Sampling Algorithms for Lattice Gaussian Codes

Antonio Campello

based on joint work with J.-C. Belfiore (Huawei Technologies France)
Discrete Gaussian Measures

\[ f : \mathbb{R}^n \to \mathbb{R}^+ \]
\[ f(\mathbf{x}) \propto e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}} \]
\[ f(\mathbf{x}) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}} \]

\[ D : \Lambda \to [0, 1] \]
\[ \Lambda \text{ is a discrete set} \]
\[ D(\mathbf{x}) \propto e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}} \]
\[ D(\mathbf{x}) = \frac{e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}}}{\sum_{\mathbf{x} \in \Lambda} e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}}} \]
Discrete Gaussian Measures

- In Computer Science (lattice-based crypto): decoding algorithms [Klein '2000], homomorphic encryption, identity-based encryption [Regev '05], complexity reductions

- In Mathematics: discrete Fourier analysis, transference theorems ([Banaszczyk '92], [Cai '03]), theta series,…

- In Communications: non-uniform signaling [Kschischang and Pasupathy '93], semantically secure codes [Ling et al. '15], capacity achieving in the AWGN [Ling and Belfiore '15], compound and ergodic fading channels , [Campello, Ling and Belfiore '16]

- In Mechanical Statistics: Maxwell-Boltzmann distribution

- …
Discrete Gaussian Measures

- In Computer Science (lattice-based crypto): decoding algorithms [Klein ‘2000], homomorphic encryption [Regev ‘05], complexity reductions

- In Mathematics: discrete Fourier analysis, transference theorems ([Banaszczyk ‘92], [Cai ‘03]), theta series, …

- In Communications: non-uniform signaling [Kschischang and Pasupathy ‘93], semantically secure codes [Ling et al. ‘15], capacity achieving in the AWGN [Ling and Belfiore ‘15], compound and ergodic fading channels, [Campello, Ling and Belfiore ‘16]

- In Mechanical Statistics: Maxwell-Boltzmann distribution
A lattice is a discrete subgroup of $\mathbb{R}^n$.

**Sampling Algorithm**

Given a lattice $\Lambda$ and a parameter $\sigma > 0$, outputs a point $x \in \Lambda$ with probability

$$D_{\Lambda, \sigma}(x) = \frac{e^{-\frac{\|x\|^2}{2\sigma^2}}}{\sum_{x \in \Lambda} e^{-\frac{\|x\|^2}{2\sigma^2}}}$$

**Non-centered version:**

$$D_{\Lambda+c, \sigma}(x) = \frac{e^{-\frac{\|x+c\|^2}{2\sigma^2}}}{\sum_{x \in \Lambda} e^{-\frac{\|x+c\|^2}{2\sigma^2}}}$$
Lattice Coding and Crypto Meeting

Motivation: Simulating Probabilistic Shaping

- Lattice codes for the Gaussian channel:
  - Transmitter maps a « message » to a lattice point \( \mathbf{x} \in \Lambda \)
  - Receiver observes a distorted version \( \mathbf{y} \sim \mathcal{N}(0, \sigma_c^2) \)
    \[ \mathbf{y} = \mathbf{x} + \mathbf{z} \]
  and guesses \( \hat{\mathbf{x}} \) in order to minimize error probability \( P(\hat{\mathbf{x}} \neq \mathbf{x}) \)

- Messages are constrained (power-constraint)
  \[ \frac{1}{n} E \left[ \| \mathbf{x} \|^2 \right] \leq P \]
Motivation: Simulating Probabilistic Shaping

- Messages are constrained (power-constraint)
  \[
  \frac{1}{n} E \left[ \|x\|^2 \right] \leq P
  \]

- Deterministic Shaping: Choose a *shaping region* \( S \subset \mathbb{R}^n \) and a code \( S \cap \Lambda \) - e.g. cube, ball, or Voronoi region of sub-lattice

- Probabilistic Shaping: Pick \( x \sim D_{\Lambda, \sigma} \) (and adjust variance)

- [Forney ’89] Coding gain *versus* shaping gain

- How to sample the lattices with best *coding gain*? (known in low dimensions)
Lattice Gaussian Sampling Problem

- Hardness: In general, as hard as finding the shortest vector in a lattice [Aggarwal et al ’14] and [Stephens-Davidowitz ’15].

- Universal algorithms (the Metropolis-Hastings-Klein algorithm) perform slow over specific lattices. E.g.: 24-dim Leech lattice and \( \sigma = 1/\sqrt{2\pi} \) requires \( 24 \times 13434 = 322416 \) calls of an uni-dimensional sampler [Wang, Ling ’14].

