Public-Key Encryption from Homogeneous CLWE

Andrej Bogdanov[*], Miguel Cueto Noval[**], Charlotte Hoffmann[***], and Alon Rosen[†]

Abstract. The homogeneous continuous LWE (hCLWE) problem is to distinguish samples of a specific high-dimensional Gaussian mixture from standard normal samples. It was shown to be at least as hard as Learning with Errors, but no reduction in the other direction is currently known.

We present four new public-key encryption schemes based on the hardness of hCLWE, with varying tradeoffs between decryption and security errors, and different discretization techniques. Our schemes yield a polynomial-time algorithm for solving hCLWE using a Statistical Zero-Knowledge oracle.

Keywords: public-key encryption, continuous learning with errors, statistical zero-knowledge, hypercontractivity, statistical-computational gaps, discrete gaussian sampling

1 Introduction

Existing public-key encryption schemes are based on relatively few hard computational problems, all from the domains of number theory [RSA78, Rab79, EG85], coding theory [McE78], lattices [AD97, Reg05], and noisy linear algebra [Ale03, ABW10]. Each of these domains yields to different tradeoffs between functionality, security, and efficiency.

In this work we explore public-key encryption based on a new type of assumption: computational hardness in statistical inference. The input of a statistical inference problem is a sequence of independent samples coming from some distribution with unknown parameters. The search (or estimation) task is to identify the parameters; the easier distinguishing (or hypothesis testing) task is to distinguish the samples from ones coming from a fixed null distribution.

Our statistical inference problem of interest is one that has attracted much algorithmic attention: learning Gaussian mixtures in high dimension. A mixture is a convex combination of k Gaussians with different means and possibly different covariance matrices. When k is constant polynomial-time learning algorithms are known [HP15, BS15] assuming sufficiently many samples are available.

[*] Chinese University of Hong Kong. Email: andrejb@cse.cuhk.edu.hk.
[**] Institute of Science and Technology Austria. Email: miguel.cuetonoval@ist.ac.at
[***] Institute of Science and Technology Austria. Email:charlotte.hoffmann@ist.ac.at
[†] Bocconi University and Reichman University. Email: alon.rosen@unibocconi.it.

Diakonikolas et al. [DKS17] showed that in general the learning problem is intractable for statistical query algorithms. Bruna et al. [BRST21] proved that even the task of distinguishing mixtures of Gaussians from standard normal samples is intractable assuming the hardness of short vectors and short bases in lattices (the GapSVP and GapSIVP problems). Gupte et al. [GVV22] recently showed the stronger claim that the hardness can be based on the Learning with Errors (LWE) problem.

The hard Gaussian mixture of [BRST21, GVV22], called the homogeneous Continuous Learning with Errors (hCLWE) distribution, consists of samples in $\mathbb{R}^n$ that have a standard normal distribution in every direction perpendicular to a secret direction $\mathbf{w} \in \mathbb{R}^n$. The distribution in direction $\mathbf{w}$ is a noisy discrete Gaussian, i.e. a mixture of "Gaussian pancakes" of standard deviation $\beta/\sqrt{\beta^2+\gamma^2} \approx \beta/\gamma$ and spacing $\gamma/(\beta^2+\gamma^2) \approx 1/\gamma$ (Figure 1.a). The (decision) hCLWE problem is to distinguish hCLWE samples from purely normal ones.

The full version of this paper [BNHR22] contains all the missing proofs.

1.1 Our contributions

In this work we construct public-key encryption that is at least as hard to break as hCLWE. The hCLWE problem not only inherits advantages of LWE (such as reduction to worst-case hardness and resistance to known quantum attacks), but is potentially more secure: hCLWE is certainly no easier than LWE and can be potentially harder.

Our constructions imply limits on the hardness of hCLWE: just as LWE, hCLWE is tractable in Statistical Zero-Knowledge. It follows that hCLWE is unlikely to be helpful for constructing encryption as secure as NP (unless NP is contained in coAM).

Four Public-Key Encryption Schemes: We present four public-key encryption schemes that offer varying tradeoffs between decryption and security errors, and use different techniques when discretizing continuous values.

The third cryptosystem of Ajtai and Dwork [AD97] already contains essentially all the ingredients needed to obtain hCLWE-based public-key encryption. Our most efficient scheme—discretized encryption—is largely based on it. We believe that our other schemes are simpler to describe, more intuitive to analyze, and offer the potential of wider applicability to other Gaussian mixtures.

Some of our schemes are based on a variant of hCLWE called $(0, 1/2)$-hCLWE. In the 1/2-hCLWE distribution, the mode in the hidden direction $\mathbf{w}$ is shifted by a relative phase of 1/2 (Figure 1.b). The hidden direction in $(0, 1/2)$-hCLWE is a labeled mixture of hCLWE and 1/2-hCLWE (Figure 1.c). Technically, $(0, 1/2)$-hCLWE is at least as hard as LWE and no harder than hCLWE.

Our first scheme ("pancake") is based on hCLWE. It has inverse polynomial decryption and constant security errors. These parameters, along with the specifics of the scheme, already suffice to prove that hCLWE can be solved in

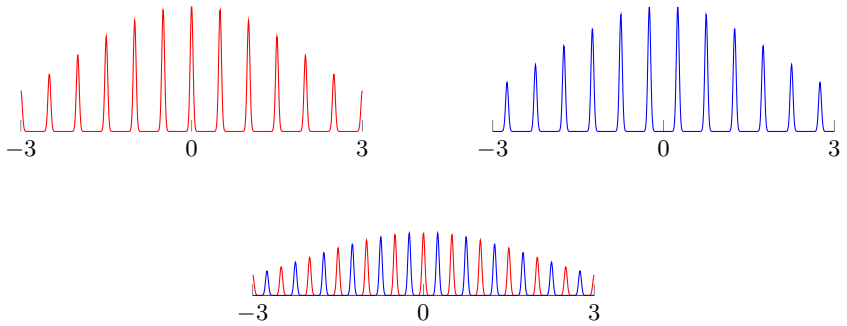

Fig. 1. Probability density function of the hidden direction in the (a) hCLWE, (b) 1/2-hCLWE, and (c) $(0, 1/2)$-hCLWE distributions with parameters $\beta = 0.05$ and $\gamma = 2$

Statistical Zero-Knowledge (SZK), and therefore is in coAM.[1] The discretization step in the scheme can be performed during encryption, and so the public key is continuous. Arguing security then necessitates proving an analog of the leftover hash lemma for Gaussian matrices, which may be of independent interest.

One could in principle rely on standard techniques to reduce decryption and security errors in the first scheme [HR05] , albeit at the price of a significant loss in efficiency. Instead, we present three different ideas to reduce the errors directly.

In the second scheme ("bimodal"), we achieve perfect decryption error by publishing $(0, 1/2)$-hCLWE samples as the public key. To encrypt a 0, Bob uses samples with $z = 0$ and to encrypt a 1, he uses samples with $z = 1/2$. This eliminates the probability that a random normal ciphertext of 1 is of the form of an hCLWE sample and thus makes decryption perfect.

The third scheme ("discretized") achieves negligible security error by mapping the samples into a parallelpiped spanned by hCLWE samples; a technique due to Ajtai and Dwork [AD97]. Here the discretization step takes place already in public-key generation, allowing for the use of the standard leftover hash lemma and yielding favorable security error in comparison with the other schemes.

In the fourth scheme ("baguette") we achieve negligible decryption error assuming only hCLWE. Instead of publishing samples that have a "pancake" distribution in one direction, we sample vectors that have a pancake distribution in ℓ hidden directions. In [BRST21] the authors give a reduction from hCLWE to this hCLWE(ℓ) distribution.

The parallelepiped technique can also be applied to the fourth scheme, yielding an hCLWE-based scheme with negligible decryption and security error. We omit a formal analysis of this step as it is similar to the discretized scheme.

[1] A distinguishing problem is in class $\mathcal{C}$ if there is an algorithm in $\mathcal{C}$ that accepts at least 2/3 of the yes instances and rejects at least 2/3 of the no instances.

Scheme	Assumption	Decryption error	Security error	PK size	SK size
Pancake	hCLWE	$O(1/n)$	$1/4$	$\tilde{O}(n^3)$	n
Bimodal	$(0, 1/2)$-hCLWE	0	$1/2$	$\tilde{O}(n^3)$	n
Discretized	$(0, 1/2)$-hCLWE	0	2^{-n+2}	$\tilde{O}(n^2)$	n
Baguette	hCLWE(ℓ)	$O(1/n^\ell)$	$1/4$	$\tilde{O}(n^3)$	$n\ell$

Table 1. Comparison of our encryption schemes. If the assumption holds against time $t(n) + n^{O(1)}$ and advantage $\Omega(\epsilon(n))$ adversaries then the corresponding scheme is resilient against time $t(n)$ and advantage (security error $+ \epsilon(n)$) adversaries.

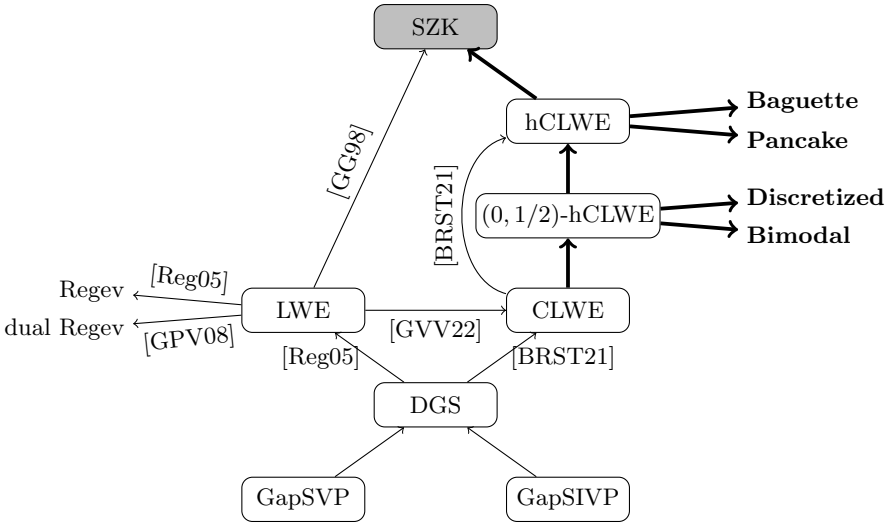

Fig. 2. Reductions between problems and encryption schemes (new results are in bold).

1.2 Related work

Bruna et al. [BRST21] show a worst-case to average-case reduction from Discrete Gaussian Sampling (DGS) to hCLWE. Their reduction factors through an intermediate problem called Continuous LWE (CLWE).

A sample from the CLWE distribution [BRST21] is of the form $(\mathbf{a}, z)$, where $\mathbf{a} \in \mathbb{R}^n$ is a vector with individual entries sampled independently from the standard normal distribution $\mathcal{N}(0, 1)$, and $z := \gamma\langle \mathbf{a}, \mathbf{w}\rangle + e \mod 1$. Here e is the noise drawn from a Gaussian distribution with mean 0 and variance β^2 for some $\beta > 0$, $\gamma > 0$ is a fixed parameter and $\mathbf{w} \in \mathbb{R}^n$ is a secret unit vector. CLWE is the problem of distinguishing multiple CLWE samples from an equal number of samples of the form $(\mathbf{a}, u)$, where u is uniform over $[0, 1)$ and independent of $\mathbf{a}$.

An hCLWE sample is a CLWE sample conditioned on $z = 0$; Bruna et al.'s reduction from CLWE to hCLWE is based on this property. We obtain an analogous reduction from CLWE to $(0, 1/2)$-hCLWE by modifying the condition on z. It is not known if there is a reduction in the opposite direction.

The CLWE problem can be viewed as a continuous analog of Regev's LWE problem [Reg05] and is at least as (quantumly) hard as the same worst-case lattice problems underlying LWE [BRST21]. Gupte et al. [GVV22] recently showed a reduction from LWE to CLWE. They in fact showed that LWE is equivalent in hardness to a variant of CLWE with a different distribution over the secrets that is supported on a discrete subset of the unit sphere. CLWE is at least as hard as this variant.

