Achieving Channel Capacity With Lattice Codes: From Fermat to Shannon

Cong Ling
Imperial College London
cling@ieee.org

May 2, 2016

1. Statement of Coding Problems in Information Theory
2. Lattices and Algebraic Number Theory
3. Coding for Gaussian, Fading and MIMO channels

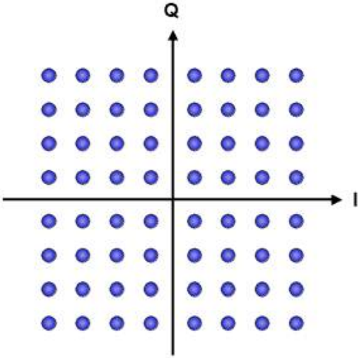
Communications in the presence of noise

- Signal vector $\mathbf{x} = [x_1, \ldots, x_n]^T$ in n-dimensional Euclidean space.
- The additive white Gaussian noise (AWGN) channel: $\mathbf{y} = \mathbf{x} + \mathbf{w}$, where signal power $P = E[\|\mathbf{x}\|^2]/n$ and noise power $= \sigma_w^2$.
- Shannon capacity (1949)

$$C = \frac{1}{2}\log(1+\rho)$$

where signal-to-noise ratio (SNR) $\rho = \frac{P}{\sigma_w^2}$.

- Shannon used random coding, but we need a concrete code to achieve the capacity.



quadrature amplitude modulation (QAM) constellation

Coding

Algebraic approach

- Hamming code
- Reed-Solomon code
- BCH code
- Algebraic-geometry code

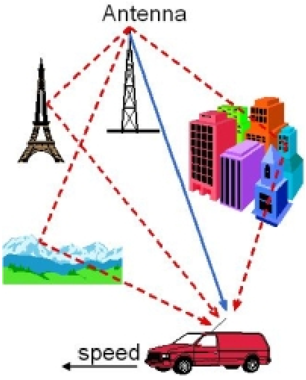
Probabilistic approach

- Convolutional code
- Turbo code
- LDPC code
- Polar code

For binary (discrete)-input channels, dream has come true with polar codes [Arikan'09] and SC-LDPC codes [Jimenez-Felstrom-Zigangirov'99].

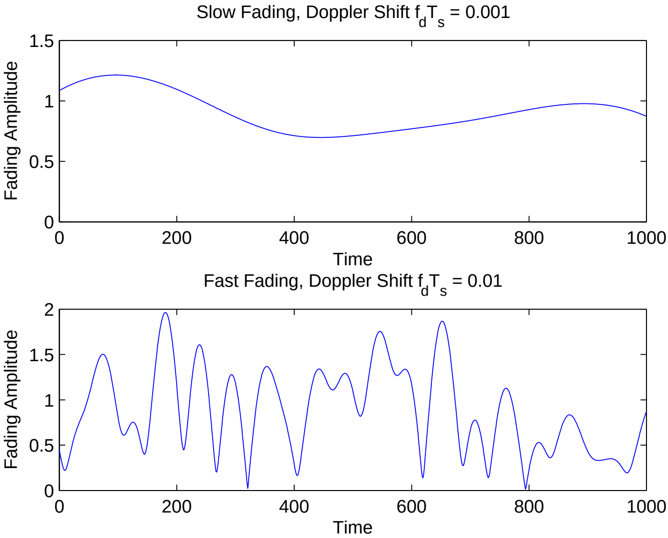
However, the question of achieving the capacity of the Gaussian channel has to be solved with lattice codes.

Multipath fading in mobile communications



- Multipath propagation in urban environment.
- Fading is multiplicative noise (large variation in signal strength)
$$y_t = h_t x_t + w_t$$
- Rayleigh fading: channel coefficient h_t is complex Gaussian.
- Time autocorrelation is modelled by a Bessel function (Jakes model) $R(\tau) = \mathrm{E}[h_t h_{t+\tau}^*] = J_0(2\pi f_d \tau)$ where $f_d = (v/c)f$ is normalized Doppler frequency shift.

Slow fading vs. fast fading



Models

- Slow fading (block fading): The fading process is nearly constant (but random) in the duration of a codeword. We need (time, frequency etc.) diversity from several independent blocks:

$$(\underbrace{h_1, h_1, \ldots, h_1}, \underbrace{h_2, h_2, \ldots, h_2}, \cdots, \underbrace{h_n, h_n, \ldots, h_n})$$

  - The length of each block is known as coherence time T.
  - Ergodicity doesn't hold due to delay constraint.
  - Capacity $C = \sum_{i=1}^{n} \log\left(1 + |h_i|^2 \rho\right)$.

Models

- Slow fading (block fading): The fading process is nearly constant (but random) in the duration of a codeword. We need (time, frequency etc.) diversity from several independent blocks:

$$(\underbrace{h_1, h_1, \dots, h_1}, \underbrace{h_2, h_2, \dots, h_2}, \cdots, \underbrace{h_n, h_n, \dots, h_n})$$

  - The length of each block is known as coherence time T.
  - Ergodicity doesn't hold due to delay constraint.
  - Capacity $C = \sum_{i=1}^{n} \log\left(1 + |h_i|^2 \rho\right)$.
- Fast fading: The fading coefficients $\{h_t\}$ are nearly independent.
  - In reality, ergodic fading is a more accurate model.
  - Capacity $C = \mathrm{E}_H \left[\log\left(1 + |h|^2 \rho\right)\right]$.

