LWE is as Hard as Continuous LWE
& a Search-to-Decision Reduction

Prince Kirpa[1]

[1]kirpa@prince.ac
[1]Computer Science, MUIC, Mahidol University

September 5, 2026

Abstract

In 2021, Bruna, Regev, Song and Tang introduced the continuous analogue of LWE, CLWE. Shortly after, Gupte, Vafa, and Vaikuntanathan showed a classical reduction from LWE to CLWE and from discrete-CLWE (secret from a discrete set) to LWE. In this paper, we show two results. The first result is a search-to-decision reduction for CLWE. The search-to-decision reduction works similarly to the standard reduction for LWE. The main difference between the reductions is the continuous nature of the CLWE secret. As a consequence, our reduction can find the secret up to a polynomial error n^{-k} for some constant k.

The second and our main result is the reduction from CLWE to discrete-CLWE. The reduction is an identity reduction, i.e. samples are passed as-is to the distinguisher. To prove the distinguisher can accept the samples we employ tools from analytic number theory. These tools allow us to show that the set $\frac{1}{r} \cdot \mathbb{Z}^n$ with fixed ℓ_2 norm r, i.e. integer vectors projected onto the unit sphere, can mimic the unit sphere S^{n-1} for some $r > r'$. By completing this missing link, we can show that LWE is equivalent to CLWE i.e. LWE $\equiv$ CLWE and existing or new results on each problem apply to the other.

1 Introduction

The learning with errors (LWE) problem introduced two decades ago in [Reg05], has been and continues to be the leader in post-quantum cryptography. LWE is a list of equations with error, in its most basic form, is defined as

$$(\mathbf{a}_i, b_i = \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i)_{i=1}^m$$

where $\mathbf{s} \in \mathbb{Z}_q^n$ is fixed, $\mathbf{a}_i \sim U(\mathbb{Z}_q^n)$ and $e \sim \chi$ where χ is a discrete distribution. The goal of the problem is to find the secret $\mathbf{s}$ given the samples $(\mathbf{a}_i, b_i)$. Since its inception many variants of the problem have been proposed, one of those is the continuous variant of LWE, continuous LWE (CLWE). Introduced relatively recently in [Bru+21], the main motivation for its creation is to connect LWE to learning gaussian mixtures. CLWE acts

as a bridge between post-quantum cryptography and theoretical machine learning. The problem is still a relatively new problem and we believe it has much to offer. In [Bru+21], a polynomial-time quantum reduction from worst-case lattice problems to CLWE was given, similar to the reduction found in [Reg05]. Furthermore, Bruna et al. resolved an open problem in learning mixtures of Gaussians and another open problem in the statistical query model concerning its computational hardness. CLWE, defined similarly to LWE, is a list of equations

$$(\mathbf{a}_i, b_i = \gamma \cdot \langle \mathbf{a}_i, \mathbf{w} \rangle + e_i \pmod 1)_{i=1}^m$$

where $\mathbf{w} \sim S^{n-1}$ is fixed, $\mathbf{a}_i \sim \mathcal{N}(\mathbf{0}, I_{n\times n})$, $e_i \sim \mathcal{N}(0, \beta^2)$ and $\gamma, \beta \in \mathbb{R}_{>0}$. Subsequently, a classical reduction was given from LWE to CLWE in [GVV22] and further applications to learning Gaussian mixtures. In [GVV22], a fully classical reduction was given for LWE to CLWE and discrete-CLWE to LWE, bypassing the quantum reduction in [Bru+21]. The reverse reduction was from discrete-CLWE to LWE and discrete-CLWE is the CLWE distribution where the vector $\mathbf{w}$ is from a discrete set, not uniformly random from a unit sphere $U(S^{n-1})$. However, the reduction from general-CLWE to discrete-CLWE was left open. The paper also improves density estimation for mixture of Gaussians which is an improvement on [Bru+21].

In this paper we improve upon [GVV22] by giving a search-to-decision reduction for general-CLWE and we show the reduction from general-CLWE to discrete-CLWE. This reduction completes the missing link and shows that LWE is equivalent to CLWE, hence results on each side applies to the other.

1.1 Search-to-Decision Reduction

Motivations. The main motivation for search-to-decision reductions, especially in cryptography, is to show that the decision version of a problem is as hard as its search version. Therefore, cryptrographers only need to focus on making the distribution of the problem indistinguishiable from its null distribution. Unlike LWE where there are many search-to-decision reductions (e.g. [Reg05], [MP12]), the reduction for CLWE was left open. By showing the reduction in this paper, cryptographers can securely build new cryptosystems based on CLWE. Informally, the theorem for the search-to-decision reduction is the following:

Theorem 1.1. *There is a polynomial time reduction from the search version to the decision version of* CLWE*, if there exists a decision oracle $\mathcal{O}$ that can distinguish samples between* CLWE *and the null distribution with non-negligible advantage.*

Technical Overview. More concretely, a decision oracle $\mathcal{O}$ is powerful enough to find the secret of a CLWE sample, i.e., the decision problem of CLWE is as hard as the search version. The reduction works similarly to the standard LWE reduction, the main difference, as the name implies, being each coordinate of the secret is continuous unlike LWE.

In LWE each coordinate of the secret vector have q possible values $0, 1, ..., q-1$ i.e. $\mathbb{Z}_q$. In CLWE each coordinate also has a range from some $[a, b]$ but infinitely many values to choose from. The reduction below shows that if we guess in small enough steps from a to b, while remaining polynomial, we can get the coordinate within some small error. Each coordinate found is off from the real coordinate by n^{-k} for some constant k.

Because the error is polynomial and not negligible, the statistical distance between samples generated with our guessed secret vector to the the true CLWE distribution is also polynomially bounded. While the error is polynomial this reduction still shows that the decision oracle is powerful enough to find a close enough vector and therefore the security of cryptographic applications built on CLWE can still be broken with the oracle.

