THEORY OF COMPUTING, Volume 14 (13), 2018, pp. 1–17
www.theoryofcomputing.org

On the Hardness of Learning With Errors with Binary Secrets

Daniele Micciancio*

Received September 17, 2017; Revised October 13, 2018; Published November 30, 2018

Abstract: We give a simple proof that the decisional Learning With Errors (LWE) problem with binary secrets (and an arbitrary polynomial number of samples) is at least as hard as the standard LWE problem (with unrestricted, uniformly random secrets, and a bounded, quasi-linear number of samples). This proves that the binary-secret LWE distribution is pseudorandom, under standard worst-case complexity assumptions on lattice problems. Our results are similar to those proved by Brakerski, Langlois, Peikert, Regev and Stehlé (STOC 2013), but provide a shorter, more direct proof, and a small improvement in the noise growth of the reduction.

1 Introduction

The Learning With Errors (LWE) problem [21, 22] plays a central role in lattice cryptography, its secure instantiation, and its most advanced applications. The usefulness of LWE in cryptography is due in large part to its pseudorandomness properties, captured by the standard decisional LWE problem defined as follows. An LWE instance is described by a matrix $\mathbf{A} \in \mathbb{Z}_q^{m\times n}$ (chosen uniformly at random) and a vector

*Research supported in part by the Defense Advanced Research Project Agency (DARPA) and the U.S. Army Research Office under the SafeWare program, and the National Science Foundation (NSF) under grant CNS-1528068. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect reflect the views, position or policy of the Government.

ACM Classification: F.2.2, F.1.3

AMS Classification: 68Q17, 52C07, 11H06

Key words and phrases: complexity theory, cryptography, pseudorandomness, lattice, learning, LWE

 DOI: 10.4086/toc.2018.v014a013

$\mathbf{b} \in \mathbb{Z}_q^m$ which may be chosen either uniformly at random, or as $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \pmod q$, where $\mathbf{s} \in \mathbb{Z}_q^n$ is a random secret and $\mathbf{e} \in \mathbb{Z}^m$ is a "small" error vector, typically chosen with independent discrete Gaussian entries of standard deviation $\sigma \approx \sqrt{n}$. The (Decisional) LWE problem asks to distinguish between these two cases.

Several variants of LWE exist in the literature, depending on how $\mathbf{s}$ and $\mathbf{e}$ are chosen, all motivated by specific cryptographic applications. In the most standard formulation of LWE, the secret $\mathbf{s} \in \mathbb{Z}_q^n$ is chosen uniformly at random. But this is often undesirable in many cryptographic applications, e. g., those making use of modulus-switching techniques, where large secrets result in substantial ciphertext quality degradation. Ideally, it would be best to choose $\mathbf{s} \in \{0,1\}^n$ as a vector with binary entries, as used for example in many Fully Homomorphic Encryption schemes (e. g., see [9, 8]). This binary-secret LWE also plays a fundamental role in theoretical studies, like the proof that LWE is leakage resilient [11], and the proof that LWE with polynomial modulus q is at least as hard as worst-case lattice problems under classical (i. e., non-quantum) reductions [6].

This last work [6] is the best currently known hardness result for binary-secret LWE, and gives a reduction from (standard) LWE with arbitrary secret in $\mathbb{Z}_q^n$, to LWE with secret in $\{0,1\}^{n \log q}$, i. e., the secret can be restricted to binary vectors at the cost of increasing the dimension[1] from n to $n \log q$. This reduction is a major part of the main result of [6] on the classical hardness of LWE, and takes a good half of that paper, going through a careful hybrid argument involving some technical ("first-is-errorless" and "extended-LWE") problem variants.

In this paper we present a direct and substantially shorter proof of this important result. In fact, while the proof of this result given in [6] is over 7 pages long, spanning multiple subsections, and involving a number of intermediate problems, our proof has a more direct structure and it is much shorter. A key insight leading to our simpler proof is the formulation of the binary LWE problem (denoted $\mathbf{LWE}_\pm$) using secrets in $\{\pm 1\}^n$, rather than $\{0,1\}^n$. This is easily seen (for odd[2] modulus q) to be equivalent to the more common $\{0,1\}$ formulation via the affine transformation $s \mapsto (2s-1)$, but has the technical advantage that all secrets have exactly the same Euclidean length, simplifying the application of discrete Gaussian convolution theorems. Given the equivalence between the two problems, we will keep referring to $\mathbf{LWE}_\pm$ informally as the binary LWE problem. Other than presenting a simpler and shorter proof, we do not claim any new results over previous work: our results, and the range of parameters for which we reduce $\mathbf{LWE}$ to $\mathbf{LWE}_\pm$, are essentially the same as in [6, Theorem 4.1], except possibly for reducing some constants, e. g., in our reduction the error grows by a factor $2\sqrt{n+1}$, while in [6, Theorem 4.1] it grows by $\sqrt{10n}$.

Given the important role played by binary LWE in many cryptographic applications, we hope that our simplified treatment will make the theoretical hardness of this problem more easily accessible, and stimulate further research.

Related work. The LWE problem with small secret was first formally considered by Applebaum, Cash, Peikert and Sahai in [3], who proved that, without loss of generality, one may assume that the secret

[1] As remarked in [6], this increase seems unavoidable, as it preserves the bit-length of the secret.

[2] Hardness results for binary LWE with *even* modulus q are easily obtained by modulus switching, i. e., scaling and (randomly) rounding each entry $x \in \mathbb{Z}_q$ of $\mathbf{A}$ or $\mathbf{b}$ $\lfloor x \cdot ((q-1)/q) \rceil_\$ \in \mathbb{Z}_{q-1}$. This increases the error roughly by an additive term $O(\|\mathbf{s}\|)$, which is small because $\mathbf{s}$ is a binary secret.

follows the same distribution as the LWE errors. This allows the secret coordinates to be as small as $\sqrt{n}$, but not as small as $\{0,1\}$. For a list of applications using LWE with small secrets see [1].

Reducing LWE to have a binary secret was first considered by Goldwasser, Kalai, Peikert and Vaikuntanathan in [11], motivated by questions in leakage-resilient cryptography, where the problem is proved hard using "noise-flooding" techniques. A stronger reduction is given by Brakerski, Langlois, Peikert, Regev and Stehlé in [6], in the context of proving classical hardness results for LWE.

A different (and much harder) problem is that of proving that LWE is computationally hard when the *error* (and not just the secret) follows the binary distribution [17, 7]. In fact, LWE with small errors can be efficiently solved when sufficiently many samples are available [4, 2, 13]. In this paper, we do not study LWE with binary errors.

Attacks against LWE with binary secret (and Gaussian errors) are considered in [1, 5]. Theoretically, the secret can be assumed binary by increasing the LWE dimension to $n \log q$ [6], but experimental results in [5] suggest that, heuristically, increasing the secret dimension by a $\log \log n$ factor may already be enough to counter the best known cryptanalytic attacks for common parameter settings.

Paper organization. In Section 2 we introduce the notation used in this paper, provide a formal definition of the LWE problem, and present some background results, including a simple lemma on the projection of discrete Gaussians (Lemma 2.6), and the construction of a gadget matrix needed in our main reduction (Lemma 2.7). The proof that LWE with binary secrets is pseudorandom is given in Section 3. Section 4 concludes with a discussion of open problems.

2 Preliminaries

We use bold lowercase letters $\mathbf{a}$ for vectors, and bold uppercase $\mathbf{A}$ for matrices. Probability distributions are denoted using calligraphic letters $\mathcal{A}$. We write vectors as columns $\mathbf{v} \in \mathbb{Z}^n = \mathbb{Z}^{n\times 1}$. The transpose of a vector or matrix $\mathbf{A}$ is denoted $\mathbf{A}^t$. We write $[\mathbf{A}_1,\ldots,\mathbf{A}_n]$ for the horizontal concatenation of matrices $\mathbf{A}_i \in \mathbb{Z}^{k\times m_i}$, and use transpose notation $[\mathbf{A}_1,\ldots,\mathbf{A}_n]^t$ for the vertical concatenation of $\mathbf{A}_1^t,\ldots,\mathbf{A}_n^t$. Let $\mathbf{e}_1,\ldots,\mathbf{e}_n$ be the standard basis of $\mathbb{Z}^n$, $\mathbf{I} = [\mathbf{e}_1,\ldots,\mathbf{e}_n]$ the $n\times n$ identity matrix, and $\mathbf{u} = \sum_i \mathbf{e}_i$ the all-ones vector. The Euclidean norm of a vector is

$$\|\mathbf{x}\| = \sqrt{\sum_i x_i^2}\,,$$

and the max norm is $\|\mathbf{x}\|_\infty = \max_i |x_i|$.

For any vector $\mathbf{z} = (z_1,\ldots,z_n) \in \mathbb{Z}^n$, we write $\mathbf{diag}(\mathbf{z}) = [z_1\cdot\mathbf{e}_1,\ldots,z_n\cdot\mathbf{e}_n]$ for the diagonal matrix with the entries of $\mathbf{z}$ along the diagonal. So, for example, $\mathbf{diag}(\mathbf{u}) = \mathbf{I}$. For any integer matrix $\mathbf{Q} \in \mathbb{Z}^{n\times m}$ and for any positive integer $k \le m$, we write $\mathbf{Q}_{[k]}$ for the matrix consisting of the first k columns of $\mathbf{Q}$, and $\mathbf{Q}_{]k[}$ for the matrix obtained by removing the first k columns from $\mathbf{Q}$. So, $\mathbf{Q} = [\mathbf{Q}_{[k]},\mathbf{Q}_{]k[}]$ where $\mathbf{Q}_{[k]} \in \mathbb{Z}^{n\times k}$ and $\mathbf{Q}_{]k[} \in \mathbb{Z}^{n\times(m-k)}$.