Models

- Slow fading (block fading): The fading process is nearly constant (but random) in the duration of a codeword. We need (time, frequency etc.) diversity from several independent blocks:

$$(\underbrace{h_1, h_1, \dots, h_1}, \underbrace{h_2, h_2, \dots, h_2}, \cdots, \underbrace{h_n, h_n, \dots, h_n})$$

  - The length of each block is known as coherence time T.
  - Ergodicity doesn't hold due to delay constraint.
  - Capacity $C = \sum_{i=1}^{n} \log\left(1 + |h_i|^2 \rho\right)$.
- Fast fading: The fading coefficients $\{h_t\}$ are nearly independent.
  - In reality, ergodic fading is a more accurate model.
  - Capacity $C = \mathrm{E}_H\left[\log\left(1 + |h|^2 \rho\right)\right]$.
- These represent two extremes of stationary fading.

Models

- Slow fading (block fading): The fading process is nearly constant (but random) in the duration of a codeword. We need (time, frequency etc.) diversity from several independent blocks:

$$(\underbrace{h_1, h_1, \ldots, h_1}, \underbrace{h_2, h_2, \ldots, h_2}, \cdots, \underbrace{h_n, h_n, \ldots, h_n})$$

  - The length of each block is known as coherence time T.
  - Ergodicity doesn't hold due to delay constraint.
  - Capacity $C = \sum_{i=1}^{n} \log\left(1 + |h_i|^2 \rho\right)$.
- Fast fading: The fading coefficients $\{h_t\}$ are nearly independent.
  - In reality, ergodic fading is a more accurate model.
  - Capacity $C = \mathrm{E}_H\left[\log\left(1 + |h|^2 \rho\right)\right]$.
- These represent two extremes of stationary fading.
- Open question: to design capacity-achieving codes over fading channels (wireless systems will operate close to capacity).

Block fading channel

- Slow fading is the realistic model in delay-constrained communication systems (4G, 5G...).
- Write down the matrix form of the channel $\mathbf{Y} = \mathbf{HX} + \mathbf{W}$, where channel matrix $\mathbf{H} = \mathrm{diag}[h_1, h_2, \ldots, h_n]$.
- Set target capacity

$$C = \log \left| \mathbf{I} + \rho \mathbf{H}^{\dagger} \mathbf{H} \right| . \tag{1}$$

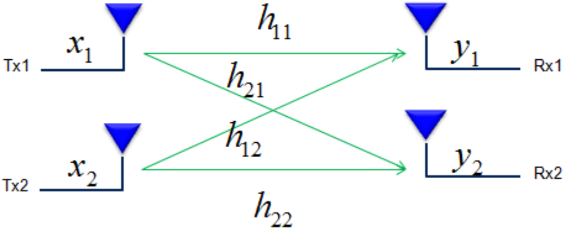
- The receiver has channel state information (CSI), while the transmitter doesn't.
- Our goal is to achieve capacity C on all channels such that (1) is true (without even knowing the distribution of $\mathbf{H}$).
- This requires a universal code on the compound channel (1), i.e., a collection of channels with the same capacity.

MIMO channel

- Capacity $\propto n$, the number of antennas.
- Channel model $\mathbf{Y} = \mathbf{HX} + \mathbf{W}$, where $\mathbf{H}$ is the $n \times n$ channel matrix, fixed (but random) in coherence time T.
- Set target capacity

$$C = \log \left| \mathbf{I} + \rho \mathbf{H}^{\dagger} \mathbf{H} \right| . \tag{2}$$

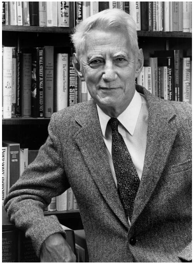
- Open question: achieve the capacity of the compound MIMO channel (2).



What are lattices?

- Lattices are regular, efficient and near-optimum arrays in the Euclidean space.
- The hexagonal lattice A_2 and FCC lattice A_3 give the densest sphere packings in dimensions 2 and 3.
- Breaking news [Viazovska (et al.)'16]: The E_8 lattice and Leech lattice are optimum for sphere packing for $n = 8$ and 24.
- Recent years see the revival of this classic area, driven by new applications, particularly coding and cryptography.

Why are lattices useful?



- Minkowski founded geometric number theory, where lattices are used to solve problems in number theory.
- Coding for the Gaussian channel is closely related to the sphere packing problem, for which lattices are near-optimum.
  - Shannon already indicated dense sphere packings to achieve channel capacity (without knowing lattices).
- Cryptographers are more interested in the hardness of lattice problems.

History of lattice coding and cryptography

- 1960s: earliest use of lattices in coding.
- 1984: lattice formulation of trellis-coded modulation.
- 1988: first book on lattice coding *Sphere Packings, Lattices and Groups*.
- 1992: ideal lattices for Rayleigh fading channels.
- 2004: capacity-achieving lattice codes for Gaussian channels.
- 2005: lattice-based space-time codes for MIMO channels.
- 1982: first use of lattices in cryptanalysis.
- 1996: first crypto scheme based on hard lattice problems.
- 2002: first book on lattice crypto *Complexity of Lattice Problems* published.
- 2005: learning with errors.
- 2006: application of ideal lattices to crypto.
- 2009: fully homomorphic encryption.
- 2012: multilinear maps.

Definition

- A lattice $\Lambda \subset \mathbb{R}^n$ is defined as

$$\Lambda = \Lambda(\mathbf{B}) = \{\mathbf{Bx} | \mathbf{x} \in \mathbb{Z}^n\}$$

  where $\mathbf{B}$ (n-by-n) is called a basis, or generator matrix. For example, the lattice $\mathbb{Z}^2$ (aka QAM) has basis $\mathbf{B} = \mathbf{I}_2$.
- May be viewed as the result of a linear transformation applied to the $\mathbb{Z}^n$ lattice (cubic lattice).
- Euclidean counterpart of a linear code (a linear code is normally defined in the Hamming space).
- The dual lattice Λ^* has basis $(\mathbf{B}^{-1})^T$. It arises in the Fourier transform of multi-dimensional signals.

