Algebraic Lattice Codes Achieve the Capacity of the Compound Block-Fading Channel

Antonio Campello
Dept. Communications et Electronique
Télécom ParisTech
Paris, France
campello@telecom-paristech.fr

Cong Ling
Department of Electrical
and Electronic Engineering
Imperial College London, U.K.
cling@ieee.org

Jean-Claude Belfiore
Mathematical and Algorithmic Sciences Lab
France Research Center
Huawei Technologies
belfiore@telecom-paristech.fr

***Abstract*—We propose a coding scheme that achieves the capacity of the compound block-fading channel with lattice decoding at the receiver. Our lattice construction exploits the multiplicative structure of number fields and their group of units to absorb ill-conditioned channel realizations. To shape the constellation, a discrete Gaussian distribution over the lattice points is applied. A by-product of our results is the proof that the lattice Gaussian distribution is capacity-achieving in the AWGN channel for any signal-to-noise ratio.**

I. Introduction

We consider a real[1] block-fading channel described by the equation

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

where $\mathbf{H} \in \mathbb{R}^{n\times n}$ is a diagonal matrix, and $\mathbf{x} \in \mathbb{R}^n$ is the input subject to the constraint $E[\mathbf{x}^{\text{tr}}\mathbf{x}] \leq nP$. The noise entries are Gaussian with zero-mean and variance σ_c^2. Assuming that the receiver has complete knowledge of $\mathbf{H}$, which is fixed for the whole transmission, the (white AWGN input) achievable rates of such model is

$$C = \frac{1}{2}\log\det\left(I + \text{SNR}\mathbf{H}^{\text{tr}}\mathbf{H}\right). \tag{2}$$

Let $\mathbb{H}$ be the set of all channel matrices with fixed capacity:

$$\mathbb{H} = \left\{\mathbf{H} \in \mathbb{R}^{n\times n} : \mathbf{H} \text{ is diagonal and } \frac{1}{2}\log\det\left(I + \text{SNR}\mathbf{H}^{\text{tr}}\mathbf{H}\right) = C\right\}. \tag{3}$$

We say that a sequence of codes is *universal* or *achieves the capacity of the compound model* for the block-fading channel if, for all $H \in \mathbb{H}$ the error probability vanishes, as $T \to \infty$, with rate R arbitrarily close to C. In this work we construct universal algebraic lattice codes for the block-fading channel.

Initial research on lattice codes for fading channels was concerned with the diversity order and minimum product distance [2]. Recently, [3] and [4] have achieved a constant gap to the capacity in the broader scope of compound MIMO channels. The work [5] showed the existence of lattice codes achieving the capacity of a fixed nonrandom MIMO channel (but not that of the compound channel - see Section IV for more details). Further, [6], [7] examined the diversity order of lattice codes, in the infinite-constellation setting, for MIMO and block-fading channels, respectively. The Poltyrev limit and dispersion on ergodic fading channels were studied in [8].

The notion of compound MIMO channels dates back to [9]. The authors provide a technique to convert traditional random codes into universal ones, under the additional assumption that the norm of $\mathbf{H}$ is bounded (see also [10]). In other words, it is shown how to achieve the compound capacity for $\mathbb{H} \cap \mathcal{S}$, where $\mathcal{S}$ is a compact set. However, the question to achieve the capacity of the compound fading channel remains open.

In this paper, we make a step towards this goal by proving that lattice codes from Generalized Construction A achieve the capacity of the compound block-fading channel over the entire space of channels (3) (the case of ergodic fading and extensions to MIMO channels will be addressed in a forthcoming journal paper). This represents an advantage of ideal lattices over the classic Gaussian random codes [9], [10] and standard Construction A [5]. This is made possible by exploiting the multiplicative structure of number fields and their group of units. Similar techniques had previously demonstrated good simulation performance in the fast fading channel with efficient decoding [11].

As in [1], we employ the lattice Gaussian distribution to shape the constellation. A technical novelty of the present work is the error probability analysis via the properties of *sub-Gaussian* random variables. Due to sub-Gaussianity, and as a corollary of our results for the block-fading channel, we were able to remove a threshold SNR assumption and simplify the analysis of the coding scheme [1].