For any integer matrix $\mathbf{Q} \in \mathbb{Z}^{k\times m}$, we write $\ker(\mathbf{Q}) = \{\mathbf{x} \in \mathbb{Z}^m \colon \mathbf{Qx} = \mathbf{0}\}$ for the kernel of $\mathbf{Q}\colon \mathbb{Z}^m \to \mathbb{Z}^k$ as an integer linear map. We say that a matrix $\mathbf{Q} \in \mathbb{Z}^{k\times m}$ is *primitive* if $\mathbf{Q}\mathbb{Z}^m = \mathbb{Z}^k$, i. e., if $\mathbf{Q}\colon \mathbb{Z}^m \to \mathbb{Z}^k$ is surjective. As a special case, a row vector $\mathbf{w}^t \in \mathbb{Z}^{1\times k}$ is primitive if and only if the greatest common divisor of its entries equals $\gcd(\mathbf{w}) = 1$.

2.1 Probabilities and asymptotics

We use standard asymptotic notation, $O(\cdot)$, $\Omega(\cdot)$ and $\omega(\cdot)$, and all asymptotics refer to a (possibly implicit) integer variable n. For example, we may write $n^{O(1)}$ for an arbitrary polynomially bounded function of n, and $n^{-\omega(1)}$ for a negligible function. Other parameters defining the size of a problem instance are always assumed to be polynomial in n. So, if $\mathbf{A} \in \mathbb{Z}^{k\times m}$ is a matrix with integer entries, the number of rows $k = n^{O(1)}$, the number of columns $m = n^{O(1)}$, and the bitsize $\max_{i,j} \log|a_{i,j}| = n^{O(1)}$ of the matrix entries are all assumed to be (at most) polynomial in n.

A probability ensemble is a sequence $\mathcal{A}_n$ of probability distributions over sets A_n, for $n \in \mathbb{N} = \{1,2,\ldots\}$. We always assume that all elements of $A_n \subseteq \{0,1\}^{\ell(n)}$ can be represented by strings of some fixed length $\ell(n)$. We write $x \leftarrow \mathcal{A}$ for the operation of sampling an element x according to distribution $\mathcal{A}$, and $\Pr\{x \leftarrow \mathcal{A}\}$ for the probability of x under $\mathcal{A}$. The uniform distribution over a set A is denoted $\mathcal{U}(A)$.

The *statistical distance* between two distributions $\mathcal{A},\mathcal{A}'$ over a set A is

$$\Delta(\mathcal{A},\mathcal{A}') = \frac{1}{2}\sum_{x\in A} |\Pr\{x \leftarrow \mathcal{A}\} - \Pr\{x \leftarrow \mathcal{A}'\}|\,.$$

Two distribution ensembles $\mathcal{A}_n,\mathcal{A}'_n$ are *statistically close* (written $\mathcal{A}_n \approx \mathcal{A}'_n$) if the statistical distance $\Delta(\mathcal{A}_n,\mathcal{A}'_n) = n^{-\omega(1)}$ is negligible. Two ensembles $\mathcal{A}_n,\mathcal{A}'_n$ are *computationally indistinguishable* if for any efficient (probabilistic polynomial-time computable) predicate $\mathcal{P}$, $\mathcal{P}(\mathcal{A}_n) \approx \mathcal{P}(\mathcal{A}'_n)$. The gap

$$\Delta(\mathcal{P}(\mathcal{A}_n),\mathcal{P}(\mathcal{A}'_n)) = |\Pr\{\mathcal{P}(\mathcal{A}_n)\} - \Pr\{\mathcal{P}(\mathcal{A}'_n)\}|$$

is called the *advantage* of $\mathcal{P}$ in distinguishing between the two distributions. An ensemble $\mathcal{A}_n$ over sets A_n is *pseudorandom* if it is computationally indistinguishable from the uniform distributions $\mathcal{U}(A_n)$. If $\mathcal{A}_n \approx \mathcal{U}(A_n)$ are statistically close, then we say that $\mathcal{A}_n$ is almost uniform or statistically pseudorandom.

We typically leave the parameter n implicit, and talk about individual distributions $\mathcal{A}$ over a single set A, but all asymptotic statements should be interpreted as referring to ensembles $\mathcal{A}_n$ parameterized by an integer n in some obvious way. For example, we may say that a distribution $\mathcal{A}$ over a set A is pseudorandom if no efficient algorithm can distinguish $\mathcal{A}$ from $\mathcal{U}(A)$ with better than negligible advantage. More precisely, an efficiently sampleable ensemble $\{\mathcal{A}_n\}_{n>0}$ over the sets $\{A_n\}_{n>0}$ is pseudorandom if any predicate $\mathcal{P}$ computable in probabilistic polynomial time $n^{O(1)}$ has at most negligible advantage

$$|\Pr\{\mathcal{P}(\mathcal{A}_n)\} - \Pr\{\mathcal{P}(\mathcal{U}(A_n))\}| \leq n^{-\omega(1)}$$

in distinguishing $\mathcal{A}_n$ from the uniform distribution $\mathcal{U}(A_n)$.

We write $\mathbb{Z}$ for the set of integers, and $\mathbb{Z}_q = \mathbb{Z}/(q\mathbb{Z})$ for the integers modulo q. We will need the following version of the leftover hash lemma, and a bound on the probability that a random vector is primitive modulo q.

Lemma 2.1 (Leftover Hash Lemma, [12]). *For any odd integer q, positive real $\varepsilon > 0$ and integers k and $n \geq \log_2(q^k/\varepsilon^2)$, the distribution $\mathcal{X} = \{(\mathbf{A},\mathbf{Az})\colon \mathbf{A} \leftarrow \mathcal{U}(\mathbb{Z}_q^{k\times n}), \mathbf{z} \leftarrow \mathcal{U}(\{\pm 1\}^n)\}$ is within statistical distance $\Delta(\mathcal{X},\mathcal{U}) \leq \varepsilon$ from the uniform distribution $\mathcal{U} = \mathcal{U}(\mathbb{Z}_q^{k\times n} \times \mathbb{Z}_q^k)$. In particular, if $n \geq k\log_2(q) + \omega(\log n)$, then $\mathcal{X} \approx \mathcal{U}$ is (statistically) pseudorandom.*

Lemma 2.2 (Primitive Vectors)**.** *For any positive integers $q = 2^{n^{O(1)}}$ and $k = \omega(\log n)$, if $\mathbf{w} \in \mathbb{Z}_q^k$ is chosen uniformly at random, then $\gcd(\mathbf{w}, q) = 1$ except with negligible probability.*

Proof. The probability that $\gcd(\mathbf{w}, q) \neq 1$ is at most

$$\sum_{p|q} p^{-k} \leq (\log q)/2^k \leq n^{O(1)}/n^{\omega(1)} = n^{-\omega(1)}$$

where the summation is over all prime factors of q. We used the fact that all prime factors are at least $p \geq 2$, and there are at most $\log_2 q$ of them. Better bounds are possible, but this crude estimate is more than enough for the purposes of this paper. □

2.2 Gaussian distributions

Let $\rho(x) = \exp(-\pi x^2)$ be the Gaussian function with total mass $\int_{x \in \mathbb{R}} \rho(x)\, dx = 1$, and $\rho_\sigma(x) = \rho(x/\sigma)$ its scaling by a factor $\sigma > 0$. For a set A, we write $\rho_\sigma(A)$ as a shorthand for $\sum_{x \in A} \rho_\sigma(A)$. The discrete Gaussian distribution of parameter σ, denoted[3] $\mathcal{D}_\sigma$, picks each integer $x \in \mathbb{Z}$ with probability proportional to $\rho_\sigma(x)$, i. e., $\Pr\{x \leftarrow \mathcal{D}_\sigma\} = \rho_\sigma(x)/\rho_\sigma(\mathbb{Z})$. The product distribution $\mathcal{D}_\sigma^k$ selects each $\mathbf{x} \in \mathbb{Z}^k$ with probability proportional to $\rho_\sigma(\mathbf{x}) = \rho_\sigma(\|\mathbf{x}\|) = \prod_i \rho_\sigma(x_i)$. If $\mathbf{x}$ and $\mathbf{y}$ are orthogonal vectors ($\mathbf{x}^t\mathbf{y} = 0$), then by the Pythagorean theorem $\rho_\sigma(\mathbf{x} + \mathbf{y}) = \rho_\sigma(\mathbf{x}) \cdot \rho_\sigma(\mathbf{y})$.

A rank-n integer lattice is the set $\Lambda = \mathbf{B}\mathbb{Z}^n \subseteq \mathbb{Z}^d$ of all integer linear combinations of n linearly independent vectors $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$ in $\mathbb{Z}^d$. The last successive minimum of a rank-n lattice Λ is the smallest positive real λ_n such that Λ contains n linearly independent vectors of length at most λ_n. Another standard quantity associated to a lattice is the smoothing parameter $\eta_\varepsilon(\Lambda)$, which is parameterized by a positive real $\varepsilon > 0$. In this paper, all we need to know about the smoothing parameter are the following two bounds.

Lemma 2.3 (See [18, Lemma 4.1] and [10, Lemma 2.4])**.** *For any lattice Λ, $\varepsilon \in (0,1)$, and vector $\mathbf{c}$ in the linear span of Λ, if $\sigma > \eta_\varepsilon(\Lambda)$, then $\rho_\sigma(\Lambda + \mathbf{c}) \in [(1-\varepsilon)/(1+\varepsilon), 1] \cdot \rho_\sigma(\Lambda)$.*

Lemma 2.4 (Smoothing Parameter Bound, [18, Lemma 3.3])**.** *For any rank-n lattice Λ and positive real $\varepsilon > 0$, the smoothing parameter is at most*

$$\eta_\varepsilon(\Lambda) \leq \sqrt{\frac{\ln(2n(1+1/\varepsilon))}{\pi}} \cdot \lambda_n(\Lambda)\,. \tag{2.1}$$

In particular, for any $\omega(\sqrt{\log n})$ function there is a negligible function $\varepsilon(n) = n^{-\omega(1)}$ such that $\eta_\varepsilon(\Lambda) \leq \omega(\sqrt{\log n}) \cdot \lambda_n(\Lambda)$.

When the smoothing parameter $\eta(\Lambda)$ is written without specifying the value of ε, it is assumed that $\varepsilon = n^{-\omega(1)}$ is an arbitrary negligible function of the asymptotic variable n. For example, the smoothing parameter of the integer lattice is $\eta(\mathbb{Z}) \leq \sqrt{\ln(2(1+1/\varepsilon)/\pi)} = \omega(\sqrt{\log n})$. We will also need the following convolution theorems for discrete Gaussians.

