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

Nico Döttling[1] and Jörn Müller-Quade[1]

Karlsruhe Institute of Technology, Karlsruhe, Germany
{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. It 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 the errors for current applications of LWE must be chosen from a gaussian distribution. However, sampling from gaussian distributions is cumbersome.
In this work we present the first worst-to-average case reduction for LWE with uniformly distributed errors, which can be sampled very efficiently. 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 is fixed in advance. Most applications of LWE can be based on the bounded variant. The proof is based on a new tool called *lossy codes*, which might be of interest in the context 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 very 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. Our resulting 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 in-approximability assumptions made in other works [GKPV10,BV11,Bra12]. We now state our main-theorem.

Theorem 1 (Main Theorem). *Let n be a security parameter and let $\sigma \in (0,1)$ be an arbitrarily small constant. Let $q = q(n)$ be a modulus and $m = m(n) = \mathsf{poly}(n)$ be a integer with $m \geq 3n$. Let $\rho = \rho(n) \in (0, 1/10)$ be such that $\rho q \geq 2n^{0.5+\sigma}m$. If 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 $\tilde{O}(n^{1+\sigma}m/\rho)$ 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 Outline of the Techniques

We will briefly outline the construction and the proof of our main results. The Learning-With-Errors Problem is basically the decoding-problem for q-ary lattices: Given a randomly chosen generator-matrix $\mathbf{A}$ and a vector $\mathbf{y}$, find the nearest lattice point (or codeword) $\mathbf{Ax}$, under the promise that $\mathbf{y}$ was generated by drawing a random point from the lattice and adding an error by some specified distribution. We want to show that this decoding-problem is hard if the error is component-wise chosen by $\mathcal{U}([-r, r])$, i.e. from the uniform distribution on some interval $[-r, r]$. Assume that we knew that there exists a distribution of *lossy* matrices $\mathbf{A'}$ such that that the decoding-problem has no unique solution if the errors come from $\mathcal{U}([-r, r])$, i.e., adding noise to a lattice-point $\mathbf{A'x}$ loses information about $\mathbf{x}$. If distinguishing such matrices from truly random matrices is hard, we can conclude that the decoding-problem must be hard for truly random matrices. Otherwise, given a decoder for random matrices we can distinguish random matrices from lossy matrices. The distinguisher samples random challenges for the decoder. If the decoder succeeds significantly often, i.e. if it outputs the same $\mathbf{x}$ that was used to sample the instance, then the given matrix must come from the random distribution, as this behavior is impossible for the lossy distribution. Thus, our task is to construct a distribution of lossy codes for the error-distribution $\mathcal{U}([-r, r])$. Our starting-point to find such a distribution is the observation that the standard LWE-problem allows us to construct pseudorandom matrices that generate lattices which contain many vectors that are significantly shorter than one would expect for lattices generated by truly random matrices. Let $\mathbf{G} \in \mathbb{Z}_q^{m \times n}$ be component-wise chosen according to a (short) discretized gaussian distribution $\bar{\Psi}_\alpha$. We want to set the parameters α and r such that the lattice generated by $\mathbf{G}$ is "bad" on average against errors from $\mathcal{U}([-r, r])$. Put differently, if $\mathbf{y} = \mathbf{Gx} + \mathbf{e}$, where $\mathbf{x}$ is chosen uniformly at random and $\mathbf{e}$ is chosen from $\mathcal{U}([-r, r])^m$, we want that, with overwhelming probability, there exist at least one more "admissible" $\mathbf{x'} \neq \mathbf{x}$ and $\mathbf{e'} \in [-r, r]^m$ such that $\mathbf{y} = \mathbf{Gx'} + \mathbf{c'}$. As $\mathbf{e}$ is distributed uniformly on the volume $[-r, r]^m$, each $\mathbf{x'}$ will have the same posterior-probability given $\mathbf{G}$ and $\mathbf{y}$. If there is at least one such

$\mathbf{x}'$, then $\mathbf{y}$ statistically hides at least one bit of information about $\mathbf{x}$ and we can implement the distinguisher sketched above. To make this lossy code pseudorandom, we *hide* the matrix $\mathbf{G}$ in a bigger matrix $\mathbf{A}$. 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 $\mathbf{B} = (\mathbf{A}' \| \mathbf{G})$ as the concatenation of $\mathbf{A}'$ and $\mathbf{G}$. $\mathbf{B}$ now contains the $\mathbf{G}$ as a submatrix. Thus, $\mathbf{B}$ has a *lossy sub-code*. As having a lossy sub-code is sufficient to be lossy, $\mathbf{A}$ is also lossy. We can randomize the generator-matrix $\mathbf{B} = (\mathbf{A}' \| \mathbf{G})$ 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 $\mathbf{A} = \mathbf{BT} = (\mathbf{A}' \| \mathbf{A}'\mathbf{T}' + \mathbf{G})$ for the same code. By the LWE-assumption (for specific parameters), the matrix $\mathbf{A}$ is pseudorandom.

Assume that $\bar{\Psi}_\alpha$ is B-bounded, where $B \ll r$. We still need to show that a matrix $\mathbf{G}$ chosen from $\bar{\Psi}_\alpha^{m \times n}$ is (with high probability) lossy for the error-distribution $\mathcal{U}([-r, r])$. By linearity, it is sufficient to show that for an $\mathbf{e}$ chosen from $\mathcal{U}([-r, r])^m$ there exist $\mathbf{x}' \neq 0$ and $\mathbf{e}' \in [-r, r]^m$ such that $\mathbf{e} = \mathbf{G} \cdot \mathbf{x}' + \mathbf{e}'$, with high probability. Then $\mathbf{e}$ can be reached from either $\mathbf{x}_1 = 0$ or $\mathbf{x}_2 = \mathbf{x}' \neq 0$ and we have established 1-bit loss (which is sufficient for the above construction).

