Secure Wireless Communications Using Algebraic Number Theory

Cong Ling
Imperial College London
cling@ieee.org

Joint work with
Laura Luzzi and Roope Vehkelahti

York 2016

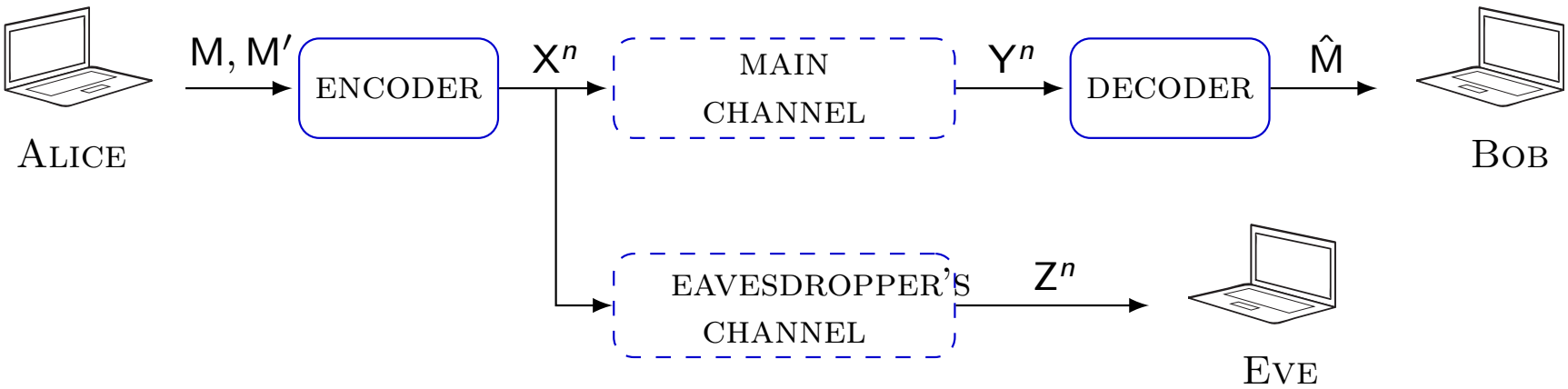
Wireless communication is vulnerable to eavesdropping

- Wireless security is a growing concern.
- Better protection is required on the physical layer.

Information theoretic security

- Classic cryptography relies on (often unproven) computational hardness assumptions; subject to computational attacks.
- In information theory security, security is guaranteed by negligible information leakage [Wyner'75].
- No keys are used, no keys to break.
- Security is obtained from physical randomness of communication channels—aka physical layer security.
- It offers additional protection of data on the physical layer, complementary to crypto on the network layer.

Wiretap channel



- M confidential message, M′ auxiliary message, X^n codeword (n: codeword length).
- Assuming symmetry and channel degradation, secrecy capacity

$$C_s = C(\text{Bob}) - C(\text{Eve}).$$

- Our problem is to construct a (coset) code to achieve C_s.

Security notions

- In information theory, information leakage is measured by mutual information:

Definition (Strong secrecy)

A wiretap code is information theoretically secure if $\mathbb{I}(\mathsf{M};\mathsf{Z}^n) \to 0$.

- In cryptography, it is measured by min-entropy or l_1 distance:

Definition (Semantic security)

A wiretap code is semantically secure if

$$\sup_{f,p_{\mathsf{M}}} \left(e^{-\mathbb{H}_\infty(f(\mathsf{M})|\mathsf{Z}^n)} - e^{-\mathbb{H}_\infty(f(\mathsf{M}))} \right) \to 0.$$

Equivalence

Definition (Distinguishing security)

A wiretap code achieves distinguishing security if

$$\max_{m,m'} \mathbb{V}(p_{Z^n|M=m}, p_{Z^n|M=m'}) \to 0.$$

- Semantic security and distinguishing security are equivalent.
- Semantic security is equivalent to strong secrecy holding for all message distributions [Bellare, Tessaro, Vardy'12].

Techniques for secrecy coding

- Coding methods
  - LDPC codes: limited success [Suresh et al.'10].
  - Polar codes: semantically secure [Mahdavifar-Vardy'11].
  - Lattice codes: semantically secure [L.-Luzzi-Belfiore-Stehle'14].
- Extractors/hash functions
  - Invertible extractors [Bellare et al.'12, Cheraghchi et al.'12, Chou et al.'14].
  - Universal hash functions [Hayashi et al.'10, Tyagi et al.'14].

Gaussian wiretap channel

- Channel model

$$\begin{cases} Y^n = X^n + W_b^n, & W_b^n \sim \mathcal{N}(0, \sigma_b^2 I_n), \\ Z^n = X^n + W_e^n, & W_e^n \sim \mathcal{N}(0, \sigma_e^2 I_n). \end{cases}$$

- Power constraint $\frac{1}{n}\mathbb{E}_{\mathcal{C}}[\|X^n\|^2] \leq P$.
- Signal-to-noise ratios (SNRs): $\rho_b = P/\sigma_b^2$, $\rho_e = P/\sigma_e^2$.
- We suppose $\sigma_e^2 > \sigma_b^2$ to have a positive secrecy capacity

$$C_s = \frac{1}{2}\log(1+\rho_b) - \frac{1}{2}\log(1+\rho_e).$$

Wiretap coding

- We use lattices to construct capacity-achieving codes for the Gaussian channel.
- Achieving the secrecy capacity of the Gaussian wiretap channel requires
  - A lattice code achieving the Shannon capacity of Bob's channel;
  - A secrecy-good lattice for Eve's channel.
- We also want explicit constructions.

Achieving Shannon capacity

- The capacity of the AWGN channel with SNR ρ is $\frac{1}{2}\log(1+\rho)$ [Shannon'1948].
- Lattice codes achieve the AWGN channel capacity:
  - Coding: A lattice Λ is AWGN-good if it achieves a vanishing error probability as long as its volume-to-noise ratio (VNR) $> 2\pi e$ [Poltyrev'94].
  - Shaping: for power constraint
    - Voronoi shaping: The shaping lattice is good for quantization [Erez-Zamir'04];
    - Gaussian shaping: Applying a discrete Gaussian distribution over an AWGN-good lattice [L.-Belfiore'14].

