Independent Metropolis-Hastings-Klein Algorithm for Lattice Gaussian Sampling

Zheng Wang and Cong Ling
Department of EEE, Imperial College London
London, SW7 2AZ, United Kingdom
Email: z.wang10, c.ling@imperial.ac.uk

***Abstract*—Sampling from the lattice Gaussian distribution is emerging as an important problem in coding and cryptography. In this paper, a Markov chain Monte Carlo (MCMC) algorithm referred to as the independent Metropolis-Hastings-Klein (MHK) algorithm is proposed for lattice Gaussian sampling, which overcomes the restriction on the standard deviation confronted by the Klein algorithm. It is proven that the Markov chain arising from the proposed MHK algorithm is uniformly ergodic, namely, it converges to the stationary distribution exponentially fast. Moreover, the rate of convergence is explicitly calculated in terms of the theta series, making it possible to predict the mixing time of the underlying Markov chain.**

***Index Terms*—Lattice Gaussian sampling, Metropolis-Hastings sampling, MCMC methods, lattice coding and decoding.**

I. Introduction

Recently, the lattice Gaussian distribution is emerging as a common theme in various research fields. In mathematics, Banaszczyk firstly applied it to prove the transference theorems for lattices [1]. In coding, lattice Gaussian distribution was employed to obtain the full shaping gain for lattice coding [2], [3], and to achieve the capacity of the Gaussian channel and the secrecy capacity of the Gaussian wiretap channel, respectively [4], [5]. In cryptography, the lattice Gaussian distribution has already become a central tool in the construction of many primitives. Specifically, Micciancio and Regev applied it to propose the lattice-based cryptosystems based on the worst-case hardness assumptions [6]. Meanwhile, it also has underpinned the fully-homomorphic encryption for cloud computing [7]. Algorithmically, lattice Gaussian sampling with a suitable variance allows to solve the shortest vector problem (SVP) and the closest vector problem (CVP); for example, it has led to efficient lattice decoding for multi-input multi-output (MIMO) systems [8], [9].

Due to the central role of the lattice Gaussian distribution playing in these fields, its sampling algorithms become an important computational problem. Unfortunately, compared to sampling from continuous Gaussian distributions, it is by no means trivial to perform the sampling even from a low-dimensional discrete Gaussian distribution. As the default sampling algorithm for lattices, Klein's algorithm [10] samples within a negligible statistical distance from the lattice Gaussian distribution if and only if the standard deviation σ is sufficiently large, namely, $\sigma \geq \omega(\sqrt{\log\, n}) \cdot \max_{1\leq i\leq n} \|\widehat{\mathbf{b}}_i\|$ [11], where n denotes the lattice dimension, $\omega(\cdot)$ is a function related to n and $\widehat{\mathbf{b}}_i$'s are the Gram-Schmidt vectors of the lattice basis $\mathbf{B}$, thereby rendering Klein's algorithm inapplicable to smaller σ. To address this issue, the Gibbs algorithm rooted in Markov chain Monte Carlo (MCMC) methods was introduced into lattice Gaussian sampling; it is the first lattice algorithm able to sample in the range that Klein's algorithm cannot reach [12]. However, the related analysis of the convergence rate for the associated Markov chain was lacking.

Basically, MCMC methods attempt to sample from the target distribution by building a Markov chain, which randomly generates the next sample conditioned on the previous samples. In this paper, we propose a new algorithm for lattice Gaussian based on the independent Metropolis-Hastings (MH) algorithm [13]. The MH algorithm makes use of a proposal distribution which suggests a possible move and then employs a acceptance-rejection rule to decide the next move. Therefore, the art of designing an efficient MH algorithm chiefly lies in choosing an appropriate proposal distribution. To this end, we use Klein's algorithm to generate the proposal distribution, leading to the new independent Metropolis-Hastings-Klein (MHK) algorithm for lattice Gaussian sampling. Moreover, the rate of convergence is analyzed and the Markov chain associated with the proposed MHK algorithm is demonstrated to be uniformly ergodic, which means it converges to its stationary distribution exponentially fast. Therefore, the mixing time of the underlying Markov chain becomes tractable.

The rest of this paper is organized as follows. Section II introduces the lattice Gaussian distribution and briefly reviews the basics of MCMC methods. In Section III, we propose the independent MHK algorithm for lattice Gaussians, followed by the demonstration of uniform ergodicity and the convergence rate analysis in Section IV.

