

Discrete Gaussian distribution is an important ingredient in the provable security of lattice-based cryptosystems. It was first employed by Micciancio and Regev (2007) in order to improve the hardness of Ajtai (1996)'s SIS problem by reducing the gaps in the approximated lattice problems. The technique was then widely adopted by subsequent lattice-based works to obtain provable security, for example, the learning with error (LWE) and ring learning with error (RLWE) problems. This discrete distribution behaves in a similar fashion as the continuous Gaussian, but with a discrete lattice support. In this section,¹ we will discuss some essential properties of discrete Gaussian distribution and how it is used to simplify and strengthen the hardness proof of SIS.

0.1 Discrete Gaussian distribution

We start by discussing some terms and intuitions about the better-understood continuous Gaussian distribution. A **Gaussian function** is a continuous function of the form

$$f(x) = a \cdot \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right).$$

Gaussian measure

The mostly common Gaussian function is the probability density function of the Gaussian distribution. For simplicity, we work with the case when $a = 1$, so we can define the **Gaussian measure** in \mathbb{R} as

$$\rho_{\sigma,c}(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right).$$

Another algebraic expression of the Gaussian measure is by using a *scale* parameter $s = \sqrt{2\pi}\sigma$. Substitute σ in the above equation and generalize the Gaussian measure to higher dimensional space \mathbb{R}^n , we get

$$\rho_{s,c}(\mathbf{x}) = \exp\left(-\frac{-\pi\|\mathbf{x}-\mathbf{c}\|^2}{s^2}\right). \quad (1)$$

Integrating the measure over \mathbb{R}^n , the total measure is²

$$\int_{\mathbf{x} \in \mathbb{R}^n} \rho_{s,c}(\mathbf{x}) \, d\mathbf{x} = s^n$$

Gaussian PDF

and hence we can define the n -dimensional (continuous) Gaussian probability density function (PDF) as

$$D_{s,c}(\mathbf{x}) = \frac{\rho_{s,c}(\mathbf{x})}{s^n}. \quad (2)$$

This is just the PDF of the Gaussian distribution that we know from probability theory. The expected squared distance from a Gaussian variable $\mathbf{x} \in \mathbb{R}^n$ to the distribution center \mathbf{c} (i.e., the variance of \mathbf{x}) is

$$E[\|\mathbf{x}-\mathbf{c}\|^2] = \sigma^2 = \frac{ns^2}{2\pi}.$$

Geometrically, most of the Gaussian samples lie in the n -dimensional ball centered at \mathbf{c} with radius $s\sqrt{n/2\pi}$.

Equation (1) and Equation (2) would still make sense if \mathbf{x} is a discrete lattice vector. In addition, a lattice L is a countable set, so the total Gaussian measure and the Gaussian PDF over the lattice are expressed slightly differently

$$\rho_{s,c}(L) = \sum_{\mathbf{x} \in L} \rho_{s,c}(\mathbf{x})$$

$$D_{s,c}(L) = \frac{\rho_{s,c}(L)}{s^n}.$$

Discrete Gaussian

Hence, we can define the **discrete Gaussian distribution** over the lattice L for all lattice vectors $\mathbf{x} \in L$ as

$$D_{L,s,c}(\mathbf{x}) = \frac{D_{s,c}(\mathbf{x})}{D_{s,c}(L)} = \frac{\rho_{s,c}(\mathbf{x})}{\rho_{s,c}(L)}.$$

¹This section is part of the work *A Tutorial Introduction to Lattice-based Cryptography and Homomorphic Encryption* by the authors Yang Li, Kee Siong Ng, Michael Purcell from the School of Computing, Australian National University @2022.

²The total measure is not 1 because the coefficient a in the Gaussian function is ignored.

This can also be interpreted as the probability of \mathbf{x} conditioning on the fact that it is a lattice vector. The numerator is the probability of \mathbf{x} being a lattice vector and follows a Gaussian distribution. The denominator is the probability of an arbitrary Gaussian random variable in \mathbb{R}^n takes on a value of a lattice vector in L .