Discrete Gaussian distribution

- Standard continuous Gaussian distribution

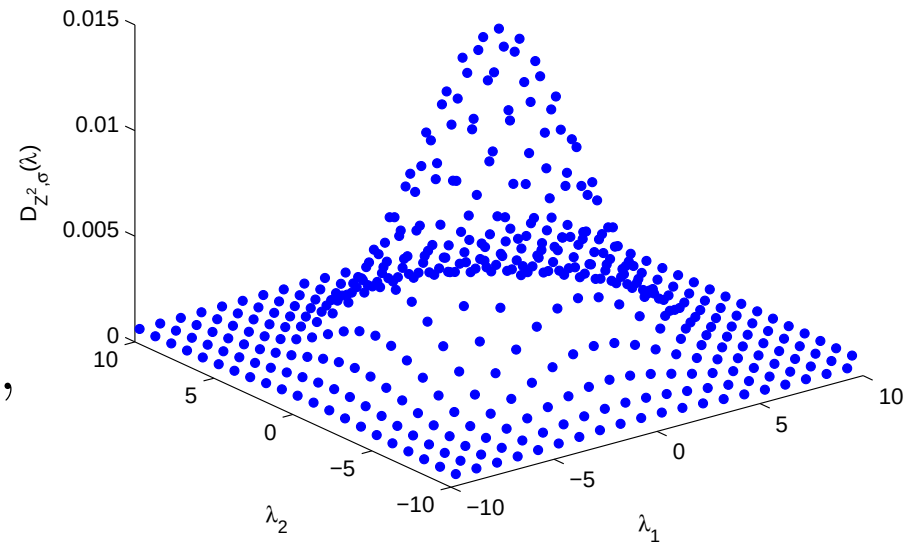
$$f_{\sigma,\mathbf{c}}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\|\mathbf{x}-\mathbf{c}\|^2}{2\sigma^2}}$$

- Discrete Gaussian distribution over lattice Λ

$$D_{\Lambda,\sigma,\mathbf{c}}(\boldsymbol{\lambda}) = \frac{f_{\sigma,\mathbf{c}}(\boldsymbol{\lambda})}{f_{\sigma,\mathbf{c}}(\Lambda)}, \quad \forall \boldsymbol{\lambda} \in \Lambda,$$

where

$$f_{\sigma,\mathbf{c}}(\Lambda) \triangleq \sum_{\boldsymbol{\lambda}\in\Lambda} f_{\sigma,\mathbf{c}}(\boldsymbol{\lambda}).$$

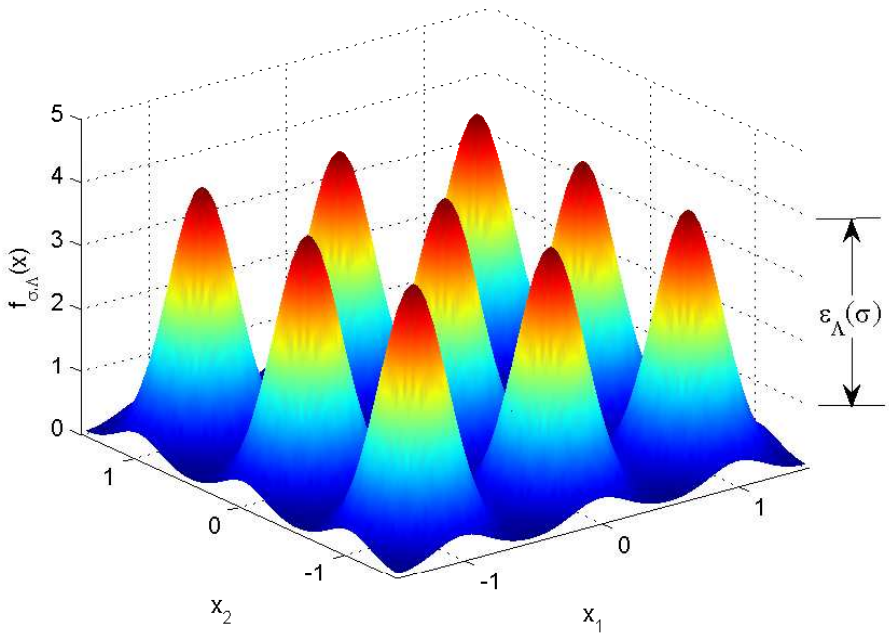


Periodic Continuous Gaussian distribution

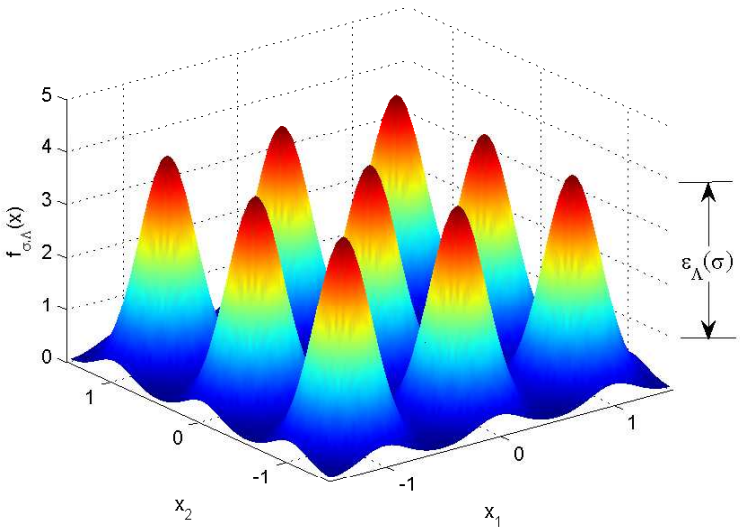
- Λ-periodic function

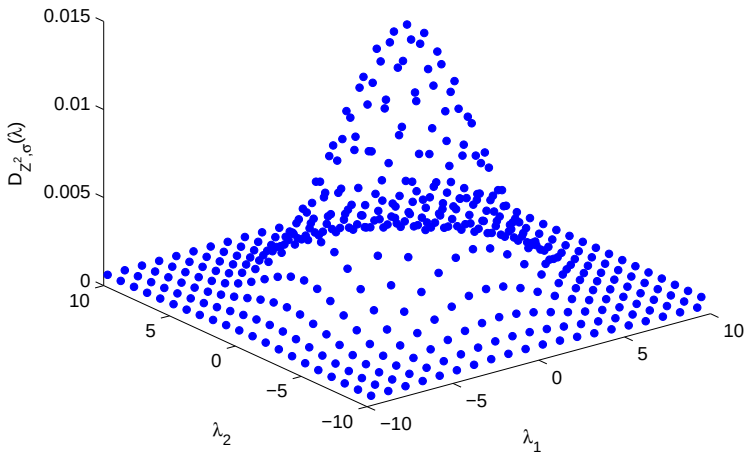
$$f_{\sigma,\Lambda}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\frac{\|\mathbf{x}-\boldsymbol{\lambda}|}{2\sigma^2}}$$

- $f_{\sigma,\Lambda}$ restricted to the quotient $\mathbb{R}^n/\Lambda$ is a (continuous) probability density.
- It arises, e.g., when Gaussian noise passes through a mod-Λ operator.
- It gets flat as σ increases.



