Continuous LWE

Joan Bruna [a] Oded Regev [a] Min Jae Song [a] Yi Tang [b]

[a]New York University
[b]University of Michigan

Motivation: Gaussian Pancakes

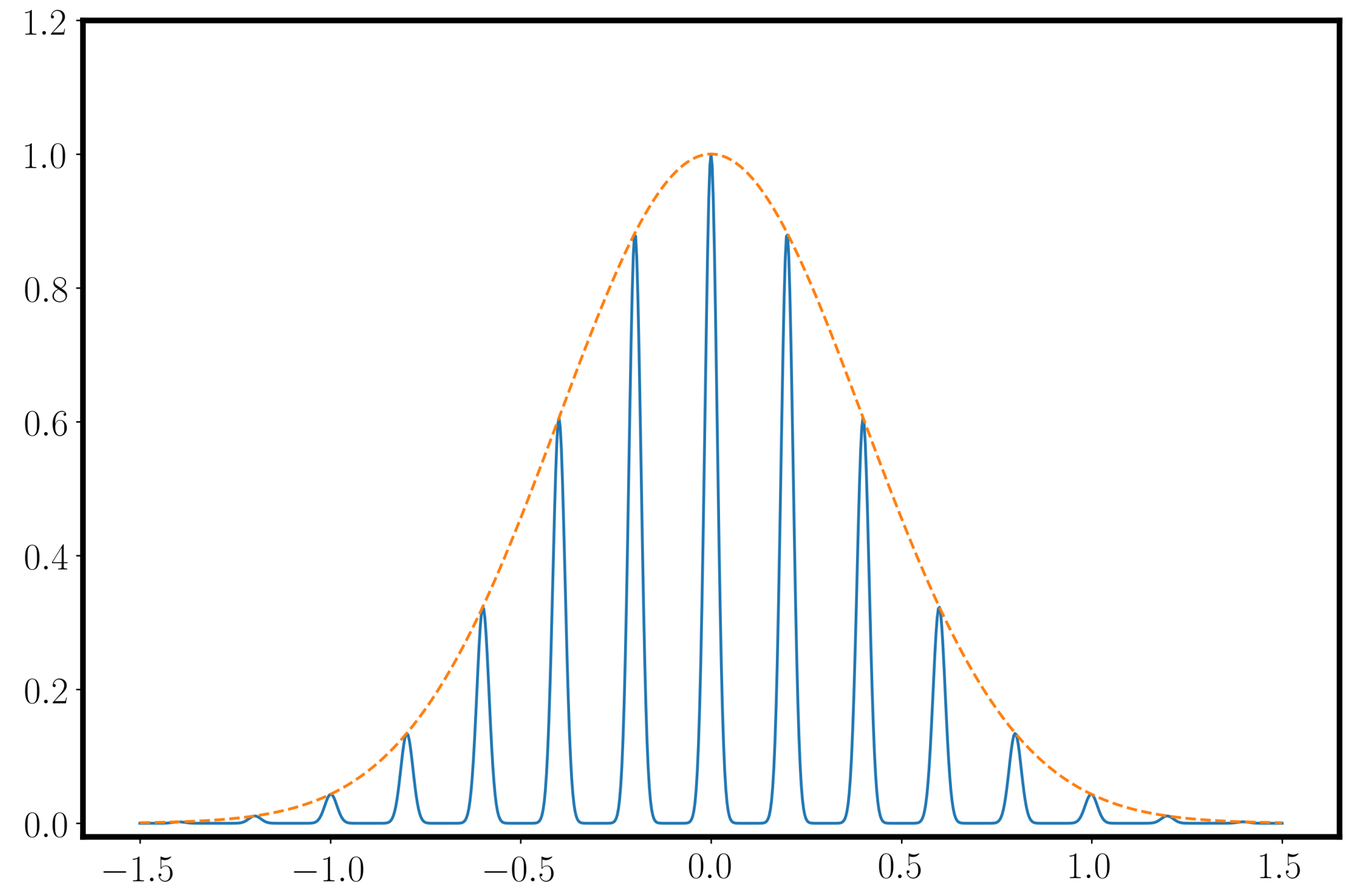

Figure 1: (Unnormalized) densities of a noisy discrete Gaussian (blue) and the standard Gaussian (orange).