1.3 CLWE, SZK, and Statistical-Computational Gaps

Several works [BR13, HWX15, BB20] uncover that hypothesis testing tasks in statistical inference tend to exhibit *statistical-computational gaps*: There is a range of sample complexities $m \in [m_{\mathrm{stat}}, m_{\mathrm{comp}}]$ for which hypothesis testing is possible, but no efficient (in terms of the length of a single sample) algorithm is known.

A striking feature of the hCLWE problem is that it is potentially intractable even when the sample complexity is unbounded, i.e., m_{comp} is infinite. Our Theorem 9.2 shows that when $m \geq \tilde{O}(n^2)$ samples are available hCLWE becomes solvable in SZK. Thus, in a world in which SZK = BPP, the computational threshold m_{comp} for hCLWE is at most $\tilde{O}(n^2)$.

In contrast, the statistical threshold for CLWE is $m_{\mathrm{stat}} = O(n)$. It is an intriguing open question whether a statistical-computational gap for hCLWE exists assuming SZK = BPP. One approach for ruling out this possibility is to design a more efficient hCLWE-based PKE scheme.

Applying the reduction from CLWE to hCLWE of Bruna et al., our result also implies that CLWE is in SZK. As their reduction does not preserve sample complexity, the resulting SZK algorithm for CLWE requires a larger number of samples.

2 Technical Overview

The messages in our encryption schemes are single bits. The distributions of encryptions of zero and one, respectively, are efficiently distinguishable with the secret key but not without it. The public keys are independent samples of the hCLWE or $(0, 1/2)$-hCLWE distributions and the secret key is the hidden direction $\mathbf{w}$ of the corresponding yes instances.

As can be seen in Figure 1, the hCLWE samples used to generate the public-key have a periodic discrete structure along the secret direction $\mathbf{w}$. Encryption is designed to retain this discrete structure in the ciphertext even though the sender is oblivious to it. Decryption calculates the correlation between the secret key $\mathbf{w}$ and the ciphertext. This correlation is close to an integer multiple of the period for encryptions of zero and (typically) far from it for encryptions of one.

2.1 "Pancake" Encryption

The first scheme (Section 4) is based on the hCLWE problem. The secret key is a random unit vector $\mathbf{w}$ and the public key is an $n \times m$ matrix $\mathbf{A}$ that consists of m hCLWE samples conditioned on the secret direction $\mathbf{w}$. To encrypt a 0, sample a uniform vector $\mathbf{t} \leftarrow \{1/\sqrt{m}, -1/\sqrt{m}\}^m$ and compute $\mathbf{At}$. To encrypt a 1, sample a standard normal vector. The ciphertext $\mathbf{c}$ is a discretization of the resulting vector using a rounding function that divides the real line into intervals ("buckets") of equal Gaussian measure.[2] To decrypt a ciphertext $\mathbf{c}$, compute $\gamma\sqrt{m}\langle \mathbf{w}, \mathbf{c}\rangle$ and output 0 if the result is close to an integer. Otherwise output 1.

The scheme has inverse polynomial decryption error since the probability of $\gamma\sqrt{m}\langle \mathbf{w}, \mathbf{c}\rangle$ being close to an integer is inverse polynomial for a random choice of $\mathbf{c}$. The main technical contribution in this scheme is the security proof, in particular Proposition 4.3. This result is an analog of the leftover hash lemma for the multiplication of Gaussian matrices with vectors with uniform vectors $\mathbf{t} \leftarrow \{1/\sqrt{m}, -1/\sqrt{m}\}^m$ which shows that the security error is $1/2$ for our choice of parameters.

2.2 "Bimodal" Encryption

In the second scheme (Section 6) we introduce the following changes: We base the scheme on the $(0, 1/2)$-hCLWE problem and publish two matrices $(\mathbf{A}_0, \mathbf{A}_1)$ as the public key. The matrix $\mathbf{A}_0$ consists of hCLWE samples conditioned on $\mathbf{w}$ and $\mathbf{A}_1$ consists of $1/2$-hCLWE samples conditioned on $\mathbf{w}$. To encrypt a 0, do the same as in the pancake scheme with the matrix $\mathbf{A}_0$. To encrypt a 1, do exactly the same with $\mathbf{A}_1$. To decrypt, check if $\gamma\sqrt{m}\langle \mathbf{w}, \mathbf{c}\rangle \mod 1$ is closer to 0 or to $1/2$. Replacing one hCLWE matrix by two $(0, 1/2)$-hCLWE matrices yields perfect decryption error for all but negligibly many choices of the public key. The security error however remains constant.

2.3 "Discretized" Encryption

The third scheme (Section 7) has perfect decryption for all but an inverse polynomial fraction of public keys and negligible security error. To achieve this we make use of the parallelepiped technique due to Ataj and Dwork [AD97] to obtain uniform matrices from $(0, 1/2)$-hCLWE samples.

We change the secret key to $\mathbf{B}^T\mathbf{w}$, where $\mathbf{B}$ is a square matrix whose columns are hCLWE samples. The public key $(\mathbf{A}_0, \mathbf{A}_1)$ again consists of 2 matrices: A matrix $\mathbf{A}_0$ that is obtained by mapping hCLWE samples into the parallelepiped $\mathcal{P}(\mathbf{B})$ spanned by the columns of $\mathbf{B}$, and a matrix $\mathbf{A}_1$ that is obtained in the same way but with $1/2$-hCLWE samples mapped to $\mathcal{P}(\mathbf{B})$. This mapping into the parallelepiped transforms Gaussian vectors in $\mathbb{R}$ into uniform vectors in $\mathcal{P}(\mathbf{B})$,

[2] In the body of the paper we use the notation $1/\gamma' = \gamma/(\beta^2 + \gamma^2)$ for the period of the hCLWE hidden direction. As the difference between $1/\gamma'$ and $1/\gamma$ is small we make no distinction between the two in this overview.

while preserving the pancakes in the secret direction. An additional rounding step discretizes the matrices $\mathbf{A}_0, \mathbf{A}_1$.

To encrypt a bit b, sample a vector $\mathbf{t}$ with uniform entries in $\{-1, 1\}$ and set $\mathbf{c} := \mathbf{A}_b\mathbf{t} \mod q$. To decrypt, check if $\gamma\langle \mathbf{B}^T w, \mathbf{c}/q\rangle \mod 1$ is closer to 0 or to $1/2$. For all but an inverse polynomial fraction of choices of the matrix $\mathbf{B}$ this scheme has perfect correctness. Security follows from the classical leftover hash lemma [IZ89] since the matrices $\mathbf{A}_0$ and $\mathbf{A}_1$ are uniform and discrete.

2.4 "Baguette" Encryption

The fourth scheme (Section 8) is based on the hCLWE(ℓ) problem, which is potentially harder than $(0, 1/2)$-hCLWE. We achieve negligible decryption error by modifying our first scheme as follows: Instead of publishing samples that have a pancake distribution in only one hidden direction, we publish a matrix $\mathbf{A}$ of samples that have a pancake distribution in $\log n$ many hidden directions, i.e. we replace the Gaussian pancakes with "Gaussian Baguettes". To encrypt 0, sample a uniform $\mathbf{t} \leftarrow \{1/\sqrt{m}, -1/\sqrt{m}\}^m$ and compute $\mathbf{At}$, and to encrypt 1, sample a standard normal vector. Discretization is identical to the first scheme.

To decrypt, multiply the ciphertext with a matrix that consists of all hidden directions. If all of the entries in the resulting vectors are close to an integer, output 0, otherwise output 1. While the probability that the inner product of the ciphertext of 1 with one secret direction is close to an integer is polynomial, the probability that this happens for all of the $\log n$ directions is negligible. The security error of this scheme remains constant but could be amplified either by a standard approach or by the above parallelepiped method.

2.5 SZK membership

Our SZK membership proof of hCLWE is established by reduction to the complete problem statistical distance: hCLWE samples are mapped to a distribution that is far from uniform over some discrete set, while standard normal samples are mapped to a distribution that is close to uniform. The two distributions are obtained by pancake encrypting a zero under an actual public key and a random placebo. Completeness and soundness then follow from the functionality and security of pancake encryption.[3]

We find it instructive to directly describe the distributions resulting from this reduction. Our Proposition 4.3 can be interpreted as saying that random $\pm 1/\sqrt{m}$ linear combinations of $m = \tilde{\Theta}(n^2)$ standard Gaussian samples fill up space evenly: For every set of sufficiently large Gaussian measure, the fraction of linear combinations that lands in the set is approximately equal to its measure. Thus if $\mathbb{R}^n$ is partitioned into suitably many regions of equal Gaussian measure, the induced distribution on the regions is close to uniform. In contrast, if there

[3] By relying on discretized encryption instead we can prove the stronger claim of coNISZK membership [GSV99] and improve the sample complexity. Details will be spelled out in the final version.

are periodic gaps in some (unknown) direction like in the hCLWE distribution, the linear combinations of samples are concentrated on few regions and the induced distribution is far from uniform.

An intriguing question left open by our work is if SZK membership also holds for aperiodic mixtures of Gaussians such as the ones underlying the statistical query lower bound of Diakonikolas et al. [DKS17].

3 The (homogeneous) CLWE distribution

Definition 3.1 (CLWE Distribution). *Given a dimension n and parameters $\beta, \gamma > 0$, and a unit vector $\mathbf{w} \in \mathbb{R}^n$, samples $(\mathbf{y}, z) \in \mathbb{R}^n \times [0,1)$ from the* CLWE *distribution $\mathcal{A}_{\mathbf{w},\beta,\gamma,n}$ are generated as follows:*

1. *Sample $\mathbf{y} \leftarrow \mathcal{N}_n(0,1)$.*
2. *Sample $e \leftarrow \mathcal{N}(0,\beta^2)$.*
3. *Output $(\mathbf{y}, \gamma\langle \mathbf{w}, \mathbf{y}\rangle + e \mod 1)$.*

Definition 3.2 (CLWE Distinguishing Problem). *For real numbers $\beta, \gamma > 0$ and $n \in \mathbb{N}$, the (average-case) distinguishing problem* $\mathrm{CLWE}_{\beta,\gamma,n}$ *asks to distinguish between $\mathcal{A}_{\mathbf{w},\beta,\gamma,n}$ for a uniform vector $\mathbf{w} \in \mathbb{R}^n$ and $\mathcal{N}_n(0,1) \times \mathcal{U}$, where $\mathcal{U}$ is the uniform distribution on $[0,1)$.*

Definition 3.3 (hCLWE **Distribution).** *Given a dimension n, parameters $\beta, \gamma > 0$, and a unit vector $\mathbf{w} \in \mathbb{R}^n$, samples $\mathbf{y} \in \mathbb{R}^n$ from the* hCLWE *distribution $\mathcal{H}_{\mathbf{w},\beta,\gamma,n}$ are generated as follows:*

1. *The pancake: Sample $k \in \mathbb{Z}$ with probability proportional to $\exp(-k^2/(2\gamma^2 + 2\beta^2))$.*
2. *The noise: Sample e from $\mathcal{N}(0,\beta'^2)$, where $\beta'^2 = \beta^2/(\gamma^2+\beta^2)$.*
3. *The rest: Sample $\mathbf{w}^\perp$ as $\mathcal{N}_{n-1}(0,1)$ on the subspace orthogonal to $\mathbf{w}$.*
4. *Output $\mathbf{w}^\perp + (k/\gamma' + e)\mathbf{w}$, where $1/\gamma' = \gamma/(\gamma^2+\beta^2)$.*

Definition 3.4 (hCLWE Distinguishing Problem). *For real numbers $\beta, \gamma > 0$ and $n \in \mathbb{N}$, the (average-case) distinguishing problem* $\mathrm{hCLWE}_{\beta,\gamma,n}$ *asks to distinguish between $\mathcal{H}_{\mathbf{w},\beta,\gamma,n}$ for a uniform vector $\mathbf{w} \in \mathbb{R}^n$ and $\mathcal{N}_n(0,1)$.*

The (s,ε) homogeneous CLWE (hCLWE(s,ε)) assumption [BRST21] postulates that for a random $\mathbf{w}$, a hCLWE oracle cannot be distinguished in size s from an oracle that outputs $\mathcal{N}(0,1)$ samples on $\mathbb{R}^n$ with advantage ε. As evidence Bruna, Regev, Song, and Tang show a polynomial-time quantum reduction from the problem of sampling a discrete gaussian of width $O(\sqrt{n}/\beta)$ times the smoothing parameter assuming $\gamma \geq 2\sqrt{n}$. Specifically, if γ and β are polynomial in n then it is plausible that hCLWE holds with s and $1/\varepsilon$ exponential in n. Note that they define the standard normal distribution as $\mathcal{N}(0,1/(2\pi))$ instead of $\mathcal{N}(0,1)$.

