By Fourier Inversion, $\sum_{x \in \mathbb{Z}} f(x) = \phi(0) = \sum_{y \in \mathbb{Z}} \hat{\phi}(y) = \sum_{y \in \mathbb{Z}} \hat{f}(y)$

Applied to $\rho_s(x)$, we have $\sum_{x \in \mathbb{Z}} \rho_s(x) = \sum_{y \in \mathbb{Z}} \hat{\rho}_s(y) = \sum_{y \in \mathbb{Z}} s \cdot \rho_{1/s}(y) = s \sum_{y \in \mathbb{Z}} \rho_{1/s}(y)$

Preimage sampling with a gadget trapdoor. Suppose we know R where $AR = G$ (and A is statistically close to uniform).

Starting point: Sampling from $D_{\mathbb{Z},s,c}$. Use rejection sampling.

1) Sample $x \leftarrow \mathbb{Z} \cap [c - s \cdot t(n),\ c + s \cdot t(n)]$. We will set $t(n) = \omega(\sqrt{\log n})$.

2) Output x w.p. $\rho_{s,c}(x)$. Otherwise reject.

By Gaussian tail bounds, when $s > \omega(\sqrt{\log \lambda})$

$$\Pr[x \leftarrow D_{\mathbb{Z},s,c} : |x - c| \geq t \cdot s] \leq 2e^{-\pi t^2} \quad (1 - \mathrm{negl}(\lambda))$$

Setting $t = \omega(\sqrt{\log \lambda})$, w.p. $1 - \mathrm{negl}(\lambda)$, x will lie in the interval $[c - s \cdot t(n),\ c + s \cdot t(n)]$.

} Truncated discrete Gaussian is statistically close to discrete Gaussian

By construction, rejection sampling algorithm outputs each x w.p. proportional to $\rho_{s,c}(x)$. Thus, this algorithm samples from truncated version of $D_{\mathbb{Z},s,c}$, which is statistically close to the desired distribution.

Algorithm will terminate with overwhelming probability after $t(n) \cdot \omega(\sqrt{\log n})$ iterations:

- $x \in [c - s,\ c + s]$ w.p. $\frac{2s+1}{2st+1} > \frac{2s}{2s(t+1)} = \frac{1}{t+1}$ since $s > 1$
- For $x \in [c - s,\ c + s]$, algorithm outputs it w.p. at least $\rho_{s,c}(c+s) = e^{-\pi} = O(1)$

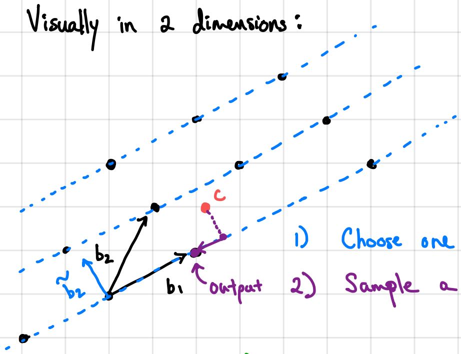
} By Chernoff bound, algorithm terminates after $t(n) \cdot \omega(\log \lambda)$ iterations w.p. $1 - \mathrm{negl}(\lambda)$

[Tricky problem: constant-time algorithm for discrete Gaussian sampling]

To sample from $D_{\mathcal{L},s,c}$ for an arbitrary $\mathcal{L} = \mathcal{L}(B)$, we proceed as follows:

1. Let $v_n \leftarrow 0$, $c_n \leftarrow c$. For $i = n, n-1, \ldots, 1$: (indexing will be convenient for analysis)
   (a) Compute $c_i' \leftarrow c_i^T \tilde{b}_i / \tilde{b}_i^T \tilde{b}_i \in \mathbb{R}$ and $s_i' \leftarrow s / \|\tilde{b}_i\|$
   (b) Sample $z_i \leftarrow D_{\mathbb{Z}, s_i', c_i'}$
   (c) Update $c_{i-1} \leftarrow c_i - z_i b_i$ and $v_{i-1} \leftarrow v_i + z_i b_i$
2. Output v_0:

Visually in 2 dimensions:



1) Choose one of these planes (direction along $\tilde{b}_2$) according to discrete Gaussian

2) Sample a discrete Gaussian along $\tilde{b}_1$

<u>Proof idea</u>: Smoothing ensures that distribution over the choice of plane for each dimension only depends on <u>distance</u> from center to the plane. Does <u>not</u> get affected by translation within the plane.
(See below for formal argument from [GPV08])

By construction, $v_0 \in \mathcal{L}(B)$ since $z_1, \ldots, z_n \in \mathbb{Z}$

We show that above algorithm outputs samples from $D_{\mathcal{L},s,c}$:

<u>Lemma.</u> Let $v = \sum_{i \in [n]} z_i b_i$ be output of above algorithm. Then, we can write

$$v - c = \sum_{i \in [n]} (z_i - c_i') \tilde{b}_i$$

Proof. Let $\pi_i : \mathbb{R}^n \to \mathrm{span}(b_1, \dots, b_i)$ be the projection from $\mathbb{R}^n$ onto the subspace spanned by $b_1, \dots, b_i$.

