Lattice Coding and its Applications in Communications

Alister Burr
University of York
alister.burr@york.ac.uk
Introduction to lattices
- Definition; Sphere packings; Basis vectors; Matrix description

Codes and lattice codes
- Shaping region; Nested lattices

Lattice constructions
- Construction A/D, LDLC codes; construction from Gaussian/Eisenstein integers

Lattice encoding and decoding
- Problems of shaping; LDLC decoding; Construction A decoding

Lattices in multi-user networks: Compute and forward
What is a lattice?

- A **lattice** is defined as:
  - the (infinite) set of points in an $n$-dimensional space given by all linear combinations with integer coefficients of a **basis** set of up to $n$ linearly independent vectors.

- It can be defined in terms of a **generator matrix** $G$, whose columns are the basis vectors:

\[
\Lambda = \{ \lambda = Gx : x \in \mathbb{Z}^n \} \]
A sphere packing is an arrangement of non-overlapping hyperspheres of equal radius in $N$-dimensional space.

We are often interested in the packing density $\eta$ or $\delta_n$ of a packing:
- the proportion of space occupied by spheres.

Dense sphere packings are often lattice packings:
- have sphere centres on lattices.
<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Lattice</th>
<th>Packing density</th>
<th>Kissing number</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Hexagonal</td>
<td>$\frac{1}{6} \pi \sqrt{3} = 0.91$</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>BCC/FCC/HCP</td>
<td>$\frac{1}{6} \pi \sqrt{2} = 0.74$</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>D4</td>
<td>$\frac{1}{16} \pi^2 = 0.62$</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>E8</td>
<td>$\frac{1}{384} \pi^4 = 0.25$</td>
<td>240</td>
</tr>
<tr>
<td>24</td>
<td>E24 (Leech)</td>
<td>$\frac{\pi^{12}}{12!} = 0.0019$</td>
<td>196 560</td>
</tr>
</tbody>
</table>
The **Voronoi region** of a lattice point is the region of the $N$-dimensional space closer to that point than to all other lattice points.

Voronoi region of red point shown shaded.
Introduction to lattices

Codes and lattice codes
  - Shaping region
  - Nested lattices

Lattice constructions

Lattice encoding and decoding

Lattices in multi-user networks: Compute and forward
i.e. forward error-correcting (FEC) codes

A code is a finite set of codewords of length $n$

- Code contains $M$ codewords – encodes $\log_2(M)$ bits

where a codeword is a sequence of $n$ symbols, usually drawn from a finite alphabet of size $q$

- we will often assume the alphabet is a Galois field ($\mathbb{F}_q$ or $\text{GF}(q)$) or a ring ($\mathcal{R}(q)$)

In a communication system the codewords must be translated into signals of length $nT$

- representing the variation in time of some quantity, such as electromagnetic field strength

Each code symbol is typically modulated to some specific real or complex value of this variable
Example

Message: 

Encode

Codeword: 

Modulate

Signal:

01111001

13212302

\[
s(t)
\]

\[
T
\]

\[
2T
\]

\[
3T
\]

\[
NT
\]
- Each coded signal can then be represented as a point in $N$-D **signal space**
  - where modulated values of symbols provide the $n$ coordinate values
- Code is represented by ensemble of points in signal space
- Noise on channel equivalent to vector $z$ in signal space
- Decoder chooses closest point
- Error probability determined by **minimum Euclidean distance** between signal space points
A **lattice code** is then defined by the (finite) set of lattice points within a certain region:

- the **shaping region**
- ideally a hypersphere centred on the origin
- this limits the maximum signal energy of the codewords

- Lattice may be offset by adding some vector
If the lattice is viewed as a sphere packing, then the minimum Euclidean distance must be twice the sphere radius.

- Signal power $S$ proportional to radius$^2$ of shaping region.
- The greater the packing density, the greater $M$ for given signal power.
- Radius$^2$ of packed spheres proportional to maximum noise power.
Hence for low error probability, noise power $N \leq r_S^2$

Radius of signal space at receiver containing signal plus noise is $\sqrt{S + N}$

Volume of $n$-D sphere of radius $r$ is $V_n r^n$

Hence max. no. of codewords in code

$$M \leq \frac{V_n (S + N)^{n/2}}{V_N r_S^{N/2}} \leq \left( \frac{S + N}{N} \right)^{n/2}$$

$$\frac{\log_2 M}{n} \leq \frac{1}{2} \log_2 \left( 1 + \frac{S}{N} \right)$$
Define fine lattice $\Lambda_C$ for the code
- plus a **coarse lattice** $\Lambda_S$ which is a sub-lattice of $\Lambda_C$
Then use a Voronoi region $V_S$ of the coarse lattice as the shaping region
- Modulo-$\Lambda_S$ operation
  - for any point $P \not\in V_S$ find $P$
  - $(\lambda \in \Lambda_S) \in V_S$
Wireless signals consist of a sine wave carrier at the transmission frequency (MHz – GHz)