It can be shown that all hCLWE versions with different variances are equivalent by rescaling the samples and the problem parameters γ and β. In particular

hCLWE with normal distribution $\mathcal{N}(0, 1/(2\pi))$ and problem parameters γ and β is equivalent to hCLWE with normal distribution $\mathcal{N}(0,1)$ and problem parameters $\gamma/\sqrt{2\pi}$ and $\beta/\sqrt{2\pi}$. We will always work with the $\mathcal{N}(0,1)$ distribution for which $\gamma \geq \sqrt{n}$ is sufficient.

4 Scheme 1: Pancake Encryption

The first encryption scheme relies on the hCLWE assumption and has polynomial decryption- and constant security error. It is the basis for all of the following encryption schemes that achieve better error bounds but either rely on an assumption that is potentially easier to break and/or incur a blow-up in the key size. Furthermore, this scheme enables us to prove that hCLWE is in the complexity class SZK. Before presenting the scheme, we define a rounding function that we will need to discretize the ciphertexts of the scheme.

4.1 Rounding into buckets of equal measure

We use of the following Gaussian rounding function $\mathrm{round}_r \colon \mathbb{R} \to \{1, \ldots, r\}$ given by

$$\mathrm{round}_r(x) = \lceil r \cdot \mu((-\infty, x)) \rceil,$$

where μ is the standard Gaussian measure on the line. In words, partition $\mathbb{R}$ into r intervals ("buckets") $J_1, J_2, \ldots, J_r$ of equal Gaussian measure, and set $\mathrm{round}_r(x)$ to be the unique i such that $x \in J_i$. We extend the definition over $\mathbb{R}^n$ coordinate-wise, i.e. $\mathrm{round}_r(x_1, \ldots, x_n) = (\mathrm{round}_r(x_1), \ldots, \mathrm{round}_r(x_n))$.

Some of the buckets are very wide (at least two of them are infinite!) so the rounding will cause encryption errors with some probability. We will argue that this is an unlikely event using the following regularity property of round_r. The *width* of an interval $J = (a, b)$ is $b - a$.

Proposition 4.1. *For every $0 < \alpha < 1$ and all r such that $r^{1-\alpha} \geq 19$, the number of i for which the width of $J_i = \mathrm{round}_r^{-1}(i)$ exceeds $r^{-\alpha}$ is at most $2r^{\alpha}/\sqrt{\ln r^{1-\alpha}} + 2$.*

The k widest intervals capture a k/r fraction of the probability mass μ at the tails of the normal distribution. If t is chosen so that $\mu((-\infty, t) \cup (t, \infty)) = k/r$ then the next widest interval is of the form (t', t) and t' is uniquely determined by the constraint $\mu((t', t)) = 1/r$. Using suitable analytic approximations for the normal CDF the maximum width $t - t'$ of all remaining intervals can be bounded by $r^{-\alpha}$ when $k = \lfloor 2r^{\alpha}/\sqrt{\ln r^{1-\alpha}} + 2 \rfloor$.

4.2 The encryption scheme

The scheme is parametrized by $\gamma > 0$; $\beta > 0$; $r > 0$ and $n, m \in \mathbb{Z}$.

- The secret key is a uniformly random unit vector $\mathbf{w} \in \mathbb{R}^n$.

- The public key is a matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ whose columns are independent hCLWE samples from $\mathcal{H}_{\mathbf{w},\beta,\gamma,n}$.
- To encrypt a 0, sample a vector $\mathbf{t} \in \{-1/\sqrt{m}, +1/\sqrt{m}\}^m$ uniformly at random and output $\mathbf{c} := \mathrm{round}_r(\mathbf{At})$.
- To encrypt a 1, sample $\mathbf{c} \leftarrow \{1, 2, \ldots, r\}^n$ at random and output $\mathbf{c}$.
- To decrypt a ciphertext $\mathbf{c}$, take any $\mathbf{z}$ such that $\mathrm{round}_r(\mathbf{z}) = \mathbf{c}$, compute $\gamma'\sqrt{m}\langle \mathbf{w}, \mathbf{z} \rangle \mod 1$ and check if it is in the interval $(-1/2n, 1/2n)$. If yes, output 0, else output 1.

Theorem 4.2. *Let* $\gamma = \sqrt{n}$,$\beta = (40000n^{3/2}\log(n))^{-1}$,$r = (40000n^3\log(n))^{5/3}$ *and* $m = 10^8 \log(n)^2 n^2$. *Assuming* hCLWE(s, ε), *the scheme has decryption error* $O(1/n) + \varepsilon$ *and security error at most* $1/4 + 2\varepsilon$.

We prove correctness and security of the scheme separately. We will assume that $\mathbf{w}$ and $\mathbf{A}$ have infinite precision. In Section 4.5 we argue that $O(\log n)$ bits of precision are sufficient.

4.3 Correctness

There are two sources of error in this encryption scheme: key generation error and encryption error. While the key generation error is negligible, the encryption error may be noticeable.

We will call a public key $\mathbf{A}$ *good* if in all its column samples the noise e has magnitude at most $\sqrt{n}\beta$. By hCLWE(s, ε) and a union bound, a public key is good except with probability $m/e^n + \varepsilon$.

The following two claims show that the scheme is correct.

Claim. Assuming hCLWE(s, ε) where s is the complexity of rounding, the probability that $\mathrm{Dec}(\mathbf{w}, \mathrm{Enc}(\mathbf{A}, 0)) \neq 0$ is at most $1/2n + \varepsilon$ for all but a fraction of $m/e^n + \varepsilon$ choices of $\mathbf{A}$.

Claim. The probability that $\mathrm{Dec}(\mathbf{w}, \mathrm{Enc}(\mathbf{A}, 1)) \neq 1$ is at most $3/2n$.

4.4 Security

We show that the above scheme has constant security error by the following argument:

1. Under the hCLWE(s, ε) assumption, $(\mathbf{A}, \mathrm{Enc}(\mathbf{A}, b))$ is ε-indistinguishable from $(\mathbf{N}, \mathrm{Enc}(\mathbf{N}, b))$ for both $b = 0$ and $b = 1$, where $\mathbf{N}$ is a $n \times m$ matrix with i.i.d. entries sampled from $\mathcal{N}(0, 1)$.
2. The distributions $(\mathbf{N}, \mathrm{Enc}(\mathbf{N}, 0))$ and $(\mathbf{N}, \mathrm{Enc}(\mathbf{N}, 1))$ are 1/4-statistically close.
3. It follows that the distributions $(\mathbf{A}, \mathrm{Enc}(\mathbf{A}, 0))$ and $(\mathbf{A}, \mathrm{Enc}(\mathbf{A}, 1))$ are at most $(1/4 + 2\varepsilon)$-indistinguishable.

The first claim follows directly from the hCLWE assumption using the fact that the encryption is an efficiently computable function of the public-key. To prove the second claim (Proposition 4.5) we will argue that for each possible set (bucket) S that is the of the form $\mathrm{round}_r^{-1}(\mathbf{c})$, the random variable $\Pr[\mathbf{Nt} \in S|\mathbf{N}]$ is unlikely to deviate from its mean $\mathbb{E}\left[\Pr[\mathbf{Nt} \in S|\mathbf{N}]\right] = \Pr[\mathbf{g} \in S]$ by much, where $\mathbf{g}$ is a standard normal vector. Then by a union bound over all the buckets we can say that with high probability over the choice of $\mathbf{N}$ the statistical distance between the two distributions is small (given $\mathbf{N}$). Recall that $\mu(S) = \Pr[\mathbf{g} \in S]$ is the standard Gaussian measure over $\mathbb{R}^n$.

Proposition 4.3. *Let $\mathbf{N}$ be an $n \times m$ matrix of independent $\mathcal{N}(0,1)$ random variables, $\mathbf{t}$ a random m-dimensional $\{-1/\sqrt{m}, +1/\sqrt{m}\}$ vector, and S be any event in $\mathbb{R}^n$. Assuming $\mu(S) \geq \exp(-\sqrt{m}/4e)$, we have*

$$\mathrm{Var}\left[\Pr[\mathbf{Nt} \in S|\mathbf{N}]\right] \leq 4e\mu(S)^2 \ln(1/\mu(S))/\sqrt{m}.$$

Proof. Using the definition $\mathrm{Var}[Z] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2$ for any random variable Z we get:

$$\mathrm{Var}\big[\Pr[\mathbf{Nt} \in S|\mathbf{N}]\big] = \Pr[\mathbf{Nt} \in S \text{ and } \mathbf{Nt'} \in S] - \Pr[\mathbf{Nt} \in S]\Pr[\mathbf{Nt'} \in S], \quad (1)$$

where $\mathbf{t}, \mathbf{t'}$ are two independent copies of a random $\pm 1/\sqrt{m}$-valued m-dimensional vector. Let $X = (X_1, \ldots, X_n) = \mathbf{Nt}$ and $X = (X'_1, \ldots, X'_n) = \mathbf{Nt'}$. Conditioned on $\mathbf{t}$ and $\mathbf{t'}$, each pair (X_i, X'_i) is a correlated Gaussian pair (independent of the others) with covariance matrix $\mathbb{E}[X_i^2] = \mathbb{E}[X'^2_i] = 1$, $\mathbb{E}[X_i X'_i] = \rho$, where $\rho = \langle \mathbf{t}, \mathbf{t'} \rangle$ is the inner product of the vectors $\mathbf{t}$ and $\mathbf{t'}$. By contractivity we get

$$\Pr[\mathbf{Nt} \in S \text{ and } \mathbf{Nt'} \in S] \leq \Pr[\mathbf{Nt} \in S]^{1/(1+|\rho|)} \Pr[\mathbf{Nt'} \in S]^{1/(1+|\rho|)}$$

for fixed choices of $\mathbf{t}$ and $\mathbf{t'}$. The quantities $\Pr[\mathbf{Nt} \in S]$ and $\Pr[\mathbf{Nt'} \in S]$ are simply the Gaussian measure $\mu(S)$ of the bucket S, so (1) gives

$$\mathrm{Var}\big[\Pr[\mathbf{Nt} \in S|\mathbf{N}]\big] \leq \mathbb{E}[\mu(S)^{2/(1+|\rho|)} - \mu(S)^2] = \mathbb{E}\big[\mu(S)^{-2|\rho|/(1+|\rho|)} - 1\big]\mu(S)^2. \quad (2)$$

The expectation here is taken over the choice of $\rho = \langle \mathbf{t}, \mathbf{t'} \rangle = (Z_1 + \cdots + Z_m)/m$, where Z_i are i.i.d. ± 1. If we further use $\mu(S) \leq 1$ and $|\rho| \geq 0$, we get that

$$\mathbb{E}\big[\mu(S)^{-2|\rho|/(1+|\rho|)} - 1\big] \leq \mathbb{E}[\mu(S)^{-2|\rho|}] - 1.$$

We further bound this expression by using the following claim:

Claim. $\mathbb{E}[\mu^{-2|\rho|}] \leq \sum_{k=0}^{\infty} (es)^k$, where $s = (2 \ln 1/\mu)/\sqrt{m}$.

By our assumption $\mu(S) \geq \exp(-\sqrt{m}/4e)$, we have $0 \leq es \leq 1/2$ so we get $\sum_k (es)^k = 1/(1-es) \leq 1 + 2es$. Plugging into (2) we get the proposition.

Using Proposition 4.3 we can now bound the statistical distance between $(\mathbf{N}, \mathrm{round}_r(\mathbf{Nt}))$ and $(\mathbf{N}, \mathrm{round}_r(\mathbf{g}))$ which are basically encryptions of 0 and 1 with a standard normal matrix instead of a public key. Security of the scheme then follows from the fact that under the hCLWE assumption $\mathbf{N}$ is indistinguishable from a public key.