We show that $\forall j \in [n]$:

$$v - v_j - \pi_j(c_j) = \sum_{i \in [j]} (z_i - c_i')\tilde{b}_i$$

For $j = 0$, claim is immediate ($v = v_0$, $\pi_0(c_0) = 0$).

Suppose claim holds for $j = k-1$. Then,

components along $b_1, \dots, b_{k-1}$

component along $\tilde{b}_k$

$$v_0 - v_k - \pi_k(c_k) = v_0 - v_{k-1} + z_k b_k - (\pi_{k-1}(c_k) + c_k'\tilde{b}_k) \qquad [v_k = v_{k-1} - z_k b_k]$$

$$= (v_0 - v_{k-1}) + z_k b_k - \pi_{k-1}(c_{k-1}) - \pi_{k-1}(z_k b_k) - c_k'\tilde{b}_k \qquad [c_k = c_{k-1} + z_k b_k]$$

$$= (v_0 - v_{k-1} - \pi_{k-1}(c_{k-1})) + z_k \underbrace{(b_k - \pi_{k-1}(b_k))}_{\tilde{b}_k} - c_k'\tilde{b}_k$$

$$= \sum_{i \in [k-1]} (z_i - c_i')\tilde{b}_i + (z_k - c_k')\tilde{b}_k \qquad \text{[inductive hypothesis]}$$

Claim now holds by considering $j = n$: $v_n = 0$, $c_n = c$ ($\pi_n(c) = c$ since B is a basis for $\mathbb{R}^n$).

Lemma. Suppose $s > \|\tilde{B}\| \cdot \omega(\sqrt{\log n})$. For any $v \in \mathcal{L}(B) = \sum_{i \in [n]} z_i b_i$, algorithm outputs v with probability

$$\rho_{s,c}(v) \prod_{i \in [n]} \frac{1}{\rho_{s_i', c_i'}(\mathbb{Z})}$$

Proof. We compute the probability that algorithm samples $z_1, \dots, z_n$:

$$\Pr[\text{algorithm outputs } v] = \prod_{i \in [n]} D_{\mathbb{Z}, s_i', c_i'}(z_i) = \frac{\prod_{i \in [n]} \rho_{s_i', c_i'}(z_i)}{\prod_{i \in [n]} \rho_{s_i', c_i'}(\mathbb{Z})}$$

$$\rho_s\Big(\sum_{i \in [n]} (z_i - c_i')\tilde{b}_i\Big) = \exp\Big(-\pi \Big\|\sum_{i \in [n]} (z_i - c_i')\tilde{b}_i\Big\|_2^2 / s^2\Big)$$

$$\Big\|\sum_{i \in [n]} (z_i - c_i')\tilde{b}_i\Big\|_2^2 = \sum_{i,j \in [n]} (z_i - c_i')(z_j - c_j')\tilde{b}_i^T b_j = \sum_{i \in [n]} (z_i - c_i)^2 \|\tilde{b}_i\|^2$$

by orthogonality

$$\Rightarrow \rho_s\Big(\sum_{i \in [n]} (z_i - c_i')\tilde{b}_i\Big) = \prod_{i \in [n]} \rho_s\big((z_i - c_i')\|\tilde{b}_i\|\big)$$

Now, observe that

$$\prod_{i \in [n]} \rho_{s_i', c_i'}(z_i) = \prod_{i \in [n]} \rho_s\big((z_i - c_i') \cdot \|\tilde{b}_i\|\big) = \rho_s\Big(\sum_{i \in [n]} (z_i - c_i')\tilde{b}_i\Big) = \rho_s(v - c) = \rho_{s,c}(v)$$

since $\rho_{s/k, c}(x) = \exp((x-c)^2 k^2 / s^2) = \rho_s((x-c)k)$

$\tilde{b}_i$ are pairwise orthogonal and ρ_s is rotationally invariant

by previous lemma

Theorem (Gentry-Peikert-Vaikuntanathan). There is an efficient algorithm that takes a basis B of a lattice $\mathcal{L} = \mathcal{L}(B)$, a coset $c + \mathcal{L}$ and a Gaussian width parameter $s \geq \|\tilde{B}\| \cdot \omega(\sqrt{\log n})$ and outputs a sample whose distribution is statistically close to $D_{\mathcal{L}, s, c}$

Proof. Follows by combining above algorithm with sampling algorithm for integers.