Sine waves can be modulated in both amplitude and phase

- hence the signal corresponding to each modulated symbol is 2-D
- also conveniently represented as a complex value
- typically represented on a phasor diagram

Hence wireless signals can be represented in 2n dimensions

- or n complex dimensions
Introduction to lattices

Codes and lattice codes

Lattice constructions
  - Constructions A and D,
  - LDLC codes
  - Construction from Gaussian and Eisenstein integers

Lattice encoding and decoding

Lattices in multi-user networks: Compute and forward
For practical purposes in communications, we require lattices in very large numbers of dimensions

- typically 1000, 10 000, 100 000…

Lattices of this sort of dimension most easily constructed using FEC codes such as LDPC and turbocodes

Most common constructions encountered are called Constructions A and D (Conway and Sloane)

- Construction A based on a single code
- Construction D is multilevel, based on a nested sequence of codes
Construction A

- Start with a $q$-ary linear code $C$ with generator matrix $G_C$
- The set of vectors $\lambda$ such that $\lambda \mod q$ is a codeword of $C$ form a Construction A lattice from $C$:

\[ \Lambda = \{ \lambda : \lambda \mod q \in C \} \]

- Alternatively we can write:

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

- The generator matrix of the lattice:

\[
G = \begin{bmatrix}
G_C & 0 \\
qI_{n-k} & 0
\end{bmatrix}
\]

- Note that minimum distance is limited by $q$
Let \( C_0 \subseteq C_1 \subseteq C_2 \ldots \subseteq C_a \) be a family of linear binary codes

- where \( C_0 \) is the \((n, n)\) code and \( C_\ell \) is an \((n, k_\ell)\) code

Then the lattice is defined by:

\[
\Lambda = \left\{ \lambda : \lambda = z + \sum_{l=1}^{a} \sum_{j=1}^{k_l} d_{j,l} \frac{c_{j,l}}{2^{l-1}} \right\}
\]

where \( z \in 2^n \mathbb{Z}^n \), \( c_{j,l} \) is the \( j^{th} \) basis codeword of \( C_\ell \), and \( d_{j,l} \in \{0,1\} \) denotes the \( j^{th} \) data bit for the \( \ell^{th} \) code.

![Diagram of a circuit with multiple inputs and a summation output.](image)
Low density lattice codes

- Uses the principle of LDPC codes:
  - Define generator matrix such that its inverse $H = G^{-1}$ is sparse
  - Then decode using sum-product algorithm (message passing) as in LDPC decoder
- However elements of $H$ and $G$ are reals (or complex) rather than binary
  - Messages are no longer simple log-likelihood ratios
- Ideally use nested lattice code
  - i.e. shaping region is Voronoi region of a coarse lattice
Gaussian and Eisenstein integers

- Construction A/D and LDLC result in real lattices
  - can exploit Gaussian/Eisenstein integers to construct complex lattices
- Gaussian and Eisenstein integers form the algebraic equivalent in complex domain of the ring of integers
- Can construct complex constellations from them which form complex lattices
- Gaussian integers are the set of complex numbers with integer real and imaginary parts, denoted \( \mathbb{Z}[i] = a + bi, \ a, b \in \mathbb{Z} \).

- They form a ring on ordinary complex arithmetic.

- Hence operations in the ring exactly mirror operations in signal space.

- Also form a lattice.
Consider *fine* and *coarse* lattices, $\Lambda_f$ and $\Lambda_c$, both based on Gaussian integers

$$\Lambda_c \subset \Lambda_f$$

Here we assume that each point in the coarse lattice is a point in the fine multiplied by some Gaussian integer $q$

- i.e. the coarse is a scaled and rotated version of the fine
- and the fine is just the Gaussian integers

We then define our constellation as consisting of those Gaussian integers which fall in the Voronoi region of the coarse lattice
- e.g. \( q = 2 + i \)
- Blue points are fine lattice
- Red points are coarse lattice
- Fundamental region is region closer to origin than any other coarse lattice point
- Hence constellation is green points, inc origin

\[ V_c(0) \]
The fundamental region is surrounded by regions corresponding to \( q, qi, -q \) and \(-qi\).