- In Communications: sampling from special lattices (constructed from error correcting codes, having decomposition as union of cosets, etc…).

- Applications: towards Gaussian shaping, lattice decoding.

- Insights between lattice Gaussian codes and theta series.
Gaussian Measures: One Dimensional

\[ \sigma = 2.5 \]
Gaussian Measures: One Dimensional

\[ \sigma = 1.5 \]
Gaussian Measures: One Dimensional

\[ \sigma = 0.5 \]
Gaussian Measures: One Dimensional

\[ \sigma = 0.1 \]
Gaussian Measures: One Dimensional

\[ \sigma = 5 \times 10^3 \]
Wrong Idea: Generate $x \sim \mathcal{N}(0, \sigma^2)$ and output $\lfloor x \rfloor$
One Dimensional Sampler

Rejection Algorithm [Brakerski et al. ’13]

Set \( \mathcal{I} = \{c - l, c - (l - 1), \ldots, 1 - c, c, \ldots, c + l\} \) and calculate

\[
p_{\mathcal{I}} = D_{\sigma^2, \mathbb{Z} + c}(\mathcal{I})
\]

\[
p'(i) = D_{\sigma^2, \mathbb{Z} + c}(i)/p_{\mathcal{I}}, \, i \in \mathcal{I}
\]

1) With probability \( p_{\mathcal{I}} \) sample on the finite distribution in \( \mathcal{I} \)

2) With probability \( (1 - p_{\mathcal{I}}) \) sample on \( \mathcal{I}^c \) by a rejection principle:

Sampling on \( \mathcal{I}^c \):
Choose between positive or negative side. Ex:

(+) Generate \( y \) continuous Gaussian in \([c + l, +\infty]\)

Output \( x = [y - c] + c \) with prob.

\[
\frac{e^{-x^2/2\sigma^2}}{e^{-y^2/2\sigma^2}}
\]

Otherwise Repeat
Definition:

\[ \Theta_{\Lambda+c}(q) := \sum_{y \in \Lambda+c} q \|y\|^2 \]

\[ \Theta_{\Lambda+c}(\tau) := \sum_{y \in \Lambda+c} e^{-\pi \tau \|y\|^2} = \sum_{x \in \Lambda} e^{-\pi \tau \|x+c\|^2}. \]

Important easily numerically calculated one-dimensional theta series:

\[ \theta_2(\tau) := \sum_{m=-\infty}^{\infty} q^{(m+1/2)^2}, \theta_3(\tau) := \sum_{m=-\infty}^{\infty} q^{m^2}. \]

\[ \Theta_{\mathbb{Z}+c}(\tau) = \sum_{m=-\infty}^{\infty} e^{-\pi \tau (m+c)^2} = \tau^{-2} \sum_{m=-\infty}^{\infty} e^{2\pi i mc - \pi m^2 / \tau} = \tau^{-2} \theta_3(\pi c | i \tau^{-1}) \]
Lattices and Theta Series

Important properties:

\[
\Theta_{\Lambda_1 \oplus \Lambda_2}(\tau) = \Theta_{\Lambda_1}(\tau) \Theta_{\Lambda_2}(\tau)
\]

\[
\Theta_{\alpha \Lambda}(\tau) = \Theta_{\Lambda}(\alpha^2 \tau)
\]

\[
\Theta_{\Lambda_1 \cup \Lambda_2}(\tau) = \Theta_{\Lambda_1}(\tau) + \Theta_{\Lambda_2}(\tau)
\]

Example: Theta Series of \( \mathbb{Z}^n \)

\[
\Theta_{\mathbb{Z}^n}(\tau) = \Theta_{\mathbb{Z}}(\tau)^n = \theta_3(\tau)^n
\]
Hexagonal lattice

\[ A_2 = \left\{ (x_1, x_2) \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} : x_1, x_2 \in \mathbb{Z} \right\} \]

\[
\Theta_{A_2}(\tau) = \theta_3(\tau)\theta_3(3\tau) + \theta_2(\tau)\theta_2(3\tau)
\]
From Theta Series to Sampling

- Hexagonal lattice

\[ A_2 = (\mathbb{Z} \oplus \sqrt{3}\mathbb{Z}) \cup \left( \mathbb{Z} \oplus \sqrt{3}\mathbb{Z} + \left( \frac{1}{2}, \frac{\sqrt{3}}{2} \right) \right) \]

\[ \Theta_{A_2}(\tau) = \theta_3(\tau)\theta_3(3\tau) + \theta_2(\tau)\theta_2(3\tau) \]
From Theta Series to Sampling

