q, \mathcal{U}([-\rho q, \rho q]))$ with non-negligible probability. Then there exists an efficient quantum-algorithm that approximates the decision-version of the shortest vector problem (GAPSVP) and the shortest independent vectors problem (SIVP) to within $\tilde{O}(n^{5/2} m/\rho)$ in the worst case.*

Using Theorem 2, we can establish the hardness of the decisional LWE problem with error-distribution $\mathcal{U}([-\rho q, \rho q])$.

Corollary 9. *Let n be a security parameter. Let $q = q(n) = \mathsf{poly}(n)$ be a prime integer and m and ρ as in Theorem 5. Assume there exists a PPT-distinguisher that distinguishes $\mathsf{LWE}(n, m, q, \mathcal{U}([-\rho q, \rho q]))$ with non-negligible advantage. Then there exists an efficient quantum-algorithm that approximates the decision-version of the shortest vector problem (GAPSVP) and the shortest independent vectors problem (SIVP) to within $\tilde{O}(n^{5/2} m/\rho)$ in the worst case.*

6 Conclusion

This work presented the first worst-to-average-case reduction for an LWE variant with uniformly distributed errors thereby answering a question from Micciano and Mol from Crypto 2011. The factor of this worst-to-average-case connection depends on the number of samples given to the adversary and we have to use a bounded LWE assumption where this number is fixed in advance. All applications of LWE the authors know of can be based on this bounded variant. The main ingredient for the proof is a new tool called lossy codes, i.e., codes which lose information when decoding noisy code words. An interesting open question is, if these techniques carry over to hardness assumptions for binary codes.

References

[ABB10] Shweta Agrawal, Dan Boneh, and Xavier Boyen. Lattice Basis Delegation in Fixed Dimension and Shorter-Ciphertext Hierarchical IBE. In *CRYPTO*, pages 98–115, 2010.

[ACPS09] Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast Cryptographic Primitives and Circular-Secure Encryption Based on Hard Learning Problems. In *CRYPTO*, pages 595–618, 2009.

[BGV12] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In *ITCS*, pages 309–325, 2012.

[BPR12] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom Functions and Lattices. In *EUROCRYPT*, pages 719–737, 2012.

[Bra12] Zvika Brakerski. Fully Homomorphic Encryption without Modulus Switching from Classical GapSVP. In *CRYPTO*, pages 868–886, 2012.

[BV11] Zvika Brakerski and Vinod Vaikuntanathan. Efficient Fully Homomorphic Encryption from (Standard) LWE. In *FOCS*, pages 97–106, 2011.

[CHKP10] David Cash, Dennis Hofheinz, Eike Kiltz, and Chris Peikert. Bonsai Trees, or How to Delegate a Lattice Basis. In *EUROCRYPT*, pages 523–552, 2010.

[DORS08] Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. *SIAM J. Comput.*, 38(1):97–139, 2008.

[GKPV10] Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Robustness of the Learning with Errors Assumption. In *ICS*, pages 230–240, 2010.

[GPV08] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *STOC*, pages 197–206, 2008.

[LPR10] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On Ideal Lattices and Learning with Errors over Rings. In *EUROCRYPT*, pages 1–23, 2010.

[MM11a] Daniele Micciancio and Petros Mol. Pseudorandom Knapsacks and the Sample Complexity of LWE Search-to-Decision Reductions. In *CRYPTO*, pages 465–484, 2011.

[MM11b] Daniele Micciancio and Petros Mol. Pseudorandom Knapsacks and the Sample Complexity of LWE Search-to-Decision Reductions. *IACR Cryptology ePrint Archive*, 2011:521, 2011.

[MP12] Daniele Micciancio and Chris Peikert. Trapdoors for Lattices: Simpler, Tighter, Faster, Smaller. In *EUROCRYPT*, pages 700–718, 2012.

[Pei09] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In *STOC*, pages 333–342, 2009.

[Pie12] Krzysztof Pietrzak. Subspace LWE. In *TCC*, pages 548–563, 2012.

[PW08] Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. In *STOC*, pages 187–196, 2008.

[Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC*, pages 84–93, 2005.