The Gaussian pancakes distribution is a noisy discrete Gaussian (blue) in one hidden direction. In other $n-1$ directions, the distribution is standard Gaussian (orange).
We also call this distribution the **homogeneous Continuous Learning with Errors (hCLWE)** distribution, for reasons we explain later.
Our work is motivated by the following open question by [BLPR19]:

> *"Can poly-time algorithms distinguish the Gaussian pancakes distribution from the standard Gaussian in high-dimensions?"*

We answer this in the **negative**.

Previous Work: SQ-Hardness of Gaussian Pancakes

Distinguishing Gaussian pancakes from the standard Gaussian is **SQ-hard**.

- **Def.** A *statistical query (SQ) algorithm* accesses the input distribution only indirectly from noisy expectations. It can query the distribution with any bounded function $f : \mathbb{R}^n \to [-1, 1]$, and receive a noisy version of $\mathbb{E}[f(x)]$, instead of getting individual samples.
- **Thm [DKS17].** Statistical query (SQ) algorithms cannot distinguish Gaussian pancakes from the standard Gaussian using polynomially many queries (even with exponentially small noise).
- **Thm [BLPR19].** Still SQ-hard when you have **multiple** discrete directions (Gaussian "baguettes").

Notice that all previous hardness results apply only to SQ algorithms. Of course, SQ algorithms are powerful and capture many known methods, but the question of whether the hardness of this distinguishing task extends beyond SQ algorithms was open.

Implications of Hardness of Gaussian Pancakes

- **[DKS17]**: Improperly learning (= density estimation) **Gaussian mixtures** is **SQ-hard**, even for a mixture with nearly non-overlapping components.
- **[BLPR19]**: Learning **robust classifiers*** is **SQ-hard**, even when they exist, are learnable information-theoretically with polynomially many samples, and learning a non-robust classifier is easy.

* Robust in the sense that the classifier is not vulnerable to small input perturbations.

Our Result: Hardness of Gaussian Pancakes

Distinguishing Gaussian pancakes, with spacing $1/\gamma$ less than $1/(2\sqrt{n})$, from the standard Gaussian with accuracy **slightly (inverse-polynomially) better than chance** is computationally hard, unless there are polynomial-time **quantum** algorithms for fundamental worst-case lattice problems.

* Thickness β can be any inverse polynomial.