II. Lattice Gaussian Distribution

Let $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n] \subset \mathbb{R}^n$ consist of n linearly independent vectors. The n-dimensional lattice Λ generated by $\mathbf{B}$ is defined by

$$\Lambda = \{\mathbf{Bx} : \mathbf{x} \in \mathbb{Z}^n\}, \tag{1}$$

where $\mathbf{B}$ is known as the lattice basis. We define the Gaussian function centered at $\mathbf{c} \in \mathbb{R}^n$ for standard deviation $\sigma > 0$ as

$$\rho_{\sigma,\mathbf{c}}(\mathbf{z}) = e^{-\frac{\|\mathbf{z}-\mathbf{c}\|^2}{2\sigma^2}}, \tag{2}$$

for all $\mathbf{z} \in \mathbb{R}^n$. When $\mathbf{c}$ or σ are not specified, we assume that they are $\mathbf{0}$ and 1 respectively. Then, the *discrete Gaussian distribution* over Λ is defined as

$$D_{\Lambda,\sigma,\mathbf{c}}(\mathbf{x}) = \frac{\rho_{\sigma,\mathbf{c}}(\mathbf{Bx})}{\rho_{\sigma,\mathbf{c}}(\Lambda)} = \frac{e^{-\frac{1}{2\sigma^2}\|\mathbf{Bx}-\mathbf{c}\|^2}}{\sum_{\mathbf{x}\in\mathbb{Z}^n} e^{-\frac{1}{2\sigma^2}\|\mathbf{Bx}-\mathbf{c}\|^2}} \quad (3)$$

for all $\mathbf{Bx} \in \Lambda$, where $\rho_{\sigma,\mathbf{c}}(\Lambda) \triangleq \sum_{\mathbf{Bx}\in\Lambda} \rho_{\sigma,\mathbf{c}}(\mathbf{Bx})$.

Obviously, an intuition of $D_{\Lambda,\sigma,\mathbf{c}}(\mathbf{x})$ suggests that a lattice point $\mathbf{Bx}$ closer to $\mathbf{c}$ will be sampled with a higher probability. Therefore, sampling from lattice Gaussian can be naturally used in solving the CVP (where $\mathbf{c}$ is the query point) and SVP (where $\mathbf{c} = \mathbf{0}$) in lattices, and because of this, Klein's algorithm that samples from a Gaussian-like distribution was originally designed for lattice decoding [10]. As shown in Algorithm 1, the operation of Klein's algroithm has polynomial complexity $O(n^2)$ excluding QR decomposition. More precisely, by sequentially sampling from the 1-dimensional conditional Gaussian distribution $D_{\mathbb{Z},\sigma_i,\widetilde{x}_i}$ in a backward order from x_n to x_1, the Gaussian-like distribution arising from Klein's algorithm is given by

$$P_{\text{Klein}}(\mathbf{x}) = \prod_{i=1}^{n} D_{\mathbb{Z},\sigma_i,\widetilde{x}_i}(x_i) = \frac{\rho_{\sigma,\mathbf{c}}(\mathbf{Bx})}{\prod_{i=1}^{n} \rho_{\sigma_i,\widetilde{x}_i}(\mathbb{Z})}, \quad (4)$$

where $P_{\text{Klein}}(\mathbf{x})$ has been demonstrated in [11] to be close to $D_{\Lambda,\sigma,\mathbf{c}}(\mathbf{x})$ within a negligible statistical distance if

$$\sigma \geq \omega(\sqrt{\log\, n}) \cdot \max_{1\leq i\leq n} \|\widehat{\mathbf{b}}_i\|. \quad (5)$$

As for sampling in the range $\sigma < \omega(\sqrt{\log\, n}) \cdot \max_{1\leq i\leq n}\|\widehat{\mathbf{b}}_i\|$, MCMC methods have become an alternative solution, where the discrete Gaussian distribution $D_{\Lambda,\sigma,\mathbf{c}}$ is viewed as a complex target distribution lacking direct sampling methods. By establishing a Markov chain that randomly generates the next state based on the previous states, MCMC is capable of sampling from the target distribution of interest, thereby removing the restriction on σ in lattice Gaussian distributions [12].

As a special case of the MH algorithm, Gibbs sampling employs 1-dimensional conditional distributions to build the Markov chain, where all the other variables in the distribution are unchanged in each Markov move. In [12], a flexible block-based Gibbs algorithm was proposed for lattice Gaussian distributions. Compared to the standard Gibbs algorithm that constructs the Markov chain by only considering univariate sampling at each time, it performs the sampling over multiple elements within a block to enhance the convergence performance of the Markov chains.

