Achieving Capacity and Security in Wireless Communications With Lattice Codes

(Invited Paper)

Cong Ling
Department of Electrical and Electronic Engineering
Imperial College London
London, UK
Email: cling@ieee.org

***Abstract*—Based on lattice Gaussian distributions and ideal lattices, we present a unified framework of lattice coding to achieve the channel capacity and secrecy capacity of wireless channels in the presence of Gaussian noise. The standard additive white Gaussian-noise (AWGN) channel, block fading channel, and multi-input multi-output (MIMO) fading channel are considered, which form a hierarchy of increasingly challenging problems in coding theory. To achieve channel capacity, we apply Gaussian shaping to a suitably defined good lattice for channel coding. To achieve secrecy capacity, we use a secrecy-good lattice nested with a coding lattice.**

I. Introduction

The lattice Gaussian distribution has emerged as a common theme in diverse areas. In mathematics, it was used to prove the transference theorems of lattices. In cryptography, it underpins lattice-based cryptosystems based on the worst-case hardness assumptions and fully-homomorphic encryption. In communications, lattice Gaussian distribution was applied to shaping of lattice codes.

Coincidentally, ideal lattices are also used in both areas. These are highly-structured lattices constructed from the ideals of the ring of integers of a number field, with a multiplicative structure and succinct representation. While the connection between lattices and number fields dates back to Minkowski and was used to build dense lattices for the additive white Gaussian noise (AWGN) channel a long time ago, Belfiore was the first to exploit the multiplicative structure of ideal lattices in Rayleigh fading channels [1]. In cryptography, ideal lattices not only improve the efficiency of lattice-based cryptosystems up to a competitive level [2], but also offer a natural tool for fully-homomorphic encryption where both additions and multiplications are performed.

More recently, we defined the *flatness factor* associated with the lattice Gaussian distribution and derived its many properties [3, 4]. With this new tool, we are now able to answer/address several major open questions in coding theory. For example, Erez and Zamir [5] proposed nested lattice codes achieving the capacity of the power-constrained AWGN channel, where a quantization-good lattice serves as the shaping lattice while an AWGN-good lattice serves as the coding lattice (dithering is also required). In [4], we proposed *lattice Gaussian coding*, where the codebook has a discrete Gaussian distribution over an AWGN-good lattice; this technique considerably simplifies the design of [5]. As another example, in [3] we used the lattice Gaussian distribution to achieve *semantic security* over the Gaussian wiretap channel, which led to the notion of *secrecy-good lattices*. In both cases, we do not need a shaping lattice or a dither. Using ideal lattices and division algebras, we are able to extend this framework to single-antenna and multi-input multi-output (MIMO) fading channels.

In this expository paper, which is based on a similar paper presented at IZS 2014 focusing on the AWGN channel [6][1] and more recent works on fading channels [7, 8], we aim to present a unified framework of lattice coding for capacity and security in wireless communications. Nevertheless, the paper also contains certain new results on fading wiretap channels. In Section II, we review lattice Gaussian distributions, the flatness factor and ideal lattices, where we generalize several definitions to the complex setting. Section III is devoted to achieving capacity of the AWGN, fading and MIMO channels. Section IV gives coding schemes for wiretap channels, where the fine code is a Gaussian-shaped lattice achieving the capacity of the legitimate channel, and the coarse code is a secrecy-good lattice which ensures the information leakage on the eavesdropper's channel is negligible. Section V outlines the prospect of lattice codes in network information theory. In the paper, we try to shed light on the commonality of the schemes for capacity and for secrecy.

Throughout this paper, we use the natural logarithm, denoted by $\log$, and information is measured in nats.

II. Background

A. Channel Model

In the general form, our framework is able to tackle the compound MIMO channel; specializing this model we will obtain the block fading channel and the AWGN channel. More precisely, we consider an $n \times n$ MIMO channel described by the equation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{1}$$

where $\mathbf{H} \in \mathbb{C}^{n\times n}$ is the channel matrix, and $\mathbf{x} \in \mathbb{C}^n$ is the input subject to the power constraint $E[\mathbf{x}^\dagger \mathbf{x}] \leq nP$. The

[1]Note that authors retain copyright of their work at International Zurich Seminar on Communications (IZS), http://www.izs.ethz.ch/.

noise entries of $\mathbf{w}$ are circularly symmetric complex Gaussian with zero-mean and variance σ_w^2. The signal-to-noise ratio (SNR) per receive antenna is defined by $\mathsf{SNR} = nP/\sigma_w^2$. Assuming that the receiver has complete knowledge of $\mathbf{H}$ (but the transmitter does not have CSIT), which is fixed during a whole transmission block. Consider the set $\mathbb{H}$ of all channel matrices with fixed white-input capacity C:

$$\mathbb{H} = \left\{\mathbf{H} \in \mathbb{C}^{n\times n} : \log\det\left(\mathbf{I} + \mathsf{SNR}\,\mathbf{H}^\dagger\mathbf{H}\right) = C\right\}. \quad (2)$$

This can be viewed as a compound channel with capacity C. The compound channel model (2) arises in several important scenarios in communications, such as the outage formulation in the open-loop mode and broadcast [9].