Fourier duals





- Discrete and continuous versions of lattice Gaussian are the Fourier dual of each other.
- Fourier series expansion on the dual lattice Λ^*

$$f_{\sigma,\Lambda}(\mathbf{x}) = \frac{1}{V(\Lambda)} \sum_{\boldsymbol{\lambda}^* \in \Lambda^*} \hat{f}_\sigma(\boldsymbol{\lambda}^*) e^{j2\pi\langle \boldsymbol{\lambda}^*, \mathbf{x}\rangle}$$

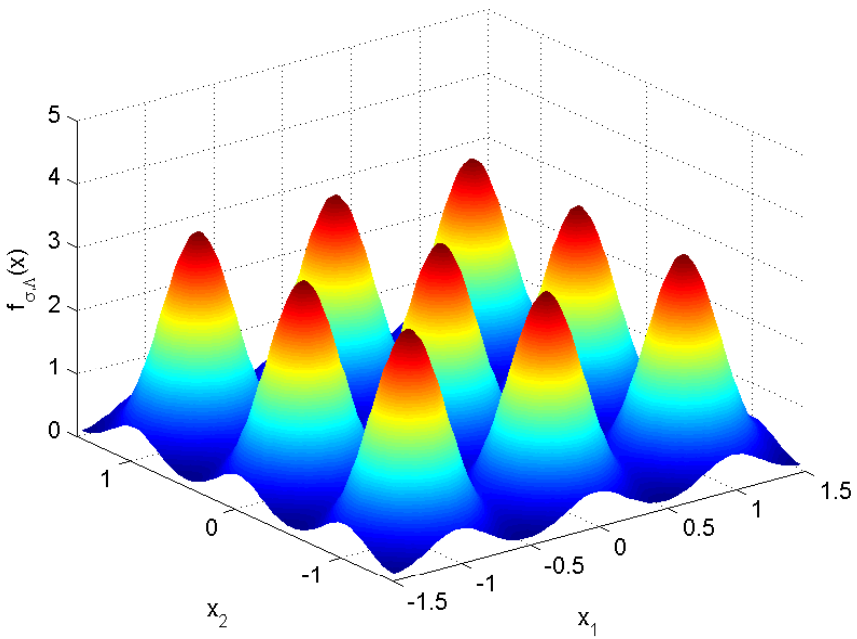
$$\hat{f}_\sigma(\mathbf{y}) = \int f_\sigma(\mathbf{x}) e^{-j2\pi\langle \mathbf{x}, \mathbf{y}\rangle} d\mathbf{x} = e^{-2\pi^2\sigma^2\|\mathbf{y}\|^2}$$

Periodic Continuous Gaussian distribution

- Λ-periodic function

$$f_{\sigma,\Lambda}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\frac{\|\mathbf{x}-\boldsymbol{\lambda}|}{2\sigma^2}}$$

- $f_{\sigma,\Lambda}$ restricted to the quotient $\mathbb{R}^n/\Lambda$ is a (continuous) probability density.
- It arises, e.g., when Gaussian noise passes through a mod-Λ operator.
- It gets flat as σ increases.

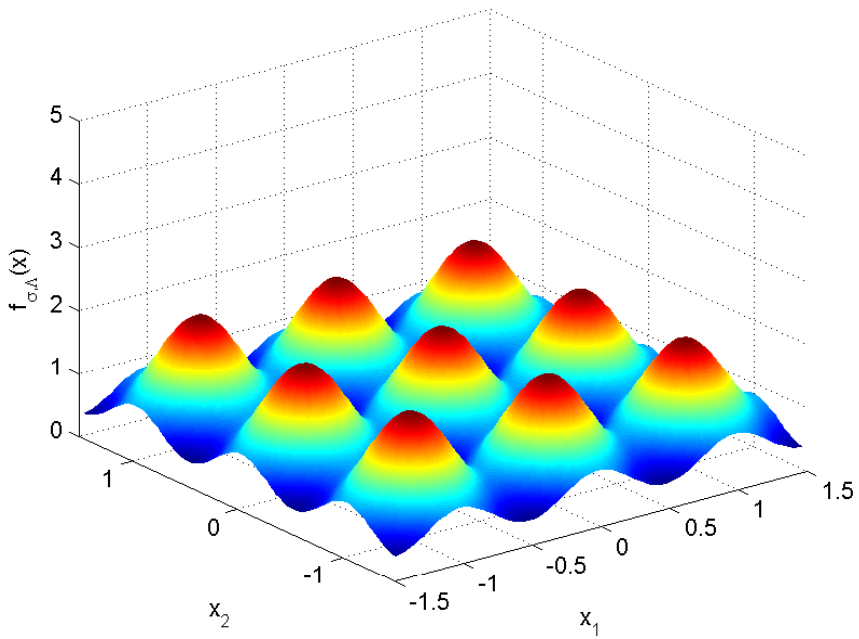


Periodic Continuous Gaussian distribution

- Λ-periodic function

$$f_{\sigma,\Lambda}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\frac{\|\mathbf{x}-\boldsymbol{\lambda}|}{2\sigma^2}}$$

- $f_{\sigma,\Lambda}$ restricted to the quotient $\mathbb{R}^n/\Lambda$ is a (continuous) probability density.
- It arises, e.g., when Gaussian noise passes through a mod-Λ operator.
- It gets flat as σ increases.

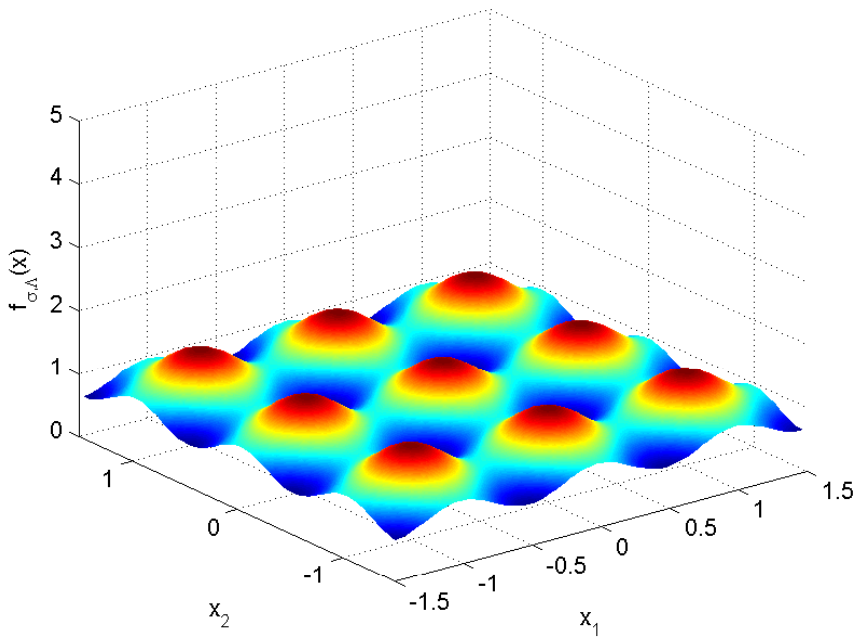


Periodic Continuous Gaussian distribution

- Λ-periodic function

$$f_{\sigma,\Lambda}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\frac{\|\mathbf{x}-\boldsymbol{\lambda}|}{2\sigma^2}}$$

- $f_{\sigma,\Lambda}$ restricted to the quotient $\mathbb{R}^n/\Lambda$ is a (continuous) probability density.
- It arises, e.g., when Gaussian noise passes through a mod-Λ operator.
- It gets flat as σ increases.