[3] In the literature, $\mathcal{D}_\sigma$ is often used for the continuous Gaussian distribution over the real numbers $\mathbb{R}$, while the discrete Gaussian is denoted $\mathcal{D}_{\mathbb{Z},\sigma}$. Since here we do not use continuous Gaussians, for brevity we use $\mathcal{D}_\sigma$ to denote the discrete Gaussian distribution over the integers.

Lemma 2.5 (Convolution, [17, Theorem 3]). *For any primitive vector $\mathbf{v} \in \mathbb{Z}^m$ and positive reals $\sigma_i \geq \sqrt{2}\|\mathbf{v}\|_\infty \eta(\mathbb{Z})$, if $y_i \leftarrow \mathcal{D}_{\sigma_i}$ for $i = 1, \ldots, m$, then the sum $y = \sum_i v_i \cdot y_i$ is statistically close to $\mathcal{D}_\sigma$, where $\sigma = \sqrt{\sum_i (v_i \sigma_i)^2}$.*

Lemma 2.6 (Gaussian Projection). *For any primitive matrix $\mathbf{T} \in \mathbb{Z}^{k \times m}$, positive reals $\alpha, \sigma > 0$, and negligible $\varepsilon = n^{-\omega(1)}$, if $\mathbf{T} \cdot \mathbf{T}^t = \alpha^2 \cdot \mathbf{I}$ and $\eta(\ker(\mathbf{T})) \leq \sigma$, then $\mathbf{T}(\mathcal{D}^m_\sigma) \approx \mathcal{D}^k_{\alpha\sigma}$.*

Proof. Let $\mathbf{y} \in \mathbb{Z}^k$ be an arbitrary integer vector, and let $\mathbf{x} \in \mathbb{Z}^m$ be such that $\mathbf{Tx} = \mathbf{y}$. By linearity, any other $\mathbf{z} \in \mathbb{Z}^m$ maps to $\mathbf{Tz} = \mathbf{y}$ if and only if $\mathbf{z} \in \mathbf{x} + \ker(\mathbf{T})$. So, by definition, the probability of $\mathbf{y} = \mathbf{Tx}$ under $\mathbf{T}(\mathcal{D}^m_\sigma)$ is proportional to $\rho_\sigma(\mathbf{x} + \ker(\mathbf{T}))$. Let $\mathbf{x}_1 = \mathbf{T}^t\mathbf{y}/\alpha^2 \in \mathbb{R}^m$ and $\mathbf{x}_0 = \mathbf{x} - \mathbf{x}_1 \in \mathbb{R}^m$, so that $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1$, and $\mathbf{x}_0$ is orthogonal to the rows of $\mathbf{T}$. It follows that $\mathbf{x}_1$ is orthogonal to $\mathbf{x}_0$ and $\ker(\mathbf{T})$. Therefore, $\rho_\sigma(\mathbf{x} + \ker(\mathbf{T})) = \rho_\sigma(\mathbf{x}_1) \cdot \rho_\sigma(\mathbf{x}_0 + \ker(\mathbf{T}))$. Since $\mathbf{x}_0$ belongs to the linear span of $\ker(\mathbf{T})$, and $\sigma \geq \eta(\ker(\mathbf{T}))$, by Lemma 2.3 the Gaussian mass $\rho_\sigma(\mathbf{x}_0 + \ker(\mathbf{T}))$ is essentially independent of $\mathbf{x}_0$, up to a negligible relative error. So, up to this error, the probability of $\mathbf{y}$ is proportional to $\rho_\sigma(\mathbf{x}_1)$. Finally, we observe that $\|\mathbf{x}_1\|^2 = \mathbf{y}^t\mathbf{T}\mathbf{T}^t\mathbf{y}/\alpha^4 = \|\mathbf{y}\|^2/\alpha^2$, and therefore $\rho_\sigma(\mathbf{x}_1) = \rho_\sigma(\|\mathbf{y}\|/\alpha) = \rho_{\alpha\sigma}(\mathbf{y})$. This proves that $\mathbf{T}(\mathcal{D}^m_\sigma)$ is statistically close to the discrete Gaussian distribution $\mathcal{D}^k_{\alpha\sigma}$. □

2.3 A gadget matrix construction

Our main proof requires an integer matrix satisfying some special properties. In the following lemma, we state the required properties and give a simple construction. We recall that notation $\mathbf{Q}_{[n]}$ (resp. $\mathbf{Q}_{]n[}$) stands for the matrix obtained by taking (resp. dropping) the first n columns of a matrix $\mathbf{Q}$. In particular, $\mathbf{Q}_{]1[}$ is the matrix $\mathbf{Q}$ without its first column.

Lemma 2.7. *There is an efficiently computable matrix $\mathbf{Q} \in \mathbb{Z}^{n \times (2n+3)}$ such that $\mathbf{Q}_{[n]}$ is invertible, $\mathbf{u}^t\mathbf{Q}_{[n]} = \mathbf{e}^t_1$, the vector $\mathbf{v}^t = \mathbf{u}^t\mathbf{Q}_{]n[}$ has norm $\|\mathbf{v}\| = 2\sqrt{n}$, $\|\mathbf{v}\|_\infty = 2$, and the matrix $\mathbf{T} = \mathbf{Q}_{]1[}$ satisfies $\mathbf{T}(\mathcal{D}^{2n+2}_\sigma) \approx \mathcal{D}^n_{2\sigma}$ for all $\sigma \geq \omega(\sqrt{\log n})$.*

Proof. Define the matrix

$$\mathbf{X} = \sum_{i=1}^{n-1} (\mathbf{e}_{i+1} - \mathbf{e}_i) \cdot \mathbf{e}^t_i = \begin{bmatrix} -1 & & \\ 1 & \ddots & \\ & \ddots & -1 \\ & & 1 \end{bmatrix} \in \mathbb{Z}^{n \times (n-1)}.$$

The idea is to start with the square matrix $\bar{\mathbf{Q}} = \bar{\mathbf{Q}}_{[n]} = [\mathbf{e}_1, \mathbf{X}]$, which is unitriangular (i. e., triangular, with unit elements along the diagonal, and, therefore, invertible), and it satisfies $\mathbf{u}^t\bar{\mathbf{Q}} = \mathbf{e}^t_1$. We would like to use Lemma 2.6 to analyze the distribution $\bar{\mathbf{Q}}_{]1[}(\mathcal{D}^m_\sigma) = \mathbf{X}(\mathcal{D}^m_\sigma)$. However, $\mathbf{X}$ is not primitive and does not satisfy the property $\mathbf{X}\mathbf{X}^t = \alpha^2\mathbf{I}$ required by Lemma 2.6 because adjacent rows of $\mathbf{X}$ have scalar product -1. Other pairs of rows are orthogonal, so $\mathbf{X}\mathbf{X}^t$ is tridiagonal (i. e., with nonzero entries only on or immediately next to the main diagonal), but not diagonal. To fix this, we extend $\bar{\mathbf{Q}}$ to

$\tilde{\mathbf{Q}} = [\bar{\mathbf{Q}}, \mathbf{Y}] = [\mathbf{e}_1, \mathbf{X}, \mathbf{Y}]$ with a block of $(n-1)$ coordinates

$$\mathbf{Y} = \sum_{i=1}^{n-1} (\mathbf{e}_{i+1} + \mathbf{e}_i) \cdot \mathbf{e}_i^t = \begin{bmatrix} 1 & & \\ 1 & \ddots & \\ & \ddots & 1 \\ & & 1 \end{bmatrix} \in \mathbb{Z}^{n\times(n-1)}$$

where adjacent rows have scalar product 1, and cancel out with $\mathbf{X}$. This time $\tilde{\mathbf{Q}}_{]1[} = [\mathbf{X}, \mathbf{Y}]$ has pairwise orthogonal rows, but the first and last rows have a different norm than the rest. So, $[\mathbf{X}, \mathbf{Y}][\mathbf{X}, \mathbf{Y}]^t$ is diagonal, but it is still not a scalar matrix $\alpha^2\mathbf{I}$. We complete the construction by adding 4 more columns to make each row of $\mathbf{Q}_{]1[}$ contain precisely 4 nonzero ± 1 entries. Our final construction is

$$\mathbf{Q} = [\mathbf{e}_1, \mathbf{X}, -\mathbf{e}_n, \mathbf{Y}, \mathbf{e}_n, \mathbf{e}_1, \mathbf{e}_1]$$

where the position and sign of the new columns have been chosen to highlight the (square) unitriangular blocks $\tilde{\mathbf{X}} = [\mathbf{X}, -\mathbf{e}_n], \tilde{\mathbf{Y}} = [\mathbf{Y}, \mathbf{e}_n] \in \mathbb{Z}^{n\times n}$. Notice that $\tilde{\mathbf{Y}} = \tilde{\mathbf{X}} + 2\mathbf{I}$, and therefore the two blocks commute, i. e., $\tilde{\mathbf{X}}\tilde{\mathbf{Y}} = \tilde{\mathbf{Y}}\tilde{\mathbf{X}}$.