Definition 1 ([14]). *A Markov chain with stationary distribution $\pi(\cdot)$ is ergodic if*

$$\lim_{t\to\infty} \|P^t(\mathbf{x},\cdot) - \pi(\cdot)\|_{TV} = 0, \quad (6)$$

where $P^t(\mathbf{x};\cdot)$ denotes the row of the transition matrix $\mathbf{P}$ for t Markov moves and $\|\cdot\|_{TV}$ represents the total variation distance.

Although *ergodicity* implies asymptotic convergence to stationarity, it does not say anything about the rate of this convergence. One qualitative convergence rate of our concern in this context is referred to as *uniform ergodicity*.

Algorithm 1 Klein's Algorithm

Input: $\mathbf{B}, \sigma, \mathbf{c}$
Output: $\mathbf{Bx} \in \Lambda$
1: let $\mathbf{B} = \mathbf{QR}$ and $\mathbf{c}' = \mathbf{Q}^T\mathbf{c}$
2: **for** $i = n, \ldots, 1$ **do**
3: let $\sigma_i = \frac{\sigma}{|r_{i,i}|}$ and $\widetilde{x}_i = \frac{c'_i - \sum_{j=i+1}^{n} r_{i,j}x_j}{r_{i,i}}$
4: sample x_i from $D_{\mathbb{Z},\sigma_i,\widetilde{x}_i}$
5: **end for**
6: return $\mathbf{Bx}$

Definition 2 ([14]). *A Markov chain having stationary distribution $\pi(\cdot)$ is uniformly ergodic if there exists $0 < \delta < 1$ and $M < \infty$ such that for all $\mathbf{x}$*

$$\|P^t(\mathbf{x},\cdot) - \pi(\cdot)\|_{TV} \leq M(1-\delta)^t. \quad (7)$$

Obviously, the value of the exponential decay coefficient δ is the key to determine the convergence rate. As M is a constant, a salient feature of uniform ergodicity is that the convergence rate does not depend on the initial state $\mathbf{x}$.

As a parameter which measures the time required by a Markov chain to get close to the stationary distribution, the *mixing time* is defined by [15].

$$t_{\text{mix}}(\epsilon) = \min\{t : \max\|P^t(\mathbf{x},\cdot) - \pi(\cdot)\|_{TV} \leq \epsilon\}. \quad (8)$$

III. Independent MHK Algorithm

In this section, we present the conventional MH sampling in MCMC and give the proposed independent MHK algorithm for lattice Gaussian sampling. Note that the Markov chain that we are concerned with here has a countably infinite state space, i.e., the lattice Λ.

In [13], the original Metropolis algorithm was extended to a general scheme known as the Metropolis-Hastings algorithm. Let us consider a target invariant distribution π together with a candidate proposal distribution $q(\mathbf{x},\mathbf{y})$. Given the current state $\mathbf{x}$ for Markov chain $\mathbf{X}_t$, a state candidate $\mathbf{y}$ for the next Markov move $\mathbf{X}_{t+1}$ is generated from the proposal distribution $q(\mathbf{x},\cdot)$. Then the acceptance ratio α is computed by

$$\alpha(\mathbf{x},\mathbf{y}) = \min\left\{1, \frac{\pi(\mathbf{y})q(\mathbf{y},\mathbf{x})}{\pi(\mathbf{x})q(\mathbf{x},\mathbf{y})}\right\}, \quad (9)$$

and $\mathbf{y}$ will be accepted as the new state by $\mathbf{X}_{t+1}$ with probability α. Otherwise, $\mathbf{x}$ will be retained by $\mathbf{X}_{t+1}$ with probability $1-\alpha$. In this way, a Markov chain $\{\mathbf{X}_0, \mathbf{X}_1, \ldots\}$ is established with the transition probability $P(\mathbf{x},\mathbf{y})$ as follows:

$$P(\mathbf{x},\mathbf{y}) = \begin{cases} q(\mathbf{x},\mathbf{y})\alpha(\mathbf{x},\mathbf{y}) & \text{if } \mathbf{y} \neq \mathbf{x}, \\ 1 - \sum_{\mathbf{z}\neq\mathbf{x}} q(\mathbf{x},\mathbf{z})\alpha(\mathbf{x},\mathbf{z}) & \text{if } \mathbf{y} = \mathbf{x}. \end{cases} \quad (10)$$

In MH algorithms, the proposal distribution $q(\mathbf{x},\mathbf{y})$ can be any fixed distribution from which we can easily draw samples. To this end, many variations of MH algorithms with different

configurations of $q(\mathbf{x},\mathbf{y})$ were proposed and a very special one among them is the independent MH algorithm where [16]

$$q(\mathbf{x},\mathbf{y}) = q(\mathbf{y}). \tag{11}$$

Clearly, the candidate state $\mathbf{y}$ generated for $\mathbf{X}_{t+1}$ does not depend on the previous state $\mathbf{x}$ and this method originally appeared as an alternative to rejection sampling and importance sampling [13]. However, how to sample the candidate state $\mathbf{y}$ tends to be difficult.