Future Work. Is there a search-to-decision reduction for CLWE where the secret found is negligibly close to the real secret?

1.2 CLWE to Discrete-CLWE

Motivations. In this paper, we show a reduction from CLWE to discrete-CLWE completing the connection from LWE to CLWE and shows that LWE is as hard as CLWE. Proving this means new results for CLWE also passes to LWE. Showing the equivalence relation LWE $\equiv$ CLWE, shows that the continuous nature of CLWE is no harder than the discrete nature of LWE. The main theorem stated informally is the following:

Theorem 1.2 (Informal Main Theorem)**.** *Let $\mathcal{S}$ be the uniform distribution over all vectors in $\mathbb{Z}^n$ with fixed norm r. There exists a large enough r' such that for all r greater than r', the distinguisher for* CLWE *with secrets from $\frac{1}{r} \cdot \mathcal{S}$ and the null distibution, is powerful enough to distinguish samples from* CLWE *with secrets over the distribution S^{n-1} and the null distribution.*

Technical Overview. The difficulty of the reduction lies in the lack of obvious ways to manipulate the CLWE samples to follow discrete-CLWE. In order to do this we can look at the problem from another angle and realize the problem is an analogue of the equidistribution of integer points on the unit sphere S^{n-1}. Intuitively, the number of vectors in the set of all integer vectors with fixed ℓ_2 norm r goes to infinity as $r \to \infty$. More concretely, the number of integer vectors with fixed norm r grows asymptotically to [Wai21]

$$r^{n-2}.$$

Hence, the tools of analytic number theory can do the heavy lifting for this proof. By using the theorem of equidistribution of lattice points on the unit sphere in [Iwa97], we can show that when $r > r'$ for some r' the distribution $\frac{1}{r} \cdot \mathcal{S}$ is "close" enough to S^{n-1}. Once we show this we can show that the samples sent to the distinguisher are as if they are discrete-CLWE samples added with negligible error. More concretely, let $\mathbf{w} \sim U(S^{n-1})$, the CLWE samples are

$$(\mathbf{a}, b = \gamma \cdot \langle \mathbf{a}, \mathbf{w} \rangle + e \pmod 1).$$

We can show the samples above are the same as

$$(\mathbf{a}, b = \gamma \cdot \langle \mathbf{a}, \mathbf{w}' \rangle + e + 2^{-n} \pmod 1)$$

where $\mathbf{w}' \in \frac{1}{r} \cdot \mathcal{S}$. This shows a surprising connection between two different fields, where a reduction problem when looked at differently, is essentially an analytic number theory problem.

Note that our reduction is only to the discrete set of *all* vectors in $\mathbb{Z}^n$ with fixed norm r and not any discrete set like in [GVV22].

Future Work. With this reduction we have showed that LWE $\geq$ CLWE and therefore results in CLWE extend to LWE. Are there applications in theoretical machine learning that can apply to LWE and its applications?

2 Preliminaries

Unless otherwise specified, all vectors are row vectors. We let $\mathbf{u}_{n,i}$ denote the n-dimensional vector with all entries equal to 0, except the i-th entry, which is 1 (i.e. a standard basis vector). For a vector $\mathbf{x}$ in $\mathbb{R}^n$ we define the ℓ_p norm to be $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ and for $p = \infty$ we define the infinity norm of $\mathbf{x}$ as $\|\mathbf{x}\|_\infty = \max_{i\in[n]} |x_i|$.

Let $I_{n\times n}$ be the $n \times n$ identity matrix. For $n \in \mathbb{N}$ we let S^{n-1} be the $(n-1)$-dimensional unit sphere or equivalently all unit vectors in $\mathbb{R}^n$. The symbol $\mathbb{T}_q$ is the torus scaled up by q where $\mathbb{T}_q = \mathbb{R}/q\mathbb{Z} = [0, q)$, with $\mathbb{T} := \mathbb{T}_1$ denoting the standard torus.

We say a function is non-negligible in n if the function is at least n^{-k} for some constant $k > 0$. A function $f(n)$ in negligible in n if $\lim_{n\to\infty} n^k f(n) = 0$ for some constant $k > 0$. We let $\epsilon(n)$ be the negligible function 2^{-n}.

2.1 Statistics

We let $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ be the multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ having a pdf proportional to $e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}$.

For arbitrary $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^n$, we define the Gaussian function as $\rho_{s,\mathbf{c}}(\mathbf{x}) = \exp(-\pi \|(\mathbf{x} - \mathbf{c})/s\|^2)$ and we omit $\mathbf{c}$ when $\mathbf{c} = \mathbf{0}$.

Definition 2.1 (Distinguisher's advantage)**.** *We say a distinguisher $\mathcal{A}$ has advantage ϵ in distinguishing two distributions $\mathcal{D}_1, \mathcal{D}_2$ defined as*

$$\left| \Pr_{x\sim\mathcal{D}_1}[\mathcal{A}(x) = 1] - \Pr_{x\sim\mathcal{D}_2}[\mathcal{A}(x) = 1] \right| = \epsilon$$

over any internal randomness of $\mathcal{A}$.

Definition 2.2 (Statistical distance)**.** *For two distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ over $\mathbb{R}^n$ with density functions ϕ_1 and ϕ_2, respectively, we define the* statistical distance *between them as*

$$\Delta(\mathcal{D}_1, \mathcal{D}_2) = \frac{1}{2} \int_{\mathbb{R}^n} |\phi_1(\mathbf{x}) - \phi_2(\mathbf{x})| d\mathbf{x}.$$

Furthermore, for two random variables $X \sim \mathcal{D}_1$ and $Y \sim \mathcal{D}_2$, we define $\Delta(X, Y)$ to be equal to $\Delta(\mathcal{D}_1, \mathcal{D}_2)$.