We already know that $\mathbf{Q}_{[n]} = [\mathbf{e}_1, \mathbf{X}]$ is invertible, $\mathbf{u}^t\mathbf{Q}_{[n]} = \mathbf{e}_1^t$, and it is immediate to verify that the vector $\mathbf{v}^t = \mathbf{u}^t\mathbf{Q}_{]n[}$ satisfies $\|\mathbf{v}\| = 2\sqrt{n}$ and $\|\mathbf{v}\|_\infty = 2$. It remains to analyze $\mathbf{T}(\mathcal{D}_\sigma^{2n+2})$, where

$$\mathbf{T} = \mathbf{Q}_{]1[} = [\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}, \mathbf{e}_1, \mathbf{e}_1]\,.$$

This matrix is primitive because it starts with a unitriangular block, and it satisfies $\mathbf{T}\mathbf{T}^t = 4\mathbf{I}$ by construction. In order to apply Lemma 2.6, and conclude that $\mathbf{T}(\mathcal{D}_\sigma^{2n+2}) \approx \mathcal{D}_{2\sigma}$, we only need to bound the smoothing parameter of $\Lambda = \ker(\mathbf{T})$. This lattice is defined by a system $\mathbf{T}\mathbf{x} = \mathbf{0}$ of n linearly independent equations in $2n+2$ variables. So, Λ is a rank-$(n+2)$ lattice. Moreover, it contains $(n+2)$ vectors of length at most 2 given by the columns of the matrix

$$\mathbf{V} = \begin{bmatrix} \tilde{\mathbf{Y}} & \mathbf{e}_1 & \\ -\tilde{\mathbf{X}} & -\mathbf{e}_1 & \\ & 1 & 1 \\ & 1 & -1 \end{bmatrix} \in \mathbb{Z}^{(2n+2)\times(n+2)}.$$

The columns are linearly independent because the matrix

$$\mathbf{W} = \begin{bmatrix} \mathbf{I} & \mathbf{I} & & \\ & & 1 & 1 \\ & & 1 & -1 \end{bmatrix} \in \mathbb{Z}^{(n+2)\times(2n+2)}$$

satisfies $\mathbf{W}\mathbf{V} = 2\mathbf{I}$. So $\mathbf{V}$ has rank $n+2$. This proves that $\lambda_{n+2}(\Lambda) \le 2$, and therefore, by Lemma 2.4, $\eta(\Lambda) \le \omega(\sqrt{\log n}) \le \sigma$. So, all hypotheses of Lemma 2.6 are satisfied and $\mathbf{T}(\mathcal{D}_\sigma^{2n+2}) \approx \mathcal{D}_{2\sigma}^n$. □

2.4 Computational problems and LWE

All computational problems considered in this paper are decision problems about pseudorandom distributions. Specifically, for any distribution ensemble $\mathcal{A}_n$ over sets A_n, the $\mathcal{A}_n$*-assumption* is the assumption that $\mathcal{A}_n$ is pseudorandom, and the $\mathcal{A}_n$*-problem* is the computational problem of distinguishing $\mathcal{A}_n$ from the uniform distribution $\mathcal{U}(A_n)$ with non-negligible advantage. So, all problems will be implicitly specified simply by defining an appropriate set of distributions $\mathcal{A}_n$.

A reduction between (the decision problems associated to) two distributions $\mathcal{A}_n$ and $\mathcal{A}'_n$ over sets A_n and A'_n (from $\mathcal{A}_n$ to $\mathcal{A}'_n$) is an efficient (probabilistic polynomial-time) algorithm that solves problem $\mathcal{A}_n$ (i. e., distinguishes $\mathcal{A}_n$ from the uniform distribution with non-negligible advantage) given access to any oracle that solves $\mathcal{A}'_n$ with (possibly different, but still) non-negligible advantage. In the simplest settings (e. g., see Lemmas 2.12 and 2.13) a reduction may be specified just by an efficient (probabilistic polynomial-time computable) function φ such that $\varphi(\mathcal{A}_n) \approx \mathcal{A}'_n$ and $\varphi(\mathcal{U}(A_n)) \approx \mathcal{U}(A'_n)$. Most of our reductions are more complex, and make use of hybrid arguments (see Lemma 2.9) that require oracle calls on distributions other than $\mathcal{A}'_n$ or $\mathcal{U}(A'_n)$.

In this paper, it is convenient to consider a version of the Learning With Errors (LWE) problem where the secret is a matrix $\mathbf{S}$, rather than a vector, defined as follows.

Definition 2.8. For any positive integers q,n,k,m and real σ, let $\mathbf{LWE}(q,n\times k,m,\sigma)$ be the LWE distribution with modulus q, number of samples m, secret dimension $n\times k$, and error parameter σ, i. e., the distribution of

$$[\mathbf{A},\mathbf{AS}+\mathbf{E}] \in \mathbb{Z}_q^{m\times(n+k)}$$

obtained by picking $\mathbf{A} \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times n})$ and $\mathbf{S} \leftarrow \mathcal{U}(\mathbb{Z}_q^{n\times k})$ uniformly at random, and $\mathbf{E} \leftarrow \mathcal{D}_\sigma^{m\times k}$ with discrete Gaussian distribution.

When $k=1$, the secret is just a vector $\mathbf{s} \in \mathbb{Z}_q^n$, and this is the standard version of LWE, which we write $\mathbf{LWE}(q,n,m,\sigma)$ instead of $\mathbf{LWE}(q,n\times 1,m,\sigma)$. The m rows of the LWE can be viewed as random noisy labeled samples from a hard-to-learn linear function defined by the secret $\mathbf{S}$. Worst-case to average-case reductions [21, 19, 6, 20] support the conjecture that the LWE problem is hard for an arbitrary (polynomially bounded) number of samples $m = n^{O(1)}$, and some reductions require this extra flexibility. (E. g., the LWE search-to-decision reduction in [21], but see also [16] for a sample-preserving reduction.) This version of the problem is denoted $\mathbf{LWE}(q,n,\sigma)$. The modulus q is always assumed to have bit-size polynomial in n (i. e., $\log_2 q \le n^{O(1)}$), but in most cryptographic applications it is just a small polynomial (e. g., $q \le n^2$), and integers modulo q are represented with $O(\log n)$ bits.

The vector and matrix variants of LWE are easily seen to be equivalent via a standard hybrid argument.

Lemma 2.9. *There is a polynomial-time reduction from* $\mathbf{LWE}(q,n,m,\sigma)$ *to* $\mathbf{LWE}(q,n\times k,m,\sigma)$.

Proof. The intuition behind the proof is that the LWE distribution with secret matrix $\mathbf{S} \in \mathbb{Z}_q^{n\times k}$ may be regarded as k copies of the standard LWE distribution with secret vectors given by the columns of $\mathbf{S}$, all using the same public random $\mathbf{A}$. More technically, the reduction considers the sequence of hybrid distributions $\mathcal{A}_i = (\mathbf{A},[\mathbf{AS}_i+\mathbf{E}_i,\mathbf{B}_i])$ where $\mathbf{A} \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times n})$, $\mathbf{S}_i \leftarrow \mathcal{U}(\mathbb{Z}_q^{n\times i})$, $\mathbf{E}_i \leftarrow \mathcal{D}_\sigma^{m\times i}$ and $\mathbf{B}_i \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times(k-i)})$, for $i=0,\ldots,k$. Each pair of neighboring hybrids $\mathcal{A}_i,\mathcal{A}_{i+1}$ can be generated by starting from an $\mathbf{LWE}(q,n,m,\sigma)$ challenge sample $(\mathbf{A},\mathbf{b})$, and then computing $\mathcal{A} = (\mathbf{A},[\mathbf{AS}+\mathbf{E},\mathbf{b},\mathbf{B}])$

where $\mathbf{S} \leftarrow \mathcal{U}(\mathbb{Z}_q^{n\times i})$, $\mathbf{E} \leftarrow \mathcal{D}_\sigma^{m\times i}$ and $\mathbf{B} \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times(k-i-1)})$. The resulting distribution equals $\mathcal{A} = \mathcal{A}_i$ if $\mathbf{b}$ is random, and $\mathcal{A} = \mathcal{A}_{i+1}$ if $\mathbf{b} = \mathbf{As}+\mathbf{e}$ is pseudorandom. So, any distinguisher with advantage ε against $\mathbf{LWE}(q, n\times k, m, \sigma)$ will achieve advantage ε/k against $\mathbf{LWE}(q,n,m,\sigma)$. $\square$

Definition 2.10. The $\mathbf{LWE}_{0,1}(q,n,m,\sigma)$ distribution (and associated decision problem and pseudorandomness assumption) is defined just like $\mathbf{LWE}(q,n,m,\sigma)$, except that the secret $\mathbf{s} \leftarrow \mathcal{U}(\{0,1\}^n)$ is chosen with random binary entries.

Definition 2.11. The $\mathbf{LWE}_{\pm}(q,n,m,\sigma)$ distribution (and associated decision problem and pseudorandomness assumption) is defined just like $\mathbf{LWE}(q,n,m,\sigma)$, except that the secret $\mathbf{s} \leftarrow \mathcal{U}(\{\pm 1\}^n)$ is chosen with random unit entries.

We remark that $\mathbf{LWE}_{0,1}$ and $\mathbf{LWE}_{\pm}$ could also be generalized to secret matrices $\mathbf{S}$, and proved equivalent to the single-vector version exactly as in Lemma 2.9. But this is not used in this paper, so, for simplicity, we only define the secret-vector version of the problems. The next two lemmas show that $\mathbf{LWE}_{0,1}$ and $\mathbf{LWE}_{\pm}$ are essentially the same problem. We remark that the lemmas are even more general than stated, and they apply to $\mathbf{LWE}$ problems with arbitrary error distribution, not just discrete Gaussians. All parameters (including the number of samples, and the exact error distribution) are preserved by the reductions, showing that the two problems are equivalent in a very strong sense.