Now, we present the proposed independent MHK algorithm, where Klein's algorithm is used to generate the multi-dimensional proposal distribution. As shown in Algorithm 2, it consists of three basic steps:

1) *Sample from the independent proposal distribution through Klein's algorithm to obtain the candidate state* $\mathbf{y}$ *for* $\mathbf{X}_{t+1}$,

$$q(\mathbf{x},\mathbf{y}) = q(\mathbf{y}) = \frac{\rho_{\sigma,\mathbf{c}}(\mathbf{By})}{\prod_{i=1}^{n}\rho_{\sigma_i,\widetilde{y}_i}(\mathbb{Z})}, \tag{12}$$

where $\mathbf{y} \in \mathbb{Z}^n$.

2) *Calculate the acceptance ratio* $\alpha(\mathbf{x},\mathbf{y})$

$$\alpha(\mathbf{x},\mathbf{y}) = \min\left\{1, \frac{\pi(\mathbf{y})q(\mathbf{y},\mathbf{x})}{\pi(\mathbf{x})q(\mathbf{x},\mathbf{y})}\right\} = \min\left\{1, \frac{\pi(\mathbf{y})q(\mathbf{x})}{\pi(\mathbf{x})q(\mathbf{y})}\right\}, \tag{13}$$

where $\pi = D_{\Lambda,\sigma,\mathbf{c}}$.

3) *Make a decision for* $\mathbf{X}_{t+1}$ *based on* $\alpha(\mathbf{x},\mathbf{y})$ *to accept* $\mathbf{X}_{t+1} = \mathbf{y}$ *or not.*

In principle, the Markov chain produced by the proposed algorithm is inherently reversible with respect to π, since

$$\begin{aligned}\pi(\mathbf{x})P(\mathbf{x},\mathbf{y}) &= \pi(\mathbf{x})q(\mathbf{x},\mathbf{y})\alpha(\mathbf{x},\mathbf{y}) \\ &= \min\{\pi(\mathbf{x})q(\mathbf{y}), \pi(\mathbf{y})q(\mathbf{x})\} \\ &= \pi(\mathbf{y})P(\mathbf{y},\mathbf{x}),\end{aligned} \tag{14}$$

where the assumption $\mathbf{y} \neq \mathbf{x}$ is sufficient because the above equation holds trivially in the case of $\mathbf{y} = \mathbf{x}$. Meanwhile, for $\pi = D_{\Lambda,\sigma,\mathbf{c}}$, it is also easy to verify that the underlying Markov chain is irreducible and aperiodic. Because ergodicity always holds for any Markov chain that are irreducible, aperiodic and reversible [15], we arrive at the following Lemma:

Lemma 1. *Given the invariant lattice Gaussian distribution* $D_{\Lambda,\sigma,\mathbf{c}}$*, the Markov chain induced by the independent MHK algorithm is ergodic:*

$$\lim_{t\to\infty} \|P^t(\mathbf{x};\cdot) - D_{\Lambda,\sigma,\mathbf{c}}(\cdot)\|_{TV} = 0 \tag{15}$$

for all states $\mathbf{x} \in \mathbb{Z}^n$.

IV. Convergence Analysis

In this section, we firstly demonstrate that the proposed independent MHK algorithm is uniformly ergodic. Then, the exponential decay coefficient δ is analyzed, leading to a quantitative estimate of the mixing time of the Markov chain.