The discrete Gaussian distribution is commonly used nowadays to introduce randomness in the proof of lattice problems and lattice-based cryptosystems. Unlike a uniform distribution over a space (e.g., the way uniformity was proved in Ajtai’s SIVP $_\gamma$ to SIS problem), Gaussian distribution does not have sharp boundaries, which is useful when smoothing a distribution over a space. More precisely, given a Gaussian distribution $\rho_{s,c}(s)$ whose center is a lattice point (i.e., $c \in L$), if random samples from this distribution are taken modulo the lattice fundamental region, the resulting samples will induce a distribution within the fundamental region. Whether or not such a distribution is close to the uniform distribution depends on the scale s of the Gaussian distribution. Obviously, the larger s is, the closer the induced distribution is to uniform.

To give a quantitative threshold on how large s needs to be, Micciancio and Regev (2007) introduced the smoothing parameter. As the name suggests, the purpose of this parameter is to measure the minimum Gaussian noise magnitude, so that if the noise is added to a lattice \mathbb{Z}^n , the lattice is “blurred” to almost a uniform distribution over \mathbb{R}^n . It is, however, formally defined in terms of the Gaussian measure on the dual lattice. For the rest of this section, we assume $\epsilon(n) > 0$ (or just $\epsilon > 0$ if the context is clear) is a negligible function of the space dimension n .

Smoothing parameter **Definition 0.1.1.** *The smoothing parameter of an n -dimensional lattice L , denoted $\eta_\epsilon(L)$, is the smallest s such that the Gaussian measure $\rho_{1/s}(L^* \setminus \{0\}) \leq \epsilon$.*

In other words, the Gaussian measure gives almost all weights to the origin in the dual lattice. Note $\rho_{1/s}(L^* \setminus \{0\})$ is a well-defined decreasing function of s with the range $(0, \infty)$. More precisely, $\lim_{s \rightarrow \infty} \rho_{1/s}(L^* \setminus \{0\}) = 0$ because when $s \rightarrow \infty$ its inverse $\frac{1}{s} \rightarrow 0$, which implies the Gaussian measure puts almost all weights on 0. Conversely, $\lim_{s \rightarrow 0} \rho_{1/s}(L^* \setminus \{0\}) = \infty$.

Next, we relate the smoothing parameter to two standard lattice quantities. We state the results and the intuitions without proving them.

$\lambda_1(L^*)$ **Lemma 0.1.2.** *The smoothing parameter of an n -dimensional lattice L satisfies $\eta_\epsilon(L) \leq \frac{\sqrt{n}}{\lambda_1(L^*)}$.*

It is not difficult to see this inverse relationship between $\eta_\epsilon(L)$ and $\lambda_1(L^*)$. By intuition, the smaller $\lambda_1(L)$ is the smaller $\eta_\epsilon(L)$ needs to be and vice versa. As explained in the previous section, $\lambda_1(L)$ and $\lambda_1(L^*)$ are in an inverse relationship. So the larger $\lambda_1(L^*)$ is the smaller $\lambda_1(L)$ is which requires a smaller $\eta_\epsilon(L)$. The following lemma relates the smoothing parameter to the successive minima of a lattice.

$\lambda_n(L)$ **Lemma 0.1.3.** *The smoothing parameter of an n -dimensional lattice L satisfies*

$$\eta_\epsilon(L) \leq \sqrt{\frac{\ln(2n(1 + 1/\epsilon))}{\pi}} \cdot \lambda_n(L).$$

0.2 Fourier transform of discrete Gaussian distribution

So far, we have defined the discrete Gaussian distribution and the smoothing parameter, and how they are essential for sampling random lattice noise. In this subsection, we will state the theorems that guarantee the uniformity of the discrete Gaussian when its scale is at least as large as the smoothing parameter and its similarity to the continuous Gaussian distribution. We also include some key mathematical tools to prove these results, including the Fourier transform of the Gaussian measure and the Poisson summation formula. Many of these results are taken from the lecture notes *The Gaussians Distribution* by Daniele Micciancio for the course *Lattice Algorithms and Applications*, Winter 2016.