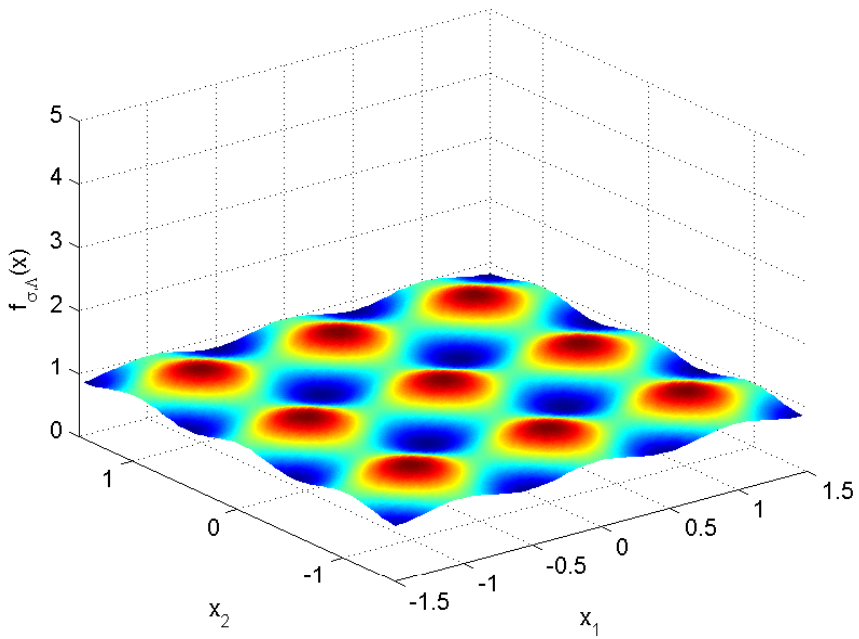


Periodic Continuous Gaussian distribution

- Λ-periodic function

$$f_{\sigma,\Lambda}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\frac{\|\mathbf{x}-\boldsymbol{\lambda}|}{2\sigma^2}}$$

- $f_{\sigma,\Lambda}$ restricted to the quotient $\mathbb{R}^n/\Lambda$ is a (continuous) probability density.
- It arises, e.g., when Gaussian noise passes through a mod-Λ operator.
- It gets flat as σ increases.



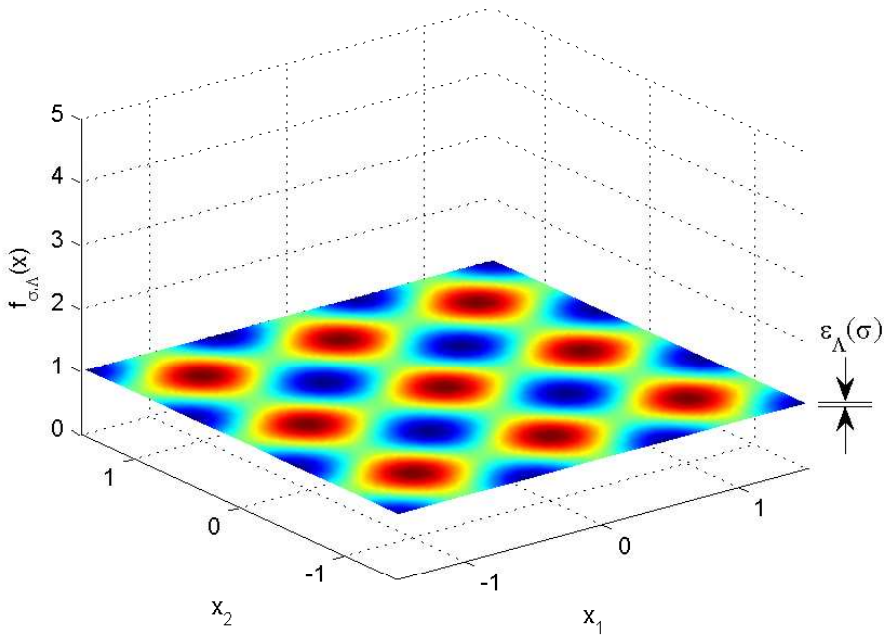
Flatness factor

Definition (Flatness factor (L.-Luzzi-Belfiore'12))

For a lattice Λ and for a parameter σ, the flatness factor is defined by:

$$\epsilon_\Lambda(\sigma) \triangleq \max_{\mathbf{x}\in\mathcal{R}(\Lambda)} |V(\Lambda) f_{\sigma,\Lambda}(\mathbf{x}) - 1|$$

where $\mathcal{R}(\Lambda)$ is a fundamental region.



$\epsilon_\Lambda(\sigma)$ quantifies the maximum variation of $f_{\sigma,\Lambda}(\mathbf{x})$. It is the other facet of the smoothing parameter [Micciancio,Regev'05].

Flatness factor

Definition (Flatness factor)

$$\epsilon_\Lambda(\sigma) = \left(\frac{\gamma_\Lambda(\sigma)}{2\pi}\right)^{\frac{n}{2}} \Theta_\Lambda\left(\frac{1}{2\pi\sigma^2}\right) - 1$$

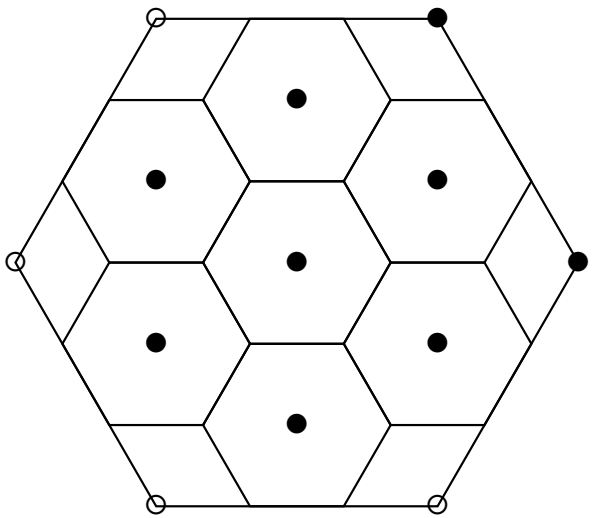
where $\Theta_\Lambda(\tau)$ is the theta series and $\gamma_\Lambda(\sigma) = \frac{(V(\Lambda))^{\frac{2}{n}}}{\sigma^2}$ is the volume-to-noise ratio (VNR).

- From the Minkowski-Hlawka theorem, there exist sequences of lattices whose flatness factors vanish exponentially as $n \to \infty$ if VNR $< 2\pi$.

Wiretap lattice code

- $\Lambda_e \subset \Lambda_b$ chain of lattices in $\mathbb{R}^n$, such that [L.-Luzzi-Belfiore'12]
  - Λ_b AWGN-good $\Rightarrow$ reliability for Bob
  - Λ_e secrecy-good $\Rightarrow$ secrecy against Eve

  Nesting ratio: $|\Lambda_b/\Lambda_e| = \lceil e^{nR_s} \rceil$, corresponding to secrecy rate R_s.



- Each confidential message $m = 1, \ldots, \lceil e^{nR_s} \rceil$ is associated to a coset $\Lambda_e + \boldsymbol{\lambda}_m, \quad \boldsymbol{\lambda}_m \in [\Lambda_b/\Lambda_e]$

The key idea

- Forget about shaping for now.
- Given message m, Alice samples a lattice point uniformly at random from a coset $\Lambda_e + \boldsymbol{\lambda}_m$.
- Due to the channel noise, Eve observes the periodic distribution

$$\frac{1}{(\sqrt{2\pi}\sigma_e)^n} \sum_{\boldsymbol{\lambda} \in \Lambda_e + \boldsymbol{\lambda}_m} e^{-\frac{\|\mathbf{z}-\boldsymbol{\lambda}\|^2}{2\sigma_e^2}}.$$

- If the flatness factor $\epsilon_{\Lambda_e}(\sigma_e)$ is small, it will be close to a uniform distribution, regardless of message m.
- Problem: one cannot really sample a point uniformly from a lattice (or its coset).