Lemma 2.12. *For any odd integer q, there is a polynomial-time reduction from the $\mathbf{LWE}_{0,1}(q,n,m,\sigma)$ problem to the $\mathbf{LWE}_{\pm}(q,n,m,\sigma)$ problem.*

Proof. On input an $\mathbf{LWE}_{0,1}$ instance $(\mathbf{A},\mathbf{b})$, the reduction outputs $\varphi(\mathbf{A},\mathbf{b}) = (\mathbf{A}/2, \mathbf{b}' = \mathbf{b} - (\mathbf{A}/2)\mathbf{u})$ where $\mathbf{A}/2$ is computed modulo q, and $\mathbf{u} = (1,\ldots,1) \in \mathbb{Z}_q^n$. Notice that, since q is odd, the factor 2 is invertible modulo q, and $\mathbf{A}/2$ is uniformly distributed. If $\mathbf{b}$ is uniform, then $\mathbf{b}'$ is also uniform. On the other hand, if $\mathbf{b} = \mathbf{As}+\mathbf{e}$, then $\mathbf{b}' = (\mathbf{A}/2)\mathbf{s}' + \mathbf{e}$ where $\mathbf{s}' = 2\mathbf{s} - \mathbf{u}$ is uniformly random in $\{\pm 1\}^n$. $\square$

Lemma 2.13. *For any odd integer q, there is a polynomial-time reduction from the $\mathbf{LWE}_{\pm}(q,n,m,\sigma)$ problem to the $\mathbf{LWE}_{0,1}(q,n,m,\sigma)$ problem.*

Proof. On input an $\mathbf{LWE}_{\pm}$ instance $(\mathbf{A},\mathbf{b})$, the reduction outputs $\varphi(\mathbf{A},\mathbf{b}) = (2\mathbf{A}, \mathbf{b}' = \mathbf{b} + \mathbf{Au})$ where $\mathbf{u} \in \{1\}^n$ is the all-ones vector. Notice that, since q is odd, the factor 2 is invertible modulo q, and $2\mathbf{A}$ is uniformly distributed. If $\mathbf{b}$ is uniform, then $\mathbf{b}'$ is also uniform. On the other hand, if $\mathbf{b} = \mathbf{As}+\mathbf{e}$, then $\mathbf{b}' = (2\mathbf{A})\mathbf{s}' + \mathbf{e}$ where $\mathbf{s}' = (\mathbf{s}+\mathbf{u})/2$ is uniformly random in $\{0,1\}^n$. $\square$

3 Pseudorandomness of binary LWE

In this section we present a proof that the binary-secret LWE distribution $\mathbf{LWE}_{\pm}$ is pseudorandom. The idea is to define a simple (efficiently computable) randomized transformation φ with the following properties:

- If the input to φ is uniformly distributed, then the output $\varphi(\mathcal{U})$ equals (or is statistically close to) the binary LWE distribution $\mathbf{LWE}_{\pm}(q,n,m,\hat{\sigma})$ for some $\hat{\sigma}$.

- There are two pseudorandom distributions $\mathcal{B}, \hat{\mathcal{B}}$ such that $\varphi(\mathcal{B})$ equals (or is statistically close to) $\hat{\mathcal{B}}$.

Since φ is efficiently computable, the pseudorandomness of $\mathcal{B}$ implies that $\varphi(\mathcal{U}) \approx \mathbf{LWE}_{\pm}(q,n,m,\hat{\sigma})$ is computationally indistinguishable from $\varphi(\mathcal{B}) \approx \hat{\mathcal{B}}$. By transitivity, since $\hat{\mathcal{B}}$ is pseudorandom, it follows that $\mathbf{LWE}_{\pm}(q,n,m,\hat{\sigma})$ is also pseudorandom.

As our aim is to give a reduction from the standard LWE problem to binary LWE, we set $\mathcal{B}$ and $\hat{\mathcal{B}}$ to two pseudorandom distributions related to LWE. Specifically, we use the distributions

$$\begin{aligned}\mathcal{B} &= \{(\mathbf{AS}+\mathbf{E})^t \mid \mathbf{A} \leftarrow \mathcal{U}(\mathbb{Z}_q^{(n-1)\times k}),\ \mathbf{S} \leftarrow \mathcal{U}(\mathbb{Z}_q^{k\times m}),\ \mathbf{E} \leftarrow \mathcal{D}_\sigma^{(n-1)\times m}\} \text{ and} \\ \hat{\mathcal{B}} &= \{(\hat{\mathbf{A}}\hat{\mathbf{S}}+\hat{\mathbf{E}})^t \mid \hat{\mathbf{A}} \leftarrow \mathcal{U}(\mathbb{Z}_q^{(n+1)\times (k+1)}),\ \hat{\mathbf{S}} \leftarrow \mathcal{U}(\mathbb{Z}_q^{(k+1)\times m}),\ \hat{\mathbf{E}} \leftarrow \mathcal{D}_{2\sigma}^{(n+1)\times m}\}\end{aligned}$$

for some σ related to $\hat{\sigma}$. In other words $\mathcal{B}$ and $\hat{\mathcal{B}}$ are the (transposed) "label" component of the LWE distributions $\mathbf{LWE}(q,k\times m,n-1,\sigma)$ and $\mathbf{LWE}(q,(k+1)\times m,n+1,2\sigma)$. Notice that any distinguisher between $\mathcal{B}$ and the uniform distribution can be immediately transformed into an LWE distinguisher that on input $(\mathbf{A},\mathbf{B}=\mathbf{AS}+\mathbf{E})$ simply discards $\mathbf{A}$, and then runs the original distinguishing procedure on $\mathbf{B}^t$. So, $\mathcal{B}$ is pseudorandom under the standard LWE assumption, and similarly for $\hat{\mathcal{B}}$.

Before getting into the details of the transformation, notice the difference between the high level structure of the proof presented here, and a typical reduction between variants of LWE. A typical reduction would map *standard* LWE samples to *binary* LWE samples, and *uniform* samples to *uniform* samples. Here, instead, on the one hand the *standard* LWE distribution is mapped again to a *standard* LWE distribution (with slightly different parameters). On the other hand, the *uniform* distribution is mapped to *binary* LWE.

Our randomized transformation φ is shown in Figure 1. The transformation uses, as randomness, both a *uniform* secret vector $\mathbf{s}$ and a *binary* secret vector $\mathbf{z}$. Informally, the intuition is that by *simultaneously* multiplying by $\mathbf{s}$ (on the *left*) and by $\mathbf{z}$ (on the *right*), the same transformation is able to produce (depending on how the input $\mathbf{B}$ was chosen) either

- a binary LWE distribution with secret $\mathbf{z}$ (when $\mathbf{B}$ is uniform), or
- a (transposed[4]) standard LWE distribution with secret $[\mathbf{s},\mathbf{S}^t]^t$ (when $\mathbf{B}=(\mathbf{AS}+\mathbf{E})^t$).

Intuitively, one may think of φ as mapping $\mathbf{B}$ to $[\mathbf{B},\mathbf{Bz}+\mathbf{e}]$. So, when $\mathbf{B}$ is uniformly random, φ outputs the binary LWE distribution by construction. On the other hand, if $\mathbf{B}=(\mathbf{AS}+\mathbf{E})^t=\mathbf{S}^t\mathbf{A}^t+\mathbf{E}^t$, the transformation outputs $[\mathbf{B},\mathbf{Bz}+\mathbf{e}]=\mathbf{S}^t[\mathbf{A}^t,\mathbf{A}^t\mathbf{z}]+[\mathbf{E}^t,\mathbf{E}^t\mathbf{z}+\mathbf{e}]$, which looks like a standard (transposed) LWE label matrix. In fact, by the Leftover Hash Lemma, one may argue that $[\mathbf{A}^t,\mathbf{A}^t\mathbf{z}]$ is statistically close to a uniformly distributed matrix. Unfortunately, the error matrix $[\mathbf{E}^t,\mathbf{E}^t\mathbf{z}+\mathbf{e}]$ does not follow the Gaussian distribution[5] required by LWE. So, in order to address this and other technical difficulties, the actual transformation φ is a bit more complex. The details of the transformation are somewhat technical,

[4]The LWE distributions $\mathcal{B}$ and $\hat{\mathcal{B}}$ are transposed to allow for the multiplication of the uniform secret $\mathbf{s}$ on the left.

[5]In particular, the second part of the error matrix $\mathbf{E}^t\mathbf{z}+\mathbf{e}$ will typically have much larger entries than $\hat{\mathbf{E}}$, and also be somehow correlated to the first part $\mathbf{E}^t$. One may try to address the error imbalance by noise flooding techniques (i. e., by adding a large perturbation term to the first part of the output $\mathbf{B}$), but this would result in much larger noise and still not remove correlations.

Transformation $\varphi(\mathbf{B};\mathbf{z},\mathbf{s},\mathbf{a},\mathbf{e},\mathbf{G})$:	Randomness:	Dimensions:
Input: B	$\mathbf{z} \leftarrow \mathcal{U}(\{\pm 1\}^n)$	$\mathbf{B} \in \mathbb{Z}_q^{m\times(n-1)}$
$\mathbf{x} = \mathbf{s} + \mathbf{e}$	$\mathbf{s} \leftarrow \mathcal{U}(\mathbb{Z}_q^m)$	$\mathbf{x} \in \mathbb{Z}_q^m$
$\mathbf{Y} = [\mathbf{s}, \mathbf{s}\cdot\mathbf{a}^t + \mathbf{B}]$	$\mathbf{a} \leftarrow \mathcal{U}(\mathbb{Z}_q^{n-1})$	$\mathbf{Q} \in \mathbb{Z}_q^{n\times(2n+3)}$
$\mathbf{X} = [\mathbf{Y},\mathbf{G}]\mathbf{Q}^t \cdot \mathbf{diag}(\mathbf{z})$	$\mathbf{e} \leftarrow \mathcal{D}_{2\sigma}^m$	$\mathbf{Y} \in \mathbb{Z}_q^{m\times n}$
Output $[\mathbf{X},\mathbf{x}]$	$\mathbf{G} \leftarrow \mathcal{D}_\sigma^{m\times(n+3)}$	$\mathbf{X} \in \mathbb{Z}_q^{m\times n}$

Figure 1: Transformation proving the pseudorandomness of binary LWE, where $\mathbf{Q}$ is the matrix specified in Lemma 2.7.

and they are primarily motivated by all the cancellations needed for the proof to work and obtain the proper LWE Gaussian error distribution.

