# Discrete vs. Continuous Assumptions

- LPN $\to$ LWE: noisy linear equations over $\mathbb{Z}_q$; worst-case lattice foundation.
- LWE $\to$ CLWE $\to$ hCLWE: a hidden direction and Gaussian pancakes.
- Signal detectable in principle $\not\Rightarrow$ detectable in polynomial time.

## Learning parity with noise (LPN)

LPN is a conjecturally hard learning problem over $\mathbb{F}_2$, using only fast bit operations. Low-noise LPN hardness can yield PKE.

Unknown $\mathbf{s} \in \mathbb{F}_2^n$, given independent $(\mathbf{a}_i, b_i)$, predict $\mathbf{a}^\top \mathbf{s}$ on a fresh $\mathbf{a}$:

$$\mathbf{a} \leftarrow \mathbb{F}_2^n, \qquad b = \mathbf{a}^\top \mathbf{s} \pmod 2.$$

Unlike real-valued regression, both inputs and labels are bits and arithmetic is modulo 2.

- Without noise, $m$ equations give $\mathbf{A}\mathbf{s} = \mathbf{b}$.
- For $m = n + \lambda$ uniform samples, failure probability is at most $2^{-\lambda}$.

If $\mathrm{rank}(\mathbf{A}) = n$, Gaussian elimination recovers $\mathbf{s}$.

$$\Pr[\mathrm{rank}(\mathbf{A}) < n] \le \sum_{\mathbf{v} \in \mathbb{F}_2^n \setminus \{0\}} \Pr[\mathbf{A}\mathbf{v} = 0] = (2^n - 1)2^{-m} < 2^{-\lambda}, \quad m = n + \lambda.$$

LPN changes only the labels, introducing independent random bit flips.

**Definition 1** (Learning parity with noise (LPN))**.** *Fix a noise rate $\eta \in (0, 1/2)$ and choose a secret $\mathbf{s} \leftarrow \mathbb{F}_2^n$. An LPN sample $(\mathbf{a}, b)$ is generated as*

$$\mathbf{a} \leftarrow \mathbb{F}_2^n, \qquad b = \mathbf{a}^\top \mathbf{s} + e \pmod 2,$$

*where $e$ is sampled from $\mathrm{Ber}(\eta)$: $\Pr[e = 1] = \eta$ and $\Pr[e = 0] = 1 - \eta$.*

- Search LPN: recover $\mathbf{s}$; decisional LPN: distinguish from uniform pairs.
- Each sample is a noisy linear equation; $m$ samples give $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$.

Write $H_2(\eta) = -\eta \log_2 \eta - (1 - \eta) \log_2(1 - \eta)$.

- With sufficiently many samples the secret is still determined in principle.
- Each noisy label carries roughly $1 - H_2(\eta)$ bits about $\mathbf{s}$.
- Information threshold: $m > \dfrac{n}{1 - H_2(\eta)}$. For constant $\eta < 1/2$, need $\Theta(n)$ samples.

As a sanity check consider the following attacks:

**Attack 1:** *Guess corrupted equations.*

- About $\eta m$ errors yield $\binom{m}{\eta m} \approx 2^{mH_2(\eta)}$ patterns.
- Exponential in $n$ for $m = \Theta(n)$ and constant $\eta$.

**Attack 2:** *Choose $n$ independent equations, solve, and test* $\mathbf{s}$ *against remaining samples.*

- The chosen equations are all clean with probability $(1-\eta)^n \approx e^{-\eta n}$ for small $\eta$.
- Gives poly$(n)$ many trials if $\eta = O(\log n/n)$, but superpoly many if $\eta n = \omega(\log n)$.