We start this subsection by stating two key properties of Gaussian distribution. The reader can safely skip the rest of this subsection if the proofs of these properties are not essential. Recall that any vector $t \in \mathbb{R}^n$ in the span of a lattice L is uniquely identifiable by a lattice vector v and a (translation of) vector $w \in F$ in the lattice fundamental domain F . This gives rise to a way of reducing an arbitrary vector in \mathbb{R}^n to a vector within F by taking $w = t \bmod F$ the vector modulo the fundamental domain. The next lemma addresses the near uniformity of the distribution over F induced by applying this modulo operation.

Near uniformity **Lemma 0.2.1.** *Let L be an n -dimensional lattice and $D_{s,\mathbf{c}}$ be a Gaussian distribution with arbitrary scale $s \geq \eta_\epsilon(L)$ and center $\mathbf{c} \in \mathbb{R}^n$, the statistical distance between $D_{s,\mathbf{c}} \bmod F$ and a uniform distribution $U(F)$ over the fundamental region F is*

$$\Delta(D_{s,\mathbf{c}} \bmod F, U(F)) \leq \frac{\epsilon}{2}.$$

It can be proved that for $s > 0$ the statistical distance $\Delta \leq \rho_{1/s}(L^* \setminus \{\mathbf{0}\})$. The lemma then follows since the Gaussian's scale is at least as large as the smoothing parameter. We do not go through the proof here (see (Micciancio and Regev, 2007) Lemma 4.1). It uses the Fourier transform of the Gaussian function and the properties of the discrete Gaussian described before.

The next lemma proves a similar behaviour of the discrete and continuous Gaussian distributions when the scale of discrete Gaussian is sufficiently large.

Similar to continuous Gaussian **Lemma 0.2.2.** *Let $D_{L,s,\mathbf{c}}$ be a discrete Gaussian distribution over an n -dimensional lattice L with arbitrary scale $s \geq 2\eta_\epsilon(L)$ and center $\mathbf{c} \in \mathbb{R}^n$. For $0 < \epsilon < 1$, the following are satisfied*

$$\begin{aligned} \|E_{\mathbf{x} \sim D_{L,s,\mathbf{c}}}[\mathbf{x} - \mathbf{c}]\|^2 &\leq \left(\frac{\epsilon}{1-\epsilon}\right)^2 s^2 n, \\ E_{\mathbf{x} \sim D_{L,s,\mathbf{c}}}[\|\mathbf{x} - \mathbf{c}\|^2] &\leq \left(\frac{1}{2\pi} + \frac{\epsilon}{1-\epsilon}\right)^2 s^2 n. \end{aligned}$$

The first inequality suggests that on expectation the random samples from $D_{L,s,\mathbf{c}}$ are close to the discrete Gaussian center, with the distance at most $s\sqrt{n}$. It ensures that if the discrete Gaussian is centered at the origin, then the lattice vectors sampled from this distribution will have small norms. The second inequality is consistent with the continuous Gaussian's variance (i.e., $\frac{ns^2}{2\pi}$) as discussed in the previous subsection.

Fourier transform According to its definition, the Gaussian measure over \mathbb{R}^n is integrable, that is, $\int_{\mathbf{x} \in \mathbb{R}^n} \rho_{s,\mathbf{c}}(\mathbf{x}) \, d\mathbf{x} = s^n < \infty$, so it has the Fourier transform

$$\hat{\rho}_{s,\mathbf{c}}(\mathbf{y}) = \int_{\mathbf{x} \in \mathbb{R}^n} \rho_{s,\mathbf{c}}(\mathbf{x}) e^{-2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} \, d\mathbf{x}. \quad (3)$$

This is the most common convention of the Fourier transform of an integrable function. The Fourier transform of the Gaussian is important for proving the Poisson summation formula as well as proving some properties of the discrete Gaussian distribution. It has several important properties, the most important one is that the Fourier transform of the Gaussian measure is itself. We state the result below for a Gaussian measure centered at the origin.

Lemma 0.2.3. *The Gaussian function centered at the origin $\rho_s(\mathbf{x}) = e^{-\pi\|\mathbf{x}\|^2/s^2}$ equals its Fourier transform times a scaling factor, that is, $\rho_s(\mathbf{x}) = s^n \hat{\rho}_s(\mathbf{x})$.*

In the special case when $s = 1$, the lemma states that the Gaussian function is its own Fourier transform. We go through the proof here in order to get familiar with the Fourier transform of Gaussian functions.