Algorithm 2 Independent Metropolis-Hastings-Klein Algorithm for Lattice Gaussian Sampling

Input: $\mathbf{B}, \sigma, \mathbf{c}, \mathbf{X}_0$
Output: samples from the target distribution $\pi = D_{\Lambda,\sigma,\mathbf{c}}$

```
for t =1,2, ..., do
    let x denote the state of X_{t-1}
    generate y by the proposal distribution q(x, y) in (12)
    calculate the acceptance ratio α(x, y) in (13)
    generate a sample u from the uniform density U[0, 1]
    if u ≤ α(x, y) then
        let X_t = y
    else
        X_t = x
    end if
    if Markov chain has reached stationarity then
        output the state of X_t
    end if
end for
```

A. Uniform Ergodicity

Lemma 2. *In the independent MHK algorithm for lattice Gaussian sampling, there exists* $\delta > 0$ *such that*

$$\frac{q(\mathbf{x})}{\pi(\mathbf{x})} \geq \delta, \tag{16}$$

for $\mathbf{x} \in \mathbb{Z}^n$.

Proof. Using (3) and (4), we have

$$\begin{aligned}\frac{q(\mathbf{x})}{\pi(\mathbf{x})} &= \frac{\rho_{\sigma,\mathbf{c}}(\mathbf{Bx})}{\prod_{i=1}^{n}\rho_{\sigma_i,\widetilde{x}_i}(\mathbb{Z})} \cdot \frac{\rho_{\sigma,\mathbf{c}}(\mathbf{\Lambda})}{\rho_{\sigma,\mathbf{c}}(\mathbf{Bx})} \\ &= \frac{\rho_{\sigma,\mathbf{c}}(\mathbf{\Lambda})}{\prod_{i=1}^{n}\rho_{\sigma_i,\widetilde{x}_i}(\mathbb{Z})} \qquad (17) \\ &\overset{(a)}{\geq} \frac{\rho_{\sigma,\mathbf{c}}(\mathbf{\Lambda})}{\prod_{i=1}^{n}\rho_{\sigma_i}(\mathbb{Z})} \qquad (18)\end{aligned}$$

where (a) follows from the bound $\rho_{\sigma_i,\tilde{x}}(\mathbb{Z}) \leq \rho_{\sigma_i}(\mathbb{Z}) \triangleq \sum_{j\in\mathbb{Z}} e^{-\frac{1}{2\sigma_i^2}j^2}$ [6].

As can be seen clearly, the right-hand side (RHS) of (18) is completely independent of $\mathbf{x}$, meaning it can be expressed by a constant δ determined by basis $\mathbf{B}$, center $\mathbf{c}$ and standard deviation σ. Therefore, the proof is completed. □

We then arrive at the main Theorem to show the uniform ergodicity of the proposed algorithm.

Theorem 1. *Given the invariant lattice Gaussian distribution* $D_{\Lambda,\sigma,\mathbf{c}}$*, the Markov chain established by the independent MHK algorithm is uniformly ergodic:*

$$\|P^t(\mathbf{x},\cdot) - D_{\Lambda,\sigma,\mathbf{c}}(\cdot)\|_{TV} \leq (1-\delta)^t \tag{19}$$

for all $\mathbf{x} \in \mathbb{Z}^n$.

Proof. To start with, let us recall the *coupling technique* [17] shown below,

$$\|\mathcal{L}(\mathbf{X}) - \mathcal{L}(\mathbf{Y})\|_{TV} \leq P(\mathbf{X} \neq \mathbf{Y}), \tag{20}$$

where $\mathbf{X}$ and $\mathbf{Y}$ denote two random multivariables defined over the state space $\mathbb{Z}^n$ with probability distributions $\mathcal{L}(\mathbf{X})$ and $\mathcal{L}(\mathbf{Y})$ respectively.

According to (20), the variation distance $\|\cdot\|_{TV}$ between two random variables is upper bounded by the probability that they are unequal. Therefore, assume two Markov chain copies $\{\mathbf{X}_t\}$ and $\{\mathbf{X}'_t\}$ and each of them marginally follows the updating rules by $P(\mathbf{x},\cdot)$ and $\pi(\cdot)$ for all t, then we have

$$\|P^t(\mathbf{x},\cdot)-\pi(\cdot)\|_{TV} \le P(\mathbf{X}_t \neq \mathbf{X}'_t). \tag{21}$$

On the other hand, based on (12) and (13), the transition probability $P(\mathbf{x},\mathbf{y})$ of the independent MHK algorithm are given by

$$P(\mathbf{x},\mathbf{y})=\begin{cases}\min\left\{q(\mathbf{y}),\frac{\pi(\mathbf{y})q(\mathbf{x})}{\pi(\mathbf{x})}\right\} & \text{if } \mathbf{y}\neq\mathbf{x},\\ q(\mathbf{x})+\sum\limits_{\mathbf{z}\neq\mathbf{x}}\max\left\{0,q(\mathbf{z})-\frac{\pi(\mathbf{z})q(\mathbf{x})}{\pi(\mathbf{x})}\right\} & \text{if } \mathbf{y}=\mathbf{x}.\end{cases} \tag{22}$$