This work is organized as follows. In Section II we introduce the notation and some main definitions, including that of good universal infinite lattices. In sections III and IV we show how to achieve the infinite compound capacity with algebraic lattices, and argue that traditional mod-p lattices fail to do so. In Section V the lattice Gaussian distribution is used to achieve the power-constrained model capacity.

II. Notation and Initial Definitions

The channel equation (1) after T uses can be written in matrix form:

$$\mathbf{Y}_{n\times T} = \mathbf{H}_{n\times n}\mathbf{X}_{n\times T} + \mathbf{Z}_{n\times T} \tag{4}$$

[1] All our results extend naturally to the complex model with some appropriate adaptations.

or in vectorized form

$$\mathbf{y}_{nT} = \mathcal{H}_{nT}\mathbf{x}_{nT} + \mathbf{z}_{nT}, \tag{5}$$

where $\mathcal{H} = I_T \otimes \mathbf{H}$, and I_T is the identity matrix $T \times T$. We omit the indices when it is clear from the context. From now on, $\|\mathbf{H}\| = \sqrt{\text{trace}(\mathbf{H}^{\text{tr}}\mathbf{H})}$ denotes the Frobenius norm of $\mathbf{H}$.

A. Lattice Codes

Let Λ be a full-rank lattice in $\mathbb{R}^n$. The *Voronoi region* of a point $\mathbf{x} \in \Lambda$ is defined as

$$\mathcal{V}_\Lambda(\mathbf{x}) \triangleq \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x}-\mathbf{y}\| \leq \|\bar{\mathbf{x}}-\mathbf{y}\| \text{ for all } \bar{\mathbf{x}} \in \Lambda\}.$$

Throughout the text, we write $\mathcal{V}_\Lambda = \mathcal{V}_\Lambda(\mathbf{0})$. The volume of Λ is defined as the volume of its Voronoi region and denoted by $V(\Lambda)$. Given $\sigma > 0$, the *volume-to-noise ratio* (VNR) of a lattice is defined as $\gamma_\Lambda(\sigma) = V(\Lambda)^{2/n}/\sigma^2$.

For $\sigma > 0$ and $\mathbf{c} \in \mathbb{R}^n$, we define the Gaussian distribution of variance σ^2 centered at $\mathbf{c} \in \mathbb{R}^n$ as

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

for all $\mathbf{x} \in \mathbb{R}^n$. For convenience, we write $f_\sigma(\mathbf{x}) = f_{\sigma,\mathbf{0}}(\mathbf{x})$.

We also consider the Λ-periodic function

$$f_{\sigma,\Lambda}(\mathbf{x}) = \sum_{\lambda\in\Lambda} f_{\sigma,\lambda}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\lambda\in\Lambda} e^{-\frac{\|\mathbf{x}-\lambda\|^2}{2\sigma^2}}, \tag{6}$$

for all $\mathbf{x} \in \mathbb{R}^n$. Observe that $f_{\sigma,\Lambda}$ restricted to the Voronoi region $\mathcal{V}_\Lambda$ is a probability density.

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

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

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

Lemma 1 in [1] shows that each component of $\mathbf{x} \sim D_{\Lambda-\mathbf{c},\sigma}$ has an average power always less than σ^2. Discrete and continuous Gaussian distributions share similar properties, if the *flatness factor* is small.

Definition 1 (Flatness factor [12])**.** *For a lattice Λ and for a parameter σ, the flatness factor is defined by:*

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

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

B. Infinite Compound Model

Since our scheme is divided in two parts, shaping and coding, we first define a compound model for the infinite constellation, analogous to the Poltyrev limit [13] for Gaussian channels. Let

$$\mathbb{H}_\infty = \left\{\mathbf{H} \in \mathbb{R}^{n\times n} : \mathbf{H} \text{ diagonal and} \det \mathbf{H}^{\text{tr}}\mathbf{H} = D\right\}, \tag{7}$$

where D is a positive constant. Consider a lattice $\Lambda \subset \mathbb{R}^{n\times T}$ with vectors written in matrix-form (4). The error probability of Λ, given $\mathbf{H}$, is denoted by $P_e(\Lambda, \mathbf{H})$.