We treat the boundaries of the latter two as belonging to the fundamental region.

- Use this to allocate certain boundary points to constellation.

This also leads to an alternative definition of the fundamental region:

\[
V_c(0) = \left\{ \lambda \in \mathbb{C} : -\frac{|q|^2}{2} \leq \Re[\lambda]\Re[q] + \Im[\lambda]\Im[q] < \frac{|q|^2}{2} \right\}
\]

\[
\& \leq -\Re[\lambda]\Im[q] + \Im[\lambda]\Re[q] < \frac{|q|^2}{2}
\]
We can establish *isomorphisms* between these constellations and either fields or rings.

An isomorphism is a one-to-one (or *bijective*, and hence invertible) mapping between the constellation \( \mathcal{C} \) and the ring \( \mathcal{R} \):

\[
\lambda = \mathcal{M}(s), \quad \lambda \in \mathcal{C}, \quad s \in \mathcal{R} \\
\quad s = \mathcal{M}^{-1}(\lambda), \quad \lambda \in \mathcal{C}, \quad s \in \mathcal{R}
\]

such that the operations on the ring are equivalent to those on the constellation:

\[
\mathcal{M}(s_1 \otimes s_2) = \mathcal{M}(s_1) \mathcal{M}(s_2) \quad \mathcal{M}(s_1 \oplus s_2) = \mathcal{M}(s_1) + \mathcal{M}(s_2)
\]

It turns out that if \( q \) is a *Gaussian prime*, then the constellation is isomorphic to a field, otherwise it is isomorphic to a ring.

Size of field/ring is \(|q|^2\).
Lattice construction

- This isomorphism can be used to construct a complex lattice from a code based on the field or ring
  - in a manner equivalent to Construction A
    \[ \Lambda = \left\{ \lambda : \lambda = z + M(c), z \in q \mathbb{Z} [i]^n, c \in C \left( \frac{\mathbb{F}}{q} \right) \right\} \]
  - that is, we encode a data sequence in the field \( \mathbb{F}_q \) using the code \( C \) (over \( \frac{\mathbb{F}}{|q|^2} \))
  - then map the resulting symbols to the complex constellation using the mapping based on the isomorphism
  - then combine with a lattice of Gaussian integers scaled by \( q \)
- Eisenstein integers
  - Set of complex values with similar properties to Gaussian integers
  - Hexagonal structure may result in denser lattices
  - Note:
    \[ \omega = \frac{1 + i \sqrt{3}}{2} = e^{2\pi i/3} \]
Outline

- Introduction to lattices
- Codes and lattice codes
- Lattice constructions
- Lattice encoding and decoding
  - Problems of shaping
  - Construction A/D decoding
  - LDLC decoding
- Lattices in multi-user networks: Compute and forward
Ideally the shaping region should be as close as possible to a hypersphere

- provides **shaping gain** up to 1.5 dB compared to hypercube shaping

Nested lattice shaping gives a good approximation to this

First multiply data vector by generator matrix

- this may generate region of lattice of arbitrary shape

Then apply modulo-lattice operation:

- decode to coarse lattice, and subtract resulting coarse lattice vector

In practice this decoding operation may be difficult

- may use hypercube shaping as simpler alternative
Construction A decoding

- Generally can be carried out with decoder for underlying code $C$
- Applying $\text{mod}_q$ operation regenerates codeword of $C$
  - then decode this codeword
  - can then recover specific point in $\mathbb{Z}^n$
- Note that in practice we use non-binary codes ($q > 2$)
  - because $q = 2$ limits minimum distance and hence coding gain
- Typically use LDPC or turbocodes to achieve good performance
  - hence need non-binary sum-product or BCJR decoder
  - messages are probability distribution of $q$ symbol values
Use multilevel decoding approach based on component codes
- decode codes $C_a$, $C_{a-1}$, … $C_1$ in succession

Component codes may usually be binary

May require iterative approach
- c.f. multilevel coded modulation
LDLC decoding

- Code structure designed for sum-product decoding, cf LDPC
  - using *factor graph*

- However symbol values are now continuous variables (reals)
  - hence messages should be probability density functions
  - requires compact means of representing PDF in decoder

- May use Fourier or Karhunen-Loeve basis representation
  - or Gaussian mixture model
- Introduction to lattices
- Codes and lattice codes
- Lattice constructions
- Lattice encoding and decoding
- Lattices in multi-user networks
  - Wireless physical-layer network coding
  - Compute and forward
Physical layer in multi-user networks