Corollary 4.4. *Let* round *be any discrete-valued function on* $\mathbb{R}^n$ *such that the value* $\mu(\text{round}^{-1}(\mathbf{c})) \geq \alpha$ *for all* $\mathbf{c}$ *in the range of* round. *Then the statistical distance between* $(\mathbf{N}, \text{round}(\mathbf{Nt}))$ *and* $(\mathbf{N}, \text{round}(\mathbf{g}))$ *is at most* $\sqrt{4e\ln(1/\alpha)/\sqrt{m}}$.

Proof. We will assume $\alpha \geq \exp(-\sqrt{m}/4e)$ for otherwise $\sqrt{4e\ln(1/\alpha)/\sqrt{m}} \geq 1$ and the claim is true. Fix $\mathbf{c}$ and let $S = \text{round}^{-1}(\mathbf{c})$. Applying the Cauchy-Schwarz inequality to Proposition 4.3 we have

$$\mathbb{E}\big|\Pr[\mathbf{Nt} \in S|\mathbf{N}] - \mu(S)\big| \leq \sqrt{\frac{4e\ln(1/\mu(S))}{\sqrt{m}}} \cdot \mu(S).$$

In particular, if $\mu(\text{round}^{-1}(\mathbf{c})) \geq \alpha \geq \exp(-\sqrt{m}/4e)$ for every $\mathbf{c}$, then

$$\begin{aligned}
&\Delta((\mathbf{N}, \text{round}(\mathbf{Nt})); (\mathbf{N}, \text{round}(\mathbf{g}))) \\
&= \frac{1}{2}\mathbb{E}\left[\sum_{\mathbf{c}}\big|\Pr[\text{round}(\mathbf{Nt}) = \mathbf{c}|\mathbf{N}] - \Pr[\text{round}(\mathbf{g}) = \mathbf{c}|\mathbf{N}]\big|\right] \\
&\leq \frac{1}{2}\sum_{\mathbf{c}}\sqrt{\frac{4e\ln(1/\mu(\text{round}^{-1}(\mathbf{c})))}{\sqrt{m}}} \cdot \mu(\text{round}^{-1}(\mathbf{c})) \\
&\leq \sqrt{\frac{e\ln(1/\alpha)}{\sqrt{m}}}\sum_{\mathbf{c}}\mu(\text{round}^{-1}(\mathbf{c})),
\end{aligned}$$

which is at most the desired expression as the summation equals $\mu(\mathbb{R}^n) = 1$.

Proposition 4.5. *The distributions* $(\mathbf{N}, \text{Enc}(\mathbf{N}, 0))$ *and* $(\mathbf{N}, \text{Enc}(\mathbf{N}, 1))$ *are* $1/4$*-statistically close for a matrix* $\mathbf{N}$ *of independent standard Gaussians.*

Proof. By construction, $\mu(\text{round}_r^{-1}(b)) = r^{-n}$ for all b. By Corollary 4.4 the statistical distance between encryptions is then at most $\sqrt{4e\ln r^n/\sqrt{m}}$ which is at most $1/4$ by our choice of parameters.

Corollary 4.6. *Assuming* $\text{hCLWE}(s, \varepsilon)$, $(\mathbf{A}, \text{Enc}(\mathbf{A}, 0))$ *and* $(\mathbf{A}, \text{Enc}(\mathbf{A}, 1))$ *are* $(s - \text{poly}(n), 1/4 + 2\varepsilon)$*-indistinguishable where* $\mathbf{A}$ *is the public key matrix.*

Proof. Let $\mathbf{N}$ be a random normal matrix. By $\text{hCLWE}(s, \varepsilon)$, $(\mathbf{A}, \text{Enc}(\mathbf{A}, b))$ and $(\mathbf{N}, \text{Enc}(\mathbf{N}, b))$ are $(s - \text{poly}(n), \varepsilon)$-indistinguishable for both $b = 0$ and $b = 1$. By Proposition 4.5, $(\mathbf{N}, \text{Enc}(\mathbf{N}, 0))$ and $(\mathbf{N}, \text{Enc}(\mathbf{N}, 1))$ are $(\infty, 1/4)$-indistinguishable. The corollary follows from the triangle inequality.

4.5 Precision

As we are working with real numbers it is also necessary to discuss how precision can affect the scheme. We denote by ρ the positive integer that determines the precision and for $\rho = \omega(\log n)$ the distance between the real value and the one obtained as a result of the approximation errors is negligible. This guarantees that decryption is not affected (up to a negligible fraction).

5 The s-hCLWE and $(0, 1/2)$-hCLWE Distributions

In this section we introduce two distributions that are indistinguishable from $\mathcal{N}_n(0,1)$ (i.e. n-dimensional vectors with i.i.d. entries from $\mathcal{N}(0,1)$) by the CLWE assumption: the s-hCLWE and the $(0,1/2)$-hCLWE distributions. Samples from the s-hCLWE distribution are CLWE samples $(\mathbf{y}_i, z_i)$ with $z_i = s$. Note that by definition the 0-hCLWE distribution is just the hCLWE distribution. Samples from the $(0,1/2)$-hCLWE distribution are CLWE samples $(\mathbf{y}_i, z_i)$ with $z_i \in \{0, 1/2\}$. We obtain them by flipping a coin and, depending on the outcome, generating either an hCLWE sample or a 1/2-hCLWE sample. In the next two encryption schemes ("bimodal" in Section 6 and "discretized" in Section 7) we use samples from the $(0,1/2)$-hCLWE distribution to construct the public key.

To argue that these two distributions are indistinguishable from $\mathcal{N}_n(0,1)$, we give a reduction from CLWE to both distributions. We also give a reduction from 1/2-hCLWE to hCLWE for completeness even though it is not needed in the rest of the paper.

5.1 The s-hCLWE Distribution

We begin by formally defining the distribution and then we show that there exists a reduction from CLWE.

Definition 5.1 (s-hCLWE Distribution). *For a unit vector $\mathbf{w} \in \mathbb{R}^n$, real numbers $\beta, \gamma > 0$, $n \in \mathbb{N}$ and $s \in [0,1]$, samples $\mathbf{y} \in \mathbb{R}^n$ for the s-hCLWE distribution $\mathcal{H}^s_{\mathbf{w},\beta,\gamma,n}$ are generated as follows:*

1. *Sample $k \in \mathbb{Z} + s$ with probability proportional to $\exp(-k^2/(2\gamma^2 + 2\beta^2))$.*
2. *Sample $e \leftarrow \mathcal{N}(0, \beta'^2)$, where $\beta'^2 := \beta^2/(\gamma^2 + \beta^2)$.*
3. *Sample $\mathbf{v}$ as $\mathcal{N}_{n-1}(0,1)$ from the subspace orthogonal to $\mathbf{w}$.*
4. *Output $\mathbf{y} := \mathbf{v} + (k/\gamma' + e)\mathbf{w}$, where $\gamma' := (\gamma^2 + \beta^2)/\gamma$.*

It follows from the definition that hCLWE corresponds to the case $s = 0$. When $s = 0$, we write $\mathcal{H}_{\mathbf{w},\beta,\gamma,n}$ instead of $\mathcal{H}^0_{\mathbf{w},\beta,\gamma,n}$. The s-hCLWE distinguishing problem is to distinguish between s-hCLWE samples and standard normal ones.

Definition 5.2 (s-hCLWE Distinguishing Problem). *For real numbers $\beta, \gamma > 0$, $n \in \mathbb{N}$ and $s \in [0,1]$, the (average-case) distinguishing problem s-hCLWE$_{\beta,\gamma,n}$ asks to distinguish between $\mathcal{H}^s_{\mathbf{w},\beta,\gamma,n}$ for a uniform unit vector $\mathbf{w} \in \mathbb{R}^n$ and $\mathcal{N}_n(0,1)$.*

We do not consider the worst-case formulation of this problem as it is equivalent to the average-case one. The proof is analogous to [BRST21, Claim 2.22] for hCLWE and CLWE.

We now proceed to compare s-hCLWE to hCLWE and CLWE. First of all, using rejection sampling it is possible to obtain s-hCLWE samples from CLWE samples. This result follows from [BRST21, Lemma 4.1], which shows this for the case $s = 0$. Let $\mathcal{A}_{\mathbf{w},\beta,\gamma,n}$ denote the distribution of CLWE samples.

Lemma 5.3. *For a unit vector* $\mathbf{w} \in \mathbb{R}^n$*, real numbers* $\beta, \gamma > 0$*,* $n \in \mathbb{N}$ *and* $s \in [0,1]$*, there exists a probabilistic algorithm that runs in time poly*$(n, 1/\delta)$ *and that on input* $\delta \in (0,1)$ *and samples from* $\mathcal{A}_{\mathbf{w},\beta,\gamma,n}$*, outputs samples from* $\mathcal{H}^s_{\mathbf{w},\sqrt{\beta^2+\delta^2},\gamma,n}$.

Proof. The same proof as the one of Lemma 4.1 in [BRST21] with $g_0(z) \coloneqq \sum_{k\in\mathbb{Z}} \rho_\delta(z+s+k)$.

If we take $\delta = \beta/\sqrt{2}$, we obtain as a corollary the following reduction:

Proposition 5.4. *For* $s \in [0,1]$*,* $n \in \mathbb{N}$ *and real numbers* $\beta = \beta(n), \gamma = \gamma(n) > 0$ *such that* β *is the inverse of a polynomial in* n*, there exists a polynomial-time reduction from* $\mathrm{CLWE}_{\beta/\sqrt{2},\gamma,n}$ *to* s-$\mathrm{hCLWE}_{\beta,\gamma,n}$.

Now that we have given a reduction from CLWE to s-hCLWE it is a natural question to ask whether there is a reduction from s-hCLWE to CLWE. However, we do not know if this is possible for any value of s.

5.2 The $(0, 1/2)$-hCLWE Distribution

We now define the $(0,1/2)$-hCLWE distribution, which is the distribution on which the following two encryptions schemes are based. Afterwards we show that there is a reduction from CLWE to $(0,1/2)$-hCLWE.

Definition 5.5 ($(0,1/2)$-hCLWE **Distribution).** *For a unit vector* $\mathbf{w} \in \mathbb{R}^n$ *and real numbers* $\beta, \gamma > 0$*,* $n \in \mathbb{N}$ *, samples* $(\mathbf{y}, z) \in \mathbb{R}^n \times \{0, 1/2\}$ *for the* $(0,1/2)$-hCLWE *distribution* $\mathcal{H}^{(0,1/2)}_{\mathbf{w},\beta,\gamma,n}$ *are generated as follows:*

1. *Sample* $z \leftarrow \{0, 1/2\}$.
2. *Sample* $\mathbf{y} \leftarrow \mathcal{H}^z_{\mathbf{w},\beta,\gamma,n}$.
3. *Output* $(\mathbf{y}, z)$.

Definition 5.6 ($(0,1/2)$-hCLWE **Distinguishing Problem).** *For real numbers* $\beta, \gamma > 0$ *and* $n \in \mathbb{N}$ *, the (average-case) problem* $(0,1/2)$-$\mathrm{hCLWE}_{\beta,\gamma,n}$ *asks to distinguish between* $\mathcal{H}^{(0,1/2)}_{\mathbf{w},\beta,\gamma,n}$ *for a uniform unit vector* $\mathbf{w} \in \mathbb{R}^n$ *and* $\mathcal{N}_n(0,1) \times \mathcal{U}(\{0,1/2\})$.

Lemma 5.7. *For a unit vector* $\mathbf{w} \in \mathbb{R}^n$*,* $n \in \mathbb{N}$ *and real numbers* $\beta, \gamma > 0$ *, there exists a probabilistic algorithm that runs in time* $\mathrm{poly}(n, 1/\delta)$ *and that on input* $\delta \in (0,1)$ *and samples from* $\mathcal{A}_{\mathbf{w},\beta,\gamma,n}$*, outputs samples from* $\mathcal{H}^{(0,1/2)}_{\mathbf{w},\sqrt{\beta^2+\delta^2},\gamma,n}$.

Proof. We first sample $z \leftarrow \{0,1/2\}$ uniformly at random. By Lemma 5.3 we can obtain a sample $\mathbf{y}$ from $\mathcal{H}^z_{\mathbf{w},\sqrt{\beta^2+\delta^2},\gamma,n}$ using samples from $\mathcal{A}_{\mathbf{w},\beta,\gamma,n}$ in time $\mathrm{poly}(n, 1/\delta)$ and $(\mathbf{y}, z)$ is a sample from $\mathcal{H}^{(0,1/2)}_{\mathbf{w},\sqrt{\beta^2+\delta^2},\gamma,n}$.