Gaussian sampling

- Alice actually samples X_m^n from lattice Gaussian distribution

$$\mathsf{X}_m^n \sim D_{\Lambda_e+\boldsymbol{\lambda}_m,\sigma_s}.$$

Lemma (Regev'09)

Let $\tilde{\sigma} = \frac{\sigma_s\sigma}{\sqrt{\sigma_s^2+\sigma^2}}$ and $\sigma_s' = \sqrt{\sigma_s^2+\sigma^2}$. If $\varepsilon = \epsilon_\Lambda(\tilde{\sigma}) < \frac{1}{2}$, the continuous distribution resulting from adding Gaussian noise of variance σ^2 to a discrete Gaussian $D_{\Lambda-\mathbf{c},\sigma_s}$ is uniformly close to $f_{\sigma_s'}(\mathbf{x})$ with L^∞ distance bounded by 4ε.

- It implies that if $\epsilon_{\Lambda_e}(\tilde{\sigma}_e) < \frac{1}{2}$, then:

$$\mathbb{V}\left(p_{\mathsf{Z}^n|\mathsf{M}}(\cdot|m), f_{\sigma_s'}\right) \leq 4\epsilon_{\Lambda_e}(\tilde{\sigma}_e).$$

- Eve's received signals converge to the same Gaussian distribution $f_{\sigma_s'}$. This already gives *distinguishing security*.

Secrecy-good lattice

- Mutual information for Eve is bounded as follows [L.-Luzzi-Belfiore-Stehle'14]:

$$\mathbb{I}(\mathsf{M};\mathsf{Z}^n) \leq 8\varepsilon_n n R_s - 8\varepsilon_n \log 8\varepsilon_n. \tag{1}$$

- A sequence of lattices $\Lambda^{(n)}$ is *secrecy-good* if

$$\epsilon_{\Lambda^{(n)}}(\sigma) = e^{-\Omega(n)}, \quad \forall \gamma_{\Lambda^{(n)}}(\sigma) < 2\pi. \tag{2}$$

- Strong secrecy has been obtained. Since we make no *a priori* assumption on the distribution of m, it also achieves *semantic security*.
- Under mild conditions, the secrecy rate

$$R_s < \frac{1}{2}\log(1+\rho_b) - \frac{1}{2}\log(1+\rho_e) - \frac{1}{2} \tag{3}$$

is achievable, which is within a half nat (possible to reduce) from the secrecy capacity.

Constructions of secrecy-good lattices

- The Minkowski-Hlawka theorem assumes random lattices.
- To make wiretap coding really work, we need concrete secrecy-good lattices.
- Polar lattice: build lattices from polar codes [Liu-Yan-L.'14]
  - Use polarization to bound mutual information $e^{-\tilde{O}(\sqrt{n})}$, similarly to polar codes.
  - Have nothing to do with theta series.
- Unimodular lattices [Lin-L.-Belfiore'14]
  - The theory of their theta series is well established, so amenable to a deep analysis of the flatness factor.
  - Information leakage is $O(e^{-cn})$: good for secrecy!
  - However, such lattices are not so explicit, let alone decoding.

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.

Dual ideal of $\mathcal{O}_K$

- Denote by Δ_K the (absolute) discriminant of the number field.

Codifferent (dual ideal)

$$\mathcal{O}_K^\vee = \{x \in K : \mathrm{Tr}(x\mathcal{O}_K) \subseteq \mathbb{Z}\}.$$

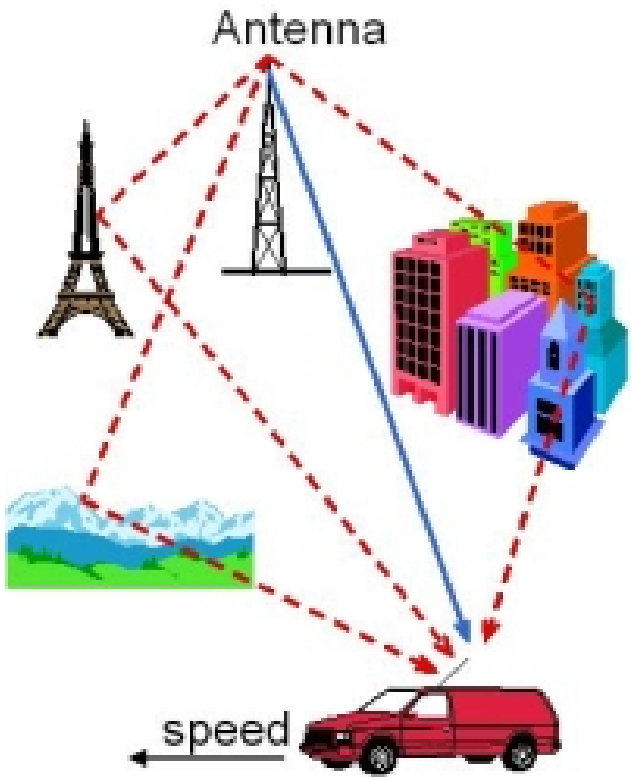
The codifferent is a fractional ideal, and its norm $N(\mathcal{O}_K^\vee) = 1/\Delta_K$. It embeds as the complex conjugate of the dual lattice Λ^* (in fact as Λ^* itself for CM fields).

- Its inverse is called the *different* whose norm is Δ_K.

For an ideal $\mathcal{I} \subseteq \mathcal{O}_K$, its inverse ideal is defined as

$$\mathcal{I}^{-1} = \{x \in K : x\mathcal{I} \subseteq \mathcal{O}_K\}.$$

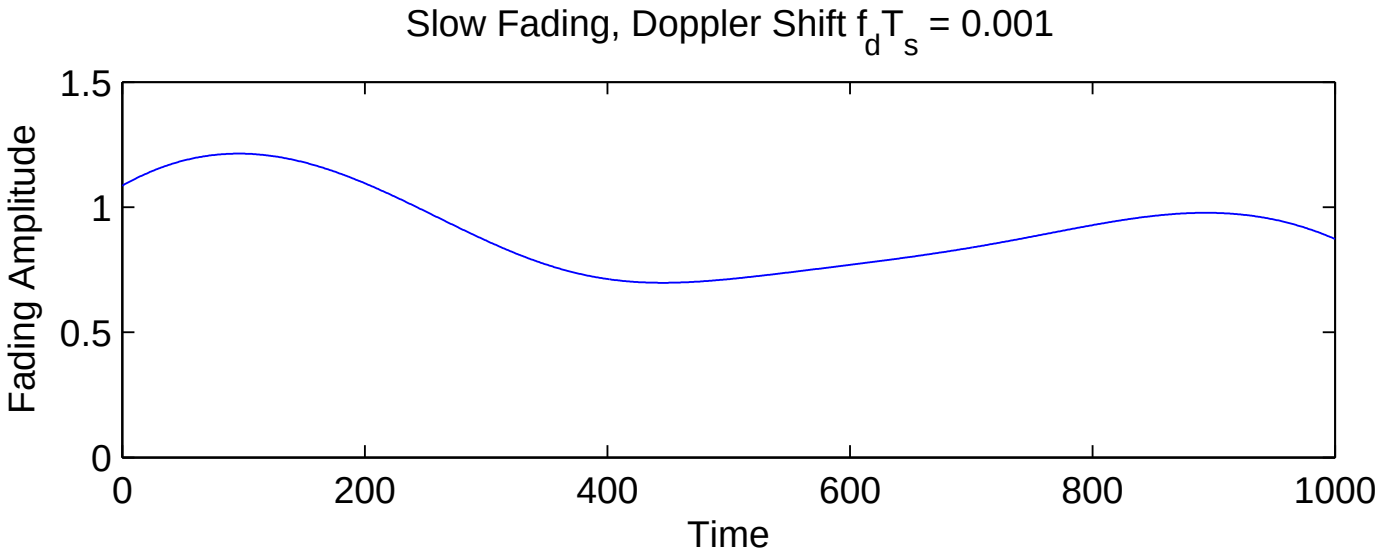
Multipath fading in mobile communications