Fundamental regions

- A fundamental region of a lattice is a piece that tiles the Euclidean space without any overlap or gap
- The volume of a fundamental region is called the fundamental volume

$$V(\Lambda) = \det(\Lambda) = |\det(\mathbf{B})|$$

- Fundamental parallelotope

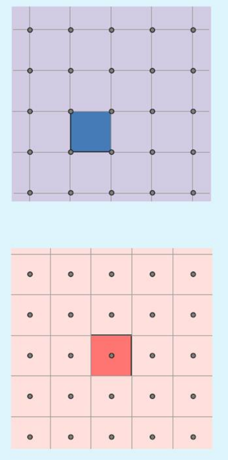
$$\mathcal{P} = \{\mathbf{y} \mid \mathbf{y} = \mathbf{B}\mathbf{x}, 0 \leq x_i < 1\}$$

- Voronoi region: the nearest-neighbor decoding region

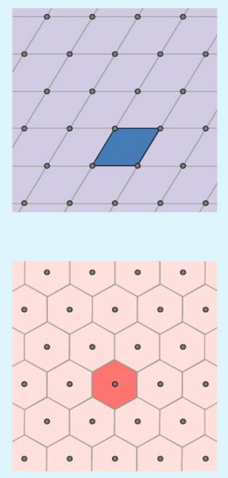
$$\mathcal{V} = \{\mathbf{y} \mid \|\mathbf{y}\| < \|\mathbf{y} - \mathbf{x}\|, \forall \mathbf{x} \in \Lambda\}$$

Examples of fundamental parallelotope and Voronoi region

Square lattice $\mathbb{Z}^2$



Hexagonal lattice A_2



An ensemble of random lattices (Loeliger ensemble)

- Consider the following family of q-ary lattices for all matrices $\mathbf{A} \in \mathbb{Z}_q^{k \times n}$,

$$\Lambda_q(\mathbf{A}) \quad = \quad \{\mathbf{y} \in \mathbb{Z}^n : \mathbf{y} = \mathbf{A}^T\mathbf{s} \mod q \text{ for } \mathbf{s} \in \mathbb{Z}^k\}$$

- These are lattices from Construction A, given the generator matrix $\mathbf{A}$ of the code.
- For a lattice vector $\mathbf{x}$,

$$\mathbf{x} \mod q = \mathbf{c} \in \mathcal{C}$$

  for a linear (n, k) code $\mathcal{C} \subset \mathbb{Z}_q^n$.
- Construction A is an important bridge between lattice theory and coding theory.

Minkowski-Hlawka Theorem

- Loeliger ensemble is discrete version of classic Minkowski-Hlawka-Siegel emsemble:

$$\mathcal{L} = \{\Lambda(\mathbf{B}) : \mathbf{B} \in \mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})\}$$

- There exist dense lattices in Loeliger ensemble as $n, q \to \infty$ [Lolieger'97]

$$\lambda_1(\Lambda) \approx \sqrt{\frac{n}{2\pi e}} V^{1/n}(\Lambda)$$

  - Its proof is based on Shannon's random coding.
  - Rogers' proof of Minkowski-Hlawka Theorem is a special case with $k = 1$.
- Good lattices can be generated from random codes.
  - Good for coding.
  - Good for quantization.
  - Good for secrecy.
  - ...

Variations

- Similarly, Construction A may be defined with the parity-check matrix:

$$\Lambda_q^{\perp}(\mathbf{A}) \quad = \quad \{\mathbf{y} \in \mathbb{Z}^n : \mathbf{A}\mathbf{y} = \mathbf{0} \mod q\}$$

- Some special lattices/generalizations
  - Cyclic codes $\Rightarrow$ cyclic lattices
  - Negacyclic codes $\Rightarrow$ negacyclic lattices
  - Double-circulant codes: aka NTRU in crypto; quasi-cyclic codes
  - Modulo a multi-dimensional lattice (D_4, E_8), ideal $\mathfrak{q}$ of a number field...

Quest for structured lattices

Construction A

Good news: dense lattices (for coding [Loeliger'97]); hard to decode (for crypto [Regev'05]).
Bad news: hard to decode (for coding); inefficient (for both)[a].

[a] Being efficient means quasi-linear complexity; n is several hundreds to thousands in practice.

- In coding: (the study of transmitting information)
  - Gaussian channels:
    - Classic approach: dense lattices.
    - Practical approach: Constructions A, D etc. from good codes.
  - Fading channels: ideal lattices from algebraic number fields.
  - MIMO channels: division algebras over number fields.
- In crypto: (the study of hiding information)
  - Cyclic lattices, ideal lattices.
  - More efficient than general lattices.
  - New functionalities: homomorphic encryption, code obfuscation...

Algebraic number theory

- Fermat's Last Theorem (1637): When $n > 2$,

$$x^n + y^n = z^n$$

  has no nontrivial solutions $x, y, z \in \mathbb{Z}$.
- It was in the Guinness Book of World Records for "most difficult mathematical problems".
- Historically gave rise to algebraic number theory:

$$(x^p + y^p) = \prod_{i=0}^{p-1} (x + \zeta_p y)$$

- Kummer proved the theorem for all regular primes ($p \nmid h_p$ of a cyclotomic number field).
- Finally settled by Andrew Wiles in 1994, 3.5 centuries later.