[vLW01] J.H. van Lint and R.M. Wilson. *A Course in Combinatorics*. Cambridge University Press, 2001.

A Additional Proofs

Lemma. *Let* $\mathbf{E}$ *be chosen from* $\bar{\Psi}_\alpha^{m\times n}$*, where* $m \geq 2n$ *and* $\alpha q \geq 1/\sqrt{\pi}$*. Then* $\mathbf{E}$ *has full rank, except with probability* $n \cdot e^{(1/2-\ln(2))\frac{n}{2}}$*.*

Proof. Let V be a d-dimensional subspace of $\mathbb{R}^m$ (for $d \leq n$) and $\mathbf{e}$ be distributed by $\bar{\Psi}_\alpha^m$. We will first bound the probability that $\mathbf{e} \in V$. The vector $\mathbf{e}$ is obtained by rounding a sample $\mathbf{e}'$ of the distribution $\Phi_{\alpha q}^m$ component-wise to the nearest integer. As for all i $|e_i - e_i'| \leq 1/2$ it holds that $\|\mathbf{e}-\mathbf{e}'\|_2 \leq \sqrt{m}/2$ and therefore $\mathsf{dist}(V,\mathbf{e}) \geq \mathsf{dist}(V,\mathbf{e}') - \sqrt{m}/2$. Consequently, $\mathsf{dist}(V,\mathbf{e}') > \sqrt{m}/2$ is sufficient for $\mathbf{e} \notin V$. As the distribution $\Phi_{\alpha q}^m$ is rotationally symmetric, we can decompose $\mathbf{e}'$ into $\mathbf{e}' = \mathbf{e}'_V + \mathbf{e}'_{V^\perp}$, where $\mathbf{e}'_V$ is distributed according to $\Phi_{\alpha q}^{\dim(V)}$ in an orthonormal basis of V and $\mathbf{e}'_{V^\perp}$ is distributed according to $\Phi_{\alpha q}^{\dim(V^\perp)}$ in an orthonormal basis of $V^\perp$. Thus, $\mathsf{dist}(V,\mathbf{e}') = \|\mathbf{e}'_{V^\perp}\|_2$. We can use a bound for the chi-squared distribution to derive a lower bound for $\|\mathbf{e}'_{V^\perp}\|_2$. Specifically, as $\sqrt{m}/2 \leq 1/2 \cdot \sqrt{2\pi}\alpha q\sqrt{m-d}$ (which is easily implied by $\alpha q \geq 1/\sqrt{\pi}$ and $m \geq 2n \geq 2d$), it holds that

$$\Pr[\|\mathbf{e}'_{V^\perp}\|_2 \leq \sqrt{m}/2] \leq \Pr[\|\mathbf{e}'_{V^\perp}\|_2 \leq 1/2 \cdot \sqrt{2\pi}\alpha q\sqrt{m-d}] \leq \mathbf{e}^{(1/2-\ln(2))\frac{m-d}{2}}.$$

Thus it holds that $\Pr[\mathbf{e} \in V] = \Pr[\mathsf{dist}(V,\mathbf{e}) = 0] = \Pr[\|\mathbf{e}'_{V^\perp}\|_2 \leq \sqrt{m}/2] \leq e^{(1/2-\ln(2))\frac{m-d}{2}}$. Let $\mathbf{e}_1,\ldots,\mathbf{e}_n$ be the columns of $\mathbf{E}$. We can now bound the the probability that $\mathbf{E}$ is rank-deficient by

$$\begin{aligned}\Pr[\mathsf{rank}(\mathbf{E}) < n] &= \Pr[\exists i : \mathbf{e}_i \in \mathsf{span}(\mathbf{e}_1,\ldots,\mathbf{e}_{i-1})] \leq \sum_i \Pr[\mathbf{e}_i \in \mathsf{span}(\mathbf{e}_1,\ldots,\mathbf{e}_{i-1})] \\ &\leq n \cdot e^{(1/2-\ln(2))\frac{m-d}{2}} \leq n \cdot e^{(1/2-\ln(2))\frac{n}{2}}.\end{aligned}$$

□