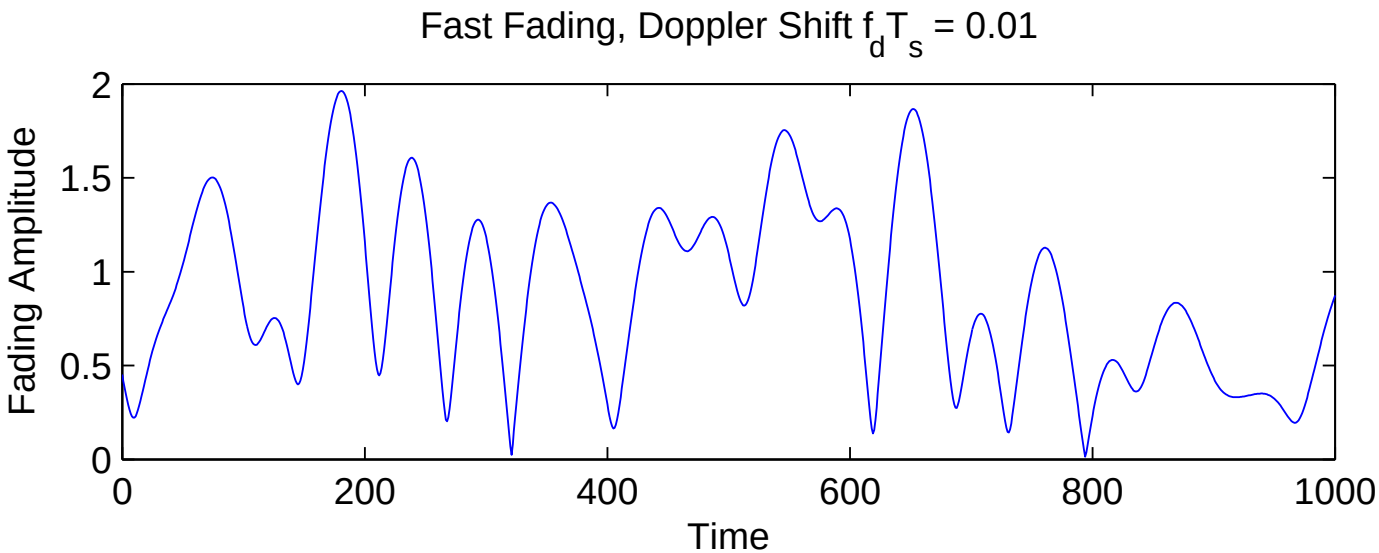


- Multipath propagation in urban environment.
- Fading is multiplicative noise (large variation in signal strength)
$$y_t = h_t x_t + w_t$$
- Rayleigh fading: channel coefficient h_t is complex Gaussian[a].
- Time autocorrelation is modelled by a Bessel function (Jakes model)
$R(\tau) = \mathbb{E}[h_t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

[a] Henceforth, signals will be complex.

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 (n: the number of blocks, or degree of the field):

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

  - 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 (n: the number of blocks, or degree of the field):

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

  - 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 $C = \mathbb{E}_H\left[\log\left(1 + |h|^2 \rho\right)\right]$.

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 (n: the number of blocks, or degree of the field):

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

  - 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 $C = \mathbb{E}_H\left[\log\left(1 + |h|^2 \rho\right)\right]$.
- 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 (n: the number of blocks, or degree of the field):

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

  - 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 $C = \mathbb{E}_H\left[\log\left(1 + |h|^2 \rho\right)\right]$.
- These represent two extremes of stationary fading.
- Open question: to design capacity-achieving codes over fading channels (5G telecom systems will operate close to capacity).

Coding for fading channel

- Good lattices for the Gaussian channel usually have rather poor performance in the fading channel.
- The construction of good lattice codes for the fading channel exploits the deep connection between algebraic number theory and lattices [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 today[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-L.'16]

Work in 1990s

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

$$y_i = h_i x_i + w_i$$

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

- Pairwise error probability

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

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, Bob's error probability

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

  where $l_{\mathbf{x}}$ is the number of nonzero elements (the sum is take over a shaping region).
- Design criteria rephrased (Oggier, Viterbo'04):
  - The minimum norm $\mathrm{N}_{\min} = \min_{x \neq 0, x \in \mathcal{I}} |\mathrm{N}(x)|$ should be maximized, subject to normalization (recall algebraic norm $\mathrm{N}(x) \triangleq \prod_{i=1}^{n} \sigma_i(x)$).
  - Full diversity comes as a byproduct.

Block fading channel

- Write down the matrix form[2] 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

$$C = \log \det(\mathbf{I} + \rho \mathbf{H}^\dagger \mathbf{H}). \tag{4}$$

- The receiver has channel state information (CSI), while the transmitter doesn't.
- Our goal is to achieve capacity C on all channels such that (4) is true (without even knowing the distribution of $\mathbf{H}$).
- This requires a universal code on the compound channel (4), i.e., a collection of channels with the same capacity.

[2] This formulation applies to MIMO channels where $\mathbf{H}$ is a full matrix.

Coding

- We need coding over time.

Coding

- We need coding over time.
- Recall the system model

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

Coding

- 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 $\gamma_{(\mathbf{I}_T \otimes \mathbf{H})\Lambda}(\sigma_w) > \pi e$ and all $\mathbf{H}$ 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 $\gamma_{(\mathbf{I}_T \otimes \mathbf{H})\Lambda}(\sigma_w) > \pi e$ and all $\mathbf{H}$ s.t. $|\mathbf{H}| = D$, $P_e(\Lambda, \mathbf{H}) \to 0$ as $T \to \infty$.

- Of course these lattices are AWGN-good (special case $\mathbf{H} = \mathbf{I}$).

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 $K = \mathbb{Q}$.

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 $K = \mathbb{Q}$.
- Existence of fading-good lattices is proven by using a random ensemble of codes (Minkowski-Hlawka theorem).

Ergodic Fading Channel

- In this case, the degree of extension $n = T$ (length).

Ergodic Fading Channel

- In this case, the degree of extension $n = T$ (length).