Number fields

- A number field K is a finite field extension of $\mathbb{Q}$, i.e., a field which is a $\mathbb{Q}$-vector space of finite dimensions. The dimension $[K:\mathbb{Q}]$ is called the degree of K.
- Any number field can be built by adding a primitive element θ to $\mathbb{Q}$, i.e., $K=\mathbb{Q}(\theta)$ (in fact, θ is an algebraic integer).
- An algebraic integer in a number field K is an element $\alpha \in K$ which is a root of a monic irreducible polynomial with integer coefficients. Such a polynomial is called the minimum polynomial of α.
- Example: $\mathbb{Q}(\sqrt{2}) = \{x + y\sqrt{2} | x, y \in \mathbb{Q}\}$ is a number field of degree 2, i.e., a quadratic field.
- Example: If ζ_m is a primitive mth root of unity, the number field $\mathbb{Q}(\zeta_m)$ is called a cyclotomic field.

Ring of integers

- The ring of integers $\mathcal{O}_K$ of a number field K is the set of all algebraic integers of K.
- Example: $\mathbb{Z}[\sqrt{2}] = \{x + y\sqrt{2} | x, y \in \mathbb{Z}\}$ is the ring of integers of $\mathbb{Q}(\sqrt{2})$.
- Example: For the mth cyclotomic number field $\mathbb{Q}(\zeta_m)$ of degree $n = \varphi(m)$, the ring of integers is given by

$$\mathbb{Z}[\zeta_m] = \mathbb{Z} + \mathbb{Z}\zeta_m + \ldots + \mathbb{Z}\zeta_m^{n-1} \cong \mathbb{Z}[X]/\langle \Phi_m(X) \rangle.$$

- There exists an integral basis $\{\omega_i\}_{i=1}^n$ of K such that any element of $\mathcal{O}_K$ can be uniquely written as $\sum_{i=1}^n a_i \omega_i$ with $a_i \in \mathbb{Z}$ for all i.
- We can get an algebraic lattice from $\mathcal{O}_K$.

Canonical embedding

- Let θ_i for $i = 1, \ldots, n$ be the distinct roots of the minimum polynomial of θ. There are exactly n embeddings $\sigma_i : K \to \mathbb{C}$, defined by $\sigma_i(\theta) = \theta_i$, for $i = 1, \ldots, n$.
- When we apply the embedding σ_i to an arbitrary element x of K, $x = \sum_{k=1}^{n} a_k \theta^k$, $a_k \in \mathbb{Q}$, we get

$$\sigma_i(x) = \sigma_i(\sum_{k=1}^{n} a_k \theta^k) = \sum_{k=1}^{n} \sigma_i(a_k)\sigma_i(\theta)^k = \sum_{k=1}^{n} a_k \theta_i^k$$

- Let r_1 be the number of embeddings with image in $\mathbb{R}$, and $2r_2$ the number of embeddings with image in $\mathbb{C}$ so that $r_1 + 2r_2 = n$.
- Canonical (Minkowski) embedding

$$\sigma(x) = (\sigma_1(x), \ldots, \sigma_{r_1}(x), \Re(\sigma_{r_1+1}(x)), \ldots, \Im(\sigma_{r_1+r_2}(x))) \in \mathbb{R}^n.$$

From $\mathcal{O}_K$ to lattice

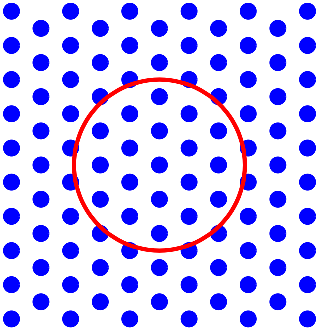
- If we take the ring of integers $\mathcal{O}_K$, we obtain a lattice with canonical embedding.
- Let $\{\omega_i\}_{i=1}^n$ be an integral basis of K. The n vectors $v_i = \sigma(\omega_i) \in \mathbb{R}^n$ form a basis of an algebraic lattice $\Lambda = \Lambda(\mathcal{O}_K) = \sigma(\mathcal{O}_K)$, whose generator matrix is given by

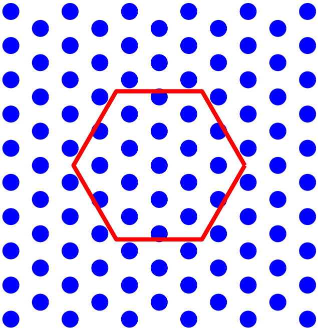
$$\mathbf{M} = \begin{pmatrix} \sigma_1(\omega_1) & \cdots & \sigma_{r_2}(\omega_1) & \Re\sigma_{r_1+1}(\omega_1) & \cdots & \Im\sigma_{r_1+r_2}(\omega_1) \\ \sigma_1(\omega_2) & \cdots & \sigma_{r_2}(\omega_2) & \Re\sigma_{r_1+1}(\omega_2) & \cdots & \Im\sigma_{r_1+r_2}(\omega_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_1(\omega_n) & \cdots & \sigma_{r_2}(\omega_n) & \Re\sigma_{r_1+1}(\omega_n) & \cdots & \Im\sigma_{r_1+r_2}(\omega_n) \end{pmatrix}.$$

- We can get more lattices $\Lambda' \subset \Lambda$ from ideals $\mathcal{I} \subseteq \mathcal{O}_K$, which are called ideal lattices.

Lattice coding

- A lattice code is a code constructed from a lattice in the Euclidean space. Thus, it is naturally suited to a Gaussian channel.
- Since a lattice is infinite, shaping is needed to obtain a code of given rate. The common practice is to apply a finite shaping region (cubic, spherical, or Voronoi).