If we take $\delta = \beta/\sqrt{2}$, we obtain as a corollary the following result:

Proposition 5.8. *For* $n \in \mathbb{N}$ *and real numbers* $\beta = \beta(n), \gamma = \gamma(n) > 0$ *such that* β *is the inverse of a polynomial in* n*, there exists a polynomial-time reduction from* $\mathrm{CLWE}_{\beta/\sqrt{2},\gamma,n}$ *to* $(0,1/2)$-$\mathrm{hCLWE}_{\beta,\gamma,n}$.

5.3 A reduction from 1/2-hCLWE to hCLWE

Finally, we show that there exists a reduction from 1/2-hCLWE to hCLWE (with slightly different parameters) to get a finer understanding of the relative hardness of these phased hCLWE problems. We obtain the reduction by constructing samples from $\mathcal{H}_{\mathbf{w},\sqrt{2}\beta,\sqrt{2}\gamma,n}$ using samples from $\mathcal{H}^{1/2}_{\mathbf{w},\beta,\gamma,n}$.

Lemma 5.9. *For a unit vector $\mathbf{w} \in \mathbb{R}^n$, $n \in \mathbb{N}$, real numbers $\beta, \gamma > 0$ such that $\gamma > \sqrt{n}$, and independent random variables Y_1, Y_2 with distribution $\mathcal{H}^{1/2}_{\mathbf{w},\beta,\gamma,n}$, the distribution of $(Y_1 - Y_2)/\sqrt{2}$ is e^{1-n}-statistically close to $\mathcal{H}_{\mathbf{w},\sqrt{2}\beta,\sqrt{2}\gamma,n}$.*

This gives the following result:

Proposition 5.10. *For $n \in \mathbb{N}$ and real numbers $\beta = \beta(n), \gamma = \gamma(n) > 0$, there exists a polynomial-time reduction from $1/2\text{-hCLWE}_{\beta/\sqrt{2},\gamma/\sqrt{2},n}$ to $\text{hCLWE}_{\beta,\gamma,n}$.*

6 Scheme 2: Bimodal Encryption

In this section we modify the "pancake" scheme from Section 4 to achieve perfect correctness. Note that the decryption error in this scheme can be at least polynomial since the pancakes have polynomial width in the secret direction. This is due to the fact that the hCLWE assumption can be broken whenever the error distribution has exponentially small width as was shown in [BRST21]. A random normal vector therefore "hits" a pancake with probability $1/\text{poly}(n)$. If we encrypt a 1 with such a vector, decryption fails. A standard approach to amplify the decryption error is sending multiple independent ciphertexts of the same message [DNR04]. This amplification increases the size of the ciphertext and the security error since a potential adversary only needs to be successful in decrypting one of the ciphertexts. Instead, we modify the encryption process of the bit 1. We introduce the following two changes:

- The public key consists of two matrices. A matrix $\mathbf{A}_0$ whose columns are independent hCLWE samples and a matrix $\mathbf{A}_1$ whose columns are independent 1/2-hCLWE samples. The samples from both matrices are obtained from the same secret direction $\mathbf{w}$.
- To encrypt a 0, take the matrix $\mathbf{A}_0$ and perform the same encryption as in the first scheme. To encrypt a 1, do exactly the same but with the matrix $\mathbf{A}_1$.

In Section 4 we have already seen that the decryption of Enc(0) is $1/\text{poly}(n)$-close to $0 \mod 1$. We show that in our modified scheme the decryption of Enc(1) is $1/\text{poly}(n)$ to $1/2$ so the scheme has perfect correctness. Security of the scheme follows by Proposition 4.5 and the triangle inequality.

6.1 The encryption scheme

The scheme is parametrized by $\gamma > 0$, $\beta > 0$, $n \in \mathbb{Z}$,$r > 0$ and $m \in \mathbb{Z} \setminus 2\mathbb{Z}$ an odd integer.

- The secret key is a uniformly random unit vector $\mathbf{w} \in \mathbb{R}^n$.
- The public key is a pair of matrices $(\mathbf{A}_0, \mathbf{A}_1) \in \mathbb{R}^{n\times m} \times \mathbb{R}^{n\times m}$. The columns of $\mathbf{A}_0$ are independent hCLWE samples and the columns of $\mathbf{A}_1$ are independent 1/2-hCLWE samples.
- To encrypt a bit $b \in \{0, 1\}$, compute $\mathbf{c} := \mathrm{round}_r(\mathbf{A}_b\mathbf{t})$, where $\mathbf{t} \leftarrow \{-1/\sqrt{m}, 1/\sqrt{m}\}^m$ is sampled uniformly at random. Check if all of the entries of $\mathbf{c}$ correspond to a bucket of width less than $1/(5\sqrt{nm}\gamma')$. If yes, output $\mathbf{c}$. If no, output b.
- To decrypt a ciphertext $\mathbf{c}$, take any $\mathbf{z}$ such that $\mathrm{round}_r(\mathbf{z}) = \mathbf{c}$, compute $\gamma'\sqrt{m} \cdot \langle \mathbf{w}, \mathbf{z}\rangle \mod 1$ and check if it is closer to 0 or closer to 1/2. In the former case output 0 in the latter case output 1.

The continuous quantities $\mathbf{w}, \mathbf{A}_0, \mathbf{A}_1$ are represented with $O(\log n)$ bits of precision. As the precision analysis is analogous to the one for pancake encryption we omit it.

Theorem 6.1. *Let* $\gamma = \sqrt{n}$, $\beta = (40000n^{5/2}\log(n)^2)^{-1}$, $r = (40000n^3\log(n))^{5/3}$ *and* $m = 10^8 n^2 \log(n)^2$. *Assuming* $(0, 1/2)$-hCLWE(s, ε) *we have that for all but a fraction of* $2^{-\Omega(n)}$ *choices of the public key the scheme has perfect correctness and security error at most* $1/2 + 1/n^2 + 3\varepsilon$.

We prove correctness and security of the scheme separately.

6.2 Correctness

We call a public key good if the norm of the noise vector is less than $m\beta'$ in both matrices. This holds except with probability $2^{-\Omega(n)}$. During the construction of the public key it can be efficiently tested if a public key is good by checking if the absolute value of the generated noise value is small enough.

Claim. If the public key is good, the scheme has perfect correctness.

6.3 Security

There are two sources of security error in this scheme:

1. If at least one of the entries of the ciphertext corresponds to a bucket of width larger than $1/(5\sqrt{nm}\gamma')$, the encryption algorithm outputs the plaintext in the clear.
2. If the above event does not happen, the ciphertexts of 0 and of 1 are $1/2+2\varepsilon$-indistinguishable.

Claim. Let $\mathbf{A}_b \in \mathbb{R}^{n\times m}$ be a matrix whose columns consist either of independent hCLWE-samples or of independent 1/2-hCLWE samples. Let $\mathbf{t} \leftarrow \{-1/\sqrt{m}, 1/\sqrt{m}\}^m$ be sampled uniformly at random. Assuming hCLWE(s,ε) and 1/2-hCLWE(s,ε), where s is the complexity of rounding, the probability that any entry of the vector $\mathbf{c} := \mathrm{round}_r(\mathbf{A}_b\mathbf{t})$ corresponds to a bucket of width larger than $1/(5\sqrt{m}\gamma')$ is at most $1/n^2 + \varepsilon$.

Proof. First consider a matrix $\mathbf{A}$ with i.i.d. entries from $\mathcal{N}(0,1)$. Since $\|\mathbf{t}\| = 1$ we get that $\mathbf{At}$ is a vector with i.i.d. entries in $\mathcal{N}(0,1)$. By Proposition 4.1 we know that the number of intervals of length larger than $1/(5\sqrt{nm}\gamma')$ is at most $10\sqrt{nm}\gamma'/\sqrt{\ln(r/(5\sqrt{nm}\gamma'))}+2$, so the probability that any entry lands in such a bucket is at most

$$\frac{10n\sqrt{nm}\gamma'}{r\sqrt{\ln(r/(5\sqrt{nm}\gamma'))}} + \frac{2n}{r} \leq \frac{\gamma' n\sqrt{nm}+2n}{r} \leq \frac{1}{n^2}.$$

The claim follows from the fact that the matrices $\mathbf{A}_0$ and $\mathbf{A}_1$ are ε-indistinguishable from $\mathbf{A}$ and the rounding function being efficiently computable.

Remark 6.2. Note that we can avoid the above event by rejection sampling the public key. Since $\mathbf{t}$ is a unit vector, the absolute value of the inner product of any vector $\mathbf{a}$ with $\mathbf{t}$ is bounded by the norm of $\mathbf{a}$. This means that we can avoid the event that an entry of the ciphertext $\mathbf{c}$ corresponds to a wide bucket by rejection sampling the matrices $\mathbf{A}_0, \mathbf{A}_1$: As long as the rows of these matrices have small enough norm, the entries of the vector $\mathbf{A}_b\mathbf{t}$ will not land in a wide bucket for both $b \in \{0,1\}$. We omit a formal analysis of this optimization because the main security issue is not the rounding error but the probability of distinguishing ciphertexts of 0 and 1 as is shown by the next claim.

Claim. The distributions $(\mathbf{N}_0, \mathbf{N}_1, \mathrm{Enc}(\mathbf{N}_0, 0))$ and $(\mathbf{N}_0, \mathbf{N}_1, \mathrm{Enc}(\mathbf{N}_1, 1))$ are 1/2-statistically close for matrices $\mathbf{N}_0, \mathbf{N}_1$ of independent standard Gaussians.

Proof. By Proposition 4.5 we have

$$\Delta((\mathbf{N}_0, \mathbf{N}_1, \mathrm{Enc}(\mathbf{N}_b, b)), (\mathbf{N}_0, \mathbf{N}_1, \mathbf{g})) \leq 1/4,$$

where $\mathbf{g}$ is a vector with i.i.d. entries sampled uniformly from $\{1, 2, \ldots, r\}$ and $b \in \{0,1\}$. By the triangle inequality we follow that

$$\Delta((\mathbf{N}_0, \mathbf{N}_1, \mathrm{Enc}(\mathbf{N}_0, 0)), (\mathbf{N}_0, \mathbf{N}_1, \mathrm{Enc}(\mathbf{N}_1, 1))) \leq 1/2.$$

Corollary 6.3. *Assuming* $(0,1/2)$-hCLWE(s,ε)*, the distributions* $(\mathbf{A}_0, \mathbf{A}_1, \mathrm{Enc}(\mathbf{A}_0, 0))$ *and* $(\mathbf{A}_0, \mathbf{A}_1, \mathrm{Enc}(\mathbf{A}_1, 1))$ *are* $(s-\mathrm{poly}(n), 1/2+2\varepsilon)$*-indistinguishable where* $\mathbf{A}_0, \mathbf{A}_1$ *are the public key matrices.*

Proof. Let $\mathbf{N_0}, \mathbf{N_1}$ be standard normal matrices. By $(0,1/2)$-hCLWE(s,ε), the distributions $(\mathbf{A}_0, \mathbf{A}_1\mathrm{Enc}(\mathbf{A}_b, b))$ and $(\mathbf{N}_0, \mathbf{N}_1, \mathrm{Enc}(\mathbf{N}_b, b))$ are $(s-\mathrm{poly}(n), \varepsilon)$-indistinguishable for both $b = 0$ and $b = 1$. By Claim 6.3, $(\mathbf{N}_0, \mathbf{N}_1, \mathrm{Enc}(\mathbf{N}_0, 0))$ and $(\mathbf{N}_0, \mathbf{N}_1, \mathrm{Enc}(\mathbf{N}_1, 1))$ are $(\infty, 1/2)$-indistinguishable. The corollary follows from the triangle inequality.

7 Scheme 3: Discretized Encryption

In this section we describe an encryption scheme based on CLWE that has negligible soundness error and perfect correctness for all but a fraction of $1/\mathrm{poly}(n)$ many public keys. The scheme is inspired by the encryption scheme in [AD97] which also achieves negligible soundness error but only polynomial decryption error. We reduce this decryption error by applying their techniques to the bimodal encryption scheme from Section 6 which is based on $(0, 1/2)$-hCLWE. Alternatively, it could be applied to the baguette encryption scheme presented in Section 8 which would yield a scheme based on hCLWE. An important concept from [AD97] is the parallelepiped technique which enables us to transform continuous Gaussian samples into uniform ones. We first describe the technique before we present the encryption scheme and prove its correctness and security.