The desired distribution can be written as

$$D_{\mathcal{L}, s, c}(v) = \rho_{s,c}(v) \cdot Q^{-1}$$

for some normalization constant $Q \in \mathbb{R}$. By the previous lemma, the algorithm outputs $v \in \mathcal{L}(B)$ w.p.

$$\rho_{s,c}(v) \cdot \frac{1}{\prod_{i \in [n]} \rho_{s_i', c_i'}(\mathbb{Z})}$$

problem: c_i' could be correlated with v so this is not a fixed normalization constant

Now, $\eta(\mathbb{Z}) \le \lambda_n(\mathbb{Z}) \cdot \omega(\sqrt{\log n}) = \omega(\log n)$. When $s \ge \|\tilde{B}\| \cdot \omega(\sqrt{\log n})$, then $s_i' = s / \|\tilde{b}_i\| \ge \omega(\sqrt{\log n}) = \eta$

Thus, $\rho_{s_i', c_i'}(\mathbb{Z}) \in [1-\text{negl.}, 1+\text{negl.}] \cdot \rho_s(\mathbb{Z})$, which is a quantity that is <u>independent</u> of v and c. Thus, the algorithm outputs v with probability proportional to $\rho_{s,c}(v)$, as required.

<u>Implication</u>: To sample from $D_{\mathcal{L},s,c}$, need a basis B for $\mathcal{L} = \mathcal{L}(B)$ where $\|\tilde{B}\| \le s / \omega(\sqrt{\log n})$. Need a <u>short</u> basis to sample preimages.

<u>Next</u>: Sampling discrete Gaussians with a gadget trapdoor.

Suppose $AR = G$ where R is short. Sampling pre-image for A is easy: to solve $Ax = y$, set $x = R \cdot G^{-1}(y)$.
To <u>sample</u> a pre-image of A, candidate approach is to sample $z \leftarrow D_{\mathcal{L}_y^\perp(G)}$ and output $x = Rz$.

Since $G = g^T \otimes I_n$, it suffice to sample from $\mathcal{L}_y^\perp(g^T)$. Now $\mathcal{L}^\perp(g^T)$ can be described by the following short basis:

$$B = \begin{pmatrix} 2 & & & \\ -1 & 2 & & \\ & -1 & \ddots & \\ & & & 2 \\ & & -1 & 2 \end{pmatrix} \in \mathbb{Z}_q^{t \times t} \quad \text{where } t = \log q$$

Observe

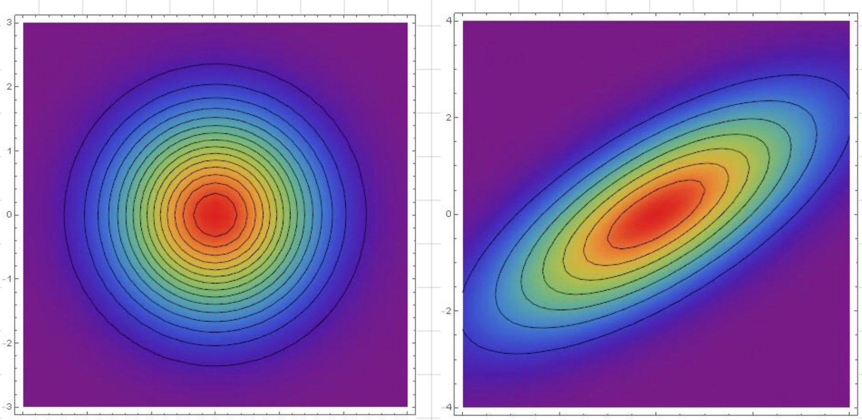
$$(1 \;\; 2 \;\; 4 \;\; \cdots \;\; 2^{\log q}) \cdot \begin{pmatrix} 2 & & & \\ -1 & 2 & & \\ & -1 & \ddots & \\ & & -1 & 2 \end{pmatrix} = 0 \pmod q$$

[this is when q is power of two, similar construction possible for non-power-of-two as well]

Gram-Schmidt norm of this basis is very short: $\tilde{B} = 2I_n$ so $\|\tilde{B}\| = 2$. Can use GPV to sample from $D_{\mathcal{L}_u^\perp(g^T), s, c}$ whenever $s > \sqrt{\omega(\log n)}$. Procedure is also very simple since $\tilde{B} = 2I_n$.

GVW allows us to sample $x \leftarrow D_{\mathbb{Z}^m, s}$ such that $Gx = y$ for any $y \in \mathbb{Z}_q^n$. What about the distribution of Rx.
Certainly $ARx = Gx = y$, but is Rx still a discrete Gaussian?

Yes... but discrete Gaussian is <u>not</u> spherical. Resulting distribution is discrete Gaussian with covariance $s^2 RR^T$.



Spherical Gaussian centered at 0

Distribution after rescaling samples by $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$

→ Is this problematic?