AWGN-good lattices

- The issues of shaping and coding are largely separable.
- Consider (infinite) lattice coding over the AWGN channel with noise variance σ_w^2.
- For an n-dimensional lattice Λ, define the volume-to-noise ratio (VNR) by

$$\gamma_\Lambda(\sigma_w) \triangleq \frac{(V(\Lambda))^{\frac{2}{n}}}{\sigma_w^2}$$

- The error probability is given by $P_e = \mathbb{P}\{\mathsf{W}^n \notin \mathcal{V}(\Lambda)\}$.
- A sequence of lattices $\Lambda^{(n)}$ of increasing dimension n is *AWGN-good* if for a fixed VNR $\gamma_\Lambda(\sigma_w) > 2\pi e$, P_e vanishes in n [Poltyrev'94].
- This is the best possible performance, achieved only if the Voronoi region is approximately a sphere.

Constructions from number fields

- The connection between coding and dense sphere packing is well known [Conway-Sloane'93].
- Lattices from cyclotomic fields [Craig'78]: Let $p = n + 1$ be a prime. Take a principal ideal $\mathcal{I} = ((1 - \zeta_p)^m)$ in $\mathbb{Q}(\zeta_p)$, for some m. $\mathcal{I}$ yields a lattice under canonical embedding.
- Lattices from class field towers [Martinet'78]: Embedding of $\mathcal{O}_{K_i}$ in the tower; densest lattices having been constructed.

Remark

Number field constructions yield dense lattices, but have not achieved the Minkowski-Hlawka bound.

Constructions from codes

- Constructions A, B, C, D... [Leech-Sloane'71].
- Existence of AWGN-good lattices from Construction A [Loeliger'97].
- Shannon-theoretic construction [Forney et al.'00]: A lattice from Construction A is AWGN-good if the code achieves capacity of the mod-q channel.

Remark

Forney et al.'s construction yields AWGN-good lattices, but are not necessarily dense.

Achieving capacity with lattice codes

- Lattice codes achieve the $\frac{1}{2}\log(1+\rho)$ capacity of the Gaussian channel.
- This also forms the basis of the recent surge in applications to network information theory.
- Erez and Zamir's scheme [2004]
  - Voronoi shaping: the coarse lattice is good for quantization.
  - It also requires dithering.
  - The existence proof is again based on random lattices.
- Polar lattice
  - Gaussian shaping: Applying a discrete Gaussian distribution over an AWGN-good lattice [Ling-Belfiore'14].
  - Explicit construction from polar codes, $O(n\log n)$ decoding complexity [Yan-Liu-Ling-Wu'15].

Coding for fading channel

- Good Codes for the Gaussian channel usually have rather poor performance in the fading channel.
- The construction of good lattice codes for the fading channel exploits algebraic number theory [Belfiore et al. 1990s].
- A powerful tool is ideal theory for the rings of algebraic integers, leading to the construction of ideal lattice codes.
- However, capacity-achieving codes for fading channels are still unavailable[1].
- Record is a constant gap to capacity [Ordentlich-Erez'13, Luzzi-Vehkalahti'15].

[1] In the special case of i.i.d fading, capacity is acheived with polar lattices [Liu-Ling'16]

Work in 1990s

- Consider the fast (i.i.d.) Rayleigh fading channel

$$y_i = h_i x_i + w_i$$

where $(x_1, \ldots, x_n) = \mathbf{x} \in \mathbb{R}^n$ is the codeword, h_i's are i.i.d. fading coefficients of the Rayleigh distribution.

- Pairwise error probability

$$P(\mathbf{x} \to \hat{\mathbf{x}}) \leq \frac{1}{2} \prod_{i: x_i \neq \hat{x}_i} \frac{8\sigma_w^2}{(x_i - \hat{x}_i)^2} = \frac{1}{2} \frac{(8\sigma_w^2)^l}{\prod_{i: x_i \neq \hat{x}_i} (x_i - \hat{x}_i)^2}$$

if the two codewords differ in l positions.

- Design criteria
  - Maximize the diversity order $\min\{l\} = \min_{\mathbf{x} \neq \hat{\mathbf{x}}} |\{i : x_i \neq \hat{x}_i\}|$. Full diversity order n is desired.
  - Maximize the product distance: $d_{p,\min} = \min_{\mathbf{x} \neq \hat{\mathbf{x}}} \prod_{i: x_i \neq \hat{x}_i} |x_i - \hat{x}_i|$.

Ideal lattice code

- Now, suppose an ideal lattice Λ built from ideal $\mathcal{I} \subseteq \mathcal{O}_K$ is used as the coding lattice.
- By the union bound and geometric uniformity, error probability

$$P_e \leq \sum_{\mathbf{x}\in\Lambda\backslash\mathbf{0}} P(\mathbf{0} \to \mathbf{x}) = \sum_{\mathbf{x}\in\Lambda\backslash\mathbf{0}} \frac{1}{2} \frac{(8\sigma_w^2)^{l_\mathbf{x}}}{\prod_{i:x_i\neq 0} |x_i|^2}$$

  where $l_\mathbf{x}$ is the number of nonzero elements (the sum is take over a shaping region).
- Design criteria rephrased (Oggier, Viterbo'04):
  - Maximize the diversity order $\min\{l\} = \min_{\mathbf{x}\in\Lambda\backslash\mathbf{0}} |\{i : x_i \neq 0\}|$. K should be totally real to achieve full diversity n.
  - The minimum norm $\mathrm{N}_{\min} = \min_{x\neq 0, x\in\mathcal{I}} |\mathrm{N}(x)|$ should be maximized (recall algebraic norm $\mathrm{N}(x) \triangleq \prod_{i=1}^n \sigma_i(x)$).

Capacity?

- Our goal for block fading channels is to achieve capacity of compound channel (with diagonal $\mathbf{H}$)

$$C = \log \left| \mathbf{I} + \rho \mathbf{H}^{\dagger} \mathbf{H} \right| .$$

Capacity?