**Attack 3:** *Gaussian elimination.*

- For an invertible $n \times n$ subsystem $\mathbf{A}_0$, $\widehat{\mathbf{s}} = \mathbf{A}_0^{-1}\mathbf{b}_0 = \mathbf{s} + \mathbf{A}_0^{-1}\mathbf{e}_0 \pmod 2$.
- A typical row of $\mathbf{A}_0^{-1}$ is dense; each bit of $\widehat{\mathbf{s}}$ combines $t = \Theta(n)$ noisy labels.
- For independent $e_i \leftarrow \mathrm{Ber}(\eta)$,

$$\Pr\left[\sum_{i=1}^{t} e_i = 1 \pmod 2\right] = \frac{1-(1-2\eta)^t}{2}.$$

- For constant $\eta$, the parity of these noise bits differs from uniform by $\frac{1}{2}(1-2\eta)^t = 2^{-\Omega(n)}$.

More sophisticated algorithms cancel variables in blocks rather than guessing errors:

- At constant noise: $2^{O(n/\log n)}$ time and samples.
- With polynomially many samples: $2^{O(n/\log\log n)}$ time.
- No polynomial-time solver is known for LPN at constant noise or even $\eta = n^{-1/2}$.

Cryptographic implications:

- Hardness of LPN with $m = 2n$ and $\eta = n^{-1/2}$ yields public-key encryption (PKE).
- For constant noise, PKE from polynomial-time LPN hardness remains open.
- General one-way functions are not known to imply public-key encryption.

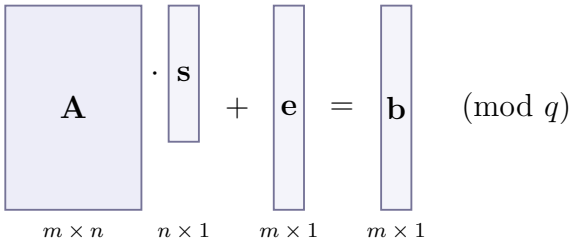
# Learning with errors (LWE)

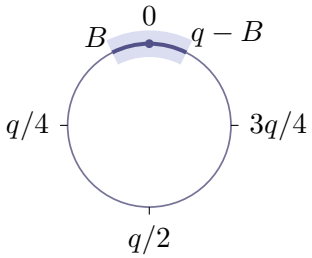
LPN hides a linear function with noisy bits. LWE uses small-magnitude noise and modular arithmetic over $\mathbb{Z}_q$.

**Definition 2** (Learning with errors (LWE))**.** *Fix an integer $q$ and secret $\mathbf{s} \leftarrow \mathbb{Z}_q^n$. Independent samples:*

$$\mathbf{a} \leftarrow \mathbb{Z}_q^n, \qquad b = \mathbf{a}^\top \mathbf{s} + e \pmod q, \qquad e \sim \textit{discrete Gaussian of width } \alpha q.$$

For $m$ samples, stack $\mathbf{a}_i^\top$ as the rows of $\mathbf{A} \in \mathbb{Z}_q^{m\times n}$; let $\mathbf{b} = (b_i)$ and $\mathbf{e} = (e_i)$. Then:






Shaded: $e \in [-B, B] \pmod q$; wraps at 0.

For $B = C\alpha q\sqrt{\log n} < q/2$: $\Pr[|e| \le B] \ge 1 - n^{-\Omega(1)}$.

- Search: recover $\mathbf{s}$; decision: distinguish $(\mathbf{a}, b)$ from uniform pairs in $\mathbb{Z}_q^n \times \mathbb{Z}_q$.
- No noise: $\mathbf{b} = \mathbf{A}\mathbf{s} \pmod q$ (linear algebra, e.g. prime $q$).
- No modular reduction: $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$ (linear regression).
- With modular reduction and noise: for $q = 2^k$, reduction retains only the lowest $k$ bits of the integer inner product, while small error blurs the lowest of those bits. Roughly, the label reveals a noisy band of middle bits. "Small" in $\mathbb{Z}_q$ means close to 0 modulo $q$.