Consider a slightly simpler problem, namely when $\mathbf{G}$ only consists of a single column $\mathbf{g}$. We first observe that errors $\mathbf{e}$ drawn from $\mathcal{U}([-r, r])^m$ show the following *typical* behavior: If we draw $\mathbf{g}$ according to $\bar{\Psi}_\alpha^m$, then there is a substantial chance (over the choice of $\mathbf{g}$) that it holds $\mathbf{e} - \mathbf{g} \in [-r, r]^m$. An $\mathbf{e}$ drawn from $\mathbf{U}([-r, r])^m$ has this property with high probability. To see this, note that there are not too many components e_i of $\mathbf{e}$ that have distance less than B from the boundaries of the interval $[-r, r]$. Call a component e_i with this property (i.e. $e_i \notin [-r + B, r - B]$) bad. For each component e_i, the probability e_i is bad is B/r. Thus, the expected number of components e_i too close to the boundaries is $m \cdot B/r$, which is < 1 for an appropriate choice of m, B and r. We can use a tail-bound for this type of Bernoulli-distribution to show that with overwhelming probability, the number of bad components e_i of $\mathbf{e}$ is less than $\log(n)/2$. Now fix an $\mathbf{e}$ with less than $\log(n)/2$ bad components. For each good component e_i of $\mathbf{e}$, it holds that $e_i + g_i \in [-r, r]$ as g_i is $B-$*bounded*. For all bad components e_i, the probability that $e_i - g_i \in [-r, r]$ is at least $1/2$, as $\bar{\Psi}_\alpha$ is symmetric and thus there is only a $1/2$ chance that g_i *goes the wrong way*. All together, it holds that $\mathbf{e} + \mathbf{g} \in [-r, r]^m$ with probability at least $\left(\frac{1}{2}\right)^{\log(n)/2} = \frac{1}{\sqrt{n}}$, which is substantial.

Now, return to the original problem. Fix an $\mathbf{e}$ that is typical in the above sense. As $\mathbf{G}$ has n columns $\mathbf{g}_1, \ldots, \mathbf{g}_n$ independently chosen from $\bar{\Psi}_\alpha^m$, the probability that there is at least one $\mathbf{g}_i$ such that $\mathbf{e} - \mathbf{g}_i \in [-r, r]^m$ is at least $1 - e^{-\sqrt{n}}$, which is overwhelming. Thus there exists an $\mathbf{x}' \neq 0$ (which is the i-th unit-vector) such that $\mathbf{e} = \mathbf{G}\mathbf{x}' + \mathbf{e}'$ and we are done.

1.3 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. Applebaum et al. [AIK11] noticed, that if the LWE-modulus q is super-polynomial, then the gaussian error-distribution can be "overridden" by a sufficiently (super-polynomially) wider rectangular uniform distribution. This however requires the underlying worst-case lattice-problems to be hard to approximate to within a super-polynomial factor. In [BPR12] an 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. To establish worst-case hardness of LWR, [BPR12] need to assume that the underlying worst-case lattice-problems are hard to approximate to within a super-polynomial factor. Bellare et al. [BKPW12] construct identity-based lossy trapdoor functions based on the hardness of the decisional-linear problem and LWE. The LWE-based construction of lossy trapdoor functions in [BKPW12] has some similarities with the lossy-codes construction in this work, though the technical details and analysis are incomparable. 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.

1.4 Concurrent Independent Work

Concurrently and independently of our work, Micciancio and Peikert [MP13] established a worst-case connection for LWE with short uniform errors. Specifically, [MP13] shows that a family of instantiations of LWE with short uniform errors, at most linear number of samples and polynomial modulus are as hard as approximating standard worst-case lattice problems to withing a factor of $\tilde{O}(\sqrt{n}q)$. For instance, their result can be instantiated with binary errors and $n \cdot (1 + \Omega(1/\log(n)))$ samples or polynomial errors (n^ϵ for some small ϵ) and a linear number of samples ($m = (1 + \epsilon/3)n$).

The main similarity of [MP13] and our work is on a conceptual level. In both [MP13] and our work a lossiness-argument is essential to establish the main result. To prove their result, Micciancio and Peikert restate the LWE problem in terms of SIS (Short Integer Solution) functions. This formulation states that, given a randomly drawn SIS-function $\mathbf{H} \in \mathbb{Z}_q^{(m-n)\times m}$ and $\mathbf{y} = \mathbf{H}\mathbf{x}$, where $\mathbf{x}$ is drawn from an input-distribution χ^n, it is hard to find $\mathbf{x}$. They establish lossiness of a pseudorandom SIS function-family by counting the number of elements in the image of functions $\mathbf{H}$ chosen from that family (on average). For appropriate parameter-choices, they can conclude that the image $\mathbf{H}(X)$ contains noticeably

fewer elements than the domain X, thus $\mathbf{H}$ must be lossy for the uniform distribution on its domain X.

For comparison, in the language of [MP13] our results might be restated as follows. We first construct pseudorandom SIS-functions $\mathbf{H}$ with domain $[-r,r]^m$ that have short vectors $\mathbf{g}_1,\ldots,\mathbf{g}_k$ (drawn from a gaussian distribution) in their kernel. Next, we show that elements $\mathbf{e}$ randomly chosen from $[-r,r]^m$ are *well behaved* in the sense that (with overwhelming probability) there exists a $\mathbf{g}_i$ such that $\mathbf{e}+\mathbf{g}_i \in [-r,r]^m$. Thus, $\mathbf{e}$ and $\mathbf{e}+\mathbf{g}_i$ form a collision for $\mathbf{H}$ (as $\mathbf{H}(\mathbf{e}+\mathbf{g}_i) = \mathbf{H}\mathbf{e}+\mathbf{H}\mathbf{g}_i = \mathbf{H}\mathbf{e}$) and we conclude that $\mathbf{H}$ loses at least 1-bit of information on the uniform distribution on $[-r,r]^m$.