One way to gain additional insight into the construction is to notice that the transformation $\varphi(\mathbf{B})$ always outputs a pair $[\mathbf{X},\mathbf{x}]$ such that $\mathbf{X}\mathbf{z} = \mathbf{s} + \mathbf{G}\mathbf{v} \approx \mathbf{s} + \mathbf{e} = \mathbf{x}$. (See proof of Claim 3.2 for details.) So, distribution $\hat{\mathcal{B}} = \varphi(\mathcal{B})$ will also satisfy this property with high probability: there must be a small vector $\hat{\mathbf{z}} \in \{\pm 1\}^{n+1}$ such that $(\hat{\mathbf{A}}\hat{\mathbf{S}} + \hat{\mathbf{E}})^t\hat{\mathbf{z}} \approx \mathbf{0}$. This shows that the pseudorandom matrix $\hat{\mathbf{B}} = (\hat{\mathbf{A}}\hat{\mathbf{S}} + \hat{\mathbf{E}})^t$ is already somehow close to a binary LWE instance because there is a ± 1 combination of the first n columns of $\hat{\mathbf{B}}$ that is close to the last column. In fact, something very similar can be proved directly, without using φ: matrix $\hat{\mathbf{A}}^t$ maps a set $\{0,1\}^{n+1}$ of size $2^n > q^{k+1}$ to a set $\mathbb{Z}_q^{k+1}$ of size q^{k+1}. So, by the pigeon-hole principle, there exist two binary inputs such that $\hat{\mathbf{A}}^t\hat{\mathbf{z}}_\mathbf{0} = \hat{\mathbf{A}}^t\hat{\mathbf{z}}_\mathbf{1}$, or, equivalently, a small vector $\hat{\mathbf{z}} = \hat{\mathbf{z}}_\mathbf{0} - \hat{\mathbf{z}}_\mathbf{1}$ (with $\|\mathbf{z}\|_\infty = 1$) such that $\hat{\mathbf{B}}\hat{\mathbf{z}} = \hat{\mathbf{E}}^t\hat{\mathbf{z}} \approx \mathbf{0}$. An informal interpretation of this argument (which, in fact, is closely related to the proof that LWE is robust with respect to the secret distribution [11]) is that matrix $\hat{\mathbf{A}}^t$ hashes the binary secret $\hat{\mathbf{z}}$ to an almost uniform (smaller dimensional) secret $\hat{\mathbf{A}}^t\hat{\mathbf{z}}$ with entries in $\mathbb{Z}_q$. But, as before, the problem with this intuitive approach is that the error distribution $\hat{\mathbf{E}}^t\hat{\mathbf{z}}$ is not Gaussian, and it is correlated with the secret $\hat{\mathbf{z}}$.

Our theorem below solves these technical problems using a carefully designed gadget matrix $\mathbf{Q}$ (described in Lemma 2.7) which efficiently adjusts the error distribution using some extra randomness $\mathbf{G}$. Notice how, in the process of transforming LWE into binary LWE, the number of samples $n-1$ in the presumed hard $\mathbf{LWE}(q, k\times m, n-1, \sigma)$ instance becomes the size n of the secret in the final binary LWE instance $\mathbf{LWE}_\pm(q,n,m,\hat{\sigma})$. Similarly, the number of columns m (i. e., the number of parallel LWE instances) in the presumed hard $\mathbf{LWE}(q,(k+1)\times m, n+1, 2\sigma)$ instance becomes the number of samples in the final binary LWE instance.

Theorem 3.1. *Assume the distributions* $\mathbf{LWE}(q, k\times m, n-1, \sigma)$ *and* $\mathbf{LWE}(q,(k+1)\times m, n+1, 2\sigma)$ *are pseudorandom. If* $q \le 2^{n^{O(1)}}$, $\sigma \ge \omega(\sqrt{\log n})$, $k \ge \omega(\log n)$, *and* $n \ge (k+1)\cdot\log_2(q) + \omega(\log n)$, *then the distribution* $\mathbf{LWE}_\pm(q,n,m,\hat{\sigma})$ *is also pseudorandom for* $\hat{\sigma} = 2\sigma\sqrt{n+1}$.

Proof. We use $\mathbf{Z}$ as a shorthand for the diagonal matrix $\mathbf{diag}(\mathbf{z})$. We first show that the transformation φ maps the uniform distribution to the binary LWE distribution.

Claim 3.2. *If* $\mathbf{B} \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times(n-1)})$ *is chosen uniformly at random, then* $\varphi(\mathbf{B}) \in \mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^m$ *is statistically close to the* $\mathbf{LWE}_\pm(q,n,m,\hat{\sigma})$ *distribution.*

Proof. We show that for any fixed values of $\mathbf{a} \in \mathbb{Z}_q^{n-1}$ and $\mathbf{z} \in \{\pm 1\}^n$, the output of the transformation $[\mathbf{X}, \mathbf{x}] = \varphi(\mathbf{B})$ is statistically close to the $\mathbf{LWE}_{\pm}$ distribution with secret $\mathbf{z}$, i. e., $\mathbf{X} \in \mathbb{Z}_q^{m \times n}$ is uniformly random, and the conditional distribution of the noise vector $\hat{\mathbf{e}} = \mathbf{x} - \mathbf{X}\mathbf{z}$ (given $\mathbf{X}$ and $\mathbf{z}$) is statistically close to $\mathcal{D}_{\hat{\sigma}}^m$. All this is over the probability space defined by the random choice of $\mathbf{B}, \mathbf{s}, \mathbf{e}, \mathbf{G}$. The claim follows by averaging over $\mathbf{a}$ and $\mathbf{z}$.

Let $\mathbf{Q} = [\mathbf{Q}_{[n]}, \mathbf{Q}_{]n[}]$ be the matrix defined in Lemma 2.7, and recall that $\mathbf{Q}_{[n]} \in \mathbb{Z}^{n \times n}$ is invertible, $\mathbf{u}^t \mathbf{Q}_{[n]} = \mathbf{e}_1^t$, and the vector $\mathbf{v}^t = \mathbf{u}^t \mathbf{Q}_{]n[} \in \mathbb{Z}^{n+3}$ has norm $\|\mathbf{v}\| = 2\sqrt{n}$, $\|\mathbf{v}\|_\infty \leq 2$. Since $\mathbf{s}$ and $\mathbf{B}$ are uniformly random, the matrix $\mathbf{Y}$ is also uniformly distributed, and independent of $\mathbf{e}$ and $\mathbf{G}$. Since $\mathbf{Q}_{[n]}^t$ and $\mathbf{Z}$ are invertible, the matrix

$$\mathbf{X} = [\mathbf{Y}, \mathbf{G}][\mathbf{Q}_{[n]}, \mathbf{Q}_{]n[}]^t \mathbf{Z} = \mathbf{Y}(\mathbf{Q}_{[n]}^t \mathbf{Z}) + (\mathbf{G}\mathbf{Q}_{]n[}^t \mathbf{Z})$$

is also uniformly distributed, independently of $\mathbf{G}, \mathbf{e}$. It remains to analyze the conditional distribution of the error vector $\hat{\mathbf{e}} = \mathbf{x} - \mathbf{X}\mathbf{z}$. Using $\mathbf{Z} \cdot \mathbf{z} = \mathbf{u}$ and $\mathbf{Y}\mathbf{Q}_{[n]}^t \mathbf{u} = \mathbf{Y}\mathbf{e}_1 = \mathbf{s}$, we get $\mathbf{X}\mathbf{z} = \mathbf{Y}\mathbf{Q}_{[n]}^t \mathbf{u} + \mathbf{G}\mathbf{Q}_{]n[}^t \mathbf{u} = \mathbf{s} + \mathbf{G}\mathbf{v}$. So, the error vector equals $\hat{\mathbf{e}} = (\mathbf{s} + \mathbf{e}) - \mathbf{X}\mathbf{z} = \mathbf{e} - \mathbf{G}\mathbf{v}$. Since the entries of $\mathbf{G}$ and $\mathbf{e}$ are independent discrete Gaussians of parameter σ and 2σ (respectively), the coordinates of $\hat{\mathbf{e}}$ are independent identically distributed random variables, each following the distribution $\mathcal{D}_{2\sigma} - \sum_i v_i \mathcal{D}_\sigma$. By Lemma 2.5, this distribution is statistically close to $\mathcal{D}_{\hat{\sigma}}$ for

$$\hat{\sigma} = \sqrt{(2\sigma)^2 + \sum_i (v_i \sigma)^2} = \sigma\sqrt{4 + \|\mathbf{v}\|^2} = 2\sigma\sqrt{n+1}\,. \qquad \square$$

Next, consider the output $[\mathbf{X}, \mathbf{x}]$ when $\mathbf{B}$ follows distribution $\mathcal{B}$.

Claim 3.3. *The distribution $\varphi(\mathcal{B})$ is statistically close to $\hat{\mathcal{B}}$.*

Proof. Let $\mathbf{B} = (\mathbf{A}\mathbf{S} + \mathbf{E})^t$ for $\mathbf{A} \leftarrow \mathcal{U}(\mathbb{Z}_q^{(n-1) \times k})$, $\mathbf{S} \leftarrow \mathcal{U}(\mathbb{Z}_q^{k \times m})$ and $\mathbf{E} \leftarrow \mathcal{D}_\sigma^{(n-1) \times m}$. By linearity, we can write $\mathbf{Y} = [\mathbf{s}, \mathbf{s}\mathbf{a}^t + \mathbf{B}] = \mathbf{Y}_s + \mathbf{Y}_e$ as the sum of two matrices

$$\mathbf{Y}_s = [\mathbf{s}, \mathbf{s}\mathbf{a}^t + \mathbf{S}^t \mathbf{A}^t] \qquad \text{and} \qquad \mathbf{Y}_e = [\mathbf{0}, \mathbf{E}^t]\,.$$

Similarly, we can also decompose $\varphi(\mathbf{B}) = [\mathbf{X}, \mathbf{x}] = [\mathbf{X}_s, \mathbf{s}] + [\mathbf{X}_e, \mathbf{e}]$ as a sum where

$$\mathbf{X}_s = \mathbf{Y}_s \mathbf{Q}_{[n]}^t \mathbf{Z} \qquad \text{and} \qquad \mathbf{X}_e = [\mathbf{Y}_e, \mathbf{G}]\mathbf{Q}^t \mathbf{Z} = [\mathbf{E}^t, \mathbf{G}]\mathbf{Q}_{]1[}^t \mathbf{Z}\,.$$

Our goal is to show that $[\mathbf{X}_s, \mathbf{s}]^t = \hat{\mathbf{A}}\hat{\mathbf{S}}$ and $[\mathbf{X}_e, \mathbf{e}]^t = \hat{\mathbf{E}}$ for $\hat{\mathbf{A}}, \hat{\mathbf{S}}, \hat{\mathbf{E}}$ distributed as in the definition of $\hat{\mathcal{B}}$.