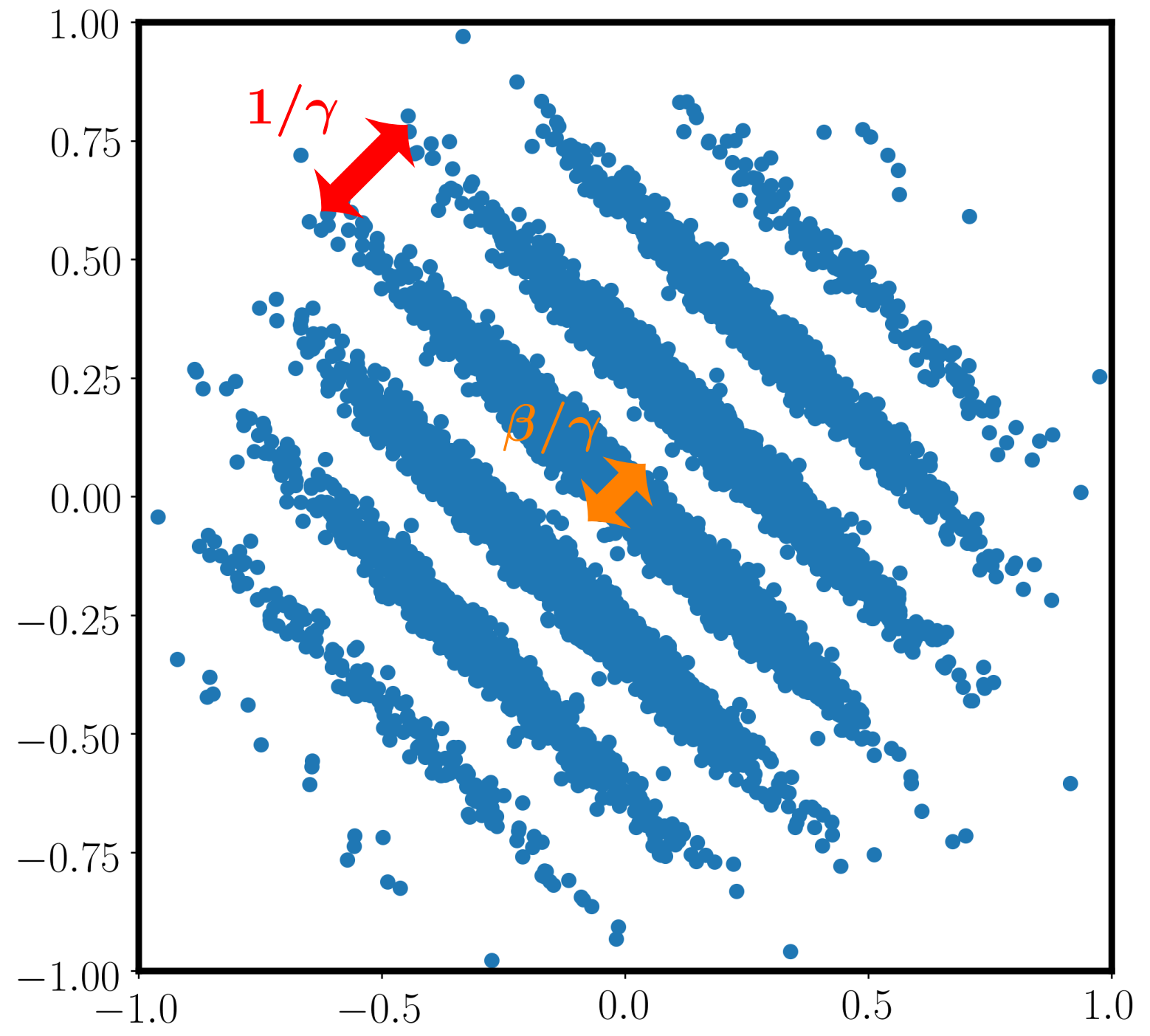

Figure 2: The hCLWE distribution is characterized by two parameters β and γ. The parameter β controls the pancake thickness and γ controls the spacing between pancakes.

This implies that assuming some worst-case lattice problems cannot be solved by polynomial-time quantum algorithms ...

- Distinguishing Gaussian pancakes/baguettes from the standard Gaussian is hard for **any** poly-time algorithm.
- Improperly learning **Gaussian mixtures** is hard for **any** poly-time algorithm, even when the mixture components are nearly non-overlapping.
- There is a binary classification task for which a robust classifier *exists*, is *learnable* info-theoretically with polynomially many samples, and a *non-robust* classifier is easy to learn, but learning a **robust classifier** is hard for **any** poly-time algorithm.

Our result is an **average-case hardness** result based on **worst-case hardness** assumptions. Only a few hardness of improper learning results are based on worst-case hardness, e.g., [KS06].

Hardness of (h)CLWE: Proof Overview

We prove a stronger hardness result, for a relaxed problem named (*inhomogeneous*) CLWE, defined below.

Def. $\mathrm{CLWE}_{\beta,\gamma}$: To decide whether the given samples of the form $(\boldsymbol{y}, z)$ with $\boldsymbol{y} \sim \mathcal{N}(0, I_n)$ have either

1. periodic "colors" z along some secret direction $w \in \mathbb{R}^n$, i.e., $z = (\gamma\langle \boldsymbol{y}, \boldsymbol{w}\rangle + e) \bmod 1$ where $e \sim \mathcal{N}(0, \beta)$, or
2. uniformly random "colors" $z \in [0, 1)$.