Theorem 2.1 (Kullback-Leibler divergence for univariate normal distributions ([BA11]))**.** *For two univariate normal distributions p and q, the KL divergence between them is*

$$D_{KL}(p||q) = \ln \frac{\sigma_1}{\sigma_0} + \frac{\sigma_0^2 + (\mu_0 - \mu_1)^2}{2\sigma_1^2} - \frac{1}{2}.$$

Theorem 2.2 (Pinsker's inequality)**.** *The statistical distance between two probability distributions P and Q is bounded by*

$$\Delta(P,Q) \leq \sqrt{\frac{1}{2} D_{KL}(P||Q)}.$$

Theorem 2.3 (PDF of a normal distribution modulo 1 ([MJ00]))**.** *The probability density function of a normal distribution modulo* 1*, where μ is the mean and σ is the standard deviation, is*

$$f_{WN}(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \exp\left[\frac{-(x-\mu+k)^2}{2\sigma^2}\right] = \sum_{n=-\infty}^{\infty} \phi(2\pi n) e^{i2\pi n x}$$

and $\phi(t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2}$.

2.2 Learning with Errors

Definition 2.3 (CLWE distribution [GVV22])**.** *Let $n, m \in \mathbb{N}$, $\gamma, \beta \in \mathbb{R}$, and let $\mathcal{A}$ be a distribution over $\mathbb{R}^n$ and $\mathcal{S}$ be a distribution over S^{n-1}. Let* $\mathrm{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)$ *be the distribution given by sampling*

$$(\mathbf{a}_i, b = \gamma \cdot \langle \mathbf{a}_i, \mathbf{w} \rangle + e_i \pmod 1)_{i=1}^m$$

where $\mathbf{a}_i \sim \mathcal{A}$, $\mathbf{w} \sim \mathcal{S}$ and $e_i \sim \mathcal{N}(0, \beta^2)$. Explicitly, for one sample, the density at $(\mathbf{y}, z) \in \mathbb{R}^n \times [0,1)$ is proportional to

$$\mathcal{A}(\mathbf{y}) \cdot \sum_{k \in \mathbb{Z}} \rho_\beta(z + k - \gamma \cdot \langle \mathbf{y}, \mathbf{w} \rangle)$$

for a fixed $\mathbf{w} \sim \mathcal{S}$. Unless otherwise specified, we always let $\mathcal{A}$ be the multivariate normal distribution $\mathcal{N}(\mathbf{0}, I_{n\times n})$ and $\mathcal{S}$ be the unit sphere S^{n-1}.

Definition 2.4. *The search version of the* CLWE *problem, denoted* $\text{S-CLWE}^m_{\gamma,\beta}$*, is given samples from the* CLWE *distribution, find the secret component $\gamma \cdot \mathbf{w}$.*

Definition 2.5. *The decision version of the* CLWE *problem, denoted* $\text{D-CLWE}^m_{\gamma,\beta}$*, is to distinguish samples from the* CLWE *distribution* $\mathrm{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)$ *from the null distribution $\mathcal{A}^{m\times n} \times \mathbb{T}^m$ with non-negligible advantage ϵ.*

2.3 Number Theory

Definition 2.6 (Positive-definite)**.** *Let M be a symmetric matrix with real entries. M is positive definite if for every nonzero column vector $\mathbf{x}$*

$$\mathbf{x}^\top M \mathbf{x} > 0.$$

Definition 2.7 (Modular and cusp forms)**.** *A modular form with level Γ and weight k is a holomorphic function $f : \mathcal{H} \to \mathbb{C}$ where $\mathcal{H}$ is the upper half-plane $\{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$ and satisfies the following conditions:*

- *For any* $\gamma \in \Gamma$, $f(\gamma(z)) = (cz+d)^k f(z)$.
- *For any* $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, *the function* $(cz+d)^{-k} f(\gamma(z))$ *is bounded as* $\mathrm{Im}(z) \to \infty$.

where $\Gamma < \mathrm{SL}_2(\mathbb{Z})$ *is a subgroup of finite index and* γ *is a matrix*

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \text{ with } \gamma(z) = \frac{az+b}{cz+d}$$

A modular form is also a cusp form if it satisfies the following:

- *For any* $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, *the function* $(cz+d)^{-k} f(\gamma(z)) \to 0$ *as* $\mathrm{Im}(z) \to \infty$.

Definition 2.8 (Spherical functions and harmonic polynomials)**.** *Spherical functions or harmonic polynomials are harmonic functions that are also homogeneous (all terms have the same degree) where a harmonic function is a smooth function* f *on* $D \subset \mathbb{R}^n$ *that satisfies the Laplace equation*

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + ... + \frac{\partial^2 f}{\partial x_n^2} = 0.$$

Theorem 2.4 (Proposition 5.5, [Iwa97])**.** *Let* $f \in S_k(\Gamma, \vartheta)$ *be a cusp form with weight* $k > 2$, *the Fourier coefficients* $a(n)$ *of the expansion of* f *are bounded by*

$$a(n) = O(n^{\frac{k}{2}-1+\sigma+\epsilon})$$

supposing that the Kloosterman sums constructed using the multiplier ϑ *for a group* Γ *satisfies*

$$\sum_{c>0} c^{-2\sigma} |S(n,n;c)| = O(n^{\epsilon})$$

for some $\frac{1}{2} \leq \sigma < 1$ *and any* $\epsilon > 0$ *with the implied constant depending on* σ, ϵ, *and* Γ.