7.1 The parallelepiped technique and $\mathbb{Z}_q$

We will make use of the parallelepiped technique introduced by Ataj and Dwork in [AD97]. Let $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n) \in \mathbb{R}^{n \times n}$ be an arbitrary matrix of rank n. We denote by $\mathcal{P}(\mathbf{B})$ the n-dimensional parallelepiped that is defined by the columns of $\mathbf{B}$, i.e.

$$\mathcal{P}(\mathbf{B}) := \left\{ \sum_{i \in [n]} \lambda_i \mathbf{b}_i : 0 \leq \lambda_i < 1 \text{ for all } i \in [n] \right\}.$$

We denote by $\mathcal{P}_q(\mathbf{B})$ the set we obtain by partitioning $\mathcal{P}(\mathbf{B})$ into q^n smaller parallelpipeds of equal volume, labelling them by vectors with entries from 0 to $q-1$ and then identifying each vector with the corresponding label, i.e.

$$\mathcal{P}_q(\mathbf{B}) := \left\{ \lfloor q\mathbf{B}^{-1}\mathbf{c} \rfloor : \mathbf{c} \in \mathcal{P}(\mathbf{B}) \right\}.$$

We will later need the following fact:

Fact 7.1. *Let $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n) \in \mathbb{R}^{n \times n}$ be an arbitrary matrix of rank n. Then $(\mathcal{P}_q(\mathbf{B}), +)$ is a group isomorphic to $\mathbb{Z}_q^n$.*

This can be seen by the following argument: We obtain $\mathcal{P}_q(\mathbf{B})$ by partitioning each vector $\mathbf{b}_i$ into q equal parts. Labelling the parts by $\{0, 1, 2, \ldots, q-1\}$ in the natural way gives an isomorphism between the q parts of $\mathbf{b}_i$ and $\mathbb{Z}_q$ for any $i \in [n]$. Fact 7.1 follows by taking the direct product of the labellings of the $\mathbf{b}_i$.

In the construction of our public key we essentially map continuous Gaussian vectors into $\mathcal{P}(\mathbf{B})$. We will need the next lemma to show that this mapping transforms them into uniformly random vectors. We denote by $\eta_\varepsilon(\mathbf{B})$ the smoothing parameter of the lattice with basis $\mathbf{B}$.

Lemma 7.1 ([MR07, Lemma 4.1]). *Let $\mathbf{B} \in \mathbb{R}^{n \times n}$ be a square matrix of rank n. For any $\varepsilon > 0$ and any $s > \eta_\varepsilon(\mathbf{B})$ the statistical distance between $\mathcal{N}_n(0, s^2)$ mod $\mathbf{B}$ and the uniform distribution over $\mathcal{P}(\mathbf{B})$ is at most $\varepsilon/2$.*

The following lemma is a special case of [MR07, Lemma 3.2].

Lemma 7.2. *For any n-dimensional lattice L with basis $B = \{b_1, b_2, \ldots, b_n\}$ we have $\eta_{2^{-n}}(B) \leq \sqrt{n} \max_i \|b_i\|$.*

7.2 The encryption scheme

The scheme is parametrized by $\gamma > 0$; $\beta > 0$; $n, m, q \in \mathbb{Z} \setminus 2\mathbb{Z}$ odd integers. We set n to be an odd integer only to clarify the description and the analysis, m and q however are always required to be odd.

- The secret key is a vector $\mathbf{B}^T\mathbf{w}$, where $\mathbf{w} \in \mathbb{R}^n$ is a uniformly random unit vector and $\mathbf{B}$ is a matrix whose columns consist of hCLWE samples, such that the smallest singular value of $\mathbf{B}$ is larger than $1/m$.
- The public key is a pair of matrices $(\mathbf{A}_0, \mathbf{A}_1) \in \mathbb{Z}_q^{n\times m} \times \mathbb{Z}_q^{n\times m}$. The columns of $\mathbf{A}_0$ and $\mathbf{A}_1$ are of the form

$$B\text{-round}(n\mathbf{a}_i \mod \mathbf{B}),$$

  where B-round $= B$-round$_q : \mathbb{R}^n \to \mathbb{Z}_q^n$ is defined as B-round$_q(a) = \lfloor q\mathbf{B}^{-1}a \rfloor$. In the case of $\mathbf{A}_0$ the vectors $\mathbf{a}_i$ are samples from the hCLWE distribution $\mathcal{H}_{\mathbf{w}.\beta,\gamma,n}$ and in the case of $\mathbf{A}_1$ they are 1/2-hCLWE samples from $\mathcal{H}^{1/2}_{\mathbf{w},\beta,\gamma,n}$.
- To encrypt a bit $b \in \{0, 1\}$, compute

$$\mathbf{c} := \mathbf{A}_b\mathbf{t} \mod q,$$

  where $\mathbf{t} \leftarrow \{-1, 1\}^m$ is sampled uniformly at random.
- To decrypt a ciphertext $\mathbf{c}$, compute

$$\gamma'\langle \mathbf{B}^T\mathbf{w}, \mathbf{c}/q\rangle \mod 1$$

  and check if it is closer to 0 or closer to 1/2. In the former case output 0 in the latter case output 1.

Remark 7.3. In the next section we will see that we require n to be an odd integer only because we need that the inner product of $\mathbf{w}$ with 1/2-hCLWE samples scaled by a factor n is approximately $1/2 \mod 1$ and not 0. One can slightly change the scheme for even values of n: Scale the samples by a factor $n+1$ instead of n. In the rest of the section we will assume that n is odd without loss of generality.

Theorem 7.4. *Set the parameters of the scheme to* $\gamma = \sqrt{n}, m = 8n\log(n), \beta = 1/n^{10}, q = n^7$. *Assuming* $(0, 1/2)$-hCLWE(s, ε) *we get that for all but a fraction of* $1/(8n^{1/2}\log(n)) + O(\varepsilon)$ *choices of the public key the scheme has perfect correctness and negligible soundness error.*

We prove correctness and soundness of the scheme separately in the next two subsections.

7.3 Correctness

We show that for all but a fraction of at most $1/(8n^{1/2}\log(n)) + \varepsilon$ choices of the key pair decryption is always correct. We denote by $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ the

columns of $\mathbf{B}$, by $\{\mathbf{a}_1^0, \ldots, \mathbf{a}_m^0\}$ the hCLWE samples used to construct $\mathbf{A}_0$ and by $\{\mathbf{a}_1^1, \ldots, \mathbf{a}_m^1\}$ the 1/2-hCLWE samples used to construct $\mathbf{A}_1$. We define $\mathbf{e} := \gamma' \mathbf{w}^T \mathbf{B} \mod 1$ which is the noise vector of the hCLWE samples $\mathbf{b}_i$. For $b \in \{0,1\}$ we define

$$\mathbf{e}_b^T := \gamma' \mathbf{w}^T \left(n\mathbf{a}_1^b, n\mathbf{a}_2^b, \ldots, n\mathbf{a}_m^b\right) - b \cdot (1/2, 1/2, \ldots, 1/2) \mod 1.$$

If $b = 0$ this is the vector where each entry is the noise value corresponding to the hCLWE sample scaled by n during the construction of $\mathbf{A}_0$. If $b = 1$ this is the noise vector we get during the construction of $\mathbf{A}_1$. We call a key pair $(\mathbf{B}^T\mathbf{w}, (\mathbf{A}_0, \mathbf{A}_1))$ *good* if the following holds:

1. $\|\mathbf{e}_0\|, \|\mathbf{e}_1\| \leq mn\beta'$;
2. $\|\mathbf{e}\| \leq n\beta'$;
3. For all $i \in [m]$ the entries of $\mathbf{a}_i^0, \mathbf{a}_i^1$ lie in the interval $\left[-n^{3/2}, n^{3/2}\right]$;
4. For all $i \in [n]$ the entries of $\mathbf{b}_i$ lie in the interval $[-n, n]$;
5. the smallest singular value of $\mathbf{B}$ is larger than $1/m$.

Note that all of these conditions can be efficiently tested during the key generation.

Claim. If the $(0, 1/2)$-hCLWE(s, ε) assumption holds, a key pair $(\mathbf{B}^T\mathbf{w}, (\mathbf{A}_0, \mathbf{A}_1))$ is good except with probability $1/(8n^{1/2}\log(n)) + O(\varepsilon)$.

For a proof of this result see the full version.

Claim. If the key-pair $(\mathbf{B}^T\mathbf{w}, (\mathbf{A}_0, \mathbf{A}_1))$ is good, decryption is correct with probability 1.

For a proof of this result see the full version.

7.4 Security

We show that encryptions of 0 and 1 are indistinguishable under the $(0, 1/2)$-hCLWE assumption by showing that the following distributions are indistinguishable for $b \in \{0, 1\}$:

1. Real_b: $(\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_b\mathbf{t} \mod q)$ is a public key of the encryption scheme together with an encryption of b.
2. Hybrid_b: $(\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_b\mathbf{t} \mod q)$ is a tuple where the columns of $\mathbf{A}_0$ and $\mathbf{A}_1$ are uniformly random vectors in $\mathbb{Z}_q^{n \times m}$.
3. Ideal: $(\mathbf{A}_0, \mathbf{A}_1, \mathbf{r})$ is the same as above but with $\mathbf{r}$ a uniformly random vector in $\mathbb{Z}_q^n$.

Real_b and Hybrid_b are computationally indistinguishable under the $(0, 1/2)$-hCLWE assumption. Hybrid_b and Ideal are statistically indistinguishable by the leftover hash lemma. In the rest of the section we formally prove the above statements. We start by showing the first claim.

Claim. Under the $(0, 1/2)$-hCLWE(s, ε) assumption the distributions Real_b and Hybrid_b are $(s - \text{poly}(n), 2^{-n+1} + \varepsilon)$-indistinguishable.

Proof. Assume that there is a distinguisher D that decides if $(\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_b\mathbf{t} \mod q)$ is from Real_b or from Hybrid_b with probability δ. We construct an algorithm D' that distinguishes between $(0, 1/2)$-hCLWE samples and random samples with probability $\delta - 2^{-n+1}$ as follows:

1. Given $\text{poly}(n)$ many $(0, 1/2)$-hCLWE samples $\{(\mathbf{y}_i, z_i)\}_{i\in[\text{poly}(n)]}$, define a matrix $\mathbf{B}$ by choosing n samples with $z_i = 0$ such that the corresponding vectors $\mathbf{y}_i$ are linearly independent. These vectors are the columns of $\mathbf{B}$.
2. Choose m samples of the form $\{(\hat{\mathbf{y}}_i, 0)\}_{i\in[m]}$ and compute

$$\mathbf{y}_i^0 = B\text{-round}\,(n\hat{\mathbf{y}}_i \mod \mathbf{B})$$

   and choose m samples of the form $\{(\tilde{\mathbf{y}}_i, 1/2)\}_{i\in[m]}$ and compute

$$\mathbf{y}_i^1 = B\text{-round}\,(n\tilde{\mathbf{y}}_i \mod \mathbf{B}),$$

   where B-round $= B$-round$_q : \mathbb{R}^n \to \mathbb{Z}_q^n$ is defined as -round$_q(\mathbf{a}) = \mathbf{B}\lfloor q\mathbf{B}^{-1}\mathbf{a}\rfloor$.
3. Let $\mathbf{A}_0$ be the matrix with columns $\mathbf{y}_i^0$ and $\mathbf{A}_1$ be the matrix with columns $\mathbf{y}_i^1$. Give $(\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_b\mathbf{t} \mod q)$ to the distinguisher D.