## Why is LWE believed to be hard?

$L(\mathbf{B}) = \{\mathbf{B}\mathbf{z} : \mathbf{z} \in \mathbb{Z}^n\}$; $\lambda_1(L)$ = length of the shortest nonzero lattice vector. In $\gamma$-GapSVP, given $(\mathbf{B}, d)$, distinguish

$$\lambda_1(L) \le d \qquad \text{from} \qquad \lambda_1(L) > \gamma d \quad \text{(promise)}.$$

**Theorem 1** (Worst-case hardness of LWE)**.** *A polynomial-time algorithm for* search LWE *gives a quantum polynomial-time algorithm for worst-case* $\widetilde{O}(n)$-GapSVP*.*

- Approx factor $\widetilde{O}(n/\alpha)$, provided $\alpha q > 2\sqrt{n}$; classical reductions other parameters.
- wc to ac: even a heuristic for random LWE samples undermines this evidence.
- $\gamma$-GapSVP is NP-hard for constant $\gamma$; the relevant $\widetilde{O}(n/\alpha)$ scale is much coarser.
- Related GapCVP: a 2026 Lean formalization claims hardness to $n^{1/2-\varepsilon}$.
- Hardness at $C\sqrt{n}$ (large $C$) would imply NP = coNP.

## Search, decision, and generation

For LWE, an efficient distinguisher can be converted into an efficient secret-recovery algorithm. *Generating* additional labeled samples is a weaker task than recovering the secret.

**Proposition 1.** *Assuming LWE is hard, a hidden representation can be hard to learn even when fresh samples using that representation can be efficiently generated.*

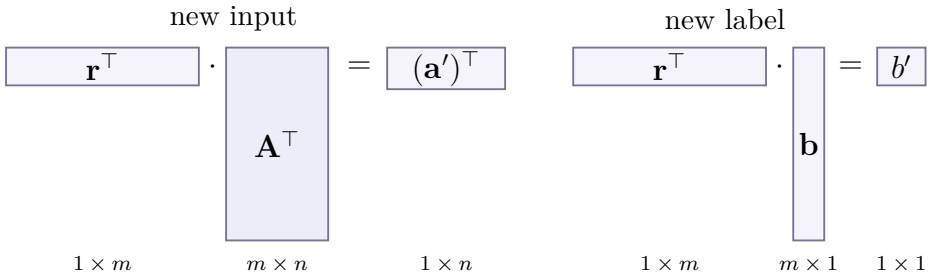
Start with samples $(\mathbf{a}_i, b_i)$, $1 \le i \le m$. For this construction, stack the inputs as *columns*: $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_m) \in \mathbb{Z}_q^{n \times m}$. Draw $\mathbf{r} \leftarrow \{0,1\}^m$ and form

$$\mathbf{a}' = \mathbf{A}\mathbf{r}, \quad b' = \mathbf{b}^\top \mathbf{r} \pmod q, \qquad \mathbf{b} = (b_1, \ldots, b_m)^\top.$$

Since $\mathbf{b} = \mathbf{A}^\top \mathbf{s} + \mathbf{e}$,

$$b' = \mathbf{r}^\top \mathbf{A}^\top \mathbf{s} + \mathbf{r}^\top \mathbf{e} = \mathbf{a}'^\top \mathbf{s} + e' \pmod q, \quad e' = \mathbf{r}^\top \mathbf{e}.$$

Typical $\mathbf{r}$ adds $\sim m/2$ errors: noise width grows by a factor about $\sqrt{m/2}$.