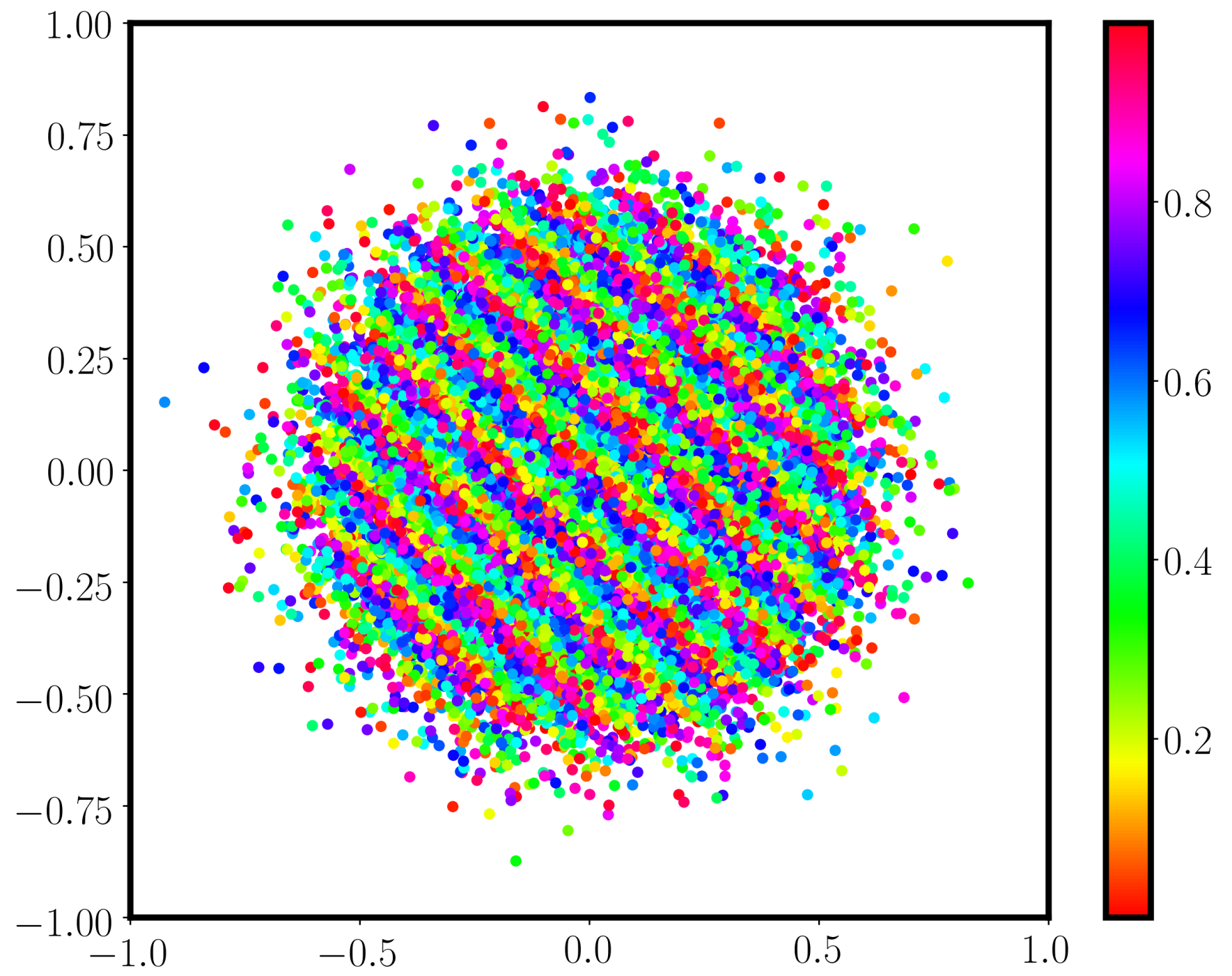

Figure 3: The CLWE distribution has periodic colors along some secret direction.

$\mathrm{hCLWE}_{\beta,\gamma}$ samples are roughly $\mathrm{CLWE}_{\beta,\gamma}$ samples with $z = 0$.

We show the hardness results by reducing worst-case lattice problems to CLWE, and reducing CLWE to hCLWE via rejection sampling by $z \approx 0$.

Hardness of CLWE. [Reg05] first gave a quantum reduction from a worst-case lattice problem called $\mathrm{GapSVP}_{O(n/\alpha)}$ (See below for a definition) to $\mathrm{LWE}_{q,\alpha}$ (for $\alpha q \geq 2\sqrt{n}$), showing the hardness of LWE.
We follow the more recent framework of [PRS17], and reduce $\mathrm{GapSVP}_{O(n/\beta)}$ to $\mathrm{CLWE}_{\beta,\gamma}$ for any polynomial $\gamma \geq 2\sqrt{n}$ and inverse-polynomial $\beta \in (0, 1)$.