Using (16) in Lemma 2, it is straightforward to check that the following relationship holds

$$P(\mathbf{x},\mathbf{y}) \ge \delta\pi(\mathbf{y}) \tag{23}$$

for all cases of $\mathbf{x},\mathbf{y}\in\mathbb{Z}^n$, which indicates all the Markov transitions have a component of size δ in common. More specifically, from the perspective of coupling, it means every Markov move gives probability at least δ of making $\mathbf{X}$ and $\mathbf{X}'$ equal, that is,

$$P(\mathbf{X}=\mathbf{X}') \ge \delta. \tag{24}$$

Therefore, during t consecutive times Markov move, the probability of $\mathbf{X}$ and $\mathbf{X}'$ not equaling to each other can be derived as

$$P(\mathbf{X}_t \neq \mathbf{X}'_t) = (1-P(\mathbf{X}=\mathbf{X}'))^t \le (1-\delta)^t, \tag{25}$$

and according to (21), we obtain

$$\|P^t(\mathbf{x},\cdot)-\pi(\cdot)\|_{TV} \le (1-\delta)^t, \tag{26}$$

completing the proof. □

Obviously, given the value of δ, the mixing time of the Markov chain can be calculated by (8) and (26), that is

$$t_{\text{mix}}(\epsilon) = \frac{\ln\epsilon}{\ln(1-\delta)} < (-\ln\epsilon)\cdot\left(\frac{1}{\delta}\right), \quad \epsilon<1 \tag{27}$$

where we use the bound $\ln(1-\delta) < -\delta$ for $0<\delta<1$. Therefore, the mixing time is proportional to $1/\delta$, and becomes $O(1)$ if $\delta\to 1$.

B. Convergence Rate

Lemma 2 shows that the ratio $q(\mathbf{x})/\pi(\mathbf{x})$ in the independent MHK sampling algorithm is lower bounded by a constant δ, thereby permitting the proof of uniform ergodicity. We further derive an explicit expression of the exponential decay coefficient δ due to its significant impact on the convergence rate, for the special case $\mathbf{c}=\mathbf{0}$.

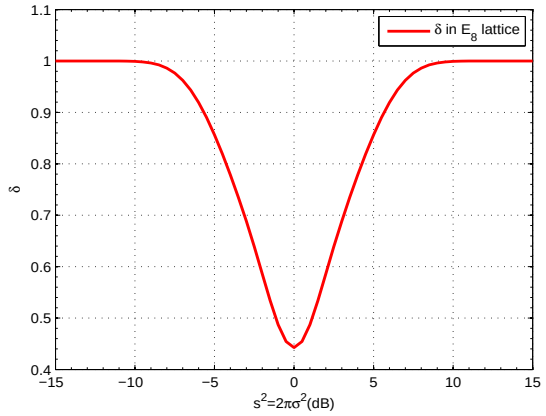

Fig. 1. Exponential decay coefficient δ of the E_8 lattice in the case of $\mathbf{c}=\mathbf{0}$.

Specifically, we have,

$$\begin{aligned}\frac{q(\mathbf{x})}{\pi(\mathbf{x})} &= \frac{\rho_{\sigma,\mathbf{0}}(\mathbf{\Lambda})}{\prod_{i=1}^n \rho_{\sigma_i,\widetilde{x}_i}(\mathbb{Z})}\\ &\overset{(a)}{\ge} \frac{\sum_{\mathbf{x}\in\mathbb{Z}^n} e^{-\frac{1}{2\sigma^2}\|\mathbf{B}\mathbf{x}\|^2}}{\prod_{i=1}^n \rho_{\sigma_i}(\mathbb{Z})}\\ &\overset{(b)}{=} \frac{\Theta_\Lambda(\frac{1}{2\pi\sigma^2})}{\prod_{i=1}^n \Theta_{\mathbb{Z}}(\frac{1}{2\pi\sigma_i^2})}\\ &\overset{(c)}{=} \frac{\Theta_\Lambda(\frac{1}{s^2})}{\prod_{i=1}^n \vartheta_3(\frac{1}{s_i^2})} = \delta.\end{aligned} \tag{28}$$