Note that in the case where the samples $\{(\mathbf{y}_i, z_i)\}_{i\in[\text{poly}(n)]}$ are $(0, 1/2)$-hCLWE samples, $(\mathbf{A}_0, \mathbf{A}_1)$ is a public key of our scheme. It remains to prove that given samples $\{(\mathbf{y}_i, z_i)\}_{i\in[\text{poly}(n)]}$, where the $\mathbf{y}_i$ are normal random vectors and the z_i are uniform in $\{0, 1/2\}$, the resulting matrices $\mathbf{A}_0, \mathbf{A}_1$ are statistically close to uniform matrices in $\mathbf{Z}_q^{n\times m}$. Lemma 7.1 says that if we sample a vector from a Gaussian distribution with standard deviation larger than $\eta_{2^{-n}}(\mathbf{B})$ and map it into $\mathcal{P}_q(\mathbf{B})$, the resulting vector is statistically close to uniform in $\mathcal{P}_q(\mathbf{B})$ and hence in $\mathbb{Z}_q^n$.

Now we only need an upper bound on the smoothing parameter in order to prove that the columns of $\mathbf{A}_0$ and $\mathbf{A}_1$ are sampled from a Gaussian with sufficiently large variance. The length of a vector with entries independently sampled from $\mathcal{N}(0, 1)$ is at most n except with probability $\sqrt{n}e^{-n}$. Hence, the smoothing parameter of $\mathbf{B}$ is at most $n^{3/2}$ by Lemma 7.2 except with probability $\sqrt{n}e^{-n}$. The entries of $\mathbf{A}_0$ and $\mathbf{A}_1$ are sampled from $\mathcal{N}(0, n^2)$. Since $n^2 > n^{3/2}$ we follow from Lemma 7.1 that $\mathbf{A}_0$ and $\mathbf{A}_1$ are 2^{-n+1}-statistically close to uniformly random matrices in $\mathbf{Z}_q^{n\times m}$.

Next we show that Hybrid_b is statistically close to Ideal, which completes the proof of soundness. This can be done using the classical leftover hash lemma [IZ89]. To this end we need to show that multiplication of a $\{-1, 1\}^m$ vector by a uniform matrix $\mathbf{H} \in \mathbb{Z}_q^{m\times n}$ is a universal family of hash functions, i.e.:

Claim. For q odd, $\mathbf{x}, \mathbf{y} \in \{-1, 1\}^m$ such that $\mathbf{x} \neq \mathbf{y}$ we have

$$\Pr_{\mathbf{H}\leftarrow\mathbb{Z}_q^{m\times n}}[\mathbf{H}\mathbf{x} = \mathbf{H}\mathbf{y} \mod q] = q^{-n}.$$

See the full version for a proof. The following is a special case of the leftover hash lemma [IZ89, Reg05]:

Lemma 7.5. *Let q be an odd integer. Let $\mathbf{H} \in \mathbb{Z}_q^{n \times m}$ be with columns chosen uniformly at random from $\mathbb{Z}_q^n$ and $\mathbf{t} \leftarrow \{-1, 1\}^m$ a uniformly random vector. Then the statistical distance of the uniform distribution on $\mathbb{Z}_q^n$ and the distribution given by multiplying $\mathbf{H}$ with $\mathbf{t}$ is at most $(q^n/2^m)^{1/4}$ w.p. $1 - (q^n/2^m)^{1/4}$.*

By our choice of parameters we have $m = 8n\log(n)$ and $q = n^7$. We follow that the statistical distance of Hybrid_0 and Hybrid_1 to Ideal is $(n^{7n}/2^{n^2})^{1/4} \leq 2^{-n}$ for large enough values of n. Hence, Hybrid_0 is at least 2^{-n+1}-close to Hybrid_1. Together with Claim 7.4 this yields that an encryption of 0 is $2^{-n+2} + 2\varepsilon$-indistinguishable from an encryption of 1.

7.5 Precision

A precision value of $\rho = O(\log n)$ guarantees that decryption is unaffected as a result of the approximations. The matrix entries of the public key are integer values.

Correctness of decryption remains unaffected and the proof is analogous to the one given for the pancake scheme in Section 4.5.

8 Scheme 4: Baguette Encryption

We now present a second approach that reduces the decryption error of the pancake scheme. The security error remains constant but could be reduced by the parallelepiped technique presented in Section 7. Instead of publishing samples that have a pancake distribution in only one secret direction, we publish samples that have a pancake distribution in multiple secret directions, i.e. samples from the hCLWE(ℓ) distribution. This is a distribution defined in [BRST21] to which the authors give a reduction from hCLWE. To decrypt we take the inner products of the ciphertext with all secret directions. If the ciphertext is an encryption of 0 all of the results are polynomially close to an integer. If the ciphertext is an encryption of 1, at least one of the results is not close to an integer with high probability since taken modulo 1 they are uniformly random values in $[0, 1)$. Before presenting the encryption scheme we formally define the hCLWE(ℓ) distribution.

8.1 The hCLWE(ℓ) distribution

Both the hCLWE(ℓ), distribution and the corresponding decision problem were introduced in [BRST21]. This problem is the extension of hCLWE to the case of ℓ hidden orthogonal directions.

Definition 8.1 (hCLWE(ℓ) Distribution). *For a matrix $\mathbf{W} = (\mathbf{w}_1 | \ldots | \mathbf{w}_\ell) \in \mathbb{R}^{n \times \ell}$ such that $\mathbf{W}^T\mathbf{W} = \mathbf{I}_\ell$, real numbers $\beta, \gamma > 0$, $n \in \mathbb{N}$ and $\ell \in \mathbb{N}$ with $0 \leq \ell \leq n$, samples $\mathbf{y} \in \mathbb{R}^n$ for the hCLWE(ℓ) distribution $\mathcal{H}_{\mathbf{W},\beta,\gamma,n,\ell}$ are generated as follows:*

1. *Sample* $k_1, \ldots, k_\ell \in \mathbb{Z}$ *independently with distribution* $\mathcal{D}_{\mathbb{Z},\gamma^2+\beta^2}$.
2. *Sample* $e_1, \ldots, e_\ell \leftarrow \mathcal{N}(0, \beta'^2)$ *independently where* $\beta'^2 := \beta^2/(\gamma^2+\beta^2)$.
3. *Sample* $\mathbf{v}$ *as* $\mathcal{N}_{n-\ell}(0,1)$ *from the subspace orthogonal to* $\mathbf{W}$.
4. *Output* $\mathbf{y} := \mathbf{v} + \sum_{i=1}^{\ell}(k_i/\gamma' + e_i)\mathbf{w}_i$ *where* $\gamma' := (\gamma^2+\beta^2)/\gamma$.

For $\ell = 0$ we get the normal distribution with covariance matrix $\mathbf{I}_n$ and for $\ell = 1$ we recover the hCLWE distribution. We refer to the columns of $\mathbf{W}$ as the hidden directions. Note that they are orthonormal vectors.

Definition 8.2 (hCLWE(ℓ) **Distinguishing Problem).** *For real numbers* $\beta, \gamma > 0$, $n \in \mathbb{N}$ *and* $\ell \in \mathbb{N}$ *with* $0 \le \ell \le n$, *the (average-case) distinguishing problem* $\mathrm{hCLWE}_{\beta,\gamma,n}(\ell)$ *asks to distinguish between* $\mathcal{H}_{\mathbf{W},\beta,\gamma,n,\ell}$ *for a uniform matrix* $\mathbf{W} \in \mathbb{R}^{n\times\ell}$ *such that* $\mathbf{W}^T\mathbf{W} = \mathbf{I}_\ell$, *and* $\mathcal{N}_n(\mathbf{0}, 1)$.

The hCLWE(ℓ)(s,ϵ) assumption postulates that the hCLWE(ℓ) distinguishing problem cannot be solved in size s with advantage ϵ. As shown in [BRST21] (Lemma 9.3.), if $n - \ell = \Omega(n^k)$ for some constant $k > 0$, there is an efficient reduction from $\mathrm{hCLWE}_{\beta,\gamma,n-\ell+1}$ to $\mathrm{hCLWE}_{\beta,\gamma,n}(\ell)$.

8.2 Encryption scheme

We now give an encryption scheme that builds on the pancake scheme from Section 4. It achieves negligible decryption error using more hidden directions instead of the $(0, 1/2)$-hCLWE distribution.

The scheme is parametrized by $\gamma > 0$; $\beta > 0$; $r > 0$, $n, \ell, m \in \mathbb{N}$ and a parameter $a > 0$ for which we will only consider two possible values, namely, $a = n$ and $a = 100$.

- The secret key is a uniformly random matrix $\mathbf{W} \in \mathbb{R}^{n\times\ell}$ such that $\mathbf{W}^T\mathbf{W} = \mathbf{I}_\ell$.
- The public key is a matrix $\mathbf{A} \in \mathbb{R}^{n\times m}$ whose columns are independently sampled from $\mathcal{H}_{\mathbf{W},\beta,\gamma,n,\ell}$.
- To encrypt 0, choose a vector $\mathbf{t} \in \{-1/\sqrt{m}, +1/\sqrt{m}\}^m$ uniformly at random and output
$$\mathbf{c} := \mathrm{round}_r(\mathbf{A}\mathbf{t}).$$
Check if all entries of $\mathbf{c}$ correspond to buckets of width less than $1/(4a\sqrt{n}\sqrt{m}\gamma')$. If yes, output $\mathbf{c}$. Otherwise, output 0.
- To encrypt 1, choose a vector $\mathbf{c} \leftarrow \{1, 2, \ldots, r\}^n$ uniformly at random. Check if all entries of $\mathbf{c}$ correspond to buckets of width less than $1/(4a\sqrt{n}\sqrt{m}\gamma')$. If yes, output $\mathbf{c}$. Otherwise, output 1.
- To decrypt a ciphertext $\mathbf{c}$, take any $\mathbf{z}$ such that $\mathrm{round}_r(\mathbf{z}) = \mathbf{c}$, compute
$$\gamma'\sqrt{m}\mathbf{W}^T\mathbf{z} \mod 1$$
and check if all ℓ entries are in $(-1/2a, 1/2a)$. If yes, output 0, else output 1.

The real matrices and vectors $\mathbf{W}, \mathbf{A}, \mathbf{t}$ are represented with $O(\log n)$ bits of precision. The precision analysis is analogous to the one done in 4.5 for pancake encryption, so we omit it.

Theorem 8.3. *Set the parameters of the scheme to* $\gamma = \sqrt{n}$, $\beta = (16{\cdot}10^4 n^3 \log(n))^{-1}$, $\ell = \log n$, $m = 10^8 n^2 \log(n)^2$, $r = (40001 n^3 \log(n))^{5/3}$ *and* $a = n$. *Assuming* hCLWE(s, ε), *the scheme has negligible decryption error and security error at most* $1/4 + 4\varepsilon$.

We prove correctness and security of the scheme separately in the next two subsections.

We are also interested in using this scheme to prove that hCLWE and hCLWE(ℓ) are in SZK (statistical zero knowledge), what is shown in Section 9 for the following choice of parameters:

$$\begin{gathered} a = 100 \\ \beta'\gamma' \ln \gamma' < 1/(4 \cdot 10^4 K n \log n) \\ \gamma' > 1 \\ m = (Kn \log n \ln \gamma')^2 \\ r = m^{10} (\gamma')^{5/3} \end{gathered} \tag{3}$$

where $K = 4 \cdot 9 \cdot 10 \cdot e \cdot 2 \cdot 5$.

8.3 Correctness

The following two claims assert that the scheme is correct.

Claim. The probability that $\mathrm{Dec}(\mathbf{W}, \mathrm{Enc}(\mathbf{A}, 0)) = 0$ over the joint choice of the public key and encryption randomness is at least

$$1 - \ell \sqrt{\frac{2\beta'^2\gamma'^2 m}{\pi}} \frac{e^{-\frac{(1/4a)^2}{2\beta'^2\gamma'^2 m}}}{1/4a}.$$

In particular,

- for the choice of parameters made in Theorem 8.3, it is at least $1 - e^{-n}$, i.e., the error is a negligible function.
- for the choice of parameters suggested in Equation 3, the probability is at least $1 - e^{-5000}$.

Claim. If $n \geq 4$, the probability that $\mathrm{Dec}(\mathbf{w}, \mathrm{Enc}(\mathbf{A}, 1)) = 1$ is at least $1 - (3/2a)^\ell - \exp(-\gamma'^2 m)$. In particular,

- for the choice of parameters made in Theorem 8.3, the probability is at least $1 - (3/2n)^{\log n} - \exp(-n^3)$, i.e., the error is negligible.
- for the choice of parameters suggested in Equation 3, the probability is at least $1 - (3/200)^\ell - \exp(-n^2)$.