**Lemma 1** (Leftover hash lemma, specialized)**.** *For independent uniform columns of* $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$*,* $\mathbf{r} \leftarrow \{0,1\}^m$*, and independent uniform* $\mathbf{u} \in \mathbb{Z}_q^n$*, if* $m \geq n \log_2 q + 2 \log_2(1/\varepsilon)$*, then* $\Delta((\mathbf{A}, \mathbf{Ar}), (\mathbf{A}, \mathbf{u})) \leq \varepsilon$.

$$\Delta(X, Y) = \frac{1}{2} \sum_x \Big| \Pr[X = x] - \Pr[Y = x] \Big|.$$

*Proof.* For fixed $\mathbf{A}$, put $p_{\mathbf{A}}(\mathbf{x}) = \Pr_{\mathbf{r}}[\mathbf{Ar} = \mathbf{x}]$. For independent uniform $\mathbf{r}, \mathbf{r}'$, a collision is certain when $\mathbf{r} = \mathbf{r}'$ (probability $2^{-m}$). Otherwise, some entry of $\mathbf{r} - \mathbf{r}'$ is $\pm 1$; conditioned on the other columns, $\mathbf{A}(\mathbf{r} - \mathbf{r}')$ is uniform in $\mathbb{Z}_q^n$ (even for composite $q$). Hence

$$\mathbb{E}_{\mathbf{A}} \sum_{\mathbf{x}} p_{\mathbf{A}}(\mathbf{x})^2 = \Pr_{\mathbf{A},\mathbf{r},\mathbf{r}'}[\mathbf{Ar} = \mathbf{Ar}'] = 2^{-m} + (1 - 2^{-m}) q^{-n}.$$

Both experiments reveal the same uniform $\mathbf{A}$. For any fixed $\mathbf{B}$, $\Pr[\mathbf{A} = \mathbf{B}] = q^{-nm}$; conditioned on $\mathbf{A} = \mathbf{B}$, the second output has distribution $p_{\mathbf{B}}$ or uniform $q^{-n}$, respectively. Thus the distance between the joint outputs averages the conditional distances over $\mathbf{A}$:

$$\begin{aligned}
\Delta((\mathbf{A}, \mathbf{Ar}), (\mathbf{A}, \mathbf{u})) &= \tfrac{1}{2} \sum_{\mathbf{B},\mathbf{x}} \Big| \Pr[\mathbf{A} = \mathbf{B}] p_{\mathbf{B}}(\mathbf{x}) - \Pr[\mathbf{A} = \mathbf{B}] q^{-n} \Big| \\
&= \tfrac{1}{2} \mathbb{E}_{\mathbf{A}} \sum_{\mathbf{x}} |p_{\mathbf{A}}(\mathbf{x}) - q^{-n}| \\
&\leq \tfrac{1}{2} \sqrt{q^n \Big( \mathbb{E}_{\mathbf{A}} \sum_{\mathbf{x}} p_{\mathbf{A}}(\mathbf{x})^2 - q^{-n} \Big)} \\
&\leq \tfrac{1}{2} \sqrt{q^n / 2^m} \leq \varepsilon.
\end{aligned}$$

First inequality: Cauchy–Schwarz in $\mathbf{x}$ ($|\langle u, v \rangle| \leq \|u\|_2 \|v\|_2$; $u_{\mathbf{x}} = 1$, $v_{\mathbf{x}} = |p_{\mathbf{A}}(\mathbf{x}) - q^{-n}|$); Jensen in $\mathbf{A}$ ($\mathbb{E}\sqrt{X} \leq \sqrt{\mathbb{E}X}$). Second: by the collision bound $\mathbb{E}_{\mathbf{A}} \sum_{\mathbf{x}} p_{\mathbf{A}}(\mathbf{x})^2 - q^{-n} \leq 2^{-m}$; third: $q^n / 2^m \leq \varepsilon^2$ by bound on $m$. □

- Uniformity holds *even when the original vectors* $\mathbf{A}$ *are revealed.*
- PKE from LWE is closely related.

# Continuous learning with errors

Replace residues by fixed-precision reals: division by $q$ moves retained residue bits to right of binary point. For $q = 2^k$, $x \bmod q = \sum_{i=0}^{k-1} b_i 2^i$ and $(x \bmod q)/q = \sum_{i=0}^{k-1} b_i 2^{i-k}$.