<u>Yes</u>: given multiple samples, can estimate covariance and this leaks R (the trapdoor)

Our goal is spherical discrete Gaussian with width s (i.e., covariance $s^2 I_n$).

<u>Key approach</u>: Gaussian convolution lemma
"Sum of two independent Gaussians is Gaussian" – analog generally holds for discrete Gaussian over lattices (see [Pei11]).

We can sample x where $Ax = y$ where x is from a Gaussian with covariance $s^2 RR^T$
To "correct" the distribution, we can sample z where $Az = 0$ and z is discrete Gaussian with covariance $\hat{s}^2 I - s^2 RR^T$

Given R where $AR = G$, goal is to sample $x \sim D_{\mathbb{Z}^m, s}$ where $Ax = y$

1. Sample perturbation $p \in \mathbb{Z}^m$ from discrete Gaussian with covariance $\hat{s}^2 I_n - s^2 RR^T$ (and mean 0)
2. Sample $u \leftarrow D_{\mathbb{Z}^m, s_1}$ where $Gu = y - Ap$ [Note: we will need that $s \gg \hat{s}$ – see analysis below]
3. Output $x = p + Ru$

Correctness: $Ax = Ap + ARu = Ap + Gu = Ap + y - Ap = y$

Security: Covariance of x is $\underbrace{\hat{s}^2 I_n - s^2 RR^T}_{p} + \underbrace{s^2 RR^T}_{Au} = \hat{s}^2 I_n$, which is independent of R $[x \sim D_{\mathbb{Z}^m, \hat{s}}]$

Can we sample a discrete Gaussian over $\mathbb{Z}^m$ with covariance $\hat{s}^2 I_n - s^2 RR^T$?

Requirement. $\hat{s}^2 I_n - s^2 RR^T$ is positive definite.

↳ In this case, we can write $\hat{s}^2 I_n - s^2 RR^T = MM^T$ (one such decomposition is given by Cholesky decomposition)

↳ Can now sample for $D_{\mathbb{Z}^m, 1}$ and scale by M

$s_1(R) = \max_{\|u\|=1} \|Ru\|$

Sufficient condition for $\hat{s}^2 I_n - s^2 RR^T$ to be positive definite: $\hat{s} \approx s \cdot s_1(R)$ where R is the largest singular value of R

quality of trapdoor is measured by $s_1(R)$

Note: Using GPV sampling algorithm, we can set $s = \omega(\sqrt{\log n})$ since $\mathcal{L}^\perp(G)$ has good basis (Gram-Schmidt norm of 2)

Recap: We will abstract the above sampling procedures into the following algorithms:

Sometimes, we let TrapGen take lattice parameters n, m, q explicitly

TrapGen(1^λ): Outputs (A, R) where $A \in \mathbb{Z}_q^{n \times m}$ is statistically close to uniform, R is short and $AR = G$.

SampleGaussian(s): Outputs $x \leftarrow D_{\mathbb{Z}^m, s}$ [rejection sampling]

SamplePre(A, R, u, s): Outputs $x \leftarrow D_{\mathcal{L}_u^\perp(A), s}$ [GPV sampling followed by perturbation]

Guarantee: if $s > s_1(R)\,\omega(\sqrt{\log m})$, then for $(A, R) \leftarrow$ TrapGen(1^λ):

if $R \in \{0,1\}^{m \times m}$, can bound $s_1(R) \lesssim m$

$$\{x \leftarrow \text{SampleGaussian}(s) : (x, Ax)\}$$
$$\overset{s}{\approx}$$
$$\{y \xleftarrow{R} \mathbb{Z}_q^m,\ x \leftarrow \text{SamplePre}(A, R, u, s) : (x, y)\}$$

GPV signatures in ROM:

could also be a short basis for A $\quad \mathbb{Z}_q^{n \times m} \quad \{0,1\}^{m \times m}$ $\quad$ bound on norm of samples from $D_{\mathbb{Z}^m, s}$: $\lesssim s \cdot \omega(\log \lambda)$ w.p. $1 - \mathrm{negl}(\lambda)$

Setup(1^λ): $(A, R) \leftarrow$ TrapGen(1^λ). Set $vk = (A, \beta)$ and $sk = (A, R, s)$ where $s, \beta = \tilde{O}(m)$

Sign(sk, m): Compute $y \leftarrow H(m) \in \mathbb{Z}_q^n$ and output $\sigma \leftarrow$ SamplePre(A, R, u, s) ← need to be careful, see HW1!

Verify(vk, m, σ): Check that $\|m\| \le \beta$ and $A \cdot \sigma = H(m)$

Security reduces to $\mathrm{ISIS}_{n,m,q,\beta}$. In the security proof, reduction algorithm does not have trapdoor. Will simulate signing queries on m by sampling $\sigma \leftarrow$ SampleGaussian(s), programming $H(m) \mapsto A\sigma$. This is statistically close to real signature distribution.