Theorem 2.5 (Corollary 11.3, [Iwa97])**.** *The representation number* $R(r,Q) = \#\{v \in \mathbb{Z}^n : Q(v) = r^2\}$ *where* $Q = \frac{1}{2}A$ *is a quadratic form and* A *is a symmetric matrix of rank* d *satisfies*

$$R(r,Q) = \Theta(r^{k-1})$$

where k *is the weight of the cusp form,* $n \geq 5$ *and the congruence* $Q(v) \equiv r \bmod (2^7|A|^3)$ *is solvable.*

3 Search-to-Decision Reduction

Theorem 3.1. *Let* $n, m \in \mathbb{N}$, $\beta, \gamma \in \mathbb{R}_{>0}$ *and* $\gamma = \tilde{O}(\sqrt{n})$. *Let* $\mathcal{S}$ *be the distribution* S^{n-1}. *If there exists a decision oracle* $\mathcal{O}$ *that can distinguish samples between* $\mathrm{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)$ *and* $D_1^{n \times m} \times U(\mathbb{T}^m)$ *with non-negligible advantage* ϵ, *then there exists a polynomial time reduction from* $\text{S-CLWE}^m_{\gamma,\beta}$ *to* $\text{D-CLWE}^m_{\gamma,\beta}$.

Proof: By Lemma 3.1, there exists a polynomial-time algorithm, which given as input the decision oracle $\mathcal{O}$ and samples $(\mathbf{a}, b = \gamma \cdot \langle \mathbf{a}, \mathbf{w} \rangle + e \pmod 1)$, returns $\mathbf{w}'$. Let CLWE be the distribution where a sample is $(\mathbf{a}, b = \gamma \cdot \langle \mathbf{a}, \mathbf{w} \rangle + e \pmod 1)$ and CLWE$'$ be the distribution where a sample is $(\mathbf{a}, b = \langle \mathbf{a}, \mathbf{w}' \rangle + e \pmod 1)$. Then by Lemma 3.2, the statistical distance between the distributions CLWE and CLWE$'$ is non-negligible and we have found the secret $\gamma \cdot \mathbf{w}$ within n^{-k} of the real secret. This completes the proof. ■

Lemma 3.1. *Let $n, m \in \mathbb{N}$, $\beta, \gamma \in \mathbb{R}_{>0}$ and $\gamma = \tilde{O}(\sqrt{n})$. There exists an efficient polynomial algorithm, where given a decision oracle $\mathcal{O}$ that can distinguish samples between* $\text{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)$ *and $D_1^{n \times m} \times U(\mathbb{T}^m)$ with non-negligible advantage ϵ, and a polynomial number of samples $(\mathbf{a}, b)$ from* $\text{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)$*, returns the secret vector $\mathbf{w}'$ non-negligibly close to $\gamma \cdot \mathbf{w}$. That is $\|\mathbf{w}' - \gamma \cdot \mathbf{w}\|_\infty < n^{-k}$ for some constant k.*

Proof: We now show a polynomial time algorithm where given $\mathcal{O}$ and samples $(\mathbf{a}, b)$ finds the secret vector $\gamma \cdot \mathbf{w}$. Let $m = n^{k_1}$ and $l = n^{k_2}$ be polynomial in n. We know that $\gamma = \tilde{O}(\sqrt{n})$ and therefore $\gamma < c_1\sqrt{n}$ for some constant $c_1 > 0$. For each coordinate in $\gamma \cdot \mathbf{w}$ the range of values are $c_1\sqrt{n} \cdot [-1, 1]$.

Let $s = n^{-k_3}$ be polynomial in n. Starting with $g = -c_1\sqrt{n}$ we pass m samples to the oracle and see whether it accepts or rejects. Repeat this step $l - 1$ more times. Then repeat this whole process with $g = -\gamma + s$ for the entire range.

We now show correctness. Let w_i be the coordinate we are finding. Recall, the m CLWE samples are defined as

$$(\mathbf{a}_j, b_j = \gamma \cdot \langle \mathbf{a}_j, \mathbf{w} \rangle + e_j \pmod 1)_{j=1}^m.$$

For each sample we modify the $\mathbf{a}$ vector by adding r to coordinate i of $\mathbf{a}$ and adding $r + g$ to b, i.e.,

$$(\mathbf{a}'_j = \mathbf{a}_j + r \cdot \mathbf{u}_{n,i}, b'_j = b_j + r \cdot g)_{j=1}^m.$$

where $r = -a_i + x$, $-a_i$ is the ith coordinate of $\mathbf{a}$ and $x \sim \mathcal{N}(0, 1)$. Observe that

$$b' = b + r \cdot g = \gamma \cdot \langle \mathbf{a}, \mathbf{w} \rangle + e + r \cdot g$$

$$= (a_1 \cdot w_1 \cdot \gamma) + ... + (a_i \cdot w_i \cdot \gamma + r \cdot g) + ... + (a_n \cdot w_n \cdot \gamma) + e$$

$$= (a_1 \cdot w_1 \cdot \gamma) + ... + (a_i \cdot w_i \cdot \gamma + r \cdot w_i \cdot \gamma - r \cdot w_i \cdot \gamma + r \cdot g) + ... + (a_n \cdot w_n \cdot \gamma) + e$$

$$= (a_1 \cdot w_1 \cdot \gamma) + ... + (\gamma w_i(a_i + r)) + ... + (a_n \cdot w_n \cdot \gamma) + e + r(g - \gamma w_i)$$

$$= \gamma \cdot \langle \mathbf{a}', \mathbf{w} \rangle + e + r(g - \gamma w_i) = \gamma \cdot \langle \mathbf{a}', \mathbf{w} \rangle + e'.$$

Notice that when $g - \gamma w_i$ is large we have added $(g - \gamma w_i)r$ where $r \sim \mathcal{N}(0, 2)$ mod 1 to the sample. By Claim 3.1, the statistical distance between $\mathcal{N}(0, 2)$ mod 1 and $U[0, 1]$ is a small constant near 0 and the oracle will reject the samples with high probability.