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. Define the *min-entropy* 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 $(\alpha q)^2/(2\pi)$, where q is given by the context. 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 $\Phi_{\alpha q}$. If $t\alpha q \geq 2$, we can derive the tail-bound $\Pr[|Y| > t\alpha q] \leq \Pr[|X| > t\alpha q - 1] \leq \Pr[|X| > t\alpha q/2] \leq e^{-\pi t^2/4}$.

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 $q\mathbb{Z}^m$ 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 $\|\mathbf{x} \mod q\|_\infty = \|\mathbf{x}\|_\infty$.

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 $(\mathbf{A}, \mathbf{Ax} + \mathbf{e})$, to find $\mathbf{x}$.

We remark that most cryptographic applications of the LWE problem require only an a-priori fixed number of samples. For those applications, the formulation of Problem 1 poses no restriction. The notable exception to this are the

KDM-secure encryption scheme in [ACPS09] and the pseudorandom functions in [BPR12]. For both schemes the number of LWE-samples required is determined adversarially. 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 2 (Worst-to-Average Case Reduction [Reg05]). *Let n be a security parameter and $q = q(n)$ be a modulus, 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 independent 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 $\mathbf{Ax} + \mathbf{e}$ 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 3 (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 1. *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 of this work, lossy codes, and show that the existence of lossy codes implies that the associated decoding problems for random codes are hard.

Definition 1 (Families of Lossy Codes). *Let n be a security parameter, let $q = q(n)$ be a modulus, let $m = m(n) = \mathsf{poly}(n)$ and $\gamma = \gamma(n)$. Let $\{\mathcal{C}_{n,m,q}\}$ be a family of distributions where $\mathcal{C}_{n,m,q}$ is defined on $\mathbb{Z}_q^{m \times n}$ and let χ be a distribution on $\mathbb{Z}_q$. Finally, let $\mathcal{U}(\mathbb{Z}_q^{m \times n})$ be the uniform distribution on $\mathbb{Z}_q^{m \times n}$. We say that $\{\mathcal{C}_{n,m,q}\}$ is γ-lossy for the error-distribution χ, if the following 3 properties hold.*

1. *$\mathcal{C}_{n,m,q}$ **is pseudorandom:** It holds that $\mathcal{C}_{n,m,q} \approx_c \mathcal{U}(\mathbb{Z}_q^{m \times n})$.*
2. *$\mathcal{C}_{n,m,q}$ **is lossy:** Let $\mathbf{A}$ be distributed according to $\mathcal{C}_{n,m,q}$, $\mathbf{y} = \mathbf{A} \cdot \tilde{\mathbf{x}} + \tilde{\mathbf{e}}$ (where $\tilde{\mathbf{x}}$ is chosen uniformly from $\mathbb{Z}_q^n$ and $\tilde{\mathbf{e}}$ is distributed according to χ^m), let $\mathbf{x}$ be chosen uniformly from $\mathbb{Z}_q^n$ and let $\mathbf{e}$ be chosen according to χ^m. Then it holds that $\Pr_{(\mathbf{A},\mathbf{y})}[H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}) \geq \gamma] \geq 1 - \mathsf{negl}(n)$.*
3. *$\mathcal{U}(\mathbb{Z}_q^{m \times n})$ **is non-lossy:** Let $\mathbf{A}$ be distributed according to $\mathcal{U}(\mathbb{Z}_q^{m \times n})$, $\mathbf{y} = \mathbf{A} \cdot \tilde{\mathbf{x}} + \tilde{\mathbf{e}}$ (where $\tilde{\mathbf{x}}$ is chosen uniformly from $\mathbb{Z}_q^n$ and $\tilde{\mathbf{e}}$ is distributed according to χ^m), let $\mathbf{x}$ be chosen uniformly from $\mathbb{Z}_q^n$ and let $\mathbf{e}$ be chosen according to χ^m. Then it holds that $\Pr_{(\mathbf{A},\mathbf{y})}[H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}) = 0] \geq 1 - \mathsf{negl}(n)$.*

Remark Notice that while we require the error-distribution χ to be efficiently samplable, we do not require the distribution $\mathcal{C}_{n,m,q}$ of lossy codes to be efficiently samplable. In our construction in the next section however, $\mathcal{C}_{n,m,q}$ will be efficiently samplable.

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

Theorem 4. *Let n be a security-parameter, let $m = m(n) = \mathsf{poly}(n)$ and let $q = q(n)$ be a modulus. Let the distribution χ on $\mathbb{Z}_q$ be efficiently samplable.*

1. *Let χ be a uniform distribution with efficiently decidable support. Then the problem $\mathsf{LWE}(n, m, q, \chi)$ is hard, given that there exists a family of 1-lossy codes $\mathcal{C}_{n,m,q} \in \mathbb{Z}_q^{m \times n}$ for the error-distribution χ.*
2. *Let $\gamma = \gamma(n) = \omega(\log(n))$. Then $\mathsf{LWE}(n, m, q, \chi)$ is hard, given that there exists a family of γ-lossy codes $\mathcal{C}_{n,m,q} \in \mathbb{Z}_q^{m \times n}$ for the error-distribution χ.*

Proof. First notice that due to the non-lossiness property of $\mathcal{U}(\mathbb{Z}_q^{m \times n})$, instances of $\mathsf{LWE}(n, m, q, \chi)$ have a unique solution, except with negligible probability. For contradiction, let $\mathcal{A}$ be a PPT-adversary that solves $\mathsf{LWE}(n, m, q, \chi)$ with non-negligible probability ϵ.

We will begin with the first statement of the theorem. Using $\mathcal{A}$, we will construct a PPT-distinguisher $\mathcal{D}$ that distinguishes $\mathcal{C}_{n,m,q}$ and $\mathcal{U}(\mathbb{Z}_q^{m \times n})$ with

non-negligible advantage. Say that a solution $\mathbf{x}$ for an instance $(\mathbf{A}, \mathbf{y})$ is valid, if $\mathbf{y} - \mathbf{A} \cdot \mathbf{x}$ is in the support of the error-distribution χ.