- Hexagonal lattice

\[ p = D_{A_2, \sigma}(\mathbb{Z} \oplus \sqrt{3}\mathbb{Z}) = \frac{\theta_3\left(\frac{1}{2\pi\sigma^2}\right)\theta_3\left(\frac{3}{2\pi\sigma^2}\right)}{\theta_3\left(\frac{1}{2\pi\sigma^2}\right)\theta_3\left(\frac{3}{2\pi\sigma^2}\right) + \theta_2\left(\frac{1}{2\pi\sigma^2}\right)\theta_2\left(\frac{3}{2\pi\sigma^2}\right)} \]

**Algorithm**

1) Throw a biased coin with probability \( p \) of heads
2) If heads, sample in the blue coset
3) If tails, sample in the red coset

- Sampling in each coset is possible by invoking the \( \mathbb{Z} \)-sampler twice.
Coset Decompositions

- Generalization to more general coset decompositions.

Construction A lattices

\[ \Lambda = 2\mathbb{Z}^n + C \]

the coset corresponding of a codeword of weight \( w \) has theta series

\[ \theta_2(4\tau)^w \theta_3(4\tau)^{n-w} \]

Suppose there are \( A_w \) codewords of given weight \( w \). The probability that a discrete distribution falls in some coset of a codeword of weight \( w \) is

\[ A_w \frac{\theta_2(4\tau)^w \theta_3(4\tau)^{n-w}}{\Theta_\Lambda(\tau)} \]

General Idea

1) Pick a weight with probability \( p_w \)
2) Pick a word of weight \( w \) uniformity at random
3) Sample in the coset \( 2\mathbb{Z}^n + c \)
The lattice $D_n$

- Construction A lattices (best coding gains dimensions 3,4,5),

\[ D_n = 2\mathbb{Z}^n + P_n \]

where $P_n$ is a parity check code

\[ \{(x_1, \ldots, x_n) \in \mathbb{F}_2^n : x_1 + \ldots + x_n \equiv 0 \mod 2\} \]

There are $\binom{n}{2l}$ vectors of weight $2l$.

The probability of picking such a coset is

\[ p_{2l} = \binom{n}{2l} \frac{\Theta_{\mathbb{Z}+\frac{1}{2}}(4\tau)^{2l} \Theta_{\mathbb{Z}+\frac{1}{2}}(4\tau)^{n-2l}}{\Theta_{D_n}(\tau)} \]
The lattice $D_n$

Algorithm

1) Pick a number $l \in \{1, \ldots, [n/2]\}$ with probability $p_{2l}$.
2) Pick a subset $\mathcal{J} \subset \{1, \ldots, n\}$ with size $2l$
3) For $j \in \mathcal{J}$
   
   $x_j \leftarrow \text{Sampler}_{Z + \frac{1}{2}}(2\tau)$
4) For $j \notin \mathcal{J}$

   $x_j \leftarrow \text{Sampler}_Z(2\tau)$

- Generalizations to shifts by vectors of type $(\alpha, \beta, \beta, \ldots, \beta)$
Coset Decompositions

- Real Constructions (A and B)

\[ \Lambda_A(C) = 2\mathbb{Z}^n + C \quad \text{and} \quad \Lambda_B(C) = 4\mathbb{Z}^n + 2P_n + C. \]

\[ \Lambda_B(C) = 2D_n + C, \text{ where } D_n = \Lambda_A(P_n) \]

- Complex Constructions (A and B)

\[ \Lambda_A(C) = \theta \mathbb{Z}[\omega]^n + C \quad \text{and} \quad \Lambda_B(C) = \theta^2 \mathbb{Z}[\omega]^n + \theta P_n + C, \]

where \( \mathbb{Z}[\omega] = \{a + b\omega : a, b \in \mathbb{Z}\} \), and \( \theta \) is a prime of norm \( p \).
**Claim:** For the aforementioned constructions, the theta series of each coset depends only on the Hamming weight of each codeword.