- Our goal for block fading channels is to achieve capacity of compound channel (with diagonal $\mathbf{H}$)

$$C = \log \left| \mathbf{I} + \rho \mathbf{H}^{\dagger} \mathbf{H} \right| .$$

- We need coding over time. Recall the system model

$$\underbrace{\mathbf{Y}}_{n \times T} = \underbrace{\mathbf{H}}_{n \times n} \underbrace{\mathbf{X}}_{n \times T} + \underbrace{\mathbf{W}}_{n \times T}$$

Capacity?

- Our goal for block fading channels is to achieve capacity of compound channel (with diagonal $\mathbf{H}$)

$$C = \log \left| \mathbf{I} + \rho \mathbf{H}^{\dagger} \mathbf{H} \right| .$$

- We need coding over time. Recall the system model

$$\underbrace{\mathbf{Y}}_{n \times T} = \underbrace{\mathbf{H}}_{n \times n} \underbrace{\mathbf{X}}_{n \times T} + \underbrace{\mathbf{W}}_{n \times T}$$

- Vectorizing this equation, we obtain

$$\underbrace{\mathbf{y}}_{nT \times 1} = \underbrace{\mathcal{H}}_{nT \times nT} \underbrace{\mathbf{x}}_{nT \times 1} + \underbrace{\mathbf{w}}_{nT \times 1}$$

where $\mathcal{H} = \mathbf{I}_T \otimes \mathbf{H}$.

Fading-good lattices

- Now we design a lattice $\Lambda \subset \mathbb{C}^{nT}$ so that $\mathbf{x} \in \Lambda$.

Fading-good lattices

- Now we design a lattice $\Lambda \subset \mathbb{C}^{nT}$ so that $\mathbf{x} \in \Lambda$.
- With Gaussian shaping, the problem boils down to finding a lattice that is good for block fading.

Fading-good lattices [Campello-Ling-Belfiore'16]

We say that a sequence of lattices Λ of increasing dimension nT is universally good for the block-fading channel if for any VNR $\gamma_{(\mathbf{I}_T \otimes \mathbf{H})\Lambda}(\sigma_w) > 2\pi e$ and all (absolute) $\mathbf{H}$ s.t. $|\mathbf{H}| = D$, $P_e(\Lambda, \mathbf{H}) \to 0$ as $T \to \infty$.

Fading-good lattices

- Now we design a lattice $\Lambda \subset \mathbb{C}^{nT}$ so that $\mathbf{x} \in \Lambda$.
- With Gaussian shaping, the problem boils down to finding a lattice that is good for block fading.

Fading-good lattices [Campello-Ling-Belfiore'16]

We say that a sequence of lattices Λ of increasing dimension nT is universally good for the block-fading channel if for any VNR $\gamma_{(\mathbf{I}_T \otimes \mathbf{H})\Lambda}(\sigma_w) > 2\pi e$ and all (absolute) $\mathbf{H}$ s.t. $|\mathbf{H}| = D$, $P_e(\Lambda, \mathbf{H}) \to 0$ as $T \to \infty$.

- If $\mathbf{H} = \mathbf{I}$, the problem reduces to that of AWGN-good lattices.

Generalized Construction A

- We resort to generalized Construction A over $\mathcal{O}_K$.

Generalized Construction A

- We resort to generalized Construction A over $\mathcal{O}_K$.

Generalized Construction A [Kositwattanarerk-Ong-Oggier'13]

Let $K/\mathbb{Q}(i)$ be a relative extension of degree n.
Let $\mathfrak{p} \subset \mathcal{O}_K$ be a prime ideal above p with norm p^ℓ. Then $\mathcal{O}_K/\mathfrak{p} \simeq \mathbb{F}_{p^\ell}$.
The $\mathcal{O}_K$-lattice Λ associated to a linear code $\mathcal{C} \subset \mathbb{F}_{p^\ell}^T$ is defined as:

$$\Lambda = \mathcal{C} + \mathfrak{p}^T.$$

Generalized Construction A

- We resort to generalized Construction A over $\mathcal{O}_K$.

Generalized Construction A [Kositwattanarerk-Ong-Oggier'13]

Let $K/\mathbb{Q}(i)$ be a relative extension of degree n.
Let $\mathfrak{p} \subset \mathcal{O}_K$ be a prime ideal above p with norm p^ℓ. Then $\mathcal{O}_K/\mathfrak{p} \simeq \mathbb{F}_{p^\ell}$.
The $\mathcal{O}_K$-lattice Λ associated to a linear code $\mathcal{C} \subset \mathbb{F}_{p^\ell}^T$ is defined as:

$$\Lambda = \mathcal{C} + \mathfrak{p}^T.$$

- It reduces to usual Construction A: $\Lambda = \mathcal{C} + p^T$ when $K = \mathbb{Q}$.

Achieving capacity

Generalized Construction A is good for fading channels

With number fields, generalized Construction A are good for block fading [Campello-Ling-Belfiore'16].
The existence of a universal lattice can be proven by Minkowski-Hlawka, i.e., averaging over random codes $\mathcal{C}$ (with $p \to \infty$).
Thanks to the unit group, the set of quantized channels is always compact (the unit group "absorbs" the channel).

Achieving capacity

Generalized Construction A is good for fading channels

With number fields, generalized Construction A are good for block fading [Campello-Ling-Belfiore'16].
The existence of a universal lattice can be proven by Minkowski-Hlawka, i.e., averaging over random codes $\mathcal{C}$ (with $p \to \infty$).
Thanks to the unit group, the set of quantized channels is always compact (the unit group "absorbs" the channel).

- With Gaussian shaping, capacity of compound fading channels is achieved.