There are two different behaviors that algorithm $\mathcal{A}$ could exhibit when receiving inputs of the form $(\mathbf{A}, \mathbf{y})$, where $\mathbf{A}$ is chosen from $\mathcal{C}_{n,m,q}$ and $\mathbf{y} = \mathbf{A}\mathbf{x}+\mathbf{e}$. In the first case, the probability that $\mathcal{A}$ outputs a valid solution $\mathbf{x}$ is negligible. In the second case, there exists a non-negligible ϵ' such that the probability that $\mathcal{A}$ outputs a valid solution $\mathbf{x}$ with probability at least ϵ'.

In the first case we can construct $\mathcal{D}$ as follows. Let $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ be $\mathcal{D}$'s input. It first samples $\mathbf{x}$ uniformly at random, samples $\mathbf{e}$ according to χ^m and sets $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$. It then runs $\mathcal{A}$ on input $(\mathbf{A}, \mathbf{y})$. If $\mathcal{A}$ outputs $\mathbf{x}$, $\mathcal{D}$ outputs 1, otherwise $\mathcal{D}$ outputs 0. Clearly, if $\mathbf{A}$ is chosen according to $\mathcal{U}(\mathbb{Z}_q^{m \times n})$, then $\mathcal{A}$ recovers $\mathbf{x}$ with probability at least ϵ. On the other hand, if $\mathbf{A}$ is chosen according to $\mathcal{C}_{n,m,q}$, then $\mathcal{A}$ recovers $\mathbf{x}$ only with negligible probability. Thus it holds that $\mathsf{Adv}(\mathcal{D}) = |\Pr[\mathcal{D}(\mathcal{U}(\mathbb{Z}_q^{m \times n})) = 1] - \Pr[\mathcal{D}(\mathcal{C}_{n,m,q}) = 1]| = \epsilon(n) - \mathsf{negl}(n)$, which is non-negligible.

In the second case, we construct $\mathcal{D}$ differently. Let $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ be $\mathcal{D}$'s input. $\mathcal{D}$ samples $\mathbf{x}$ uniformly at random, $\mathbf{e}$ according to χ^m and sets $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$. It then runs $\mathcal{A}$ on input $(\mathbf{A}, \mathbf{y})$. If $\mathcal{A}$ outputs an $\mathbf{x}' \neq \mathbf{x}$ such that $\mathbf{e}' = \mathbf{y} - \mathbf{A}\mathbf{x}'$ is in the support of χ^m, it outputs 1, otherwise 0. First, observe that such a *collision* $\mathbf{x}' \neq \mathbf{x}$ cannot exist (except with negligible probability) if $\mathbf{A}$ is chosen according to the uniform distribution $\mathcal{U}(\mathbb{Z}_q^{m \times n})$. This is due to the non-lossiness property of $\mathcal{U}(\mathbb{Z}_q^{m \times n})$. On the other hand, consider that $\mathbf{A}$ is chosen according to $\mathcal{C}_{n,m,q}$. Then it holds (with overwhelming probability) that $H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}) \geq 1$. Thus it holds (even for an unbounded $\mathcal{A}$) that $\mathcal{A}$ outputs the same $\mathbf{x}$ that was chosen by $\mathcal{D}$ with probability at most $1/2$, conditioned that $\mathcal{A}$ outputs a valid $\mathbf{x}$. Thus, conditioned that $\mathcal{A}$ gives a valid output, there is a chance of $1/2$ that $\mathcal{A}$ outputs a valid $\mathbf{x}' \neq \mathbf{x}$. As $\mathcal{A}$ gives a valid output with probability at least ϵ', $\mathcal{A}$ outputs a collision $\mathbf{x}'$ with probability at least $\epsilon'/2$. Thus $\mathcal{D}$ distinguishes $\mathcal{U}(\mathbb{Z}_q^{m \times n})$ from $\mathcal{C}_{n,m,q}$ with advantage at least $\epsilon'/2$, which is non-negligible.

We now turn to the second statement of the theorem. In this case the construction of the distinguisher $\mathcal{D}$ is straightforward. Let $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ be $\mathcal{D}$'s input. As before, $\mathcal{D}$ samples $\mathbf{x}$ uniformly at random, $\mathbf{e}$ according to χ^m, sets $\mathbf{y} = \mathbf{A}\mathbf{x}+\mathbf{e}$ and runs $\mathcal{A}$ on input $(\mathbf{A}, \mathbf{y})$. If $\mathcal{A}$ outputs $\mathbf{x}$ it outputs 1, otherwise 0. Again, if $\mathbf{A}$ was chosen from $\mathcal{U}(\mathbb{Z}_q^{m \times n})$, then $\mathcal{A}$ outputs $\mathbf{x}$ (which is in this case unique) with probability at least ϵ. On the other hand, if $\mathbf{A}$ comes from the distribution $\mathcal{C}_{n,m,q}$, then $\mathcal{A}$ finds $\mathbf{x}$ with probability at most $2^{-H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x}+\mathbf{e}=\mathbf{y})} \leq 2^{-\gamma(n)}$ (this holds with overwhelming probability in the choice of $\mathbf{A}$ and $\mathbf{y}$), which is negligible (as $\gamma(n) = \omega(\log(n))$). All together, $\mathcal{D}$ distinguishes $\mathcal{U}(\mathbb{Z}_q^{m \times n})$ from $\mathcal{C}_{n,m,q}$ with advantage at least $\epsilon - 2^{-\gamma}$, which is non-negligible.

4 Construction of Lossy Codes for Uniform Errors from Standard-LWE

We will now provide the details of the construction outlined in Section 1.2.