- Traditional role of PHY:
  - signals from elsewhere in network treated as harmful interference
  - however they may carry related information that can be exploited
Two-way relay channel

- Two terminals want to exchange data via a relay:

  ![Diagram showing two terminals A and B communicating through a relay R]

- Conventionally this would require 4 time-slots:

  ![Time-axis diagram with slots for SA, SB, R(A), R(B)]
We can do better using Wireless Physical-layer Network Coding using two phases

Assume both sources transmit BPSK:
- map data symbol ‘1’ to signal +1; ‘0’ to -1
- At relay, map signals +2 and -2 to ‘0’; 0 to ‘1’

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>a+b</th>
<th>a ⊕ b</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>+2</td>
<td>0</td>
</tr>
</tbody>
</table>
A general network model

- Model a network with $P$ layers of relays
- In general all nodes in a layer transmit simultaneously
- Each relay decodes a (linear) function of symbols from previous layer
  
  $$s_l^{(p)} = a_{1l}s_1^{(p-1)} + a_{2l}s_2^{(p-1)} + \cdots + a_{Ll}s_L^{(p-1)} = \mathbf{a}_l^T \mathbf{s}^{(p-1)}$$

- based on the combined signals they receive
- Destination extracts symbol it is interested in from outputs of functions
- Lattices provide useful signal sets
Network coding model of network

- We can relate the vector of outputs of each layer to its inputs via the matrix $A$:
  $$ s^{(p)} = A^{(p)} s^{(p-1)} $$

- We can combine these in cascade, so that:
  $$ s^{(p)} = A^{(p)} A^{(p-1)} \cdots A^{(1)} s $$

- We can write this as a single matrix relating the vector of symbol $s^D$ at relays connected to the destination:
  $$ s^D = B s $$

- We assume that the destination can (in principle) decode all symbols in its connection set
Consider relay receiving from two sources via channel $h_A, h_B$

Sources transmit codewords $c_A, c_B$ from the same fine lattice $\Lambda_C$

Received signal at relay is then:

$$x = h_A c_A + h_B c_B + w$$

Now the sum of any integer multiples of two lattice points is another lattice point

- hence if $h_A, h_B$ were integers we could decode at the relay using the same lattice decoder

Key idea is to scale received signal by scaling factor $\alpha$ so that $\alpha h_A$ and $\alpha h_B$ are approximately integers
Then:

\[ \alpha x = \alpha h_A c_A + \alpha h_B c_B + \alpha \omega \approx a_A c_A + a_B c_B \]

- where \( a_A \) and \( a_B \) are integers
- Approximation error is:

\[ (\alpha h_A - a_A) c_A + (\alpha h_B - a_B) c_B + \alpha \omega \]

- We can minimise this by choosing \( \alpha \):

\[ \alpha_{\text{MMSE}} = \frac{P \sum_i h_i a_i}{N + P \sum_i |h_i|^2} \]

- where \( P \) is signal power
- Also need to choose \( a_A \) and \( a_B \)
  - could choose such that \( a_A / a_B = h_A / h_B \)
  - but might require large \( \alpha \), and hence increase noise
- $h_A = 0.55; h_B = 1.0$
- Choose: $a_A = 1; a_B = 2; \alpha = 1.95$
- Blue points are received signal
- Red are approximated lattice
- Sum of two points from a lattice code may in general result in point outside shaping region
- Hence we apply modulo-lattice operation
  - returns a point in the original lattice code
  - so we can use the same decoder to recover sum point
- For lattice constellations isomorphic with field this operation can always be inverted
Conclusions

- Lattices can be extensively used in communications especially for \textit{lattice coding}.
- Can be shown to achieve capacity, as lattice dimension tends to infinity.
- Practical lattice constructions are based on FEC codes that can provide high dimension lattices with practical decoding algorithms.
- For wireless channels use complex lattice constellations based on Gaussian/Eisenstein integers.
- Important application is \textit{compute and forward} that applies to relay networks.
Lattice quantisation:
- Quantising signals to lattice points in high dimension can reduce mean square error.
- Applying modulo-lattice operation also allows Wyner-Ziv compression of correlated sources.

Lattice reduction aided MIMO detection
- MIMO channel may distort received signal:
  - LRA treats as a different lattice.

- Uri Erez, Shlomo Shamai (Shitz), and Ram Zamir “Achieving 1/2 \( \log(1 + \text{SNR}) \) on the AWGN channel with lattice encoding and decoding”, *IEEE Trans. Inf. Theory*, 50(10):2293–2314, October 2004.