Achieving capacity

Generalized Construction A is good for fading channels

With number fields, generalized Construction A are good for block fading [Campello-Ling-Belfiore'16].
The existence of a universal lattice can be proven by Minkowski-Hlawka, i.e., averaging over random codes $\mathcal{C}$ (with $p \to \infty$).
Thanks to the unit group, the set of quantized channels is always compact (the unit group "absorbs" the channel).

- With Gaussian shaping, capacity of compound fading channels is achieved.
- However, there is a large gap between theory (given above) and state of the art [Ordentlich-Erez'13, Luzzi-Vehkalahti'15].

Universality (for $T = 1$)

- In general, there is no guarantee that a faded lattice still has good minimum distance.

Universality (for $T = 1$)

- In general, there is no guarantee that a faded lattice still has good minimum distance.
- Nevertheless, an ideal lattices $\mathcal{I}$ is "incompressible" [Luzzi-Vehkalahti'15]: the minimum Euclidean distance of faded lattice $\mathbf{H}\mathcal{I}$

$$\begin{aligned}
\min_{\mathbf{H}:|\mathbf{H}|=D} d^2_{\min}(\mathbf{H}\mathcal{I}) &= \min_{\mathbf{H}:|\mathbf{H}|=D} \min_{\mathbf{x}\in\mathcal{I},\mathbf{x}\neq\mathbf{0}} \|\mathbf{H}\mathbf{x}\|^2 \\
&= \min_{\mathbf{x}\in\mathcal{I},\mathbf{x}\neq\mathbf{0}} nD^{2/n}(x_1\cdots x_n)^{2/n} \\
&= nD^{2/n}\mathrm{N}(\mathcal{I})^{2/n}
\end{aligned}$$

which follows from AM-GM inequality.

Universality (for $T = 1$)

- In general, there is no guarantee that a faded lattice still has good minimum distance.
- Nevertheless, an ideal lattices $\mathcal{I}$ is “incompressible” [Luzzi-Vehkalahti’15]: the minimum Euclidean distance of faded lattice $\mathbf{H}\mathcal{I}$

$$\begin{aligned} \min_{\mathbf{H}:|\mathbf{H}|=D} d_{\min}^2(\mathbf{H}\mathcal{I}) &= \min_{\mathbf{H}:|\mathbf{H}|=D} \min_{\mathbf{x}\in\mathcal{I},\mathbf{x}\neq\mathbf{0}} \|\mathbf{H}\mathbf{x}\|^2 \\ &= \min_{\mathbf{x}\in\mathcal{I},\mathbf{x}\neq\mathbf{0}} nD^{2/n}(x_1\cdots x_n)^{2/n} \\ &= nD^{2/n}\mathrm{N}(\mathcal{I})^{2/n} \end{aligned}$$

which follows from AM-GM inequality.

- However, $d_{\min} = 0$ for usual mod-q lattices.

Error probability of MIMO

- Recall channel model $\underbrace{\mathbf{Y}}_{n\times T} = \underbrace{\mathbf{H}}_{n\times n}\underbrace{\mathbf{X}}_{n\times T} + \underbrace{\mathbf{W}}_{n\times T}$.
- For a linear space-time code $\mathcal{S}$, consider pairwise error probability

$$P(\mathbf{0} \to \mathbf{X}) = \mathrm{E}_{\mathbf{H}}\left[Q\left(\frac{\|\mathbf{HX}\|_F}{\sqrt{2}\sigma_w}\right)\right] \quad (3)$$

- For Rayleigh fading,

$$P(\mathbf{0} \to \mathbf{X}) \leq \left|\mathbf{I} + \frac{\mathbf{XX}^\dagger}{4\sigma_w^2}\right|^{-n} \quad (4)$$

- If codeword matrix $\mathbf{X}$ has full rank, then high-SNR behavior

$$P(\mathbf{0} \to \mathbf{X}) \leq |\mathbf{XX}^\dagger|^{-n}\left(\frac{1}{4\sigma_w^2}\right)^{-n^2} \quad (5)$$

Non-vanishing determinant (NVD)

- Define
$$\Delta = \inf_{\mathbf{0} \neq \mathbf{X} \in \mathcal{S}} |\mathbf{X}\mathbf{X}^{\dagger}| \tag{6}$$
- Non-vanishing $\Delta > 0$ implies full diversity, in fact, optimum DMT (diveristy-multiplexing gains tradeoff) of the space-time code $\mathcal{S}$.
- Larger Δ gives larger coding gain.
- Many approaches were tried in 2000's, but $\Delta \to 0$ as constellation grows for most of them.
- Solution: construct a lattice code from division algebra [Sethuraman-Rajan'02] over a number field [Belfiore-Rekaya'03].

Division algebra

Definition

Let A be a ring and denote by A^* the set of invertible elements of A for multiplication. If $A^* = A\backslash\{0\}$, then A is referred to as a division algebra.

Hamilton's quaternions

Let $\{1, i, j, k\}$ be a basis for a vector space of dimension 4 over $\mathbb{R}$, which satisfy the relations $i^2 = -1$, $j^2 = -1$, $k^2 = -1$ and $k = ij = -ji$. Hamilton's quaternions are defined as the set

$$\mathbb{H} = \{x + yi + zj + wk \mid x, y, z, w \in \mathbb{R}\}.$$

The key link

- Rewrite a quaternion $q = x + yi + zj + wk$ as
$$q = (x + yi) + j(z - wi) = \alpha + j\beta$$
where $\alpha = x + yi$ and $\beta = z - wi$.
- It matrix representation
$$q \Longleftrightarrow \mathbf{X} = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}.$$
- Check
$$|\mathbf{X}| = |\alpha|^2 + |\beta|^2 = N(q) > 0 \quad \text{if} \quad \mathbf{X} \neq \mathbf{0}$$
The norm $N(q) = 0$ if and only if $q = 0$.
- This is the famous Alamouti code with full diversity [Alamouti'98]. It's used in DVB, WiFi and 4G.