8.4 Security

In order to analyze the security of the scheme we have to take into account the possibility that at least one of the entries of the ciphertext corresponds to a bucket of width larger than $1/(4a\sqrt{n}\sqrt{m}\gamma')$ as the encryption algorithm outputs the plaintext in the clear in that case.

Claim. Let r be such that the following inequalities are satisfied

$$r^{-3/5} \leq \frac{1}{4a\sqrt{n}\sqrt{m}\gamma'} \tag{4}$$

$$\frac{2nr^{-2/5}}{\sqrt{\ln r^{2/5}}} + \frac{2n}{r} \leq \delta(n). \tag{5}$$

Let $\mathbf{A} \in \mathbb{R}^{n\times m}$ be a matrix whose columns consist of independent hCLWE(ℓ) samples and assume hCLWE(ℓ)(s,ε) where s is the complexity of rounding and ε is a function of n. Let $\mathbf{t} \leftarrow \{-1/\sqrt{m}, 1/\sqrt{m}\}^m$ be sampled uniformly at random. The probability that any entry of the vector $\mathbf{c} := \mathrm{round}_r(\mathbf{At})$ corresponds to a bucket of width larger than $1/(4a\sqrt{n}\sqrt{m}\gamma')$ is at most $\delta(n)+\varepsilon$. For the choice of parameters made in Theorem 8.3 and in Equation 3 both conditions are satisfied for $\delta(n) = \frac{1}{24}$.

The next claim follows directly from Proposition 4.5.

Claim. If the ciphertexts are not the messages, the distributions $(\mathbf{N}, \mathrm{Enc}(\mathbf{N}, 0))$ and $(\mathbf{N}, \mathrm{Enc}(\mathbf{N}, 1))$ are $\sqrt{4e \ln r^n/\sqrt{m}}$-statistically close for a matrix $\mathbf{N}$ of independent standard Gaussians. In particular,

- for the choice of parameters made in Theorem 8.3, the distance is at most $1/\sqrt{50} < 1/4$.
- for the choice of parameters suggested in Equation 3, the distance is at most $1/3$.

Corollary 8.4. *If* hCLWE(ℓ)(s,ε) *holds, then the distributions* $(\mathbf{A}, \mathrm{Enc}(\mathbf{A}, 0))$ *and* $(\mathbf{A}, \mathrm{Enc}(\mathbf{A}, 1))$ *are* $(s - \mathrm{poly}(n), \sqrt{4e \ln r^n/\sqrt{m}} + 4\varepsilon)$*-indistinguishable where* $\mathbf{A}$ *is the public key matrix. In particular,*

- *for the choice of parameters made in Theorem 8.3, and* $\varepsilon = 1/24$*, we get* $1/4 + 4/24 < 1/2$.
- *for the choice of parameters suggested in Equation 3 and* $\varepsilon = 1/24$*, we get* $1/3 + 4/24 = 1/2$.

9 hCLWE and hCLWE(ℓ) are in SZK

In this section we prove that hCLWE and hCLWE(ℓ) are in SZK, which is the class of decision problems that admit a statistical zero-knowledge proof [GMR89]. Zero-knowledge is defined with respect to honest verifiers.

We say that a sampling problem is in SZK if there is a polynomial-time honest-verifier statistical zero-knowledge protocol that accepts at least 2/3 of the YES instances and rejects at least 2/3 of the NO instances. The choice of threshold 2/3 is operational.

Our proof consists in a reduction from hCLWE to the statistical difference problem (SD). Sahai and Vadhan proved in [SV03] that SD is complete for SZK.

Definition 9.1 (SD Problem). *The YES instances of the Statistical Difference (SD) problem are pairs of circuits (C_0, C_1) such that $\Delta(C_0, C_1) > 2/3$ and the NO instances are pairs of circuits (C_0, C_1) such that $\Delta(C_0, C_1) < 1/3$.*

Here Δ is the statistical (total variation) distance between the output distributions sampled by the circuits when instantiated with a uniformly random seed. That is, if the output space of C_0 and C_1 is some finite set Ω,

$$\Delta(C_0, C_1) = \sup_{A \subseteq \Omega} |\Pr[C_0 \in A] - \Pr[C_1 \in A]| = \frac{1}{2} \sum_{\omega \in \Omega} |\Pr[C_0 = \omega] - \Pr[C_1 = \omega]|$$

Since SD is a complete problem for the SZK class and SZK is a class closed under reductions (see [SV03]), we can study the SZK class by considering reductions to SD instead of interactive proof systems. This approach also removes any reference to zero-knowledge.

In order to show that hCLWE is in SZK, it suffices to define two circuits that satisfy the conditions of Definition 9.1.

Theorem 9.2. *Let K, K' be sufficiently large constants. If $\gamma' > 1$, $\beta'\gamma' \ln \gamma' < 1/(K' n \log n)$ and γ' is polynomially bounded, the* hCLWE$_{\beta,\gamma,n}$ *problem with $m = (Kn \log n \ln \gamma')^2$ samples is in* SZK.

Proof. Take K and r as in Equation 3, that is, $K = 4 \cdot 9 \cdot 10 \cdot e \cdot 2 \cdot 5$ and $r = m^{10}(\gamma')^{5/3}$. Let $K' = 4 \cdot 10^4 K$. Let $\mathbf{X}$ be either a valid public key $\mathbf{A} \in \mathbb{R}^{n \times m}$ or a matrix $\mathbf{N} \in \mathbb{R}^{n \times m}$ with i.i.d. entries sampled from $\mathcal{N}(0, 1)$. We define two circuits C_0, C_1 that take as input the pair $(\mathbf{t}, \mathbf{u})$ where $\mathbf{t} \in \{-1/\sqrt{m}, 1/\sqrt{m}\}^m$ and $\mathbf{u} \in \{1, 2, \ldots, r\}^n$. C_0 outputs $\mathrm{round}_r(\mathbf{Xt})$, i.e., an encryption of 0 using randomness t, while C_1 outputs $\mathbf{u}$, i.e., an encryption of 1 with randomness $\mathbf{u}$.

If $\mathbf{X} = \mathbf{A}$, by Claim 8.3 and Claim 8.3 and Claim 8.4 for $\epsilon(n) = 1/24 = \delta(n)$, the decryption error is at most $e^{-5000} + 3/200 + \exp(-n^2) + 1/24 + 1/24$. It follows that $\Delta(C_0, C_1) > 2/3$.

If $\mathbf{X} = \mathbf{N}$, then the statistical distance between C_0 and C_1 is at most 1/3 by Proposition 4.5.

We also have an analogous statement for hCLWE(ℓ).

Theorem 9.3. *Let K, K' be sufficiently large constants. If $\gamma' > 1$, $\beta'\gamma' \ln \gamma' < 1/(K' n \log n)$, γ' is polynomially bounded and $1 \le \ell \le n$,* hCLWE$_{\beta,\gamma,n}(\ell)$ *with $m = (Kn \log n \ln \gamma')^2$ samples is in* SZK.

Acknowledgements

We are grateful to Devika Sharma and Luca Trevisan for their insight and advice and to an anonymous reviewer for helpful comments.

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101019547). The first author was additionally supported by RGC GRF CUHK14209920 and the fourth author was additionally supported by ISF grant No. 1399/17, project PROMETHEUS (Grant 780701), and Cariplo CRYPTONOMEX grant.

Bibliography

[ABW10] Benny Applebaum, Boaz Barak, and Avi Wigderson. Public-key cryptography from different assumptions. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, page 171–180, New York, NY, USA, 2010. Association for Computing Machinery.

[AD97] Miklós Ajtai and Cynthia Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97, page 284–293, New York, NY, USA, 1997. Association for Computing Machinery.

[Ale03] M. Alekhnovich. More on average case vs approximation complexity. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 298–307, 2003.

[BB20] Matthew S. Brennan and Guy Bresler. Reducibility and statistical-computational gaps from secret leakage. In Jacob D. Abernethy and Shivani Agarwal, editors, *Conference on Learning Theory, COLT 2020, 9-12 July 2020, Virtual Event [Graz, Austria]*, volume 125 of *Proceedings of Machine Learning Research*, pages 648–847. PMLR, 2020.

[BNHR22] Andrej Bogdanov, Miguel Cueto Noval, Charlotte Hoffmann, and Alon Rosen. Public-key encryption from continuous lwe. Cryptology ePrint Archive, Paper 2022/093, 2022. `https://eprint.iacr.org/2022/093`.

[BR13] Quentin Berthet and Philippe Rigollet. Complexity theoretic lower bounds for sparse principal component detection. In Shai Shalev-Shwartz and Ingo Steinwart, editors, *Proceedings of the 26th Annual Conference on Learning Theory*, volume 30 of *Proceedings of Machine Learning Research*, pages 1046–1066, Princeton, NJ, USA, 12–14 Jun 2013. PMLR.

[BRST21] Joan Bruna, Oded Regev, Min Jae Song, and Yi Tang. Continuous lwe. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, pages 694–707, New York, NY, USA, 2021. Association for Computing Machinery.

[BS15] Mikhail Belkin and Kaushik Sinha. Polynomial learning of distribution families. *SIAM Journal on Computing*, 44(4):889–911, 2015.

[DKS17] Ilias Diakonikolas, Daniel M. Kane, and Alistair Stewart. Statistical query lower bounds for robust estimation of high-dimensional gaussians and gaussian mixtures. *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 73–84, 2017.

[DNR04] Cynthia Dwork, Moni Naor, and Omer Reingold. Immunizing encryption schemes from decryption errors. In *EUROCRYPT*, 2004.

[EG85] Taher El Gamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In *Proceedings of CRYPTO 84 on Advances in Cryptology*, pages 10–18, Berlin, Heidelberg, 1985. Springer-Verlag.

[GG98] Oded Goldreich and Shafi Goldwasser. On the limits of non-approximability of lattice problems. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, page 1–9, New York, NY, USA, 1998. Association for Computing Machinery.

[GMR89] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM J. Comput.*, 18(1):186–208, 1989.

[GPV08] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In Cynthia Dwork, editor, *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 197–206. ACM, 2008.

[GSV99] Oded Goldreich, Amit Sahai, and Salil P. Vadhan. Can statistical zero knowledge be made non-interactive? or on the relationship of szk and niszk. In *Advances in Cryptology - CRYPTO '99, 19th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19, 1999, Proceedings*, volume 1666 of *Lecture Notes in Computer Science*, pages 467–484. Springer, 1999.

[GVV22] Aparna Gupte, Neekon Vafa, and Vinod Vaikuntanathan. Continuous LWE is as hard as LWE & applications to learning gaussian mixtures. Cryptology ePrint Archive, Report 2022/437, 2022. `https://ia.cr/2022/437`.

[HP15] Moritz Hardt and Eric Price. Tight bounds for learning a mixture of two gaussians. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 753–760, 2015.

[HR05] Thomas Holenstein and Renato Renner. One-way secret-key agreement and applications to circuit polarization and immunization of public-key encryption. In *Proceedings of the 25th Annual International Conference on Advances in Cryptology*, CRYPTO'05, page 478–493, Berlin, Heidelberg, 2005. Springer-Verlag.

[HWX15] Bruce Hajek, Yihong Wu, and Jiaming Xu. Computational lower bounds for community detection on random graphs. In *Proceedings of*

The 28th Conference on Learning Theory, volume 40 of *Proceedings of Machine Learning Research*, pages 899–928, Paris, France, 03–06 Jul 2015. PMLR.

[IZ89] Russell Impagliazzo and David Zuckerman. How to recycle random bits. pages 248–253. IEEE, 1989.

[McE78] R. J. McEliece. A Public-Key Cryptosystem Based On Algebraic Coding Theory. *Deep Space Network Progress Report*, 44:114–116, January 1978.

[MR07] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM J. Comput.*, 37:267–302, 2007.

[Rab79] Michael O. Rabin. *Digitalized signatures and public-key functions as intractable as factorization.* MIT Laboratory for Computer Science, 1979.

[Reg05] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC*, pages 84–93, 2005. Full version in [Reg09].

[Reg09] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6), 2009.

[RSA78] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, feb 1978.

[SV03] Amit Sahai and Salil Vadhan. A complete problem for statistical zero knowledge. *J. ACM*, 50(2):196–249, mar 2003.