The compound channel demands a universal code that achieves the capacity for all members $\mathbf{H} \in \mathbb{H}$. This represents one of the most difficult problems in coding theory. Note that (2) reduces to a compound block fading channel if $\mathbf{H}$ is diagonal (here n denotes the number of blocks), and to the AWGN channel if $\mathbf{H} = \mathbf{I}$.

B. Lattice Gaussian Distribution

In this subsection, we generalize the lattice Gaussian distribution from $\mathbb{Z}$-lattices to $\mathbb{Z}[i]$-lattices[2]. Everything is formally the same as its real counterpart in [3], and the difference is a factor 2 in most cases.

An n-dimensional $\mathbb{Z}[i]$-lattice Λ in the Euclidean space $\mathbb{C}^n$ is a set defined by

$$\Lambda = \mathcal{L}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} : \mathbf{x} \in \mathbb{Z}[i]^n\}$$

where $\mathbf{B} \in \mathbb{C}^{n\times n}$ is the generator matrix. The dual lattice Λ^* of a lattice Λ is defined as the set of vectors $\mathbf{v} \in \mathbb{C}^n$ such that $\langle \mathbf{v}, \boldsymbol{\lambda}\rangle = \mathbf{v}^\dagger\boldsymbol{\lambda} \in \mathbb{Z}[i]$, for all $\boldsymbol{\lambda} \in \Lambda$. The volume of Λ is defined as that of its real equivalent: $V(\Lambda) = |\det\mathbf{B}|^2$.

For $\sigma > 0$ and $\mathbf{c} \in \mathbb{C}^n$, the continuous Gaussian distribution of covariance matrix $\boldsymbol{\Sigma}$ centered at $\mathbf{c}$ is given by

$$f_{\sqrt{\boldsymbol{\Sigma}},\mathbf{c}}(\mathbf{x}) = \frac{1}{\pi^n\det(\boldsymbol{\Sigma})}e^{-(\mathbf{x}-\mathbf{c})^\dagger\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mathbf{c})},$$

for $\mathbf{x} \in \mathbb{C}^n$. For convenience, we write $f_{\sqrt{\boldsymbol{\Sigma}}}(\mathbf{x}) = f_{\sqrt{\boldsymbol{\Sigma}},\mathbf{0}}(\mathbf{x})$.

Consider the Λ-periodic function

$$f_{\sqrt{\boldsymbol{\Sigma}},\Lambda}(\mathbf{x}) = \sum_{\boldsymbol{\lambda}\in\Lambda} f_{\sqrt{\boldsymbol{\Sigma}},\boldsymbol{\lambda}}(\mathbf{x}) = \frac{1}{\pi^n\det(\boldsymbol{\Sigma})}\sum_{\boldsymbol{\lambda}\in\Lambda} e^{-(\mathbf{x}-\mathbf{c})^\dagger\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mathbf{c})} \quad (3)$$

for all $\mathbf{x} \in \mathbb{C}^n$. Observe that $f_{\sigma,\Lambda}$ restricted to a fundamental region $\mathcal{R}(\Lambda)$ is a probability density.

We define the *discrete Gaussian distribution* over Λ centered at $\mathbf{c} \in \mathbb{C}^n$ as the following discrete distribution taking values in $\boldsymbol{\lambda} \in \Lambda$:

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

where $f_{\sqrt{\boldsymbol{\Sigma}},\mathbf{c}}(\Lambda) \triangleq \sum_{\boldsymbol{\lambda}\in\Lambda} f_{\sqrt{\boldsymbol{\Sigma}},\mathbf{c}}(\boldsymbol{\lambda}) = f_{\sqrt{\boldsymbol{\Sigma}},\Lambda}(\mathbf{c})$. Again for convenience, we write $D_{\Lambda,\sqrt{\boldsymbol{\Sigma}}} = D_{\Lambda,\sqrt{\boldsymbol{\Sigma}},\mathbf{0}}$.