We first look at the distribution of the error matrix $\hat{\mathbf{E}}^t = [\mathbf{X}_e, \mathbf{e}]$. The last column $\mathbf{e}$ is a discrete Gaussian of parameter 2σ by construction. Since $[\mathbf{E}^t, \mathbf{G}]$ has Gaussian distribution $\mathcal{D}_\sigma^{m \times (2n+2)}$, the rest of the matrix is distributed according to

$$\mathbf{X}_e^t \leftarrow \mathbf{Z}^t \mathbf{Q}_{]1[}(\mathcal{D}_\sigma^{(2n+2) \times m}) \approx \mathbf{Z}(\mathcal{D}_{2\sigma}^{n \times m}) = \mathcal{D}_{2\sigma}^{n \times m}\,,$$

for any fixed value of $\mathbf{z}$, where we have used the property $\mathbf{Q}_{]1[}(\mathcal{D}_\sigma^{2n+2}) \approx \mathcal{D}_{2\sigma}^n$ from Lemma 2.7, and the symmetry $\mathbf{Z}\mathcal{D}_{2\sigma}^n = \mathcal{D}_{2\sigma}^n$. This proves that $\hat{\mathbf{E}}$ has Gaussian distribution of parameter 2σ, and it depends only on $\mathbf{E}, \mathbf{G}$ and $\mathbf{e}$.

We now look at the distribution of $[\mathbf{X}_s, \mathbf{s}] = (\hat{\mathbf{A}}\hat{\mathbf{S}})^t$ over the random choice of $\mathbf{a}, \mathbf{s}, \mathbf{z}, \mathbf{A}$ and $\mathbf{S}$. The idea is to set $\hat{\mathbf{S}} = [\mathbf{s}, \mathbf{S}^t]^t$, so that $\hat{\mathbf{S}}$ is distributed uniformly at random over $\mathbb{Z}_q^{(k+1)\times m}$. But, in order to properly randomize $\hat{\mathbf{A}}$, we define

$$\hat{\mathbf{S}} = \mathbf{W}^{-1} \begin{bmatrix} \mathbf{s}^t \\ \mathbf{S} \end{bmatrix}$$

where $\mathbf{W}$ is a uniformly random invertible matrix in $\mathbb{Z}_q^{(k+1)\times(k+1)}$. Since $\mathbf{W}$ is invertible, $\hat{\mathbf{S}}$ is still uniformly distributed, and independent of $\mathbf{W}$. Next, define

$$\hat{\mathbf{A}} = \begin{bmatrix} \mathbf{I} \\ \mathbf{z}^t \end{bmatrix} \mathbf{Z}\mathbf{Q}_{[n]}\mathbf{H}\mathbf{W} \in \mathbb{Z}_q^{(n+1)\times(k+1)} \qquad \text{where} \qquad \mathbf{H} = \begin{bmatrix} 1 & \mathbf{0}^t \\ \mathbf{a} & \mathbf{A} \end{bmatrix} \in \mathbb{Z}_q^{n\times(k+1)}.$$

Using the identities $\mathbf{z}^t\mathbf{Z}\mathbf{Q}_{[n]} = \mathbf{u}^t\mathbf{Q}_{[n]} = \mathbf{e}_1^t$ and $\mathbf{H}\mathbf{W}\hat{\mathbf{S}} = \mathbf{H}[\mathbf{s}, \mathbf{S}^t]^t = \mathbf{Y}_s^t$, we see that our choice of $\hat{\mathbf{A}}, \hat{\mathbf{S}}$ satisfies $(\hat{\mathbf{A}}\hat{\mathbf{S}})^t = [\mathbf{X}_s, \mathbf{s}]$ as desired. All that is left to do is to prove that $\hat{\mathbf{A}}$ is statistically close to uniform, independently of $\hat{\mathbf{S}}$. We first look at $\mathbf{H}\mathbf{W}$. Let $\mathbf{w}^t$ be the first row of $\mathbf{W}$. That's also the first row of $\mathbf{H}\mathbf{W}$. The remaining rows of $\mathbf{H}\mathbf{W}$ are $[\mathbf{a}, \mathbf{A}]\mathbf{W}$. The first row $\mathbf{w}^t$ is distributed uniformly at random among all primitive vectors in $\mathbb{Z}_q^{k+1}$, i. e., all vectors such that $\gcd(\mathbf{w}, q) = 1$. So, by Lemma 2.2, the vector $\mathbf{w}$ is within negligible statistical distance from the uniform distribution over $\mathbb{Z}_q^{k+1}$. Finally, since $[\mathbf{a}, \mathbf{A}]$ is uniform by construction, and $\mathbf{W}$ is invertible, the bottom rows ($[\mathbf{a}, \mathbf{A}]\mathbf{W}$) of $\mathbf{H}\mathbf{W}$ are uniform too, and independent of $\mathbf{w}$. So, $\mathbf{H}\mathbf{W}$ is statistically close to uniform over $\mathbb{Z}_q^{n\times(k+1)}$. The matrix $\check{\mathbf{A}} = (\mathbf{Z}\mathbf{Q}_{[n]}\mathbf{H}\mathbf{W})^t$ is also statistically close to uniform (and independent of $\mathbf{z}$) because $\mathbf{Q}_{[n]}$ and $\mathbf{Z}$ are invertible. Finally, using the Leftover Hash Lemma (Lemma 2.1) and the assumption $n \geq (k+1)\log_2(q) + \omega(\log n)$, we see that $\hat{\mathbf{A}}^t = \check{\mathbf{A}}[\mathbf{I}, \mathbf{z}] = [\check{\mathbf{A}}, \check{\mathbf{A}}\mathbf{z}]$ is also statistically close to uniform. This concludes the proof that $\varphi(\mathbf{B}) = (\hat{\mathbf{A}}\hat{\mathbf{S}} + \hat{\mathbf{E}})^t$ where $\hat{\mathbf{A}}, \hat{\mathbf{S}}$ and $\hat{\mathbf{E}}$ follow the LWE distribution as in the definition of $\hat{\mathcal{B}}$. □

We are now ready to prove the theorem. It follows from pseudorandomness of the $\mathbf{LWE}(q, k\times m, n-1, \sigma)$ that the distribution $\mathcal{B}$ is computationally indistinguishable from the uniform distribution $\mathcal{U}$ over $\mathbb{Z}_q^{m\times(n-1)}$. Since φ is efficiently computable, the distributions $\varphi(\mathcal{B})$ and $\varphi(\mathcal{U})$ are also computationally indistinguishable. By Claim 3.2, $\varphi(\mathcal{U})$ is statistically close to $\mathbf{LWE}_{\pm}(q, n\times 1, m, \hat{\sigma})$. Similarly, by Claim 3.3, $\varphi(\mathcal{B})$ is statistically close to $\hat{\mathcal{B}}$. So, $\mathbf{LWE}_{\pm}(q, n\times 1, m, \hat{\sigma})$ is computationally indistinguishable from $\hat{\mathcal{B}}$. Finally, from the pseudorandomness of $\mathbf{LWE}(q, (k+1)\times m, n+1, 2\sigma)$, we know that the distribution $\hat{\mathcal{B}}$ is computationally indistinguishable from the uniform distribution over $\mathbb{Z}_q^{m\times(n+1)}$. It follows by transitivity that the binary LWE distribution $\mathbf{LWE}_{\pm}(q, n, m, \hat{\sigma})$ is computationally indistinguishable from the uniform distribution over $\mathbb{Z}_q^{m\times(n+1)}$, i. e., $\mathbf{LWE}_{\pm}(q, n, m, \hat{\sigma})$ is pseudorandom. □

The statement in the above theorem can be simplified using Lemma 2.9 to rephrase it in terms of the basic LWE problem, and by noticing that $\mathbf{LWE}(q, k, n, \sigma)$ does not get any easier when k and σ grow, or when n gets smaller.

Corollary 3.4. *Assume the distribution* $\mathbf{LWE}(q, k, n+1, \sigma)$ *is pseudorandom for some* $q \leq 2^{n^{O(1)}}$, $\sigma \geq \omega(\sqrt{\log n})$, $k \geq \omega(\log n)$, *and* $(n+1) \geq (k+1)\cdot(\log_2(q)+1)$. *Then the distribution* $\mathbf{LWE}_{\pm}(q, n, n^{O(1)}, \hat{\sigma})$ *is also pseudorandom for* $\hat{\sigma} = 2\sigma\sqrt{n+1}$.

Proof. Notice that, under the assumptions in the corollary statement,

$$n \geq (k+1)\log_2 q + k \geq (k+1)\log_2 q + \omega(\log n)$$

as required by Theorem 3.1. In order to invoke the theorem, we also need to verify the pseudorandomness conditions. Assume $\mathbf{LWE}(q,k,n+1,\sigma)$ is pseudorandom. Dropping the last two rows from the samples $[\mathbf{A},\mathbf{b}] \leftarrow \mathbf{LWE}(q,k,n+1,\sigma)$ shows that $\mathbf{LWE}(q,k,n-1,\sigma)$ is also pseudorandom. The samples $[\mathbf{A},\mathbf{b}]$ can also be mapped to $\mathbf{LWE}(q,k+1,n+1,2\sigma)$ by performing the following two operations:

- Add an extra Gaussian error term $\mathbf{e} \leftarrow \mathcal{D}^{n+1}_{\sqrt{3}\sigma}$ to $\mathbf{b}$. By Lemma 2.5, this has the effect of increasing the error rate to $\sqrt{\sigma^2 + 3\sigma^2} = 2\sigma$.
- Append an extra column $\mathbf{a}$ to $\mathbf{A}$ and add a random multiple $\mathbf{a} \cdot s$ to $\mathbf{b}$. This has the effect of extending the secret with an extra coordinate s.