Construction 1 *Let n be a security parameter, let $q = q(n)$ be a modulus, $m = m(n) = \mathsf{poly}(n)$ and $k = k(n) \leq n$. The distribution $\mathcal{C}_{n,m,q,k,\alpha}$ defined on $\mathbb{Z}_q^{m\times n}$ is specified as follows. Choose $\mathbf{A}' \in \mathbb{Z}_q^{m\times(n-k)}$ uniformly at random, choose $\mathbf{T}' \in \mathbb{Z}_q^{(n-k)\times(n-k)}$ uniformly at random and sample $\mathbf{G} \in \mathbb{Z}_q^{m\times k}$ from $\bar{\Psi}_\alpha^{m\times k}$. Output $\mathbf{A} = (\mathbf{A}'\|\mathbf{A}'\mathbf{T}' + \mathbf{G})$.*

We will show that, for certain parameter choices, Construction 1 yields a lossy code for the error-distribution $\mathcal{U}([-r, r])$. The pseudorandomness of the distribution $\mathcal{C}_{n,m,q,k,\alpha}$ can be established directly from Lemma 1, assuming the hardness of $\mathsf{LWE}(n, m, q, \bar{\Psi}_\alpha)$.

Lemma 2. *Let n be a security-parameter, let $q = q(n)$ be a modulus, let $m = m(n) = \mathsf{poly}(n)$, let $k = \lceil \beta n \rceil$ for some constant $\beta \in (0, 1)$ and let $\alpha = \alpha(n) \in (0, 1)$ with $\alpha q \geq 2\sqrt{n}$. Assuming that $\mathsf{LWE}(n, m, q, \bar{\Psi}_\alpha)$ is hard, it holds that $\mathcal{C}_{n,m,q,k,\alpha} \approx_c \mathcal{U}(\mathbb{Z}_q^{m\times n})$.*

The non-lossiness of the truly random distribution $\mathcal{U}(\mathbb{Z}_q^{m\times n})$ can be established by a simple Gilbert-Varshamov-type argument.

Lemma 3. *Let n be a security parameter, let $\tau > 0$ be a constant, $r = \mathsf{poly}(n)$, $q > (4r + 1)^{1+\tau}$ and $m > (1 + 2/\tau)n$. Let $\mathbf{A}$ be chosen from $\mathcal{U}(\mathbb{Z}_q^{m\times n})$. Then the shortest vector of the lattice $\Lambda_q(\mathbf{A})$ has length (in the $\|\cdot\|_\infty$-norm) greater than $2r$, except with negligible probability. Thus it holds that $\Pr_{(\mathbf{A},\mathbf{y})}[H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}) = 0] \geq 1 - \mathsf{negl}(n)$.*

Proof. Let $\mathbf{A} \in \mathbb{Z}_q^{m\times n}$ be chosen uniformly at random. Clearly, it holds that $H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}) = 0$ if the length of the shortest vector in $\Lambda_q(\mathbf{A})$ (in the $\|\cdot\|_\infty$-norm) is greater than $2r$. Now fix a vector $\mathbf{x} \neq 0 \in \mathbb{Z}_q^n$. Then the vector $\mathbf{A}\cdot\mathbf{x}$ is distributed uniformly at random in $\mathbb{Z}_q^m$. Thus it holds that $\Pr_{\mathbf{A}}[\|\mathbf{A} \cdot \mathbf{x}\|_\infty \leq 2r] \leq \left(\frac{4r+1}{q}\right)^m$. Thus, a union-bound yields that $\Pr[\exists x \neq 0 \in \mathbb{Z}_q^n : \|\mathbf{A}\mathbf{x}\|_\infty \leq r] \leq \frac{(4r+1)^m}{q^{m-n}}$. This expression is negligible whenever $(m - n)\log(q) - m\log(4r + 1) > \omega(\log(n))$. This is certainly the case if $r = \mathsf{poly}(n)$, $q > (4r + 1)^{1+\tau}$ and $m > (1 + 2/\tau))n$ for some constant $\tau > 0$.

We now turn to showing that $\mathcal{C}_{n,m,q,k,\alpha}$ fulfills the lossiness-requirement.

Definition 2. *We say that a vector $\mathbf{y} \in \mathbb{Z}_q^m$ is N-*ambiguous *for a matrix $\mathbf{A}$ and a distance r, if $|\{\mathbf{x} \in \mathbb{Z}_q^n | \|\mathbf{y} - \mathbf{A} \cdot \mathbf{x}\|_\infty \leq r\}| \geq N$. If $\mathbf{A}$ and r are clear by context, we just say that $\mathbf{y}$ is N-ambiguous.*

Notice that if $\mathbf{y}$ is N-ambiguous, then for every $\mathbf{z} \in \mathbb{Z}_q^n$ by linearity it holds that $\mathbf{y} + \mathbf{A}\mathbf{z}$ is also N-ambiguous. Since we want to establish lossiness for the uniform 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 n be a security parameter, let $q = q(n)$ be a modulus and let $r = r(n)$ and $N = N(n)$. Fix a matrix $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$. Let $\mathbf{y} \in \mathbb{Z}_q^m$ be N-ambiguous for the matrix $\mathbf{A}$ and distance r. Let $\mathbf{x} \in \mathbb{Z}_q^n$ be chosen uniformly at random and $\mathbf{e}$ be distributed according to $\mathcal{U}([-r,r])^m$. Then it holds that $H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}) \geq \log(N)$.*

Proof. Since $\mathbf{x}$ and $\mathbf{e}$ are drawn from uniform distributions, $p := \Pr[\mathbf{x} = \tilde{\mathbf{x}}, \mathbf{e} = \tilde{\mathbf{e}}]$ is the same for all $\tilde{\mathbf{x}} \in \mathbb{Z}_q^n$ and $\tilde{\mathbf{e}} \in [-r,r]^m$. Let $X := \{\mathbf{z} \in \mathbb{Z}_q^n | \|\mathbf{y} - \mathbf{A}\mathbf{z}\| \leq r\}$. As $\mathbf{y}$ is N-ambiguous it holds that $|X| \geq N$, thus

$$\Pr[\mathbf{A}\mathbf{x}+\mathbf{e} = \mathbf{y}] = \sum_{\mathbf{z} \in \mathbb{Z}_q^n} \Pr[\mathbf{A}\mathbf{x}+\mathbf{e} = \mathbf{y}, \mathbf{x} = \mathbf{z}] = \sum_{\mathbf{z} \in X} \Pr[\mathbf{e} = \mathbf{y}-\mathbf{A}\mathbf{z}, \mathbf{x} = \mathbf{z}] \geq p \cdot N.$$

Thus it holds for all $\mathbf{z} \in \mathbb{Z}_q^n$ that

$$\Pr[\mathbf{x} = \mathbf{z}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}] = \frac{\Pr[\mathbf{x} = \mathbf{z}, \mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}]}{\Pr[\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}]} \leq \frac{1}{N}.$$

This immediately implies $H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}) \geq \log(N)$, which concludes the proof.