[2]Extension to $\mathbb{Z}[j]$ or other imaginary quadratic fields $\mathbb{Q}(\sqrt{-d})$ is possible.

The flatness factor of a lattice Λ quantifies the maximum variation of $f_{\sqrt{\boldsymbol{\Sigma}},\Lambda}(\mathbf{x})$ for $\mathbf{x} \in \mathbb{C}^n$.

Definition 1 (Flatness factor)**.** *For a lattice Λ and for covariance matrix $\boldsymbol{\Sigma}$, the flatness factor is defined by:*

$$\epsilon_\Lambda(\sqrt{\boldsymbol{\Sigma}}) \triangleq \max_{\mathbf{x}\in\mathcal{R}(\Lambda)} \left|V(\Lambda)f_{\sqrt{\boldsymbol{\Sigma}},\Lambda}(\mathbf{x}) - 1\right|.$$

In words, $\frac{f_{\sqrt{\boldsymbol{\Sigma}},\Lambda}(\mathbf{x})}{1/V(\Lambda)}$, the ratio between $f_{\sqrt{\boldsymbol{\Sigma}},\Lambda}(\mathbf{x})$ and the uniform distribution over $\mathcal{R}(\Lambda)$, is within the range $[1 - \epsilon_\Lambda(\sqrt{\boldsymbol{\Sigma}}), 1 + \epsilon_\Lambda(\sqrt{\boldsymbol{\Sigma}})]$.

Proposition 1 (Expression of $\epsilon_\Lambda(\sqrt{\boldsymbol{\Sigma}})$)**.** *We have:*

$$\begin{aligned}\epsilon_\Lambda(\sqrt{\boldsymbol{\Sigma}}) &= \frac{V(\Lambda)}{\pi^n\det(\boldsymbol{\Sigma})}\sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\boldsymbol{\lambda}^\dagger\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}} \\ &= \sum_{\boldsymbol{\lambda}^*\in\Lambda^*} e^{-\pi^2\boldsymbol{\lambda}^\dagger\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}} - 1.\end{aligned}$$

In particular, if $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$, then

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

where $\gamma_\Lambda(\sigma) = \frac{V(\Lambda)^{1/n}}{\sigma^2}$ is the volume-to-noise ratio (VNR), and $\Theta_\Lambda(\tau) = \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\pi\tau\|\boldsymbol{\lambda}\|^2}$ is the theta series.

The following result guarantees the existence of sequences of Construction-A lattices whose flatness factors can vanish as $n \to \infty$.

Theorem 1 (Minkowski-Hlawka)**.** *$\forall\sigma > 0$ and $\forall\delta > 0$, there exists a sequence of lattices $\Lambda^{(n)}$ such that*

$$\epsilon_{\Lambda^{(n)}}(\sigma) \le (1+\delta)\cdot\left(\frac{\gamma_{\Lambda^{(n)}}(\sigma)}{\pi}\right)^n, \quad (4)$$

i.e., the flatness factor can go to zero exponentially for any fixed VNR $\gamma_{\Lambda^{(n)}}(\sigma) < \pi$. More generally, $\epsilon_\Lambda(\sqrt{\boldsymbol{\Sigma}}) \to 0$ if the generalized VNR $\gamma_{\Lambda^{(n)}}(\sqrt{\boldsymbol{\Sigma}}) = \frac{V(\Lambda)^{1/n}}{\det(\boldsymbol{\Sigma})^{1/n}} < \pi$.

The significance of a small flatness factor is two-fold. Firstly, it assures the "folded" distribution $f_{\sqrt{\boldsymbol{\Sigma}},\Lambda}(\mathbf{x})$ is flat; secondly, it implies the discrete Gaussian distribution $D_{\Lambda,\sqrt{\boldsymbol{\Sigma}},\mathbf{c}}$ is "smooth". We refer the reader to [3, 4] for more details.

The following lemma is a generalization of Regev's and is particularly useful for communications and security [8].