Here, for notational simplicity, $s=\sqrt{2\pi}\sigma$ and $s_i=\sqrt{2\pi}\sigma_i = s/\|\widehat{\mathbf{b}}_i\|$ are applied in the equations. In (a), the inequality $\rho_{\sigma_i,\tilde{x}}(\mathbb{Z}) \le \rho_{\sigma_i}(\mathbb{Z})$ is used again. Theta series Θ_Λ and Jacobi theta function ϑ_3 are applied in (b) and (c) respectively, where

$$\Theta_\Lambda(\tau) = \sum_{\lambda\in\Lambda} e^{-\pi\tau\|\lambda\|^2}, \tag{29}$$

$$\vartheta_3(\tau) = \sum_{n=-\infty}^{+\infty} e^{-\pi\tau n^2} \tag{30}$$

with $\Theta_{\mathbb{Z}}=\vartheta_3$ [18].

Now, we consider some lattices whose theta series are more understood.

Lemma 3. *The coefficient* $\delta = \frac{\Theta_\Lambda(\frac{1}{s^2})}{\prod_{i=1}^n \vartheta_3(\frac{1}{s_i^2})}$ *for an isodual lattice* Λ *has a multiplicative symmetry point at* $s=1$*, and asymptotically converges to 1 on both sides when* s *goes to* 0 *and* ∞*.*

Proof. According to the *Jacobi's formula* [19]

$$\Theta_\Lambda(\tau) = |\det(\mathbf{B})|^{-1}\left(\frac{1}{\tau}\right)^{\frac{n}{2}}\Theta_{\Lambda^*}\left(\frac{1}{\tau}\right), \tag{31}$$

where $\det(\cdot)$ denotes the determinant of a matrix and Λ^* is the dual lattice of Λ, we have

$$\Theta_\Lambda\left(\frac{1}{s^2}\right) = s^n\Theta_\Lambda(s^2), \tag{32}$$

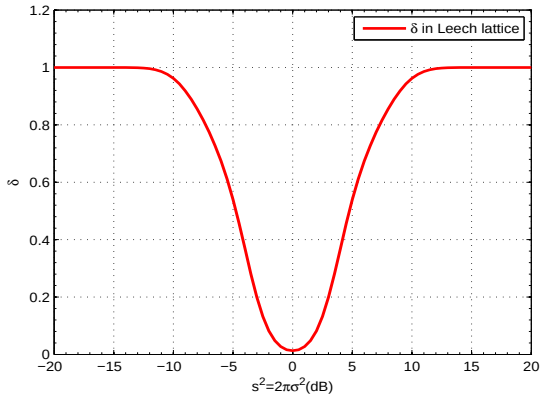

Fig. 2. Exponential decay coefficient δ of the Leech lattice in the case of $\mathbf{c} = \mathbf{0}$.

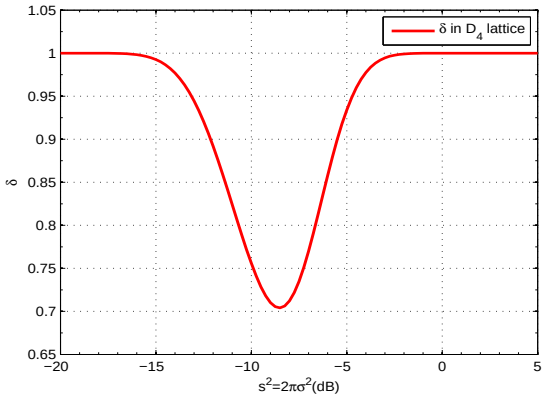

Fig. 3. Exponential decay coefficient δ of the D_4 lattice in the case of $\mathbf{c} = \mathbf{0}$.

$$\vartheta_3\left(\frac{1}{s_i^2}\right) \quad = \quad s_i\vartheta_3(s_i^2). \tag{33}$$

An isodual lattice is one that is geometrically similar to its dual. Here, we note that the theta series Θ_Λ of an isodual lattice Λ and that of its dual Λ^* are the same, i.e., $\Theta_\Lambda(\tau) = \Theta_{\Lambda^*}(\tau)$, and the volume of an isodual lattice $|\det(\mathbf{B})|$ naturally equals 1. Then from (32) and (33), the symmetry with respect to $s = 1$ can be obtained as follows,

$$\begin{aligned}
\frac{\Theta_\Lambda(\frac{1}{s^2})}{\prod_{i=1}^n \vartheta_3(\frac{1}{s_i^2})} &= \frac{s^n\Theta_\Lambda(s^2)}{\prod_{i=1}^n s_i\vartheta_3(s_i^2)} \\
&= \frac{\Theta_\Lambda(s^2)}{\prod_{i=1}^n \frac{1}{\|\widehat{\mathbf{b}}_i\|}\vartheta_3(s_i^2)} \\
&= \frac{\Theta_\Lambda(s^2)}{\frac{1}{|\det(\mathbf{B})|}\cdot\prod_{i=1}^n \vartheta_3(s_i^2)} \\
&= \frac{\Theta_\Lambda(s^2)}{\prod_{i=1}^n \vartheta_3(s_i^2)}.
\end{aligned} \tag{34}$$