The following lemma shows that if we sample $\mathbf{e}$ from $\mathcal{U}([-r,r])^m$, then with overwhelming probability $\mathbf{e}$ is such that if we add a sample $\mathbf{g}$ from an appropriately bounded distribution χ^m, $\mathbf{e} - \mathbf{g}$ is, with substantial probability over the choice of $\mathbf{g}$, also in $[-r,r]^m$. Say that a distribution χ is strictly B-bounded if the support of χ is contained in $[-B,B]$.

Lemma 5. *Let $n, m, B > 0$ be integers, let $r > (m+1)B$ and let $\epsilon < 1/2$. Let χ be a strictly B-bounded symmetrical distribution on $\mathbb{Z}$. Let $\mathbf{e}$ be chosen uniformly at random from $[-r,r]^m$ and let $\mathbf{g}$ be distributed according to χ^m. Then it holds that*

$$\Pr_{\mathbf{e}}\left[\Pr_{\mathbf{g}}[\|\mathbf{e} - \mathbf{g}\|_\infty \leq r] \geq \epsilon\right] \geq 1 - m \cdot \epsilon^{\log(r/(m \cdot B))}.$$

Proof. We will first bound the probability that it holds for more than $k = \lfloor -\log(\epsilon) \rfloor$ components e_i of $\mathbf{e}$ that $|e_i| > r - B$, i.e. that e_i is not in the interval $[-r+B, r-B]$. For $i = 1, \ldots, m$ let Z_i be a random-variable that is 1 if $|e_i| > r - B$ and 0 otherwise. As $e_1, \ldots, e_m$ are iid., $Z_1, \ldots, Z_m$ are also iid. Thus let $p = \Pr[Z_1 = 1] = \cdots = \Pr[Z_m = 1]$. As e_1 is distributed by $\mathcal{U}([-r,r])$ and $p = \Pr[Z_1 = 1] = \Pr[|e_1| > r - B]$ it holds that $(B-1)/r \leq p \leq B/r$. Set $Z = \sum_{i=1}^m Z_i$. Clearly, Z is the number of components of $\mathbf{e}$ that are not in the interval $[-r+B, r-B]$ and it is binomially distributed. We can bound the probability $\Pr[Z > k]$ by

$$\Pr[Z > k] = \sum_{i=k+1}^{m} \binom{m}{i} p^i (1-p)^{m-i} \overset{(1)}{\leq} m \underbrace{\binom{m}{k+1}}_{\leq m^{k+1}} \underbrace{p^{k+1}}_{\leq (B/r)^{k+1}} \underbrace{(1-p)^{m-k-1}}_{\leq 1}$$

$$\leq m \cdot \left(\frac{m \cdot B}{r}\right)^{k+1} \overset{(2)}{<} m \cdot \left(\frac{m \cdot B}{r}\right)^{-\log(\epsilon)} = m \cdot \epsilon^{\log(r/(m \cdot B))}.$$

Inequality (1) holds, as $\binom{m}{i}p^i(1-p)^{m-i}$ is monotonically decreasing for $i \geq \lfloor (m+1)p \rfloor \geq \lfloor (m+1)(B-1)/r \rfloor = 0$. Inequality (2) holds as $m \cdot B/r < 1$ and $k+1 > -\log(\epsilon)$.

Now, fix an $\mathbf{e}$ and assume that it holds that it holds for at most k components $e_{i_1}, \ldots, e_{i_k}$ of $\mathbf{e}$ that $|e_{i_j}| > r - B$. Let $i \in \{i_1, \ldots, i_k\}$. If $\mathsf{sgn}(g_i) = \mathsf{sgn}(e_i)$, then it holds that $|e_i - g_i| = |e_i| - |g_i| \leq |e_i| \leq r$. As χ is a symmetrical distribution, it holds that $\Pr[\mathsf{sgn}(g_i) = \mathsf{sgn}(e_i)] \geq \frac{1}{2}$. Therefore, it holds that $\Pr[|e_i - g_i| \leq r] \geq \frac{1}{2}$. For all other indexes $j \notin \{i_1, \ldots, i_k\}$ it holds that $|e_j| \leq r - B$. The triangle-inequality yields $|e_j - g_j| \leq |e_j| + |g_j| \leq r - B + B = r$. Therefore, we have that $\Pr[|e_j - g_j| \leq r] = 1$. Putting this together, we get that

$$\Pr[\|\mathbf{e} - \mathbf{g}\|_\infty \leq r] = \prod_{i=1}^{m} \Pr[|e_i - g_i| \leq r] \geq 2^{-k} \geq \epsilon.$$

All together, it holds that

$$\Pr_{\mathbf{e}}[\Pr_{\mathbf{g}}[\|\mathbf{e} - \mathbf{g}\|_\infty \leq r] \geq \epsilon] \geq 1 - m \cdot \epsilon^{\log(r/(m \cdot B))},$$

which concludes the proof.

We can now show that Construction 1 also suffices the lossiness-condition, for appropriate parameters.