Proof. We first prove the equality for 1 dimensional space \mathbb{R} . This can be done by substituting the Gaussian function into the integral of Fourier transform

$$\begin{aligned} \hat{\rho}_s(y) &= \int_{x \in \mathbb{R}} \rho_s(x) e^{-2\pi i x y} \, dx \\ &= \int_{x \in \mathbb{R}} e^{-\pi x^2/s^2} e^{-2\pi i x y} \, dx \\ &= \int_{x \in \mathbb{R}} e^{-\pi(x^2/s^2 - 2ixy)} \, dx \end{aligned}$$

Making the exponent a complete square and taking the y term out of the integral, it becomes

$$\begin{aligned}
 &= e^{-\pi y^2 s^2} \int_{x \in \mathbb{R}} e^{-\pi(x+iy s^2)^2/s^2} dx \\
 &= \rho_s(y) \int_{x \in \mathbb{R}+iy s^2} \rho_s(x) dx \\
 &= \rho_s(y) \int_{x \in \mathbb{R}} \rho_s(x) dx \\
 &= \rho_s(y)s.
 \end{aligned}$$

The second last equality is by Cauchy's Theorem and the last equality is just the total integral of the Gaussian measure over \mathbb{R} . Now we can prove the general case in \mathbb{R}^n using the above result by integrating each term in the vectors \mathbf{x} and \mathbf{y}

$$\begin{aligned}
 \hat{\rho}_s(\mathbf{y}) &= \int_{\mathbf{x} \in \mathbb{R}^n} \rho_s(\mathbf{x}) e^{-2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} d\mathbf{x} \\
 &= \int_{x \in \mathbb{R}^n} \prod_{k=1}^n e^{-\pi x_k^2/s^2} e^{-2\pi i x_k y_k} dx \\
 &= \prod_{k=1}^n \int_{x_k \in \mathbb{R}} e^{-\pi x_k^2/s^2} e^{-2\pi i x_k y_k} dx_k \\
 &= \prod_{k=1}^n \hat{\rho}_s(y_k) \\
 &= \prod_{k=1}^n s \rho_s(y_k) \\
 &= s^n \rho_s(\mathbf{y}).
 \end{aligned}$$

The third equality is by the property of integrating exponential functions. The second last equality is by the result in the 1-dimensional space. \square

The following lemma states another three important properties of the Fourier transform of integrable functions, not necessarily just the Gaussian function. It explicitly tells us how a function's Fourier transform changes when the function's input is translated or linearly transformed by a non-singular matrix. These are particularly important for our study of the discrete Gaussian distribution, as they allow us to shift back to the standard discrete Gaussian distribution D_L , that is, $D_{L,s,\mathbf{c}}$ where the scale $s = 1$ and the center $\mathbf{c} = \mathbf{0}$. We will state the lemma without proving it, as the proofs are relatively straightforward by writing out the full algebraic forms of the Fourier transform and Gaussian function.

Lemma 0.2.4. *For any integrable function $f(\mathbf{x})$ in the n -dimensional space \mathbb{R}^n , the following are satisfied:*

1. *For a non-singular matrix M , if the function $h(M\mathbf{x}) = f(\mathbf{x})$ then its Fourier transform $\hat{h}(\mathbf{y}) = \det(M) \hat{f}(M^T \mathbf{y})$.*
2. *If $h(\mathbf{x}) = f(\mathbf{x} + \mathbf{v})$, then its Fourier transform $\hat{h}(\mathbf{y}) = \hat{f}(\mathbf{y}) \cdot e^{2\pi i \langle \mathbf{y}, \mathbf{v} \rangle}$.*
3. *If $h(\mathbf{x}) = f(\mathbf{x}) \cdot e^{2\pi i \langle \mathbf{x}, \mathbf{v} \rangle}$, then its Fourier transform $\hat{h}(\mathbf{y}) = \hat{f}(\mathbf{y} - \mathbf{v})$.*

It follows from this lemma an important mathematical result, the **Poisson summation formula**. For our purpose, we only state the formula for arbitrary lattices. The formula relates the sum of an integrable function over a lattice to the sum of its Fourier transform over the dual lattice. This in return entails properties of the Gaussian measure over lattices.