Cyclic algebra

Definition

Let L/K be a Galois extension of degree n whose Galois group is cyclic with generator σ. Choose an element $0 \neq \gamma \in K$. A cyclic algebra $\mathcal{A} = (L/K, \sigma, \gamma)$ is defined as the direct sum

$$\mathcal{A} = L \oplus eL \oplus \cdots \oplus e^{n-1}L$$

where

$$e^n = \gamma \quad \text{and} \quad \lambda e = e\sigma(\lambda) \quad \lambda \in L.$$

The cyclic algebra may be viewed as a vector space over L, i.e., an element $x \in \mathcal{A}$ is written as

$$x = x_0 + ex_1 + \cdots + e^{n-1}x_{n-1} \quad \text{for} \quad x_i \in L.$$

The rule $\lambda e = e\sigma(\lambda)$ defines multiplication.

Matrix representation

An element $x \in \mathcal{A}$

$$x = x_0 + ex_1 + \cdots + e^{n-1}x_{n-1} \quad \text{for} \quad x_i \in L$$

can be represented by

$$\begin{pmatrix} x_0 & \gamma\sigma(x_{n-1}) & \gamma\sigma^2(x_{n-2}) & \dots & \gamma\sigma^{n-1}(x_1) \\ x_1 & \sigma(x_0) & \gamma\sigma^2(x_{n-1}) & \dots & \gamma\sigma^{n-1}(x_2) \\ \vdots & & \vdots & & \vdots \\ x_{n-2} & \sigma(x_{n-3}) & \sigma^2(x_{n-4}) & \dots & \gamma\sigma^{n-1}(x_{n-1}) \\ x_{n-1} & \sigma(x_{n-2}) & \sigma^2(x_{n-3}) & \dots & \sigma^{n-1}(x_0) \end{pmatrix}. \qquad (7)$$

Cyclic division algebra [Oggier-Belfiore-Viterbo'07]

If $0 \neq \gamma, \gamma^2, \dots, \gamma^{n-1} \in K$ are not a norm of some element of L, then $\mathcal{A} = (L/K, \sigma, \gamma)$ is a cyclic division algebra.

Golden code ($n = 2$, optional in WiMAX)

In general, a space-time code is an order of $\mathcal{A}$. Then $|\mathbf{X}|$ is reduced norm and $\Delta > 0$ naturally.
If $n = 2$, consider the cyclic division algebra [Belfiore-Rekaya-Viterbo'05]

$$\mathcal{A} = (L = \mathbb{Q}(i, \sqrt{5})/\mathbb{Q}(i), \sigma, i)$$

where $\sigma : \sqrt{5} \mapsto -\sqrt{5}$. The ring of integers $\mathcal{O}_L$ is given by

$$\mathcal{O}_L = \{a + b\theta \mid a, b \in \mathbb{Z}[i]\}$$

where $\theta = \frac{1+\sqrt{5}}{2}$. A codeword of the Golden code is of the form

$$\mathbf{X} = \begin{pmatrix} a + b\theta & c + d\theta \\ i(c + d\sigma(\theta)) & a + b\sigma(\theta) \end{pmatrix}$$

where $a, b, c, d \in \mathbb{Z}[i]$.

Construction A from division algebras

Generalized Construction A [Vehkalahti-Kositwattanarerk-Oggier'14]

Let Λ be the natural order of cyclic division algebra $\mathcal{A}$.
Take a two-sided ideal $\mathcal{J}$ of Λ and consider the quotient ring $\Lambda/\mathcal{J}$.
Define a reduction $\beta : \Lambda \to \Lambda/\mathcal{J}$.
For a linear code $\mathcal{C}$ over $\Lambda/\mathcal{J}$, $\beta^{-1}(\mathcal{C})$ is a lattice (in $\mathbb{C}^{n^2 T}$).

- $\Lambda/\mathcal{J}$ can be a matrix ring, skew polynomial ring...
- Nevertheless still possible to prove Minkowski-Hlawka using codes over rings [Campello-Ling-Belfiore'16].

NVD implies universality (for $T = 1$)

- Again, Λ with NVD $\Delta > 0$ is "incompressible" [Luzzi-Vehkalahti'15]: the minimum Euclidean distance of faded lattice $\mathbf{H}\Lambda$

$$\begin{aligned}\min_{\mathbf{H}:|\mathbf{H}|=D} d^2_{\min}(\mathbf{H}\Lambda) &= \min_{\mathbf{H}:|\mathbf{H}|=D}\ \min_{\mathbf{X}\in\Lambda,\mathbf{X}\neq\mathbf{0}} \|\mathbf{H}\mathbf{X}\|_F^2 \\ &= \min_{\mathbf{X}\in\Lambda,\mathbf{X}\neq\mathbf{0}} nD^{2/n}|\mathbf{X}|^{2/n} \\ &= nD^{2/n}\Delta^{2/n}\end{aligned}$$

which follows from Hadamard's inequality and AM-GM inequality.

Concluding remarks

- The coding problem for Gaussian channels has been solved.
- Algebraic number theory is an indispensable tool to design modern coding systems over fading/MIMO channels.
    - Achieve capacity of compound fading channels requires a combination of number theory and coding theory.
    - These can be viewed as concatenated codes (inner code is ideal/order; outer code is a code).
    - Extension to ergodic fading is possible.
- Emerging applications to multi-user networks:
    - Compute-and-forward
    - Interference alignment