Lemma 6. *Let n be a security-parameter, let $m = m(n) = \mathsf{poly}(n)$, let $k = k(n) = \lceil \beta n \rceil$ for some constant $\beta \in (0,1)$ and let $c \in (0,1)$ be a constant. Let $q = q(n)$ be a modulus, $\alpha = \alpha(n) \in (0,1)$, let $B = B(n)$ and assume that the distribution $\bar{\Psi}_\alpha$ is B-bounded, except with negligible probability. Finally let $r = r(n) > 0$ be such that $r \geq m \cdot Bn^c$.*

Let $\mathbf{G}$ be chosen according to $\bar{\Psi}_\alpha^{m \times k}$, let the matrix $\mathbf{A}'$ be distributed according to $\mathcal{U}(\mathbb{Z}_q^{m \times (n-k)})$, $\mathbf{T}'$ be distributed according to $\mathcal{U}(\mathbb{Z}_q^{(n-k)\times(n-k)})$ and let $\mathbf{A} = (\mathbf{A}' \| \mathbf{A}'\mathbf{T}' + \mathbf{G})$. Let $\mathbf{y} = \mathbf{A}\mathbf{x}' + \mathbf{e}'$, with $\mathbf{x}'$ chosen uniformly from $\mathbb{Z}_q^n$ and $\mathbf{e}'$ chosen from $\mathcal{U}([-r,r])^m$. Also let $\mathbf{x}$ be chosen uniformly from $\mathbb{Z}_q^n$ and $\mathbf{e}$ be chosen from $\mathcal{U}([-r,r])^m$. Then it holds that $\Pr_{(\mathbf{A},\mathbf{y})}[H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{y}) \geq 1] \geq 1 - \mathsf{negl}(n)$.

Proof. Assume first that $\mathbf{G}$ was chosen from $\chi^{m \times k}$, where χ is a symmetrical strictly B-bounded distribution. Fix an $\mathbf{e}'$ with $\Pr_{\mathbf{g}}[\|\mathbf{e}' - \mathbf{g}\|_\infty \leq r] \geq n^{-c}$, where $\mathbf{g}$ is distributed according to χ^m. Let $\mathbf{g}_1, \ldots, \mathbf{g}_k$ be the columns of $\mathbf{G}$, thus each $\mathbf{g}_i$ is distributed according to χ^m. We first show that $\mathbf{e}'$ is 2-ambiguous for the matrix $\mathbf{G}$ and distance r with high probability over the choice of $\mathbf{G}$. If there is at least one column $\mathbf{g}_i$ of $\mathbf{G}$ such that $\|\mathbf{e}' - \mathbf{g}_i\|_\infty \leq r$, then $\|\mathbf{e}' - \mathbf{G}\mathbf{x}^{(i)}\|_\infty \leq r$ (where $\mathbf{x}^{(i)}$ is the i-th unit vector) and we have that $\mathbf{e}'$ is 2-ambiguous. Here $\mathbf{x}_1 = 0$ and $\mathbf{x}_2 = \mathbf{x}^{(i)}$ are two different points satisfying $\|\mathbf{e}' - \mathbf{G} \cdot \mathbf{x}\|_\infty \leq r$.

The probability of the event that it holds for all $i = 1, \ldots, k$ that $\|\mathbf{e}' - \mathbf{g}_i\|_\infty > r$ is at most $(1 - n^{-c})^k \leq e^{-k \cdot n^{-c}} \leq e^{-\beta \cdot n^{1-c}}$. Thus we have that $\Pr[\mathbf{e}' \text{ 2-ambiguous for } \mathbf{G}] \geq 1 - e^{-\beta n^{1-c}}$. The same holds for the matrix $\mathbf{A} =$

$(\mathbf{A}'\|\mathbf{A}'\mathbf{T}'+\mathbf{G})$, as we can obtain $\mathbf{A}$ from $\mathbf{G}$ by appending extra columns and applying a basis-change. Both operations straightforwardly do not decrease the ambiguity. Therefore it holds $\Pr[\mathbf{e}' \text{ 2-ambiguous for } \mathbf{A}] \geq 1 - e^{-\beta n^{1-c}}$

Now let $\mathbf{e}'$ be distributed by $\mathcal{U}([-r,r])^m$. It holds that $r \geq m \cdot B \cdot n^c > (m+1)B$ and $\epsilon := n^{-c} < 1/2$ for sufficiently large n. Thus the above and Lemma 5 imply that $\Pr_{\mathbf{e}'}[\Pr_{\mathbf{A}}[\mathbf{e}' \text{ 2-ambiguous for } \mathbf{A}] \geq 1 - e^{-\beta n^{1-c}}] \geq 1 - m \cdot n^{-c \cdot \log(r/(m \cdot B))} = 1 - m \cdot n^{-c^2 \log(n)}$. This immediately yields $\Pr_{\mathbf{A},\mathbf{e}'}[\mathbf{e}' \text{ 2-ambiguous for } \mathbf{A}] \geq 1 - e^{-\beta n^{1-c}} - m \cdot n^{-c^2 \log(n)} = 1 - \mathsf{negl}(n)$. By linearity, this holds also if we shift $\mathbf{e}'$ by $\mathbf{A}\mathbf{x}'$ for any $\mathbf{x}' \in \mathbb{Z}_q^n$. Thus we get $\Pr_{\mathbf{A},\mathbf{y}}[\mathbf{y} \text{ 2-ambiguous for } \mathbf{A}] \geq 1 - \mathsf{negl}(n)$. Now, since $\bar{\Psi}_\alpha$ is statistically close to symmetrical strictly B-bounded distribution χ (which can be sampled by rejecting samples of $\bar{\Psi}_\alpha$ greater than B), this probability drops at most by a negligible amount if we sample $\mathbf{G}$ from $\bar{\Psi}_\alpha^{m \times k}$. Applying Lemma 4 yields $\Pr_{(\mathbf{A},\mathbf{y})}[H_\infty(\mathbf{x}|\mathbf{A}\mathbf{x}+\mathbf{e}=y) \geq 1] \geq 1 - \mathsf{negl}(n)$, which concludes the proof.