Definition 2. *We say that a sequence of lattices Λ_T of increasing dimension is universally good for the block-fading channel if for any $\gamma_{\Lambda_T}(\sigma) > \frac{2\pi e}{D^{1/n}}$ and all $\mathbf{H} \in \mathbb{H}_\infty$, $P_e(\Lambda_T, \mathbf{H}) \to 0$.*

Notice that the condition on the VNR is equivalent to $\gamma_{\mathbf{H}\Lambda_T}(\sigma) > 2\pi e$. We stress that this definition requires a sequence of lattices to be simultaneously good for *all* channels in the set. For a fixed $\mathbf{H}$, this requirement is not different from the original Gaussian channel coding problem. However, as shown in the end of Section IV traditional codes [13] fail to achieve the compound capacity.

III. Construction A from $\mathcal{O}_K$-lattices

We follow closely the construction of [14]. For an introduction to the algebraic theory used in this section, the reader is referred to [2]. We describe in the next subsection some main concepts used throughout the paper.

A. Basic Notation

We consider *algebraic number fields* $K/\mathbb{Q}$, i.e. extensions of $\mathbb{Q}$ with finite degree n. There are n homomorphisms $\sigma_1, \ldots, \sigma_n$ that embed K into $\mathbb{C}$. If the images of all these embeddings are contained in $\mathbb{R}$, we say that K is a *totally real* extension. The *ring of integers* of K is denoted by $\mathcal{O}_K$, and its invertible elements are called *units*. The map $\sigma : K \to \mathbb{R}^n$, $\sigma(x) = (\sigma_1(x), \ldots, \sigma_n(x))$ is called the *canonical embedding*. It takes $\mathcal{O}_K$ into a lattice in $\mathbb{R}^n$. The squared volume of this lattice, Δ_K, is the *discriminant* of K.

Any ideal $\mathfrak{p}$ of $\mathcal{O}_K$ can be decomposed as the product of prime ideals. Let p be a prime number and consider the decomposition $p\mathcal{O}_K = \prod_{i=1}^g \mathfrak{p}_i^{e_i}$. We say that $\mathfrak{p}_i$ is *above* p. It follows that $\mathcal{O}_K/\mathfrak{p}_i \simeq \mathbb{F}_{p^l}$, for some l. When $g = n$, $l = 1$, and we say that p *splits*.

B. Construction A

From now on, let $K/\mathbb{Q}$ be a totally real extension with $[K : \mathbb{Q}] = n$. Let $\mathfrak{p} \subset \mathcal{O}_K$ be a prime ideal above p, so that there exists an isomorphism $\phi : \mathcal{O}_K/\mathfrak{p} \to \mathbb{F}_{p^l}$. Denote by π the canonical projection $\pi : \mathcal{O}_k \to \mathcal{O}_K/\mathfrak{p}$. The $\mathcal{O}_K$-lattice Λ associated to a linear (T,k)-code $\mathcal{C} \subset \mathbb{F}_{p^l}^T$ is defined as the pre-image by $\phi\circ\pi$ of $\mathcal{C}$ (ϕ and π are applied componentwise):

$$\Lambda^{\mathcal{O}_K}(\mathcal{C}) = \pi^{-1} \circ \phi^{-1}(\mathcal{C}). \tag{8}$$

If $\mathcal{C}$ is linear, $\Lambda^{\mathcal{O}_K}(\mathcal{C})$ is a lattice and $\Lambda^{\mathcal{O}_K}(\mathcal{C})/\mathfrak{p}^T \simeq \mathcal{C}$. The associated real lattice Λ is obtained by applying (elementwise) the canonical embedding $\sigma : K \to \mathbb{R}^n$. It follows that an element $y = \sigma(\mathbf{x})$, with $\mathbf{x} \in \mathcal{O}_K^T$, belongs to Λ if and only if $(\phi \circ \pi)(\mathbf{x}) \in \mathcal{C}$. The equivalent real lattice has volume $p^{l(T-k)}\sqrt{\Delta_K}^T$.