When g is within s of $\gamma \cdot w_i$ i.e., $\gamma \cdot w_i - s \leq g \leq \gamma \cdot w_i + s$, then the error e' is

$$e' = e + r(g - \gamma \cdot w_i) = e + r \cdot \pm s = e + r \cdot \pm n^{-k_3}.$$

The statistical distance between e and e' is

$$\Delta(e, e') = \Delta(\mathcal{N}(0, \beta^2), \mathcal{N}(0, \beta^2 + 2 \cdot n^{-2k_3})).$$

By Theorem 2.1 the KL divergence between $\mathcal{N}(0, \beta^2)$ and $\mathcal{N}(0, \beta^2 + 2 \cdot n^{-2k_3})$ is

$$\ln\left(\frac{\beta^2}{\beta^2 + 2 \cdot n^{-2k_3}}\right)^{1/2} + \frac{\beta^2 + 2 \cdot n^{-2k_3}}{2\beta^2} - \frac{1}{2}$$

$$\text{Let } x = 2 \cdot n^{-2k_3}$$

$$\frac{1}{2}(-\ln(1+x) + (1+x) - 1) \approx \frac{1}{2}(-x + \frac{x^2}{2} + (1+x) - 1)$$

$$= \frac{x^2}{4} = \frac{4n^{-4k_3}}{4\beta^4} = \frac{n^{-4k_3}}{\beta^4}$$

Applying Pinsker's Inequality

$$\Delta(e, e') \leq \frac{n^{-2k_3}}{\beta^2}$$

Then for

$$k_3 > O(k_1)$$

the distinguisher will accept the samples with high probability.

For each step we repeat the number of trials l times to ensure we have the correct guess. Let the probability of the oracle accepting a guess g_i be p_i. Let $\epsilon = \frac{\Delta}{2}$ where Δ is equal to $p_h - p_{sh}$, p_h is the guess with the highest probability of being accepted, and p_{sh} the second highest. For a single guess trialed l times using the Chernoff-Hoeffding bound the absolute value of the observed $\hat{p}_i$ minus p_i being greater than ϵ is

$$P(|\hat{p}_i - p_i| \geq \epsilon) \leq 2e^{-2l\epsilon^2}.$$

To ensure the above holds for all guesses we apply the union bound giving

$$\frac{2c_1\sqrt{n}}{s} \cdot 2e^{-2l\epsilon^2}.$$

For a confidence parameter δ we want

$$\frac{2c_1\sqrt{n}}{s} \cdot 2e^{-2l\epsilon^2} \leq \delta$$

which gives

$$l \geq \frac{\ln(\frac{s\delta}{4c_1\sqrt{n}})}{2\epsilon^2} = \frac{\ln(\frac{s\delta}{4c_1\sqrt{n}})}{2\Delta^2}.$$

Hence, we can repeat l times to ensure we have the guess g where $\gamma \cdot w_i - s \leq g \leq \gamma \cdot w_i + s$ for each coordinate w_i. Once all coordinates have been found return $\mathbf{w}'$ and this concludes the proof. ■

Lemma 3.2. *Let $n, m \in \mathbb{N}$, $\beta, \gamma \in \mathbb{R}_{>0}$ and $\gamma = \tilde{O}(\sqrt{n})$. Let* $\text{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)$ *be the* CLWE *distribution where a sample is* $(\mathbf{a}, b = \gamma \cdot \langle \mathbf{a}, \mathbf{w} \rangle + e \pmod 1)$. *Let $\mathbf{w}'$ be the vector generated by Lemma 3.1 and* $\text{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)'$ *be the CLWE distribution where a sample is* $(\mathbf{a}, b = \langle \mathbf{a}, \mathbf{w}' \rangle + e \pmod 1)$. *Then, the statistical distance between the distributions* $\text{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)$ *and* $\text{CLWE}(m, \mathcal{A}, \mathcal{S}, \gamma, \beta)'$ *is bounded by a non-negligible function.*

Proof: The second component b of a CLWE$'$ sample where $\mathbf{a} \sim \mathcal{N}(\mathbf{0}, I_{n\times n})$ and $e \sim \mathcal{N}(0, \beta^2)$ is

$$\langle \mathbf{a}, \mathbf{w}' \rangle + e = a_1 \cdot w_1' + ... + a_n \cdot w_n' + e$$
$$= a_1(\gamma \cdot w_1 \pm \delta_1) + ... + a_n(\gamma \cdot w_n \pm \delta_n) + e = \gamma \cdot \langle \mathbf{a}, \mathbf{w} \rangle + e + \langle \mathbf{a}, \boldsymbol{\delta} \rangle,$$

where each δ_i is the error of each term as in Lemma 3.1 and $\delta_i \leq n^{-k}$ for some constant k. The b component of the CLWE$'$ sample is $\gamma \cdot \langle \mathbf{a}, \mathbf{w} \rangle + e + \langle \mathbf{a}, \boldsymbol{\delta} \rangle \mod 1$ and the error component is $e + \langle \mathbf{a}, \boldsymbol{\delta} \rangle$. We know $\mathbf{a} \sim \mathcal{N}(\mathbf{0}, I_{n\times n})$ and $\|\boldsymbol{\delta}\|_2 \leq c \cdot n^{-k}$ therefore $\langle \mathbf{a}, \boldsymbol{\delta} \rangle \sim \mathcal{N}(0, \|\langle \mathbf{a}, \boldsymbol{\delta} \rangle\|_2^2) \leq \mathcal{N}(0, c \cdot n^{-k})$. Then $e + \langle \mathbf{a}, \boldsymbol{\delta} \rangle \sim \mathcal{N}(0, \beta^2 + c \cdot n^{-k})$.