We will summarize the statements of Lemmas 2, 3 and 6 in the following theorem.

Theorem 5. *Let n be a security-parameter and let $\sigma \in (0,1)$, $\beta \in (0,1)$ and $\tau \geq 1$ be constants. Let $q = q(n)$, $m = m(n) = \mathsf{poly}(n)$, $k = \lceil \beta n \rceil$, $\alpha = \alpha(n) \in (0,1)$ and $r = r(n)$ be such that the following holds*

- $m \geq (1 + 2/\tau)n$
- $r \geq 2mn^{0.5+\sigma}$
- $q > (4r+1)^\tau$
- $2\sqrt{n} \leq \alpha q \leq \frac{r}{mn^\sigma}$

Then the distribution $\mathcal{C}_{n,m,q,k,\alpha}$ given in Construction 1 is 1-lossy for the error-distribution $\mathcal{U}[-r,r]$, provided that $\mathsf{LWE}(n,m,q,\bar{\Psi}_\alpha)$ is hard.

Proof. It is straightforward that this parameter-set satisfies the requirements of Lemmas 2 and 3, thus both lemmas holds. A gaussian tail-bound yields that $\bar{\Psi}_\alpha$ is $B = \alpha q n^{\sigma/2}$-bounded, except with probability $e^{-\pi n^\sigma/4}$, which is negligible. Item 4 above implies that $r \geq mn^\sigma \alpha q = mBn^{\sigma/2}$, thus setting $c = \sigma/2$ we can apply Lemma 6 and the claim follows.

4.1 LWE with Uniform Errors

Using Theorems 4 and 5 we can translate the worst-case connection for standard-LWE (Theorem 2) to LWE with uniform errors. For simplicity, we will set $\tau = 1$, $\beta = 1/2$ and $r = \rho \cdot q$, for a parameter $\rho = \rho(n) \in (0, 1/10)$.

Theorem 6 (Main Theorem). *Let n be a security parameter and let $\sigma \in (0,1)$ be an arbitrarily small constant. Let $q = q(n)$ be a modulus and $m = m(n) = \mathsf{poly}(n)$ be a integer with $m \geq 3n$. Let $\rho = \rho(n) \in (0, 1/10)$ be such that $\rho q \geq 2n^{0.5+\sigma}m$. If 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 $\tilde{O}(n^{1+\sigma}m/\rho)$ in the worst case.

Proof. Set $\alpha = \alpha(n) = \frac{\rho}{mn^{\sigma}}$. Then the requirements of Theorem 5 are fulfilled:

- For $\tau = 1$ it holds that $m \geq 3n$
- $r = \rho q \geq 2mn^{0.5+\sigma}$
- $q > 4r + 1$ is equivalent to $r < (q-1)/4$, which holds for $r = \rho q < q/10$ and $q \geq 2$.
- $\alpha q = \frac{\rho q}{mn^{\sigma}} \geq 2\sqrt{n}$ and $\alpha q = \frac{\rho q}{mn^{\sigma}} = \frac{r}{mn^{\sigma}}$.

Thus by Theorem 5 there exists a 1-lossy code $\mathcal{C}$ for the error-distribution $\mathcal{U}([-\rho q, \rho q])$, provided that $\mathsf{LWE}(n, m, q, \bar{\Psi}_{\alpha})$ is hard. The uniform distribution $\mathcal{U}([-\rho q, \rho q])^m$ clearly has efficiently decidable support, and so the first statement of Theorem 4 yields that $\mathsf{LWE}(n, m, q, \mathcal{U}[-\rho q, \rho q])$ is at least as hard as $\mathsf{LWE}(n, m, q, \bar{\Psi}_{\alpha})$. Thus, setting $\alpha = \frac{\rho}{mn^{\sigma}}$ the claim follows by Theorem 2.

Using the search-to-decision reduction of Theorem 3, we can establish the hardness of the decisional LWE problem with error-distribution $\mathcal{U}([-\rho q, \rho q])$. We therefore need to restrict q to be a polynomially small prime integer.

Corollary 1. *Let n be a security parameter and let $\sigma \in (0, 1)$ be an arbitrarily small constant. Let $q = q(n)$ be a modulus and $m = m(n) = \mathsf{poly}(n)$ be a integer with $m \geq 3n$. Let $\rho = \rho(n) \in (0, 1/10)$ be such that $\rho q \geq 2n^{0.5+\sigma}m$. If there exists a PPT-distinguisher that distinguishes $\mathsf{DLWE}(n, m, q, \mathcal{U}([-\rho q, \rho q]))$ 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 $\tilde{O}(n^{1+\sigma}m/\rho)$ in the worst case.*

5 Conclusion

This work presented the first worst-to-average-case reduction for an LWE variant with polynomial modulus and uniformly distributed errors, thereby answering a question from Micciancio 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. Overcoming this limitation poses an interesting open problem. The main ingredient in our proof is a new tool called lossy codes, i.e., codes which lose information when decoding noisy code words. Another interesting question is, if these techniques carry over to hardness assumptions for binary codes.

6 Acknowledgement

The authors would like to thank Daniele Micciancio, Chris Peikert and the anonymous reviewers of Eurocrypt 2013 for valuable comments on this work.

The authors are especially grateful to one particular reviewer who suggested the first statement of Theorem 4, thereby allowing a major simplification in the proof of Lemma 6. Nico Döttling was supported by IBM Research and Development Germany within the Homer-Project.

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.

[AIK11] Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. How to garble arithmetic circuits. In *FOCS*, pages 120–129, 2011.

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

[BKPW12] Mihir Bellare, Eike Kiltz, Chris Peikert, and Brent Waters. Identity-based (lossy) trapdoor functions and applications. In *EUROCRYPT*, pages 228–245, 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.

[MP13] Daniele Micciancio and Chris Peikert. Hardness of SIS and LWE with Small Parameters. *IACR Cryptology ePrint Archive*, 2013:069, 2013.

[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.