Generalized Construction A for ergodic fading [Vehkalahti-Kositwattanarerk-Oggier'14]

Let $K/\mathbb{Q}(i)$ be an extension with cyclic Galois group $\langle\sigma\rangle$.
Consider a prime p that splits completely and $\mathfrak{p}$ an ideal above p. There exists an isomorphism $\phi : \mathcal{O}_K/\mathfrak{p} \to \mathbb{F}_p$. Let $\pi : \mathcal{O}_K \to \mathcal{O}_K/\mathfrak{p}$ denote the projection operator. The reduction $\psi : \Lambda_K \to \mathbb{F}_p^n$ is given by

$$\psi(x, \sigma(x), \ldots, \sigma^{n-1}(x)) = ((\phi \circ \pi)(x), \ldots, (\phi \circ \pi)(\sigma^{n-1}(x))). \quad (5)$$

The construction A lattice is defined by $\Lambda_K(\mathcal{C}) = \psi^{-1}(\mathcal{C})$.

Ergodic Fading Channel

- In this case, the degree of extension $n = T$ (length).

Generalized Construction A for ergodic fading [Vehkalahti-Kositwattanarerk-Oggier'14]

Let $K/\mathbb{Q}(i)$ be an extension with cyclic Galois group $\langle\sigma\rangle$.
Consider a prime p that splits completely and $\mathfrak{p}$ an ideal above p. There exists an isomorphism $\phi : \mathcal{O}_K/\mathfrak{p} \to \mathbb{F}_p$. Let $\pi : \mathcal{O}_K \to \mathcal{O}_K/\mathfrak{p}$ denote the projection operator. The reduction $\psi : \Lambda_K \to \mathbb{F}_p^n$ is given by

$$\psi(x, \sigma(x), \ldots, \sigma^{n-1}(x)) = ((\phi \circ \pi)(x), \ldots, (\phi \circ \pi)(\sigma^{n-1}(x))). \quad (5)$$

The construction A lattice is defined by $\Lambda_K(\mathcal{C}) = \psi^{-1}(\mathcal{C})$.

- Ergodic capacity is achieved by appealing to Minkowski-Hlawka [Campello-L.-Belfiore'16].

A concrete code [Luzzi-Vehkalahti'15]

- Use $\mathcal{O}_K$ of number fields with constant root discriminant $\sqrt[n]{\Delta_K}$ from class field theory [Martinent'78].
- Denser than cyclotomic fields $K = \mathbb{Q}(\zeta_m)$ whose $\sqrt[n]{\Delta_K} = \frac{m}{\prod_{p|m} p^{1/(p-1)}} \leq n$.

A concrete code [Luzzi-Vehkalahti'15]

- Use $\mathcal{O}_K$ of number fields with constant root discriminant $\sqrt[n]{\Delta_K}$ from class field theory [Martinent'78].
- Denser than cyclotomic fields $K = \mathbb{Q}(\zeta_m)$ whose $\sqrt[n]{\Delta_K} = \frac{m}{\prod_{p|m} p^{1/(p-1)}} \leq n$.
- Gap to capacity $\approx \log\left(\frac{2G}{\pi e}\right)$ where $G \approx 92$.

A concrete code [Luzzi-Vehkalahti'15]

- Use $\mathcal{O}_K$ of number fields with constant root discriminant $\sqrt[n]{\Delta_K}$ from class field theory [Martinent'78].
- Denser than cyclotomic fields $K = \mathbb{Q}(\zeta_m)$ whose $\sqrt[n]{\Delta_K} = \frac{m}{\prod_{p|m} p^{1/(p-1)}} \leq n$.
- Gap to capacity $\approx \log\left(\frac{2G}{\pi e}\right)$ where $G \approx 92$.
- However, efficient implementation of Martinent's lattices is open.

Fading wiretap codes: Error-probability criterion

- Fading wiretap channel:

$$\mathbf{y} = \mathbf{H}_b\mathbf{x} + \mathbf{w}_b, \quad \mathbf{z} = \mathbf{H}_e\mathbf{x} + \mathbf{w}_e$$

- Eve's correct decoding probability [Belfiore-Oggier'13]

$$P_c \leq \text{const.}\rho_e^n V(\Lambda_b) \sum_{\mathbf{x}\in\Lambda_e} \prod_{i=1}^{n} \frac{1}{(1+\rho_e|x_i|^2)^2}$$

Design criterion

Minimize the inverse determinant sum $\sum_{\mathbf{x}\in\Lambda_e} \prod_{i=1}^{n} \frac{1}{(1+\rho_e|x_i|^2)^2}$.
This is a new problem related to Dedekind's zeta function, with interesting results recently obtained in [Vehkalahti-Lu-Luzzi'13, Karpuk et al.'15].

Fading wiretap codes: Flatness factor criterion

- To encode the message $m \in \mathcal{M}$, Alice samples $\mathbf{x}_m$ from $D_{\Lambda_e+\boldsymbol{\lambda}_m,\sigma_s}$.
- Eve observes discrete Gaussian $D_{\mathcal{H}_e(\Lambda_e+\boldsymbol{\lambda}_m),\mathcal{H}_e\sigma_s}$, plus noise[3].
- Goal: To make Eve's signal indistinguishable from continuous Gaussian of covariance matrix $\Sigma_0 = \sigma_s^2\mathcal{H}_e\mathcal{H}_e^\dagger + \sigma_e^2\mathbf{I}$, regardless of m

Generalized Regev lemma [Luzzi-L.-Vehkalahti, ISIT'16]

Given $\mathbf{x}_1$ sampled from discrete Gaussian distribution $D_{\Lambda+\mathbf{c},\sqrt{\Sigma_1}}$ and $\mathbf{x}_2$ sampled from continuous Gaussian distribution $f_{\sqrt{\Sigma_2}}$. Let $\Sigma_0 = \Sigma_1 + \Sigma_2$ and let $\Sigma_3^{-1} = \Sigma_1^{-1} + \Sigma_2^{-1}$. If $\epsilon_\Lambda(\sqrt{\Sigma_3}) \leq \varepsilon \leq \frac{1}{2}$, then the distribution g of $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ is close to $f_{\sqrt{\Sigma_0}}$:

$$g(\mathbf{x}) \in f_{\sqrt{\Sigma_0}}(\mathbf{x})\,[1-4\varepsilon, 1+4\varepsilon]\,.$$

[3] $\mathcal{H}_e = \mathbf{I}_T \otimes \mathbf{H}_e$ for block fading, and $\mathcal{H}_e = \mathbf{H}_e$ for block fading.

Flatness factor for correlated Gaussian

- Can define flatness factor $\epsilon_\Lambda(\sqrt{\Sigma})$ for correlated Gaussian with covariance matrix Σ.

Design criterion for Λ_e [Luzzi-L.-Vehkalahti, ISIT'16]

The flatness factor of $\mathcal{H}_e\Lambda_e$ for correlated Gaussian with covariance matrix Σ_3

$$\epsilon_{\mathcal{H}_e\Lambda_e}(\sqrt{\Sigma_3}) \to 0$$

where $\Sigma_3^{-1} = \sigma_s^{-2}(\mathcal{H}_e\mathcal{H}_e^\dagger)^{-1} + \sigma_e^{-2}\mathbf{I}$.

- Information theoretic security: Information leakage $\mathbb{I}(\mathrm{M};\mathrm{Z})$ can be bounded in terms of $\epsilon_{\mathcal{H}_e\Lambda_e}(\sqrt{\Sigma_3})$.