**Definition 3** (Continuous LWE (CLWE))**.** *Fix $\gamma, \beta > 0$ and a hidden unit vector $\mathbf{w} \leftarrow S^{n-1}$. Independent samples:*

$$\mathbf{a} \leftarrow N(0, I_n), \qquad b = \gamma \mathbf{a}^\top \mathbf{w} + e \pmod 1, \qquad e \leftarrow N(0, \beta^2).$$

- Search: recover $\mathbf{w}$; decision: distinguish from $(\mathbf{a}, u)$ with independent $u \leftarrow [0, 1)$ uniform.
- Without mod 1: ordinary regression.
- Along $\mathbf{w}$: label bands repeat with period $1/\gamma$, blurred by noise $e$ of width $\beta$.

**Theorem 2** (Hardness of CLWE)**.** *A polynomial-time algorithm for* decisional CLWE *gives a polynomial-time algorithm for* decisional LWE.

- Conjectured hard for $\gamma \gtrsim \sqrt{n}$, $\beta \in (0, 1)$, and $\gamma/\beta \leq \text{poly}(n)$.
- Near-zero noise ($\beta \approx 0$) permits polynomial-time lattice-reduction attacks.

## Homogeneous CLWE

**Recall HW1.** For $\mathbf{a} \sim N(0, I_n)$ and unit $\mathbf{w}$,

$$\mathbf{a} = t\mathbf{w} + \mathbf{z}, \quad t = \mathbf{a}^\top \mathbf{w}, \quad \mathbf{z} \sim N(0, I_n - \mathbf{w}\mathbf{w}^\top), \quad t \perp \mathbf{z}.$$

- HW1: keep $|t| \leq b$ (one slab); for the sketch set $b = \beta/\gamma$.
- CLWE: $b = (\gamma t + e) \bmod 1$. Keep $b \approx 0$, erase $b$: $\gamma t + e \approx k$, $k \in \mathbb{Z}$ (spacing $1/\gamma$, width $\sim \beta/\gamma$).
- In both cases selection depends on $t$ (and independent $e$), not $\mathbf{z}$. Hence $\mathbf{z}$ is *unchanged*.




$|t| \leq b$ spacing $1/\gamma$, width $\sim \beta/\gamma$

*Sketch:* $z_1$ is one coordinate in $\mathbf{w}^\perp$; dashed circles show original Gaussian contours; gray bands indicate horizontal width (not a hard cutoff for hCLWE).

**Definition 4** (Homogeneous CLWE (hCLWE))**.** *Given $m$ unlabeled vectors $\mathbf{a}_i \in \mathbb{R}^n$, distinguish:*

- Null: *independent* $\mathbf{a}_i \leftarrow N(0, I_n)$.
- Planted: *one hidden unit* $\mathbf{w}$*; Gaussian in every direction orthogonal to* $\mathbf{w}$*, while* $a_i^{\parallel} = \mathbf{a}_i^\top \mathbf{w}$ *lies near* $k/\gamma$*,* $k \in \mathbb{Z}$*, with peak width about* $\beta/\gamma$*.*

- Along known $\mathbf{w}$: periodic spikes are easy to detect.
- With unknown $\mathbf{w}$: the task is to find their direction.

**Proposition 2** (CLWE $\to$ hCLWE)**.** *An efficient hCLWE distinguisher yields an efficient decisional CLWE distinguisher (with suitable parameters).*

*Proof sketch.* Given $(\mathbf{a}, b)$, retain samples with $b \approx 0 \pmod 1$; then erase $b$.

- *Null:* $b$ is independent uniform, so retained $\mathbf{a} \sim N(0, I_n)$.
- *CLWE:* $\gamma t + e \approx k$ for some $k \in \mathbb{Z}$: periodic Gaussian slabs along $\mathbf{w}$. By HW1, $\mathbf{z}$ stays independent and Gaussian.