Let $e' = e + \langle \mathbf{a}, \boldsymbol{\delta} \rangle$. Then the statistical distance between e and e' is

$$\Delta(e, e') = \Delta(\mathcal{N}(0, \beta^2), \mathcal{N}(0, \beta^2 + c \cdot n^{-k})).$$

By Theorem 2.1 the KL divergence between $\mathcal{N}(0, \beta^2)$ and $\mathcal{N}(0, \beta^2 + c \cdot n^{-k})$ is

$$\ln\left(\frac{\beta^2}{\beta^2 + c \cdot n^{-k}}\right)^{1/2} + \frac{\beta^2 + c \cdot n^{-k}}{2\beta^2} - \frac{1}{2} \approx \frac{c^2 \cdot n^{-2k}}{4\beta^4}$$

By Pinsker's Inequality,

$$\Delta(e, e') \leq \frac{c \cdot n^{-k}}{\sqrt{8} \cdot \beta^2} = c \cdot n^{-k}$$

hence the statistical distance between e and e' is bounded by a polynomial function in n. ■

Claim 3.1. *The statistical distance between $\mathcal{N}(0, 2) \mod 1$ and the uniform distribution $U[0, 1]$ is a small constant.*

Proof: By Theorem 2.3, the pdf of $\mathcal{N}(0, 2) \mod 1$ is:

$$\sum_{n=-\infty}^{\infty} e^{-\frac{1}{2}2(2\pi n)^2} e^{i2\pi nx} = \sum_{n=-\infty}^{\infty} e^{-4\pi^2 n^2} e^{i2\pi nx}$$

$$\to e^{-4\pi^2 0^2} e^{i2\pi 0x} + \sum_{n=-\infty}^{-1} e^{-4\pi^2 n^2} e^{i2\pi nx} + \sum_{n=1}^{\infty} e^{-4\pi^2 n^2} e^{i2\pi nx}$$

$$1 + \sum_{n=1}^{\infty} e^{-4\pi^2 n^2} (e^{i2\pi nx} + e^{-i2\pi nx})$$

$$\text{Euler's formula} \to 1 + \sum_{n=1}^{\infty} 2\cos(2\pi nx) e^{-4\pi^2 n^2}$$

$$\approx 1 + 2\cos(2\pi x) e^{-4\pi^2}.$$

The statistical distance between $\mathcal{N}(0, 2) \mod 1$ and $U[0, 1]$ is:

$$\frac{1}{2} \int_0^1 \left| 1 + 2\cos(2\pi x) e^{-4\pi^2} - 1 \right| dx$$

$$= \frac{e^{-4\pi^2}}{2\pi} \frac{2}{\pi} \approx 7.25 \times 10^{-19}.$$

■

4 Reducing CLWE to LWE

Theorem 4.1 (Main Theorem)**.** *Let $n \geq 4, m \in \mathbb{N}$ and $\gamma, \beta \in \mathbb{R}_{>0}$. Let $\mathcal{S}$ be the uniform distribution over all vectors in $\mathbb{Z}^n$ with fixed ℓ_2 norm r. There exists an r' such that for $r \geq r'$, the statistical distance between the distributions* $\mathrm{CLWE}(m, D_1^n, \frac{1}{r} \cdot \mathcal{S}, \gamma, \beta)$ *and* $\mathrm{CLWE}(m, D_1^n, S^{n-1}, \gamma, \beta)$ *is negligible. Consequently, there is no T-time distinguisher between the distributions* $\mathrm{CLWE}(m, D_1^n, \frac{1}{r} \cdot \mathcal{S}, \gamma, \beta)$ *and $D_1^{n \times m} \times U(\mathbb{T}_1^m)$ assuming there is no T-time distinguisher between the distributions* $\mathrm{CLWE}(m, D_1^n, S^{n-1}, \gamma, \beta)$ *and $D_1^{n \times m} \times U(\mathbb{T}_1^m)$.*

Proof: The distinguisher accepts samples generated using a fixed $\mathbf{w}$ from the distribution $A := \frac{1}{r} \cdot \mathcal{S}$. Recall, our general CLWE samples ($\mathrm{CLWE}(m, D_1^n, S^{n-1}, \gamma, \beta)$) are generated in the following way:

$$(\mathbf{a}_i, b_i = \gamma \cdot \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i \pmod 1)_{i=1}^m$$

where $\mathbf{s} \in S^{n-1}$ is drawn uniformly at random from the unit sphere. Let $\mathbf{t} \in A$ be the closest vector to the secret $\mathbf{s}$, i.e.,

$$\mathbf{t} = \arg\min_{\mathbf{v} \in A} \|\mathbf{v} - \mathbf{s}\| .$$

Let e_i' be defined as

$$e_i' = \gamma \cdot \langle \mathbf{a}_i, \mathbf{s} \rangle - \gamma \cdot \langle \mathbf{a}_i, \mathbf{t} \rangle = \gamma \cdot \langle \mathbf{a}_i, \mathbf{s} - \mathbf{t} \rangle.$$

We can now use Lemma 4.1 to bound the max distance of $\mathbf{s}$ and $\mathbf{t}$. Let $\epsilon(n) := 2^{-n}$ be the negligible function. Then, there exists an r' such that for any $r \geq r'$ the distance between $\mathbf{s}$ and $\mathbf{t}$ is less than $\epsilon(n)$.

By the Cauchy–Schwarz inequality

$$|\langle \mathbf{a}, \mathbf{s} - \mathbf{t} \rangle| \leq \|\mathbf{a}\| \, \|\mathbf{s} - \mathbf{t}\| \leq \|\mathbf{a}\| \cdot \epsilon(n).$$

By standard tail bounds $\|\mathbf{a}\| = O(\sqrt{n})$ and substituting this gives

$$|e_i'| = \gamma \cdot \langle \mathbf{a}, \mathbf{s} - \mathbf{t} \rangle \leq \gamma \cdot O(\sqrt{n}) \cdot \epsilon(n) = \epsilon'(n)$$

where $\epsilon'(n)$ is a new negligible function and e_i' is negligible. Therefore,

$$b_i = (\gamma \cdot \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i) \bmod 1 = (\gamma \cdot \langle \mathbf{a}_i, \mathbf{t} \rangle + e_i + e_i') \bmod 1$$

where $\mathbf{t}$ is in the discrete set A and e_i' is negligible.