Following steps of [13] and [15, Appendix B][2], the set of such lattices satisfies, asymptotically, the Minkowski-Hlawka theorem, as $p \to \infty$. More formally, let $\lambda > 0$ be a scaling factor and $\alpha = (\lambda^{-1} p^{l(T-k)} \sqrt{\Delta_K}^T)^{1/nT}$. Consider a function $f : \mathbb{R}^{nT} \to \mathbb{R}$ bounded outside a compact support. The ensemble

$$\mathbb{L}_{K,T,k,p,\lambda} = \left\{ \frac{1}{\alpha} \Lambda_K(\mathcal{C}) : \mathcal{C} \text{ is an } (T,k,p) \text{ code} \right\} \quad (9)$$

satisfies

$$\lim_{p\to\infty} E_{\mathbb{L}_{K,T,k,p,\lambda}} \left[\sum_{\mathbf{x} \in \Lambda(\mathcal{C}) \setminus \{\mathbf{0}\}} f(\mathbf{x}) \right] = \lambda \int_{\mathbb{R}^{nT}} f(\mathbf{x}) \mathrm{d}\mathbf{x}. \quad (10)$$

All lattices in the ensemble have volume $1/\lambda$. We can also visualize their vectors in matrix form

$$\mathbf{X}_{n\times T} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix}$$

with $\mathbf{x}_i = \sigma_i(\mathbf{x}), \mathbf{x} \in \mathcal{O}_K^T$. These lattices are closed by multiplication by units , i.e., if u is a unit in $\mathcal{O}_k$, then $u\Lambda^{\mathcal{O}_K}(\mathcal{C}) = \Lambda^{\mathcal{O}_K}(\mathcal{C})$. In matrix form, this reads $\mathbf{U}\Lambda = \Lambda$, where $\mathbf{U} = \mathrm{diag}(\sigma_1(u), \ldots, \sigma_n(u))$.

IV. Coding: Poltyrev Limit for Block-Fading Channels

Suppose that $\mathbf{H} \in \mathbb{H}_\infty$ (Eq. (7)), and let $\tilde{\mathbb{H}}_\infty = \mathbb{H}_\infty / D^{1/2n}$. To achieve the infinite compound capacity, we first show how to "compactify" $\mathbb{H}_\infty$.

A. Quantizing the Channel Coefficients

We use the group of units of $\mathcal{O}_K$ to quantize the channel coefficients. Let $\mathbf{U} = \mathrm{diag}(\sigma_1(u), \ldots, \sigma_n(u))$ be the diagonal matrix corresponding to the embedding of a unit. Let $\mathcal{U}$ be the set of all possible matrices $\mathbf{U}$. For a normalized channel matrix $\tilde{\mathbf{H}} = \mathbf{H}/D^{1/2n}$, we define

$$\mathbf{U}_{\tilde{H}} = \arg\min_{\mathbf{U}\in\mathcal{U}} \left\| \tilde{\mathbf{H}}\mathbf{U}^{-1} \right\|, \quad (11)$$

with ties broken in a systematic manner. The association $\tilde{\mathbf{H}} \to \mathbf{U}_{\tilde{H}}$ defines an equivalence relation. By quotienting $\tilde{\mathbb{H}}_\infty$ by this relation, we obtain the equivalence classes associated to the error matrices $\mathbf{E}_H = \tilde{\mathbf{H}}\mathbf{U}_{\tilde{H}}^{-1}$. Let

$$\mathcal{E} = \left\{ \mathbf{E}_{\tilde{H}} = \tilde{\mathbf{H}}\mathbf{U}_{\tilde{H}}^{-1} : \tilde{\mathbf{H}} \in \tilde{\mathbb{H}} \right\} \quad (12)$$

be the set of all the possible error matrices. In what follows we argue that $\mathcal{E}$ is compact and provide bounds on the elements of $\mathcal{E}$. First recall that Dirichlet's Unit Theorem (e.g. [16, Thm 7.3]) states the existence of $u_1, \ldots, u_{n-1}$ fundamental units such that any unit in $\mathcal{O}_K$ can be written as

$$u = \zeta \prod_{i=1}^{n-1} u_i^{k_i}, \text{ where } k_i \in \mathbb{Z} \text{ and } \zeta \text{ is a root of unit.} \quad (13)$$