*Poisson
summation
formula*

Lemma 0.2.5. *For any n -dimensional lattice L and an integrable function f of L , it satisfies*

$$\sum_{\mathbf{x} \in L} f(x) = \det(L^*) \sum_{\mathbf{y} \in L^*} \hat{f}(\mathbf{y}). \tag{4}$$

A special case of the Poisson summation formula is when f is the Gaussian function. It implies that the total Gaussian measure over a lattice L equals the total Gaussian measure over its dual multiplies the dual determinant, that is,

$$\begin{aligned}\rho_{s,\mathbf{c}}(L) &= \det(L^*)\hat{\rho}_{s,\mathbf{c}}(L^*) \text{ or} \\ \hat{\rho}_{s,\mathbf{c}}(L^*) &= \det(L)\rho_{s,\mathbf{c}}(L).\end{aligned}$$

An important application of the Poisson summation formula in this context is the statement that the Gaussian measure of a lattice is maximized when centering at a lattice point.

Lemma 0.2.6. *For an n -dimensional lattice L , the Gaussian measure $\rho_{s,\mathbf{c}}(\mathbf{x})$ with the scale $s > 0$ and center $\mathbf{c} \in \mathbb{R}^n$ satisfies $\rho_{s,\mathbf{c}}(\mathbf{x}) \leq \rho_s(\mathbf{x})$.*

Proof. The key of the proof is shifting the Gaussian measure $\rho_{s,\mathbf{c}}(\mathbf{x})$ to the zero centered Gaussian $\rho_s(\mathbf{x})$, then use the property that a lattice vector and a dual lattice vector has an integer dot product.

$$\begin{aligned}\rho_{s,\mathbf{c}}(L) &= \det(L^*)\hat{\rho}_{s,\mathbf{c}}(L^*) \\ &= \det(L^*) \sum_{\mathbf{y} \in L^*} \hat{\rho}_{s,\mathbf{c}}(\mathbf{y}) \\ &= \det(L^*) \sum_{\mathbf{y} \in L^*} e^{2\pi i \langle -\mathbf{c}, \mathbf{y} \rangle} \hat{\rho}_s(\mathbf{y}) \\ &\leq \det(L^*) \sum_{\mathbf{y} \in L^*} \hat{\rho}_s(\mathbf{y}) \\ &= \rho_s(L).\end{aligned}$$

The third equality is by the second point of Lemma 0.2.4 and the equality $\hat{\rho}_{s,\mathbf{c}}(\mathbf{y}) = \hat{\rho}_s(\mathbf{y} + \mathbf{c})$. The inequality is due to the fact that $e^{2\pi i \langle -\mathbf{c}, \mathbf{y} \rangle} \leq 1$ and the maximum is obtained when the dot product $\langle -\mathbf{c}, \mathbf{y} \rangle$ is an integer, which requires the Gaussian center \mathbf{c} to be a lattice point in L . \square

0.3 Discrete Gaussian for provable security

In this subsection, we revisit the hardness proof of Ajtai's short integer solution (SIS) problem, but use the discrete Gaussian tool to reduce the gaps of the hard lattice problems. Recall that SIS is parameterized by a modulus q , number of linearly independent vectors m and a norm bound β . These parameters are often functions of the security parameter n when building a cryptosystem based on SIS. The purpose of SIS is to find a short integer vector $\mathbf{x} \in \mathbb{Z}^m$ such that

- $\|\mathbf{x}\| \leq \beta$ and
- $A\mathbf{x} = \mathbf{0} \in \mathbb{Z}_q^n$ for an arbitrary integer matrix $A \in \mathbb{Z}_q^{n \times m}$.

A guarantee of an SIS solution was stated in Lemma 5.2 in (Micciancio and Regev, 2007) with the parameter constraint $\beta(n) \geq \sqrt{m}q^{n/m}$.

Similar to Ajtai's proof, Micciancio and Regev (2007) also introduced an intermediate lattice problem - incremental guaranteed distance decoding - for a simple reduction to SIS. The standard lattice problems can be reduced to this intermediate problem relatively easier. The problem is similar to the bounded distance decoding (BDD) problem (??) but finds a lattice vector that is within a bounded distance to the target.