Lemma 1. *Given $\mathbf{x}_1$ sampled from discrete Gaussian distribution $D_{\Lambda+\mathbf{c},\sqrt{\boldsymbol{\Sigma}_1}}$ and $\mathbf{x}_2$ sampled from continuous Gaussian distribution $f_{\sqrt{\boldsymbol{\Sigma}_2}}$. Let $\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$ and let $\boldsymbol{\Sigma}_3^{-1} = \boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1}$. If $\epsilon_\Lambda(\sqrt{\boldsymbol{\Sigma}_3}) \le \varepsilon \le \frac{1}{2}$, then the distribution g of $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ is close to $f_{\sqrt{\boldsymbol{\Sigma}_0}}$:*

$$g(\mathbf{x}) \in f_{\sqrt{\boldsymbol{\Sigma}_0}}(\mathbf{x})\left[1 - 4\varepsilon, 1 + 4\varepsilon\right].$$

This lemma has profound implications. On one hand, it implies capacity, i.e., the discrete Gaussian distribution over a lattice is almost capacity-achieving if the flatness factor is small [4]. On the other hand, it implies security, i.e., Eve's signal is indistinguishable from a continuous Gaussian distribution.

C. Ideal Lattices

We refer to [10] for an introduction to algebraic number theory for coding. Consider a *relative extension* $K/\mathbb{Q}(i)$ of degree n. There are n homomorphisms $\sigma_1, \ldots, \sigma_n$ that embed K into $\mathbb{C}$. The *ring of integers* of K is denoted by $\mathcal{O}_K$, and its invertible elements are called *units*. The map $\sigma : K \to \mathbb{C}^n$, $\sigma(x) = (\sigma_1(x), \ldots, \sigma_n(x))$ is called the *canonical embedding*. It takes $\mathcal{O}_K$ into a lattice in $\mathbb{C}^n$. The volume of this lattice is given by $V(\Lambda) = 2^{-n}\sqrt{\Delta_K}$ where Δ_K is the *absolute discriminant* of K.

Under the canonical embedding, an ideal of $\mathcal{O}_K$ becomes an ideal lattice. Any ideal can be decomposed as a product of prime ideals. Let $\mathfrak{p}_i$ be a prime ideal. It follows that $\mathcal{O}_K/\mathfrak{p}_i \simeq \mathbb{F}_{p^l}$, for some prime p and some integer l.

III. Achieving Channel Capacity

A. AWGN Channel

Consider the classic AWGN channel $\mathbf{y} = \mathbf{x} + \mathbf{w}$ where the vectors have dimension T, the codeword length. In [4], we proposed a new coding scheme based on the lattice Gaussian distribution. Let Λ be an AWGN-good lattice in $\mathbb{C}^T$ of dimension T, whose error probability vanishes if the VNR $\frac{V(\Lambda)^{1/T}}{\sigma_w^2} > \pi e$. The encoder maps the information bits to points in Λ, which obey the lattice Gaussian distribution

$$\mathbf{x} \sim D_{\Lambda,\sigma_s}.$$

Since the continuous Gaussian distribution is capacity-achieving, we want the lattice Gaussian distribution to behave like the continuous Gaussian distribution (in particular $P \approx \sigma_s^2$). This can be assured by a small flatness factor. Thus, while we are concerned with the discrete distribution D_{Λ,σ_s}, we in fact require the associated periodic distribution $f_{\sigma_s,\Lambda}$ to be flat.

Since the lattice points are not equally probable a priori in the lattice Gaussian coding, we will use maximum-a-posteriori (MAP) decoding. In [3], it was shown that MAP decoding is equivalent to Euclidean lattice decoding of Λ using a scaling coefficient $\alpha = \frac{\sigma_s^2}{\sigma_s^2+\sigma_w^2}$, which is asymptotically equal to the MMSE coefficient $\frac{P}{P+\sigma_w^2}$. In fact, the error probability of the proposed scheme under MMSE lattice decoding admits almost the same expression as that of Poltyrev, with σ_w replaced by $\tilde{\sigma}_w = \frac{\sigma_s\sigma_w}{\sqrt{\sigma_s^2+\sigma_w^2}}$. To satisfy AWGN-goodness, we choose the fundamental volume $V(\Lambda)$ such that

$$V(\Lambda)^{1/T} > \pi e \tilde{\sigma}_w^2. \tag{5}$$

Meanwhile, the rate of the scheme is given by the entropy of the lattice Gaussian distribution:

$$\begin{aligned} R &\to \log(\pi e\sigma_s^2) - \frac{1}{T}\log V(\Lambda) \\ &< \log(\pi e\sigma_s^2) - \log\left(\pi e \frac{\sigma_s^2\sigma_w^2}{\sigma_s^2+\sigma_w^2}\right) \\ &= \log\left(1+\frac{\sigma_s^2}{\sigma_w^2}\right) \\ &\to \log(1+\mathsf{SNR}). \end{aligned}$$

Combining these results, we arrive at the following theorem.