By definition, it is straightforward to verify that

$$\frac{\Theta_\Lambda(\frac{1}{s^2})}{\prod_{i=1}^n \vartheta_3(\frac{1}{s_i^2})} \to 1, \text{ when } s \to 0. \tag{35}$$

Then because of the symmetry, $\frac{\Theta_\Lambda(\frac{1}{s^2})}{\prod_{i=1}^n \vartheta_3(\frac{1}{s_i^2})}$ will also asymptotically approach 1 when $s \to \infty$, completing the proof. □

Examples of the coefficient δ for the isodual E_8 and Leech lattice are shown in Fig. 1 and Fig. 2, respectively. It is worth pointing out that δ has a minimum at the symmetry point $s = 1$, namely $\sigma^2 = \frac{1}{2\pi}$. On the other hand, as for non-isodual lattices, D_4 lattice is applied to give the illustration, where the symmetry still holds but centers at $s = 0.376$. Therefore, with the exact value of δ, the explicit estimation of mixing time for the underlying Markov chain can be obtained.

ACKNOWLEDGMENT

This work was supported in part by FP7 project PHYLAWS (EU FP7-ICT 317562).

REFERENCES

[1] W. Banaszczyk, "New bounds in some transference theorems in the geometry of numbers," *Math. Ann.*, vol. 296, pp. 625–635, 1993.
[2] G. Forney and L.-F. Wei, "Multidimensional constellations–Part II: Voronoi constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 941–958, Aug 1989.
[3] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 39, pp. 913–929, May 1993.
[4] C. Ling and J.-C. Belfiore, "Achieiving the AWGN channel capacity with lattice Gaussian coding," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 5918–5929, Oct 2014.
[5] C. Ling, L. Luzzi, J.-C. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the Gaussian wiretap channel," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 6399–6416, Oct 2014.
[6] D. Micciancio and O. Regev, "Worst-case to average-case reuctions based on Gaussian measures," in *Proc. Ann. Symp. Found. Computer Science*, Rome, Italy, Oct. 2004, pp. 372–381.
[7] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *Proc. STOC*, 2009.
[8] S. Liu, C. Ling, and D. Stehlé, "Decoding by sampling: A randomized lattice algorithm for bounded distance decoding," *IEEE Trans. Inform. Theory*, vol. 57, pp. 5933–5945, Sep. 2011.
[9] Z. Wang, S. Liu, and C. Ling, "Decoding by sampling - part II: Derandomization and soft-output decoding," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4630–4639, Nov. 2013.
[10] P. Klein, "Finding the closest lattice vector when it is unusually close," in *ACM-SIAM Symp. Discr. Algorithms*, 2000, pp. 937–941.
[11] C. Gentry, C. Peikert, and V. Vaikuntanathan, "Trapdoors for hard lattices and new cryptographic constructions," in *Proc. 40th Ann. ACM Symp. Theory of Comput.*, Victoria, Canada, 2008, pp. 197–206.
[12] Z. Wang, C. Ling, and G. Hanrot, "Markov chain Monte Carlo algorithms for lattice Gaussian sampling," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Honolulu, USA, Jun. 2014, pp. 1489–1493.
[13] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.
[14] S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*. UK, Cambridge University Press, 2009.
[15] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Time*, American Mathematical Society, 2008.
[16] L. Tierney, "Markov chains for exploring posterior distributions," in *Proc. Computer Science and Statistics: 23rd Symp Interface.*, 1991.
[17] G. O. Roberts, "General state space Markov chains and MCMC algorithms," *Probability Surveys*, vol. 1, pp. 20–71, 2004.
[18] J. H. Conway and N. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer-Verlag, 1998.
[19] F. Oggier, P. Solé, and J. C. Belfiore, "Lattice codes for the wiretap Gaussian channel: Construction and analysis," *submitted to IEEE Trans. Inform. Theory,*, 2013. [Online]. Avaiable: http://arxiv.org/abs/1103.4086.