Definition 0.3.1. *Given a basis B of an n -dimensional lattice L , a set of linearly independent lattice vectors $S \subseteq L$, a target vector $\mathbf{t} \in \mathbb{R}^n$ and a real $r > \gamma(n)\lambda_n(B)$, the **incremental guaranteed distance decoding (INCGDD)** problem outputs a lattice vector $\mathbf{v} \in L$ such that $\|\mathbf{v} - \mathbf{t}\| \leq (\|S\|/g) + r$.*

Here, $\|S\|$ is the maximum length of a lattice vector in S and r is needed to guarantee a solution exists. Notice we present the definition with $r > \gamma(n)\lambda_n(B)$, where the original definition has $r > \gamma(n)\phi(B)$, where $\phi(B)$ is an arbitrary function on the lattice, such as its smoothing parameter or $\lambda_n(B)$.

For a set of vectors $S = \{\mathbf{s}_1, \dots, \mathbf{s}_n\}$, denote $P(S) = \{\sum_{i=1}^n x_i \mathbf{s}_i \mid x_i \in [0, 1)\}$ the half-opened parallelepiped generated by S . If S is a lattice basis, $P(S)$ is the fundamental region. The discrete Gaussian tool is used to prove one of the key steps of the reduction. More precisely, the discrete

Gaussian distributions gives rise to a sampling mechanism (as shown next), whose output is a pair (\mathbf{c}, \mathbf{y}) where \mathbf{c} is almost uniform in $P(B)$ and \mathbf{y} is a sample from a discrete Gaussian distribution over the lattice $L(B)$.

Lemma 0.3.2. *Given an n -dimensional lattice $L(B)$, a vector $\mathbf{t} \in \mathbb{R}^n$ and a scale $s \geq \eta_\epsilon(L)$ for some $\epsilon > 0$, there is a PPT sampling algorithm $S(B, \mathbf{t}, s)$ outputs a pair $(\mathbf{c}, \mathbf{y}) \in P(B) \times L(B)$ such that*

- *the statistical distance between \mathbf{c} 's distribution and the uniform distribution over $P(B)$ is at most $\epsilon/2$,*
- *for any vector $\hat{\mathbf{c}} \in P(B)$, the distribution of \mathbf{y} conditioning on $\mathbf{c} = \hat{\mathbf{c}}$ is a discrete Gaussian distribution $D_{L, s, \mathbf{t} + \hat{\mathbf{c}}}$.*

Proof. The sampling mechanism S starts by generating a Gaussian noise $\mathbf{r} \leftarrow D_{s, \mathbf{t}}$, then outputs $\mathbf{c} = -\mathbf{r} \bmod P(B)$ and $\mathbf{y} = \mathbf{c} + \mathbf{r}$. It implies that $\mathbf{c} \sim D_{s, -\mathbf{t}} \bmod P(B)$. By Lemma 0.2.1, we have $\Delta(D_{s, -\mathbf{t}} \bmod F, U(P(B))) \leq \epsilon/2$.

For any vector $\hat{\mathbf{c}}$ in the parallelepiped, when adding to the Gaussian noise \mathbf{r} , the result $\hat{\mathbf{c}} + \mathbf{r} \sim D_{s, \mathbf{t} + \hat{\mathbf{c}}}$ is also from Gaussian but with a shifted center. Moreover, $\mathbf{c} = \hat{\mathbf{c}}$ implies $\hat{\mathbf{c}} + \mathbf{r} = \mathbf{c} + \mathbf{r} = \mathbf{y} \in L(B)$. So the probability of \mathbf{y} conditioning on $\mathbf{c} = \hat{\mathbf{c}}$ is equivalent to conditioning on $\hat{\mathbf{c}} + \mathbf{r}$ being a lattice vector, which implies \mathbf{y} 's distribution is discrete Gaussian. \square

The output of the sampling mechanism is then used by a combining procedure to solve the INCGDD problem together with the hypothetical SIS oracle. We skip that part of the reduction and refer the reader to Lemma 5.8 and Theorem 5.9 in Micciancio and Regev (2007) for the detailed proofs.

References

- M. Ajtai. Generating hard instances of lattice problems. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pages 99–108, 1996.
- D. Micciancio and O. Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007.