[2]The proof in [15] is presented for quadratic number fields, but it can be generalized to any number field. Details are left for a journal version of this paper

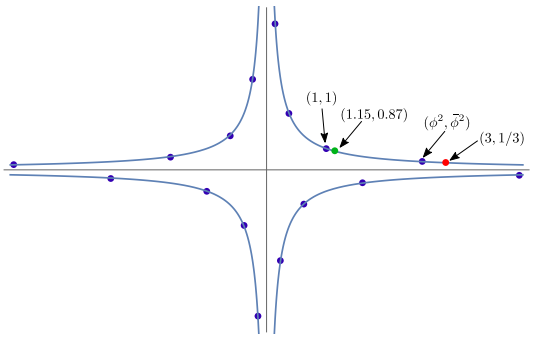

Fig. 1. Handling an ill-conditioned channel realization by the quantization of the channel space

This implies that the group of units, under the transformation

$$\ell(u) = (\log|\sigma_1(u)|, \ldots, \log|\sigma_n(u)|) \quad (14)$$

is an $(n-1)$-dimensional lattice in $\mathbb{R}^n$, contained in the hyperplane orthogonal to the vector $(1, \ldots, 1)$. The volume of this lattice, referred to as *logarithmic lattice*, is called the *regulator* of K.

Theorem 1. *For any channel matrix* $\mathbf{H}$*, there exists* $\mathbf{U} = diag(\sigma_1(u), \ldots, \sigma_n(u))$ *such that*

$$\left\| \tilde{\mathbf{H}}\mathbf{U}^{-1} \right\| \le \sqrt{n} e^{(n-1)\left(\frac{R_K}{\mathcal{V}_{n-1}}\right)^{1/(n-1)}},$$

where R_K *is the regulator and* $\mathcal{V}_n$ *is the volume of a unit* n*-dimensional Euclidean sphere.*

Proof. Write the diagonal elements of $\tilde{\mathbf{H}}$ in vector form as $\tilde{\mathbf{h}} = (\log|\tilde{h}_1|, \ldots, \log|\tilde{h}_n|)$. Let $\mathbf{v} = (\log|\sigma_1(u)|, \ldots, \log|\sigma_n(u)|)$ be the closest point in the logarithm lattice to $\tilde{\mathbf{h}}$. Let ρ be the covering radius of the logarithmic lattice. We have $(\log|\tilde{h}_i| - \log|\sigma_i(u)|) \le \left\|\tilde{\mathbf{h}} - \mathbf{v}\right\| \le \rho$,therefore

$$\sum_{i=1}^{n} |\tilde{h}_i|^2 |\sigma_i(u)^{-1}|^2 = \sum_{i=1}^{n} e^{2(\log|\tilde{h}_i| - \log|\sigma_i(u)|)} \le n e^{2\rho}.$$

For the final bound we combine three inequalities for the minimum norm and covering radius of an n-dimensional lattice: $\rho(\Lambda)\lambda_1(\Lambda^*) \le \frac{1}{2}n$ [17, Thm. 2.2], $\lambda_1(\Lambda)\lambda_1(\Lambda^*) \ge 1$ and $\lambda_1(\Lambda) \le 2(\det\Lambda)^{1/n}\mathcal{V}_n^{1/n}$. □

As a corollary, the set of error matrices $\mathcal{E}$ (Equation (12)) is compact.

Example 1. *Let* $K = \mathbb{Q}[\sqrt{5}]$*, so that* $\mathcal{O}_K = \mathbb{Z}[\phi]$*, where* $\phi = (1+\sqrt{5})/2$ *is the Golden ratio. The units of* $\mathcal{O}_K$ *are of the form* $\pm\phi^k$*,* $k \in \mathbb{Z}$*, and its embeddings in* $\mathbb{R}^2$ *are the blue dots depicted in Figure 1. After normalization, the channel realizations* h_1, h_2 *lie in the hyperbola* $h_1 h_2 = 1$*. Any realization* (h_1, h_2) *can be taken, by multiplication by an appropriate unit, to a bounded fundamental domain. This way, ill-conditioned channel realizations can be "absorbed" by the group of units.*

B. Achieving the Limit

In this subsection we prove the asymptotic goodness of the proposed algebraic lattices.