Theorem 2 (Coding theorem)**.** *Consider a lattice code whose codewords are drawn from the discrete Gaussian distribution D_{Λ,σ_s} for an AWGN-good lattice Λ. Any rate up to the channel capacity $\log(1+\mathsf{SNR})$ is achievable, while the error probability of MMSE lattice decoding vanishes exponentially fast.*

B. Block Fading Channel

Recall the coded system model

$$\underbrace{\mathbf{Y}}_{n\times T} = \underbrace{\mathbf{H}}_{n\times n}\underbrace{\mathbf{X}}_{n\times T} + \underbrace{\mathbf{W}}_{n\times T} \tag{6}$$

where $\mathbf{H}$ is diagonal and T is the coherence time (codeword length). Vectorizing this equation, we obtain

$$\underbrace{\mathbf{y}}_{nT\times 1} = \underbrace{\mathcal{H}}_{nT\times nT}\underbrace{\mathbf{x}}_{nT\times 1} + \underbrace{\mathbf{w}}_{nT\times 1} \tag{7}$$

where $\mathcal{H} = \mathbf{I}_T \otimes \mathbf{H}$. Now we design a coding lattice $\Lambda \subset \mathbb{C}^{nT}$ so that $\mathbf{x} \in \Lambda$. With Gaussian shaping, the problem of achieving capacity of compound block fading channels boils down to finding a lattice that is good for block fading.

Definition 2 (Fading-good lattices [7])**.** *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. $|\det\mathbf{H}| = D$, $P_e(\Lambda,\mathbf{H}) \to 0$ as $T \to \infty$.*

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

$$\Lambda = \mathcal{C} + \mathfrak{p}^T. \tag{8}$$

Note that it reduces to usual Construction A $\Lambda = \mathcal{C} + p^T$ when $K = \mathbb{Q}$.

The existence of a universal lattice can be proven by the Minkowski-Hlawka theorem, i.e., averaging over random codes $\mathcal{C}$ (with $p \to \infty$). The proof is tricky and relies on the unit group, which compacts the set of quantized channels. Thus, generalized Construction A is good for block fading [7]. Then, with MMSE-GDFE lattice decoding for Gaussian shaping, it can be shown that the average error probability $\mathbb{E}_\Lambda[P_e(\Lambda)]$ vanishes as long as the VNR $> \pi e$ (as $T \to \infty$):

$$\frac{|\mathbf{I}+\mathsf{SNR}\,\mathbf{H}^\dagger\mathbf{H}|^{\frac{1}{n}}V(\Lambda)^{\frac{1}{nT}}}{\sigma_s^2} > \pi e. \tag{9}$$

Thus, any rate

$$R \to n\log(\pi e\sigma_s^2) - \frac{1}{T}\log(V(\Lambda)) < \log|\mathbf{I}+\mathsf{SNR}\,\mathbf{H}^T\mathbf{H}| = C$$

is achievable. Note that the achievable rate only depends on $\mathbf{H}$ through determinant $|\mathbf{I}+\mathsf{SNR}\,\mathbf{H}^\dagger\mathbf{H}|$.

C. MIMO Fading Channel

The case of MIMO channels is more technical due to non-commutativity of the underlying algebra. Let $\mathcal{O}$ be the natural order of cyclic division algebra $\mathcal{A}$. Take a two-sided ideal $\mathcal{J}$ of $\mathcal{O}$ and consider the quotient ring $\mathcal{O}/\mathcal{J}$. Define a reduction $\beta : \mathcal{O} \to \mathcal{O}/\mathcal{J}$. For a linear code $\mathcal{C}$ over $\mathcal{O}/\mathcal{J}$, $\beta^{-1}(\mathcal{C})$ is a lattice Λ (in $\mathbb{C}^{n^2T}$). However, the quotient ring $\mathcal{O}/\mathcal{J}$ is non-commutative in general, e.g., a matrix ring, skew polynomial ring etc. Nevertheless it is still possible to prove the Minkowski-Hlawka theorem using codes over rings. Thus, there exists a sequence of lattices universally good for MIMO fading, hence achieving the capacity of compound MIMO channels. Note that recently [9] and [11] have achieved a constant gap to the capacity of compound MIMO channels.

IV. Approaching Secrecy Capacity

A. Gaussian Wiretap Channel

Now consider the Gaussian wiretap channel where Alice and Bob are the legitimate users, while Eve is an eavesdropper. The outputs $\mathbf{y}$ and $\mathbf{z}$ at Bob and Eve's ends are respectively given by

$$\mathbf{y} = \mathbf{x} + \mathbf{w}_b, \quad \mathbf{z} = \mathbf{x} + \mathbf{w}_e, \tag{10}$$

where $\mathbf{w}_b$, $\mathbf{w}_e$ are T-dimensional Gaussian noise vectors with zero mean and variance σ_b^2, σ_e^2 respectively.