Since this transformation also preserves the uniform distribution, it provides a reduction from

$$\mathbf{LWE}(q,k,n+1,\sigma) \qquad \text{to} \qquad \mathbf{LWE}(q,k+1,n+1,2\sigma)\,,$$

and proves that $\mathbf{LWE}(q,k+1,n+1,2\sigma)$ is pseudorandom. Finally, by Lemma 2.9,

$$\mathbf{LWE}(q,k\times m,n-1,\sigma) \qquad \text{and} \qquad \mathbf{LWE}(q,(k+1)\times m,n+1,2\sigma)$$

are also pseudorandom, as required by Theorem 3.1. □

Notice how Corollary 3.4 establishes the pseudorandomness of $\mathbf{LWE}_{\pm}$ for any polynomial number of samples $m = n^{O(1)}$, using, as an assumption, only the pseudorandomness of **LWE** for a fixed number $(n+1 \approx k\log q)$ of samples. (This property is also implicit in [6].) We remark that we phrased Theorem 3.1 and Corollary 3.4 asymptotically (in terms of polynomial-time distinguishers achieving at most negligible advantage $\varepsilon = n^{-\omega(1)}$) only for simplicity. All statements and proofs are easily adapted to other settings, e. g., to prove hardness of binary LWE against adversaries running in subexponential time.

4 Conclusion

We presented a simple proof that the LWE problem with binary secret of size $n = O(k\log_2 q)$ is as hard as LWE with uniformly random secret in $\mathbb{Z}_q^k$. More specifically, if LWE with secrets in $\mathbb{Z}_q^k$ and $n \approx k\log q$ samples is pseudorandom, then LWE with secrets in $\{0,1\}^n$ or $\{\pm 1\}^n$ (and an arbitrary polynomial number of samples $n^{O(1)}$) is also pseudorandom. As already observed in [6], the growth in the dimension of the secret is seemingly optimal, because it approximately preserves the bit-size of the secret, and the cost of a brute force attack. Starting from LWE with a fixed number of samples $m \approx k\log q = O(k\log k)$ (for typical modulus $q = k^{O(1)}$ polynomial in the LWE secret dimension k) is potentially useful for cryptanalysis, as it allows to generate and publish fixed-size random challenges for any value of k. (By contrast, the general LWE problem would require to give to the adversary access to an LWE sampling oracle that can be called an arbitrary number of times.) An interesting question is whether a reduction can be given starting from LWE with an even smaller number of samples, e. g., $m = O(k)$ linear in the secret dimension.

An important open problem is whether similar results can be proved for the structured variants of LWE based on algebraic lattices [15, 14]. The use of structured lattices is of primary importance to make lattice cryptography efficient in practice, and the use of LWE with binary secrets plays an important role in some applications, like Fully Homomorphic Encryption schemes [9, 8], to control the noise growth when computing on encrypted data. We remark that the use of binary secrets and errors does not seem to pose any difficulty in the setting of one-way hash functions based on structured lattices [15]. However, for LWE [21, 14], it is unclear how to adapt the proof in this paper to the algebraic lattice setting. We hope our simple proof for unstructured lattices will bring more attention to this problem, and serve as a possible starting point to establish similar results for ring LWE.

Acknowledgments. The author thanks the anonymous Theory of Computing referees and editor Oded Regev for useful comments on earlier drafts of this paper.

References

[1] MARTIN R. ALBRECHT: On dual lattice attacks against small-secret LWE and parameter choices in HElib and SEAL. In *Proc. 36th Ann. Internat. Conf. on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'17)*, pp. 103–129. Springer, 2017. [doi:10.1007/978-3-319-56614-6_4] 3

[2] MARTIN R. ALBRECHT, CARLOS CID, JEAN-CHARLES FAUGÈRE, ROBERT FITZPATRICK, AND LUDOVIC PERRET: Algebraic algorithms for LWE problems. *ACM Comm. Computer Algebra*, 49(2):62, 2015. [doi:10.1145/2815111.2815158] 3

[3] BENNY APPLEBAUM, DAVID CASH, CHRIS PEIKERT, AND AMIT SAHAI: Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In *Proc. 29th Ann. Internat. Crypto. Conf. (CRYPTO'09)*, pp. 595–618. Springer, 2009. [doi:10.1007/978-3-642-03356-8_35] 2

[4] SANJEEV ARORA AND RONG GE: New algorithms for learning in presence of errors. In *Proc. 38th Internat. Colloq. on Automata, Languages and Programming (ICALP'11)*, pp. 403–415. Springer, 2011. [doi:10.1007/978-3-642-22006-7_34] 3

[5] SHI BAI AND STEVEN D. GALBRAITH: Lattice decoding attacks on binary LWE. In *Proc. 19th Australasian Conf. on Information Security and Privacy (ACISP'14)*, pp. 322–337. Springer, 2014. [doi:10.1007/978-3-319-08344-5_21] 3

[6] ZVIKA BRAKERSKI, ADELINE LANGLOIS, CHRIS PEIKERT, ODED REGEV, AND DAMIEN STEHLÉ: Classical hardness of learning with errors. In *Proc. 45th STOC*, pp. 575–584. ACM Press, 2013. [doi:10.1145/2488608.2488680, arXiv:1306.0281] 2, 3, 8, 14

[7] JOHANNES A. BUCHMANN, FLORIAN GÖPFERT, RACHEL PLAYER, AND THOMAS WUNDERER: On the hardness of LWE with binary error: Revisiting the hybrid lattice-reduction and meet-in-the-middle attack. In *Proc. 8th Internat. Conf. on Progress in Cryptology (AFRICACRYPT'16)*, pp. 24–43. Springer, 2016. [doi:10.1007/978-3-319-31517-1_2] 3

[8] ILARIA CHILLOTTI, NICOLAS GAMA, MARIYA GEORGIEVA, AND MALIKA IZABACHÈNE: Faster fully homomorphic encryption: Bootstrapping in less than 0.1 seconds. In *Proc. 22nd Internat. Conf. on the Theory and Application of Cryptology and Information Security (ASIACRYPT'16)*, pp. 3–33. Springer, 2016. [doi:10.1007/978-3-662-53887-6_1] 2, 15

[9] LÉO DUCAS AND DANIELE MICCIANCIO: FHEW: Bootstrapping homomorphic encryption in less than a second. In *Proc. 34th Ann. Internat. Conf. on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'15)*, pp. 617–640. Springer, 2015. [doi:10.1007/978-3-662-46800-5_24] 2, 15

[10] CRAIG GENTRY, CHRIS PEIKERT, AND VINOD VAIKUNTANATHAN: Trapdoors for hard lattices and new cryptographic constructions. In *Proc. 40th STOC*, pp. 197–206. ACM Press, 2008. [doi:10.1145/1374376.1374407] 5

[11] SHAFI GOLDWASSER, YAEL TAUMAN KALAI, CHRIS PEIKERT, AND VINOD VAIKUNTANATHAN: Robustness of the learning with errors assumption. In *Proc. 1st Conf. on Innovations in Theoret. Computer Science (ITCS'10)*, pp. 230–240. Tsinghua University Press, 2010. 2, 3, 11

[12] JOHAN HÅSTAD, RUSSELL IMPAGLIAZZO, LEONID A. LEVIN, AND MICHAEL LUBY: A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999. [doi:10.1137/S0097539793244708] 4

[13] PAUL KIRCHNER AND PIERRE-ALAIN FOUQUE: An improved BKW algorithm for LWE with applications to cryptography and lattices. In *Proc. 35th Ann. Internat. Crypto. Conf. (CRYPTO'15)*, pp. 43–62. Springer, 2015. [doi:10.1007/978-3-662-47989-6_3, arXiv:1506.02717] 3

[14] VADIM LYUBASHEVSKY, CHRIS PEIKERT, AND ODED REGEV: On ideal lattices and learning with errors over rings. *J. ACM*, 60(6):43:1–43:35, 2013. Preliminary version in EUROCRYPT'10. [doi:10.1145/2535925] 15

[15] DANIELE MICCIANCIO: Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. *Comput. Complexity*, 16(4):365–411, 2007. Preliminary version in FOCS'02. [doi:10.1007/s00037-007-0234-9] 15

[16] DANIELE MICCIANCIO AND PETROS MOL: Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In *Proc. 31st Ann. Internat. Crypto. Conf. (CRYPTO'11)*, pp. 465–484. Springer, 2011. [doi:10.1007/978-3-642-22792-9_26] 8

[17] DANIELE MICCIANCIO AND CHRIS PEIKERT: Hardness of SIS and LWE with small parameters. In *Proc. 33rd Ann. Internat. Crypto. Conf. (CRYPTO'13)*, pp. 21–39. Springer, 2013. [doi:10.1007/978-3-642-40041-4_2] 3, 6

[18] DANIELE MICCIANCIO AND ODED REGEV: Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007. Preliminary version in FOCS'04. [doi:10.1137/S0097539705447360] 5

[19] CHRIS PEIKERT: Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In *Proc. 41st STOC*, pp. 333–342. ACM Press, 2009. [doi:10.1145/1536414.1536461] 8

[20] CHRIS PEIKERT, ODED REGEV, AND NOAH STEPHENS-DAVIDOWITZ: Pseudorandomness of ring-LWE for any ring and modulus. In *Proc. 49th STOC*, pp. 461–473. ACM Press, 2017. [doi:10.1145/3055399.3055489] 8

[21] ODED REGEV: On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):34:1–34:40, 2009. Preliminary version in STOC'05. [doi:10.1145/1568318.1568324] 1, 8, 15

[22] ODED REGEV: The learning with errors problem (invited survey). In *Proc. 25th Conf. on Computational Complexity (CCC'10)*, pp. 191–204. IEEE Comp. Soc. Press, 2010. [doi:10.1109/CCC.2010.26] 1

AUTHOR

Daniele Micciancio
Professor
University of California at San Diego (UCSD)
La Jolla, CA, USA
daniele@cs.ucsd.edu
http://cseweb.ucsd.edu/~daniele/

ABOUT THE AUTHOR

DANIELE MICCIANCIO got his Ph. D. in Computer Science from M.I.T. in 1998, under the supervision of Shafi Goldwasser. He is a full professor in CSE department at the University of California, San Diego (UCSD), where he has worked since 1999. He is broadly interested in theoretical computer science and cryptography. His research focuses on lattice cryptography and the complexity of lattice and coding problems.