We now run the distinguisher for $\mathrm{CLWE}(m, D_1^n, \frac{1}{r} \cdot \mathcal{S}, \gamma, \beta)$ and $D_1^{n \times m} \times U(\mathbb{T}_1^m)$. Since this is an identity reduction, we send the samples unmodified. When a sample is from $\mathrm{CLWE}(m, D_1^n, S^{n-1}, \gamma, \beta)$, we have shown above the samples are computationally indistinguishable from $\mathrm{CLWE}(m, D_1^n, \frac{1}{r} \cdot \mathcal{S}, \gamma, \beta)$. The case for a sample from $D_1^{n \times m} \times U(\mathbb{T}_1^m)$ is trivial and this completes the proof. ■

Lemma 4.1. *Let A be the set $\{\frac{\mathbf{v}}{r} \in \frac{1}{r} \cdot \mathbb{Z}^n \mid \|\mathbf{v}\|_2 = r\}$. For any $\epsilon > 0$, there exists an r' such that for any $r \geq r'$ the distance between the set A and the farthest real point $\tilde{\mathbf{q}}$ is less than ϵ. More concretely,*

$$\max_{\mathbf{q} \in S^{n-1}} \min_{\mathbf{v} \in A} \|\mathbf{v} - \mathbf{q}\|_p < \epsilon$$

for any Archimedean norm p.

Proof: Let $\tilde{\mathbf{q}}$ be the farthest real point from our set A,

$$\tilde{\mathbf{q}} = \arg\max_{\mathbf{q} \in S^{n-1}} \min_{\mathbf{v} \in A} \|\mathbf{v} - \mathbf{q}\|_p$$

where p is an arbitrary Archimedean norm.

Let F be a spherical cap with radius k and center $\tilde{\mathbf{q}}$ on S^{n-1}, such that $F \cap A = \emptyset$. Define the representation number:

$$R(r, Q) = \#\{\mathbf{v} \in \mathbb{Z}^n : Q(\mathbf{v}) = r^2\}$$

where Q is the positive-definite identity matrix. We define the representation number on the domain F as

$$R_F(r, Q) = \#\{\mathbf{v} \in \mathbb{Z}^n : Q(\mathbf{v}) = r^2, \frac{\mathbf{v}}{r} \in F\}.$$

Notice that $R_F(r, Q) = 0$.

Let B be the set $\{\frac{\mathbf{v}}{c \cdot r} \in \frac{1}{c \cdot r} \cdot \mathbb{Z}^n \mid \|\mathbf{v}\|_2 = c \cdot r\}$. Using Claim 4.1, A is a subset of B. Clearly, $|A| = R(r, Q)$ and $|B| = R(c \cdot r, Q)$. From Theorem 4.2, as r goes to infinity and satisfies its congruence condition we have:

$$R_F(r, Q) \sim |F| R(r, Q)$$

where $|F|$ is the surface area of F. Therefore, there must exist a c large enough such that $R_F(c \cdot r, Q) > 0$. Then $\tilde{\mathbf{q}}$ is no longer the farthest real point and B now has a new farthest real point $\tilde{\mathbf{q}}'$. Then we can repeat with $A = B$ and $\tilde{\mathbf{q}} = \tilde{\mathbf{q}}'$ until our distance is less than ϵ. ■

Theorem 4.2 ([Iwa97])**.** *The vectors $v \in \mathbb{Z}^n$, where $n \geq 4$ and $Q(v) = r$ are asymptotically equidistributed as $r \to \infty$ if there exists a $v \in \mathbb{Z}^n$ such that*

$$Q(v) \equiv r \pmod{2^7 |A|^3}$$

where $Q(x) = \frac{1}{2} A[x]$ and A is a symmetric, positive definite matrix.

Proof: We include the proof here for completeness. Let Q be a positive definite quadratic form with integer coefficients and define the unit ellipsoid as

$$\mathcal{E} = \{x \in \mathbb{R}^n : Q(x) = 1\}.$$

Let F be a nice domain. We define the representation number

$$R_F(r, Q) = \#\{m \in \mathbb{Z}^n : Q(m) = r, \frac{m}{\sqrt{r}} \in F\}.$$

We will show that as $r \to \infty$ and satisfying the congruence condition then

$$R_F(r, Q) \sim |F| R(r, Q)$$

where $|F|$ is the surface area of F proportional to the surface area of the ellipsoid. The above property can be redefined as an approximation with respect to any smooth function $f : \mathcal{E} \to \mathbb{C}$ as

$$R_f(r, Q) \sim \left(\int_{\mathcal{E}} f(x) dx \right) R(r, Q)$$

where

$$R_f(r, Q) = \sum_{Q(m)=r} f\left(\frac{m}{\sqrt{r}} \right), \; R(r, Q) = \#\{m \in \mathbb{Z}^n : Q(m) = r\}.$$

All functions f can be constructed from spherical functions that make a complete orthonormal system on $\mathcal{E}$. Therefore, we only need to restrict to when f is a spherical function. See Chapter 9, [Iwa97].

We split into two cases when the degree $\nu = 0$ and $\nu > 0$. Our main term is when $\nu = 0$, i.e. f is a constant. We define the value of the function f for points on the surface area F to be a constant when $\nu = 0$. When $\nu > 0$ this represents our error and f is a harmonic polynomial. We can show that $r^{\frac{\nu}{2}} R_f(r, Q)$ is a Fourier coefficient of a cusp form.