For secrecy rate R_s, we use coset coding induced by a lattice partition $\Lambda_e \subset \Lambda_b$ such that

$$\frac{1}{T}\log|\Lambda_b/\Lambda_e| = R_s.$$

The fine lattice Λ_b is the usual coding lattice for Bob, i.e., it is an AWGN-good lattice. The coarse lattice Λ_e is new, and turns out to be a secrecy-good lattice. To encode, Alice uses the secret bits to select one coset of Λ_e and transmits a random point inside this coset.

Consider a message set $\mathcal{M} = \{1, \ldots, e^{TR_s}\}$, and a one-to-one function $\phi : \mathcal{M} \to \Lambda_b/\Lambda_e$ which associates each message $m \in \mathcal{M}$ to a coset $\widetilde{\boldsymbol{\lambda}}_m \in \Lambda_b/\Lambda_e$. One could choose the coset representative $\boldsymbol{\lambda}_m \in \Lambda_b \cap \mathcal{R}(\Lambda_e)$ for any fundamental region $\mathcal{R}(\Lambda_e)$. In order to encode the message $m \in \mathcal{M}$, Alice samples $\mathbf{x}_m$ from lattice Gaussian distribution

$$\mathbf{x}_m \sim D_{\Lambda_e + \boldsymbol{\lambda}_m, \sigma_s}.$$

Let $\tilde{\sigma}_e = \frac{\sigma_s\sigma_e}{\sqrt{\sigma_s^2+\sigma_e^2}}$ and $\sigma_s' = \sqrt{\sigma_s^2+\sigma_e^2}$. Regev's Lemma (cf. Lemma 1) implies that if $\epsilon_{\Lambda_e}(\tilde{\sigma}_e) < \frac{1}{2}$, then:

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

We see that the received signals converge to the same Gaussian distribution $f_{\sigma_s'}$. This already gives *distinguishing security*, which means that, asymptotically, the channel outputs are indistinguishable for different input messages.

An upper bound on the amount of leaked information then follows.

Theorem 3 (Information leakage [3])**.** *Suppose that the wiretap coding scheme described above is employed on the Gaussian wiretap channel (10), and let $\varepsilon_T = \epsilon_{\Lambda_e}(\tilde{\sigma}_e)$. Assume that $\varepsilon_T < \frac{1}{2}$ for all T. Then the mutual information between the confidential message and the eavesdropper's signal is bounded as follows:*

$$\mathbb{I}(\mathsf{M};\mathsf{Z}) \le 8\varepsilon_T T R_s - 8\varepsilon_T \log 8\varepsilon_T. \tag{11}$$

A wiretap coding scheme is secure in the sense of *strong secrecy* if $\lim_{T\to\infty}\mathbb{I}(\mathsf{M};\mathsf{Z}) = 0$. From (11), a flatness factor $\varepsilon_T = o(\frac{1}{T})$ would be enough. In practice, an exponential decay of the information leakage is desired, and this motivates the notion of secrecy-good lattices:

Definition 3 (Secrecy-good lattices)**.** *A sequence of lattices $\Lambda^{(T)}$ is* secrecy-good *if*

$$\epsilon_{\Lambda^{(n)}}(\sigma) = e^{-\Omega(T)}, \quad \forall \gamma_{\Lambda^{(T)}}(\sigma) < \pi. \tag{12}$$

It can be shown that, under mild conditions (similar to those in [3]), the secrecy rate

$$R < \log(1+\mathsf{SNR}_b) - \log(1+\mathsf{SNR}_e) - 1 \tag{13}$$

is achievable, which is within 1 nat from the secrecy capacity. It is worth mentioning that this small gap may be fictitious, due to our proof technique.