Theorem 2. *There exists a sequence of lattices in the ensemble* (9) *universally good for the block-fading channel.*

The proof uses the techniques of [9], [10, Appendix], and consists of two parts: (i) an universal code for a finite set of channel matrices and (ii) fine quantization of the channel coefficients

Proof. (i) Suppose that we have $\mathbf{H}_1, \ldots, \mathbf{H}_L$ channel matrices. Averaging the sum of these probabilities $\mathcal{P}_e(\Lambda, \mathbf{H}_i)$ over all lattices in the ensemble, we have

$$\begin{aligned} &E_{\mathbb{L}}\left[\sum_{i=1}^{L} \mathcal{P}_e(\Lambda, \mathbf{H}_i)\right] = \sum_{i=1}^{L} E_{\mathbb{L}}\left[\mathcal{P}_e(\Lambda, \mathbf{H}_i)\right] \\ &\leq L\left(P(\mathbf{Z} \notin \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}) + \lambda \frac{\text{vol } \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}}{D^{\frac{n}{2}}}\right), \end{aligned} \tag{15}$$

where the expression in brackets in the rhs is due to Eq. (10) combined with [13, Eq. 13] (see also the proof of Theorem 3 [5]). This sum of probabilities can be made arbitrarily small as long as the threshold condition VNR $> 2\pi e/D^{1/n}$ is satisfied.

(ii) Infinite part. First notice, by writting $Y = DE_h\bar{\mathbf{X}} + \mathbf{Z}$, $\bar{\mathbf{X}} = \mathbf{U}_{\tilde{H}}\mathbf{X}$, that $\mathcal{P}_e(\Lambda, \mathbf{H}) = \mathcal{P}_e(\Lambda, D\mathbf{E}_{\tilde{H}})$. This follows since the lattices in our ensemble are closed under multiplication by a unit $\mathbf{U}$. It follows from [9, Lem. 7] that for a second matrix $\mathbf{E}_{\tilde{H}_0}$ such that $\left\|\mathbf{E}_{\tilde{H}_0}^{-1} - \mathbf{E}_{\tilde{H}}^{-1}\right\| \leq \eta$, there is a constant κ, s.t.

$$\begin{aligned} &\mathcal{P}_e(\Lambda, D\mathbf{E}_{\tilde{H}} | \mathbf{Z} \in \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}) \leq \\ &e^{nT\eta\kappa}\mathcal{P}_e(\Lambda, D\mathbf{E}_{\tilde{H}_0} | \mathbf{Z} \in \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}). \end{aligned} \tag{16}$$

Now, since the set $\mathcal{E}$ is compact, for any arbitrarily small η, we can choose $L_{\eta,n}$ large enough and matrices $\mathbf{E}_{\tilde{H}_1}, \ldots, \mathbf{E}_{\tilde{H}_L}$ such that for all $\mathbf{E} \in \mathcal{E}$, there exists i, such that $\left\|\mathbf{E}_{\tilde{H}_i}^{-1} - \mathbf{E}^{-1}\right\| \leq \eta$. Therefore, for any $H \in \mathbb{H}_\infty$, there exists i such that

$$\begin{aligned} \mathcal{P}_e(\Lambda, \mathbf{H}) \leq\ & P(\mathbf{Z} \notin \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}) \\ &+ e^{nT\eta\kappa} P_e(\Lambda, \mathbf{H}_i | \mathbf{Z} \in \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}). \end{aligned} \tag{17}$$

Taking the average over the ensemble:

$$\begin{aligned} E_{\mathbb{L}}\left[\mathcal{P}_e(\Lambda, \mathbf{H})\right] \leq\ & P(\mathbf{Z} \notin \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}) \\ &+ \lambda\, L_{\eta,n} e^{nT\eta\kappa} \frac{\text{vol } \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}}{D^{\frac{n}{2}}}. \end{aligned} \tag{18}$$