First, the Fourier coefficient of the theta function

$$\Theta(z, Q) = \sum_{m \in \mathbb{Z}^n} e(Q(m)z) = \sum_{r=0}^{\infty} R(r, Q) e(rz)$$

is $R(r, Q)$. We define the theta function with respect to f as

$$\Theta_f(z, Q) = \sum_{m \in \mathbb{Z}^n} f(m) e(Q(m)z).$$

Because f is a homogeneous harmonic polynomial with degree ν,

$$\Theta_f(z, Q) = \sum_{m \in \mathbb{Z}^n} r^{\frac{\nu}{2}} f\left(\frac{m}{\sqrt{r}} \right) e(Q(m)z) = \sum_{r=0}^{\infty} \left(\sum_{Q(m)=r} r^{\frac{\nu}{2}} f\left(\frac{m}{\sqrt{r}} \right) \right) e(rz).$$

Taking out $r^{\frac{\nu}{2}}$ and simplifying

$$\Theta_f(z, Q) = \sum_{r=1}^{\infty} r^{\frac{\nu}{2}} R_f(r, Q) e(rz)$$

which shows $r^{\frac{\nu}{2}} R_f(r, Q)$ is the Fourier coefficient. This cusp form has weight $k = \nu + \frac{n}{2}$, and a multiplier ϑ on the group $\Gamma_0(2N)$. Then we can bound the Fourier coefficients, using Theorem 2.4, to be

$$R_f(r, Q) = O(r^{\frac{n-1}{4} + \epsilon}).$$

Finally, we can prove the theorem. If $n \geq 4$ and f is a linear combination of spherical functions, then

$$R_f(r,Q) = \left(\int_{\mathcal{E}} f(x)dx\right) R(r,Q) + O(r^{\frac{n-1}{4}+\epsilon}).$$

Dividing by $R(r,Q)$

$$\frac{R_f(r,Q)}{R(r,Q)} = \left(\int_{\mathcal{E}} f(x)dx\right) + \frac{O(r^{\frac{n-1}{4}+\epsilon})}{R(r,Q)}.$$

Using Theorem 2.5 we can bound $R(r,Q)$ to be

$$R(r,Q) = \Theta(r^{k-1})$$

and taking the limit as $r \to \infty$

$$\lim_{r\to\infty} \frac{O(r^{\frac{n-1}{4}+\epsilon})}{r^{\nu+\frac{n}{2}-1}} = 0.$$

Therefore,

$$\lim_{r\to\infty} \frac{R_f(r,Q)}{R(r,Q)} = \left(\int_{\mathcal{E}} f(x)dx\right).$$

The requirement for the existence of $v \in \mathbb{Z}^n$ such that

$$Q(v) \equiv r \bmod 2^7|A|^3$$

is needed to show that

$$\delta_p(r,Q) \neq 0 \text{ for all } p \text{ dividing } 2|A|$$

where $\delta_p(r,Q)$ is defined as

$$p^{-a(n-1)}\#\{v \in \mathbb{Z}^n : Q(v) \equiv r \pmod{p^a}\}.$$

Chapter 11, [Iwa97] shows that $\delta_p(r,Q) \neq 0$ for all p not dividing $2|A|$. $\delta_p(r,Q)$ represents the p-adic local density of solutions to $Q(v) = r$. If any local density is zero then there are no global solutions, i.e. no solutions to $Q(v) = r$. ■

Claim 4.1. *Let A be the set of vectors $v \in \mathbb{Z}^n$ with fixed norm r. Let B be the set of vectors $v \in \mathbb{Z}^n$ with fixed norm $c \cdot r$ where $c \in \mathbb{Z}_{>1}$. Then*

$$\frac{1}{r} \cdot A \subset \frac{1}{c \cdot r} \cdot B.$$

Proof: Let $\mathbf{q} \in A$ and $\|\mathbf{q}\|_p = r$. Then $\|c \cdot \mathbf{q}\|_p = \sqrt[p]{c^p(q_1^p + ... + q_d^p)} = c \cdot \sqrt[p]{(q_1^p + ... + q_d^p)} = c \cdot r$. Therefore $c \cdot \mathbf{q} \in B$. Then $\frac{\mathbf{q}}{r} = \frac{c\cdot\mathbf{q}}{c\cdot r}$ and

$$\frac{1}{r} \cdot A \subset \frac{1}{c \cdot r} \cdot B.$$

■

5 Acknowledgements

Thank you Songpon Teerakanok, Pakawut Jiradilok, Kanat Tangwongsan and Panupong Cheewaphutthisakun for helpful advice and suggestions. I greatly thank Thotsaporn Thanatipanonda for his support and feedback.

References

[BA11] Dmitry I. Belov and Ronald D. Armstrong. "Distributions of the Kullback–Leibler divergence with applications". In: *British Journal of Mathematical and Statistical Psychology* 64.2 (2011), pp. 291–309. DOI: https://doi.org/10.1348/000711010X522227.

[Bru+21] Joan Bruna et al. "Continuous LWE". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. 2021, pp. 694–707.

[GVV22] Aparna Gupte, Neekon Vafa, and Vinod Vaikuntanathan. "Continuous lwe is as hard as lwe & applications to learning gaussian mixtures". In: *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2022, pp. 1162–1173.

[Iwa97] Henryk Iwaniec. *Topics in classical automorphic forms*. Vol. 17. American Mathematical Soc., 1997.

[MJ00] Kantilal Varichand Mardia and Peter E. Jupp. *Directional Statistics*. Vol. 494. Chichester: John Wiley & Sons, 2000. ISBN: 978-0-471-95333-3. DOI: 10.1002/9780470316979.

[MP12] Daniele Micciancio and Chris Peikert. "Trapdoors for Lattices: Simpler, Tighter, Faster, Smaller". In: *Advances in Cryptology – EUROCRYPT 2012*. Ed. by David Pointcheval and Thomas Johansson. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 700–718. ISBN: 978-3-642-29011-4.

[Reg05] Oded Regev. "On lattices, learning with errors, random linear codes, and cryptography". In: *Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing*. STOC '05. Baltimore, MD, USA: Association for Computing Machinery, 2005, pp. 84–93. ISBN: 1581139608. DOI: 10.1145/1060590.1060603. URL: https://doi.org/10.1145/1060590.1060603.

[Wai21] Stephen Wainger. "An Introduction to the Circle Method of Hardy, Littlewood, and Ramanujan". In: *The Journal of Geometric Analysis* 31.9 (2021), pp. 9113–9130. DOI: 10.1007/s12220-020-00579-9. URL: https://doi.org/10.1007/s12220-020-00579-9.