B. Fading Wiretap Channel

The channels for Bob and for Eve are given by

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

respectively. We fix the capacity C_e of Eve's compound channel with white inputs

$$\mathbb{H}_e = \left\{\mathbf{H}_e \in \mathbb{C}^{n\times n} : \log\det\left(\mathbf{I} + \mathsf{SNR}_e\,\mathbf{H}_e^\dagger\mathbf{H}_e\right) = C_e\right\}.$$

as well as the capacity C_b of Bob's compound channel. The secrecy capacity of compound MIMO wiretap channels with white inputs is given by [12]:

$$C_s = C_b - C_e. \tag{14}$$

Similarly to lattice coding over the Gaussian wiretap channel, we use a pair of nested lattices $\Lambda_b \subset \Lambda_e$. These lattices are built in the same manner as before:

$$\Lambda_b = \mathcal{C}_b + \mathfrak{p}^T \quad \Lambda_e = \mathcal{C}_e + \mathfrak{p}^T \tag{15}$$

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

In order to encode the message $m \in \mathcal{M}$, Alice samples $\mathbf{x}_m$ from $D_{\Lambda_e+\boldsymbol{\lambda}_m,\sigma_s}$. Similarly to (7), let $\mathcal{H}_e = \mathbf{I}_T \otimes \mathbf{H}_e$ of size nT. Eve observes a discrete Gaussian distribution $D_{\mathcal{H}_e(\Lambda_e+\boldsymbol{\lambda}_m),\mathcal{H}_e\sigma_s}$, contaminated by i.i.d. Gaussian noise of standard deviation σ_e. We would like this to be indistinguishable from a continuous Gaussian distribution of covariance matrix $\boldsymbol{\Sigma}_0 = \sigma_s^2\mathcal{H}_e\mathcal{H}_e^\dagger + \sigma_e^2\mathbf{I}$, regardless of m. By Lemma 1, we need

$$\epsilon_{\mathcal{H}_e\Lambda_e}(\sqrt{\boldsymbol{\Sigma}_3}) \to 0$$

where $\boldsymbol{\Sigma}_3^{-1} = \sigma_s^{-2}(\mathcal{H}_e\mathcal{H}_e^\dagger)^{-1} + \sigma_e^{-2}\mathbf{I}$. In other words, we want the flatness factor $\epsilon_{\mathcal{H}_e\Lambda_e}(\sqrt{\boldsymbol{\Sigma}_3}) = \varepsilon_T$ to vanish with T.

Applying Minkowski-Hlawka, we obtain

$$\begin{aligned} & \mathbb{E}_{\Lambda_e}[\epsilon_{\mathcal{H}_e\Lambda_e}(\sqrt{\mathbf{\Sigma}_3})] \\ = \quad & \frac{V(\Lambda_e)}{\pi^{nT}} \det(\sigma_s^{-2}\mathbf{I} + \sigma_e^{-2}\mathcal{H}_e^\dagger \mathcal{H}_e) \qquad (16) \\ \to \quad & \frac{V(\Lambda_e)}{(\pi\sigma_s^2)^{nT}} \det(\mathbf{I} + \mathsf{SNR}_e\, \mathbf{H}_e^\dagger \mathbf{H}_e)^T. \end{aligned}$$

Now we calculate the information leakage to Eve. If we slightly reduce the VNR of Λ_e, $\mathbb{E}_{\Lambda_e}[\epsilon_{\mathcal{H}_e\Lambda_e}(\sqrt{\mathbf{\Sigma}_3})]$ in (16) will vanish exponentially with T. Similar to the Gaussian wiretap channel (11), the mutual information between Alice and Eve is bounded for all $\mathbf{H}_b$, $\mathbf{H}_e$ as

$$\mathbb{I}(\mathsf{M};\mathsf{Z}) \le 8\varepsilon_T T R_s - 8\varepsilon_T \log(8\varepsilon_T). \qquad (17)$$

Again, it is tricky to exhibit the existence of a universal code for all $\mathbf{H}_b$, $\mathbf{H}_e$. Fortunately, thanks to the unit groups, this can be resolved by quantizing the channels in the same manner as for capacity [12].

For a vanishing flatness factor, we need the condition

$$\frac{\det(\mathbf{I} + \mathsf{SNR}_e\, \mathbf{H}_e^\dagger \mathbf{H}_e)^{1/n} V(\Lambda_e)^{\frac{1}{nT}}}{\sigma_s^2} < \pi. \qquad (18)$$

From (9) and (18), we obtain the secrecy rate

$$R_s < \log \frac{\left|\mathbf{I} + \mathsf{SNR}_b\, \mathbf{H}_b^\dagger \mathbf{H}_b\right|}{\left|\mathbf{I} + \mathsf{SNR}_e\, \mathbf{H}_e^\dagger \mathbf{H}_e\right|} - n = C_b - C_e - n,$$

which is the secrecy capacity to within a constant gap of n nats. Again, this gap may well be fictitious.

Then one may claim the existence of a universal lattice code which achieves the secrecy capacity to within n nats.