A suitable Gaussian filter gives hCLWE with adjusted noise and retains an inverse-polynomial fraction of samples. □

A Gaussian mixture draws from one of several Gaussian components, chosen according to fixed mixing weights. Idealizing $b \approx 0$ as conditioning on $b = 0$, the planted law is

$$P_{\mathbf{w}} = \sum_{k \in \mathbb{Z}} \pi_k \, N(\mu_k \mathbf{w}, \Sigma), \quad \mu_k = \frac{\gamma k}{\gamma^2 + \beta^2},$$
$$\Sigma = I_n - \frac{\gamma^2}{\gamma^2 + \beta^2} \mathbf{w}\mathbf{w}^\top, \quad \pi_k \propto \exp\left( -\frac{k^2}{2(\gamma^2 + \beta^2)} \right).$$

**Proposition 3.** *Assuming hCLWE is hard, there are Gaussian mixtures that are hard to distinguish from a standard Gaussian in polynomial time.*

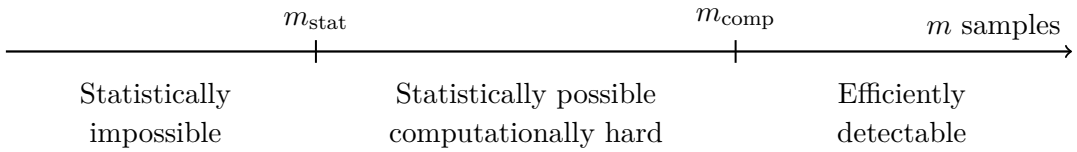
Write $\mathbf{a} = a^{\parallel}\mathbf{w} + \mathbf{a}^{\perp}$. For fixed integer $k$, the band $\gamma a^{\parallel} + e \approx k$ makes $a^{\parallel}$ approximately Gaussian, centered near $k/\gamma$ with width $\beta/\gamma$; $\mathbf{a}^{\perp}$ remains an independent standard Gaussian.

Thus each $k$ gives an approximately Gaussian component centered near $(k/\gamma)\mathbf{w}$, narrow along $\mathbf{w}$ but standard in orthogonal directions. Mixing with the probabilities of each $k$ gives the planted distribution; the null is one standard Gaussian.

## Statistical to computational gaps

Consider the hypothesis test $H_0$ = independent null samples; $H_1$ = samples sharing one hidden secret / direction. *Detection* distinguishes the cases; *recovery* reconstructs the signal.

- $m_{\text{stat}}$: sample scale where detection becomes information-theoretically possible.
- $m_{\text{comp}}$: sample scale where an efficient algorithm becomes effective.
- If $m_{\text{stat}} \ll m_{\text{comp}}$, there is a gap.
- LPN: $m_{\text{stat}} = \Theta(n)$ at constant noise.
- LWE, CLWE, hCLWE: roughly linear in $n$ up to logarithmic factors.




**Proposition 4.** *Assuming LWE (similarly LPN) is hard, some hypothesis-testing problems have no polynomial sample threshold* $m_{\text{comp}}$.

Generate polynomially many additional noisy samples from an initial collection using subset sums. Any efficient test that works after a polynomial number of samples (allowing for the increased noise) would then distinguish the original LWE samples.

For a nonperiodic planted mixture, match the first $k-1$ Gaussian moments along the hidden direction $\mathbf{w}$.

- Moments of degree $< k$ do not reveal the direction; the $k$th empirical tensor, with $n^k$ entries, can distinguish.
- Time $n^{O(k)}$; statistical-query lower bounds suggest this may be optimal in relevant models. Conjectured $m_{\text{comp}} = n^{\Theta(k)}$ vs. $m_{\text{stat}} = O(n)$.
- Other gaps: planted clique, sparse PCA, tensor PCA, planted dense subgraph; pseudorandom generators offer cryptographic examples.