Lattices and lattice problems. For a basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ of $\mathbb{R}^n$, the lattice $\mathcal{L}$ generated by the basis is the set of all *integer* linear combinations of the basis vectors. The minimum distance $\lambda_1(\mathcal{L})$ is the shortest length of nonzero lattice vectors in lattice $\mathcal{L}$.

Def. The Gap Shortest Vector Problem (GapSVP_φ): To decide whether $\lambda_1(\mathcal{L}) \leq 1$ or $\lambda_1(\mathcal{L}) > \varphi$ for a given lattice $\mathcal{L}$.

GapSVP_φ is believed to be computationally hard (even *quantumly*) for any polynomial $\varphi = \varphi(n)$.

Analogies Between LWE and CLWE

Def. The Learning With Errors problem ($\mathrm{LWE}_{q,\alpha}$): To decide whether the given samples of the form $(\boldsymbol{a}, b)$ with $\boldsymbol{a} \sim \mathbb{Z}_q^n$ have either

1. periodic b along some secret $\boldsymbol{s} \in \mathbb{Z}_q^n$, i.e., $b = (\langle \boldsymbol{a}, \boldsymbol{s}\rangle/q + e) \bmod 1$, where $e \sim \mathcal{N}(0, \alpha)$, or
2. uniformly random $b \in [0, 1)$.

* By discretizing with $b' = \lfloor q \cdot b \rceil \in \mathbb{Z}_q$, the search problem (to find secret $\boldsymbol{s}$) can be viewed as solving system of linear equations with errors over $\mathbb{Z}_q$, of the form $b' = \langle \boldsymbol{a}, \boldsymbol{s}\rangle + e'$.

$\mathrm{LWE}_{q,\alpha}$	$\mathrm{CLWE}_{\beta,\gamma}$
secret $\boldsymbol{s} \in \mathbb{Z}_q^n$	secret $\boldsymbol{w} \in \mathcal{S}^{n-1}$
samples $(\boldsymbol{a}, b)$	samples $(\boldsymbol{y}, z)$
$\boldsymbol{a} \sim \mathbb{Z}_q^n$	$\boldsymbol{y} \sim \mathcal{N}(0, \boldsymbol{I}_n)$
$b = (\langle \boldsymbol{a}, \boldsymbol{s}\rangle/q + e) \bmod 1$	$z = (\gamma\langle \boldsymbol{y}, \boldsymbol{w}\rangle + e) \bmod 1$
$e \sim \mathcal{N}(0, \alpha)$	$e \sim \mathcal{N}(0, \beta)$
$\alpha \cdot q$	inverse spacing γ

Other Results Related to CLWE

Noise is necessary for hardness.

- Noiseless CLWE can be efficiently solved with LLL (or even CLWE with exponentially small noise [SZB21]).
- Analogous to solving noiseless LWE with Gaussian elimination.
- Bypasses SQ-hardness since LLL inspects samples individually.

Subexponential algorithms for hCLWE with $\gamma = o(\sqrt{n})$.

- Simply compute covariance using $\exp(\gamma^2)$ many samples.
- Analogous to the Arora-Ge algorithm for LWE [AG11].

Gaussian pancakes with uniform-spacing is SQ-hard.

- Technically, the Gaussian pancakes of [DKS17] and [BLPR19] have non-uniform spacing between the pancakes, whereas our spacing is uniform. Our SQ-hardness result for uniform-spacing shows that we have not changed the distinguishing problem's difficulty too much by changing the spacing between the Gaussian pancakes.

Follow-up Work and Conclusion

Follow-up work by [SZB21] observes that CLWE hardness also implies hardness of learning high-dimensional cosines ("cosine neurons") of the form $f(x) = \cos(2\pi\gamma\langle w, x\rangle)$ over the Gaussian input distribution if small (inverse-polynomial) label noise is added. Hence, hardness of CLWE implies the hardness of this seemingly simple supervised improper learning task as well.

Together with our result on hardness of learning Gaussian mixtures, this shows the versatility of CLWE/hCLWE as a primitive for showing hardness of improper learning.