Now if we choose λ to be lesser than $D^{1/2T}/\sqrt{2\pi e\sigma_c^2}$, then

$$\lambda \frac{\text{vol } \mathcal{B}_{\sqrt{nT(\sigma_c^2+\varepsilon)}}}{D^{\frac{n}{2}}} \to 0 \text{ exponentially in } T.$$

Therefore, we can choose $L_{\eta,n}$, independent of T, such that the total exponent is negative, and hence the average probability of error of the ensemble can be made arbitrarily small. □

We close this section arguing that mod-p lattices [13] fail to universally achieve capacity. All mod-p lattices contain multiples of the canonical vectors (say, $p\beta\mathbf{e}_i$, where β is a scaling factor). Hence $\mathcal{V}_{\mathbf{H}\Lambda}$ is contained in the set $S = \left\{\mathbf{x} \in \mathbb{R}^{nT} : |x_1| \leq h_1\beta p/2\right\}$, and therefore for any Λ in the mod-p ensemble

$$P_e(\Lambda, \mathbf{H}) \geq P(\mathbf{z} \notin S) = P(|z_1| \geq h_1\beta p/2). \tag{19}$$

Consider now the matrix $\mathbf{H} \in \mathbb{H}_\infty$, with $h_1 = 1/p^2, h_2 = p^2$, $h_i = D^{1/(n-2)}, i = 3, \ldots, n$. It is clear that $P_e(\Lambda, \mathbf{H}) \to 1$, as $p \to \infty$, and there is no good lattice (in the sense of Def. 2) in the ensemble. This does *not* contradict [5, Thm. 3], who showed, for a *given fixed* $\mathbf{H}$, the existence of a good Λ (depending on $\mathbf{H}$), which does not imply the existence of one single sequence with vanishing probabilities for all $\mathbf{H}$. Note that this effect is prevented in our construction, since our lattices have full diversity.

V. Shaping: The Lattice Gaussian Distribution

The final transmission scheme is similar to [1]. Using a coding lattice of dimension nT from the ensemble (9), the transmitter chooses a vector $\mathbf{x}$ in Λ drawn according to a lattice Gaussian distribution D_{Λ,σ_s}. The received applies MAP decoding to recover an estimate $\hat{\mathbf{x}}$ of the sent symbol.

Consider a vector-form channel equation (5), with indices omitted. Let $\rho = \frac{\sigma_s^2}{\sigma_c^2}$. MAP decoding reads:

$$\begin{aligned} \hat{\mathbf{x}} &= \arg\max_{x\in\Lambda} p(\mathbf{x}|\mathbf{y}, \mathcal{H}) = \arg\max_{x\in\Lambda} p(\mathbf{y}|\mathbf{x}, \mathcal{H})p(\mathbf{x}) \\ &= \arg\max_{x\in\Lambda} f_{\sigma_c}(\mathbf{y} - \mathcal{H}\mathbf{x}) f_{\sigma_s}(\mathbf{x}) \\ &\overset{(a)}{=} \arg\min_{x\in\Lambda} \sigma_c^{-2}\|\mathbf{y} - \mathcal{H}\mathbf{x}\|^2 + \sigma_s^{-2}\|\mathbf{x}\|^2 \\ &\overset{(b)}{=} \arg\min_{x\in\Lambda} \|\mathbf{F}\mathbf{y} - \mathbf{R}\mathbf{x}\|^2 \end{aligned}$$

where $\mathbf{R}^{\text{tr}}\mathbf{R} = \rho^{-1}\mathcal{H}^{\text{tr}}\mathcal{H} + \mathbf{I}$ and $\mathbf{F}^{\text{tr}} = \rho^{-1}\mathcal{H}\mathbf{R}^{-1}$. In the above equation, (a) is due to the definition of D_{Λ,σ_s}, while (b) is obtained by completing the square. This coincides with the well-known MMSE-GDFE [5], except that SNR is replaced by ρ. We note that the matrices $\mathbf{F}$ and $\mathbf{R}$ are block diagonal, namely, the MMSE filter is only applied on the spatial dimension. Therefore MAP decoding is equivalent to MMSE-GDFE filtering plus lattice decoding.