Block fading channel

- Compound wiretap channel

$$\begin{aligned}\mathbb{H}_b &= \left\{\mathbf{H}_b \in \mathbb{C}^{n\times n} : \log\det\left(\mathbf{I}+\rho_b\mathbf{H}_b^\dagger\mathbf{H}_b\right) = C_b\right\}\\ \mathbb{H}_e &= \left\{\mathbf{H}_e \in \mathbb{C}^{n\times n} : \log\det\left(\mathbf{I}+\rho_e\mathbf{H}_e^\dagger\mathbf{H}_e\right) = C_e\right\}\end{aligned}$$

- Secrecy capacity

$$C_s = C_b - C_e.$$

Coding scheme

Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}(K)$. The lattices are given by

$$\begin{aligned}\Lambda_b &= \mathcal{C}_b + \mathfrak{p}^T\\ \Lambda_e &= \mathcal{C}_e + \mathfrak{p}^T\end{aligned}$$

where the codes $\mathcal{C}_e \subseteq \mathcal{C}_b$.

Secrecy rate

- Applying Minkowski-Hlawka, we obtain [L., ISTC'16]

$$\mathbb{E}_{\Lambda_e}[\epsilon_{\mathcal{H}_e\Lambda_e}(\sqrt{\Sigma_3})] = \frac{V(\Lambda_e)}{(\pi\sigma_s^2)^{nT}} \det(\mathbf{I} + \rho_e \mathbf{H}_e^\dagger \mathbf{H}_e)^T.$$

- For a vanishing flatness factor, we need the condition

$$\frac{\det(\mathbf{I} + \rho_e \mathbf{H}_e^\dagger \mathbf{H}_e)^{1/n} V(\Lambda_e)^{\frac{1}{nT}}}{\sigma_s^2} < \pi. \tag{6}$$

- Secrecy rate

$$R_s < \log \frac{\det\left(\mathbf{I} + \rho_b \mathbf{H}_b^\dagger \mathbf{H}_b\right)}{\det\left(\mathbf{I} + \rho_e \mathbf{H}_e^\dagger \mathbf{H}_e\right)} - n = C_b - C_e - n$$

Ergodic fading

- Ergodic capacity ($n = T$)

$$\frac{1}{T}\log\det\left(1+\rho_b \mathbf{H}_b^\dagger \mathbf{H}_b\right) \quad \to \quad C_b \tag{7}$$

$$\frac{1}{T}\log\det\left(1+\rho_e \mathbf{H}_e^\dagger \mathbf{H}_e\right) \quad \to \quad C_e \tag{8}$$

- Coding scheme: Take Λ from a number field with constant root discriminant $\sqrt[2n]{\Delta_K} = G$, and let $\Lambda_b = \alpha_b \Lambda$, $\Lambda_e = \alpha_e \Lambda$ [Luzzi-L.-Vehkalahti, ISIT'16].

Design criterion for Λ_e

The flatness factor for correlated Gaussian

$$\epsilon_{\mathbf{H}_e\Lambda_e}(\sqrt{\Sigma}) = \epsilon_{\sqrt{\Sigma^{-1}}\mathbf{H}_e\Lambda_e}(1) \to 0,$$

where $\Sigma^{-1} = \sigma_s^{-2}(\mathbf{H}_e^\dagger \mathbf{H}_e)^{-1} + \sigma_e^{-2}\mathbf{I}$.

Smoothing parameter

Smoothing parameter

The *smoothing parameter* $\eta_\epsilon(\Lambda)$ is the smallest $s = \sqrt{2\pi}\sigma > 0$ such that the flatness factor $\epsilon_\Lambda(\sigma) \leq \epsilon$. For a correlation matrix Σ, we say

$$\sqrt{2\pi\Sigma} \succeq \eta_\epsilon(\Lambda) \quad \text{if} \quad \epsilon_\Lambda(\sqrt{\Sigma}) \leq \epsilon.$$

- Upper bound $\eta_\epsilon(\Lambda) \leq \frac{2\sqrt{n}}{\lambda_1(\Lambda^*)}$ for $\epsilon = 2^{-2n}$ [Micciancio-Regev'05].
- Now, for a faded lattice

$$\eta_\epsilon(\sqrt{\Sigma^{-1}}\mathbf{H}_e\Lambda_e) \leq \frac{2\sqrt{n}}{\lambda_1(\sqrt{\Sigma}(\mathbf{H}_e^\dagger)^{-1}\Lambda_e^*)}. \tag{9}$$

Bounds

- Using AM-GM inequality,

$$\lambda_1(\sqrt{\Sigma}(\mathbf{H}_e^{\dagger})^{-1}\Lambda_e^*) = \alpha_e^{-1} \min_{x\in\mathcal{O}_K^{\vee}\setminus\{0\}} \left\|\sqrt{\Sigma}(\mathbf{H}_e^{\dagger})^{-1}\sigma(x)\right\|$$

$$\geq \alpha_e^{-1} \min_{x\in\mathcal{O}_K^{\vee}\setminus\{0\}} \sqrt{n}\sigma_s \prod_{i=1}^{n}\left(\frac{1}{1+\frac{\sigma_s^2}{\sigma_e^2}|h_{e,i}|^2}\right)^{\frac{1}{2n}} \prod_{i=1}^{n}|\sigma_i(x)|^{\frac{1}{n}}$$

$$\approx \frac{\sqrt{n}\sigma_s}{\alpha_e \sqrt[2n]{\Delta_K} e^{C_e/2}} = \frac{\sqrt{n}\sigma_s}{\alpha_e G e^{C_e/2}}.$$

- Substituting into (9), we require

$$\eta_\epsilon(\sqrt{\Sigma^{-1}}\mathbf{H}_e\Lambda) \lessapprox \alpha_e G e^{C_e/2}/\sigma_s \leq \sqrt{2\pi}$$

$$\Rightarrow \alpha_e \lessapprox \frac{\sqrt{2\pi}\sigma_s}{G e^{C_e}}.$$

Secrecy rate

- A sufficient condition for reliability of Λ_b is $\frac{\lambda_1}{2} > \sqrt{n}\sigma_b$ (noise radius). Hence [Luzzi-Vehkalahti'15]

$$\alpha_b \gtrapprox \frac{\sqrt{2\pi e}\sigma_s}{e^{C_b/2}}.$$

Secrecy rate [Luzzi-Ling-Vehkalahti'16]

$$R_s = \frac{1}{n}\log\frac{V(\Lambda_e)}{V(\Lambda_b)} = 2\log\frac{\alpha_e}{\alpha_b} \lessapprox C_b - C_e - \log\frac{2G^2}{\pi}$$

which is within a constant gap to secrecy capacity.

- The gap is roughly doubled compared to that to Shannon capacity C_b only.
- Such lattices are simultaneously good for capacity and security.

Concluding remarks

- Algebraic number theory is a powerful tool to design modern coding systems for wireless fading channels.
- Lattice codes from number fields achieve capacity of fading channels, without CSIT.
  - These algebraic lattices are incompressible.
- In physical layer security, Eve's channel is never known.
  - Therefore, algebraic lattices appear to be absolutely necessary.
- Extension to MIMO channels is conceptually straightforward, but more technical (Campello-L.-Belfiore, Luzzi-L.-Vehkalahti, in progress).