<table>
<thead>
<tr>
<th>Construction</th>
<th>Theta Series of a Coset</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$\theta_2(4\tau)^w \theta_3(4\tau)^{n-w}$</td>
</tr>
</tbody>
</table>
| B            | $(1/2)\theta_2(4\tau)^w \theta_3(4\tau)^{n-w}$ for $w \geq 1$  
               | $(1/2)\theta_3(4\tau)^n + (1/2)\theta_4(4\tau)^n$ for $w = 0$ |
| $A_c, \theta = 2$ | $\phi_1(4\tau)^w \phi_0(4\tau)^{n-w}$ |
| $A_c, \theta = \sqrt{-3}$ | $\phi_2(3\tau)^w \phi_0(3\tau)^{n-w}$ |
| $B_c, \theta = \sqrt{-3}$ | $(1/3)\phi_2(3\tau)^w \phi_0(3\tau)^{n-w}$ for $w \geq 1$  
                          | $(1/3)(\phi_0(3\tau)^n + 2(\phi_0(9\tau) - \phi_2(9\tau))^n)$ for $w = 0$ |

**TABLE I**  
**Theta Series of a coset $\Lambda' + c$, $\text{wt}(c) = w$, for several constructions**
Extremal even unimodular lattice in dimension 24. Theta series:

\[
\frac{1}{8} \left( \theta_2(0, \tau)^8 + \theta_3(0, \tau)^8 + \theta_4(0, \tau)^8 \right)^3 - \frac{45}{16} \theta_2(0, \tau)^8 \theta_3(0, \tau)^8 \theta_4(0, \tau)^8
\]
Density Doubling

Consider the construction B lattice $H_{24} = 2D_{24} + G_{24}$, where $G_{24}$ is the $(24, 12, 8)_2$ Golay code. The Leech lattice is

$$\Lambda_{24} = H_{24} \cup (H_{24} + a)$$

where $a = ((-3/2)^1, (1/2)^{23})$

Theta series of « first » half is already known (Construction B)

For the second half:
All cosets $2D_{24} + c + a$ have same theta series given by

$$\frac{\beta(q^4)^{24} - \alpha(q^4)^{24}}{2}$$
Algorithm

1) Throw a biased coin with probability $D_{\Lambda_{24},\sigma}(H_{24})$ of heads.

2) if the output is heads
   Sample $x \in H_{24}$ from the Construction B sampler

3) else
   Choose $c \in G_{24}$ uniformly at random
   Draw $x \in 2D_{24} + a + c$ using Dn sampler

Output $x$

Properties of Golay code for Cons. B sampler
Uses 24 calls of a uni-dimensional sampler for any $\sigma$. 
Simulating Probabilistic Shaping

- [Ling and Belfiore ’15]. Gaussian Shaping.
  - Choose a « good » lattice for coding
  - Choose a point \( x \sim D_{\Lambda+c,\sigma^2} \) to be transmitted over a Gaussian channel.

**Proposition (closed form power/rate)**

The power and rate of a lattice Gaussian code is

\[
P = -\frac{1}{n\pi} \frac{\Theta'_{\Lambda+c}(\tau)}{\Theta_{\Lambda+c}(\tau)} \quad \text{and} \quad R = -\frac{\tau}{n} \frac{\Theta'_{\Lambda+c}(\tau)}{\Theta_{\Lambda+c}(\tau)} + \frac{1}{n} \ln \Theta_{\Lambda+c}(\tau)
\]

Rate is maximized in the center distribution

Relations to modular forms
Probabilistic Shaping

- Leech Lattice Sampler: discrete Gaussian versus cubic constellation
Sampling algebraic Construction A lattices (wireless channels):

E.g.: Ring of integers of $\mathbb{Q}(\sqrt{d}), d \geq 5, d \equiv 1 \pmod{4}$

Basis for ideal lattice:

$$
\begin{pmatrix}
1 & 1 \\
\frac{1+\sqrt{d}}{2} & \frac{1-\sqrt{d}}{2}
\end{pmatrix}
$$

Has rectangular sub-lattice of index 2 generated by embedding of $\mathbb{Z}[\sqrt{d}]$ and decomposes (up to rotation) as

$$
\Lambda = \left( \sqrt{2}\mathbb{Z} \oplus \sqrt{2d}\mathbb{Z} \right) \bigcup \left( \sqrt{2}\mathbb{Z} \oplus \sqrt{2d}\mathbb{Z} + \left( \frac{\sqrt{2}}{2}, \frac{\sqrt{2d}}{2} \right) \right)
$$
Final Remarks

- How to use symmetries between well-known lattices to deriving fast discrete sampling algorithms
- New insights between lattice Gaussian codes and theta series
- Open: sampling other algebraic lattices
- Probabilistic shaping models: from a sampler to an encoder
Thank you