C. MIMO Wiretap Channel

The MIMO case is similar, using cyclic division algebra. The security proof is very much the same, except that $\mathbf{H}_b$ and $\mathbf{H}_e$ are full matrices.

V. Multi-Terminal Systems in Network Information Theory

So far we have been concerned with point-to-point channels. The structure of lattice codes is a significant advantage for their applications in wireless networks. Due to the superposition nature of wireless signals, the algebraic structure of lattice codes lead to new possibilities that were not available before. In this section, we review several applications pertaining to future networks.

A. Compute and Forward

In compute-and-forward, the relay computes a linear function of transmitted messages. Here, an AWGN-good lattice suffices (along with Gaussian shaping). All the lattices given in Section III are AWGN-good, although those lattices good for fading and MIMO channels seem to be an overkill. Note that explicit construction of AWGN-good lattices are now available, in particular, polar lattices and LDA lattices with with $O(T \log T)$ complexity.

B. Dirty Paper Coding and Broadcast Channel

In dirty paper coding, a nested lattice code is required where the fine lattice is good for channel coding and the coarse lattice is good for source coding. A construction of polar lattices achieving the *rate-distortion bound* of Gaussian sources is reported in [13]. Combing this lattice with an AWGN-good lattice yields an explicit code achieving the capacity of the dirty paper channel with $O(T \log T)$ complexity. It can be shown that such a dirty paper code in conjunction with beamforming achieves the capacity of the downlink channel [14].

C. Distributed Source Coding

Swapping the roles of the fine and coarse lattices yields a code for the Wyner-Ziv problem. Let X, Y be two jointly Gaussian sources and $X = Y + Z$, where Z is Gaussian and is independent of Y. Given side information Y, the Wyner-Ziv problem is to reconstruct X. An $O(T \log T)$-complexity scheme based on polar lattices is also reported in [13], achieving the rate-distortion bound of Wyner-Ziv coding.

Acknowledgment

This work was supported in part by Huawei Technologies. The author would like to thank Antonio Campello and Jean-Claude Belfiore for helpful discussions.

References

[1] K. Boulle and J.-C. Belfiore, "Modulation schemes designed for the Rayleigh channel," in *Proc. CISS*, Princeton, NJ, pp. 288–293, 1992.

[2] V. Lyubashevsky, C. Peikert, and O. Regev, "On ideal lattices and learning with errors over rings," *J. ACM*, vol. 60, no. 6, pp. 43:1–43:35, Nov. 2013.

[3] C. Ling, L. Luzzi, J.-C. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the Gaussian wiretap channel," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 6399–6416, Oct. 2014.

[4] C. Ling and J.-C. Belfiore, "Achieiving AWGN channel capacity with lattice Gaussian coding," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 5918–5929, Oct. 2014.

[5] U. Erez and R. Zamir, "Achieving $\frac{1}{2}$ log(1+SNR) on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[6] C. Ling and J.-C. Belfiore, "Lattice Gaussian coding for capacity and secrecy: Two sides of one coin," in *International Zurich Seminar on Communications 2014*.

[7] A. Campello, C. Ling, and J.-C. Belfiore, "Algebraic lattice codes achieve the capacity of the compound block-fading channel," in *IEEE Int. Symp. Inform. Theory (ISIT)*, 2016.

[8] L. Luzzi, C. Ling, and R. Vehkalahti, "Almost universal codes for fading wiretap channels," *CoRR*, vol. abs/1601.02391, 2016. [Online]. Available: http://arxiv.org/abs/1601.02391

[9] O. Ordentlich and U. Erez, "Precoded integer-forcing universally achieves the MIMO capacity to within a constant gap," *IEEE Trans. Inform. Theory*, vol. 61, no. 1, pp. 323–340, Jan 2015.

[10] F. Oggier and E. Viterbo, "Algebraic number theory and code design for Rayleigh fading channels," *Foundations and Trends on Communications and Information Theory*, vol. 1, pp. 336–415, 2004.

[11] L. Luzzi and R. Vehkalahti, "Almost universal codes achieving ergodic MIMO capacity within a constant gap," *CoRR*, vol. abs/1507.07395, 2015. [Online]. Available: http://arxiv.org/abs/1507.07395

[12] R. F. Schaefer and S. Loyka, "The secrecy capacity of compound Gaussian MIMO wiretap channels," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5535–5552, Oct 2015.

[13] L. Liu and C. Ling, "Polar lattices are good for lossy compression," *CoRR*, vol. abs/1501.05683, 2015. [Online]. Available: http://arxiv.org/abs/1501.05683

[14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge University Press, 2005.