To analyze the error probability, we write

$$\mathbf{y}' = \mathbf{F}\mathbf{y} = \mathbf{R}\mathbf{x} + (\mathbf{F}\mathcal{H} - \mathbf{R})\mathbf{x} + \mathbf{F}\mathbf{w} = \mathbf{R}\mathbf{x} + \mathbf{w}'$$

where $\mathbf{w}' \triangleq (\mathbf{F}\mathcal{H} - \mathbf{R})\mathbf{x} + \mathbf{F}\mathbf{w}$ can be viewed as the equivalent noise. The error probability associated with a lattice Λ is given by

$$P_e(\Lambda) = \sum_{\mathbf{x}\in\Lambda} P(\text{error}|\mathbf{x})P(\mathbf{x}) = P\left\{\mathbf{w}' \notin \mathcal{V}(\mathbf{R}\Lambda)\right\} \tag{20}$$

where the last step follows from the total probability theorem. We stress that in (20), the probability is evaluated with respect to both distributions $\mathbf{x} \sim D_{\Lambda,\sigma_s}$ and $\mathbf{w} \sim f_{\sigma_c}$.

Next, we will show that the equivalent noise $\mathbf{w}'$ is sub-Gaussian. Therefore, the error probability is exponentially

bounded above by that of a Gaussian noise, and a good ensemble as in Def. 2 will also have a vanishing probability of error for $\mathbf{w}'$. Let us recall the definition of sub-Gaussian.

Definition 3 (sub-Gaussian [18])**.** *A random variable X is sub-Gaussian with parameter $\sigma > 0$ if for all $t \in \mathbb{R}$, the moment-generating function satisfies $\mathbb{E}[e^{tX}] \leq e^{\sigma^2 t^2/2}$. The tails of X are dominated by a Gaussian of parameter σ, i.e., $\mathbb{P}(|X| \geq t) \leq 2e^{-t^2/(2\sigma^2)}$ for all $t \geq 0$. More generally, we say that a random vector $\mathbf{x}$ is sub-Gaussian (of parameter σ) if all its one-dimensional marginals $\mathbf{u}^T\mathbf{x}$ for a unit vector $\mathbf{u}$ are sub-Gaussian (of parameter σ).*

Lemma 1. *Let $\mathbf{x} \sim D_{\Lambda,\sigma}$. Then the moment generating function of $\mathbf{Ax}$ for any square matrix $\mathbf{A}$ satisfies*

$$E[e^{\mathbf{t}^{\mathrm{tr}}\mathbf{Ax}}] \leq e^{\frac{\sigma^2}{2}\|\mathbf{A}^{\mathrm{tr}}\mathbf{t}\|^2}.$$

Proof. We rewrite the moment generating function as follows:

$$\begin{aligned} f_\sigma(\Lambda) \cdot E[e^{\mathbf{t}^{\mathrm{tr}}\mathbf{Ax}}] &= \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\mathbf{x}\in\Lambda} e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2} + \mathbf{t}^{\mathrm{tr}}\mathbf{Ax}} \\ = \quad e^{\frac{\sigma^2}{2}\|\mathbf{A}^{\mathrm{tr}}\mathbf{t}\|^2} f_\sigma\left(\Lambda - \frac{\sigma}{\sqrt{2}}\mathbf{A}^{\mathrm{tr}}\mathbf{t}\right). \end{aligned}$$

Since $f_\sigma(\Lambda - \mathbf{a}) \leq f_\sigma(\Lambda)$ for any vector $\mathbf{a}$, the proof is completed. □

Lemma 2. *The equivalent noise $\mathbf{w}'$ is sub-Gaussian with parameter σ_c.*

Proof. Let us derive its moment generation function:

$$\begin{aligned} E[e^{\mathbf{t}^{\mathrm{tr}}\mathbf{w}'}] &= E[e^{\mathbf{t}^{\mathrm{tr}}((\mathbf{F}\mathcal{H}-\mathbf{R})\mathbf{x}+\mathbf{F}\mathbf{w})}] \\ &= E[e^{\mathbf{t}^{\mathrm{tr}}(\mathbf{F}\mathcal{H}-\mathbf{R})\mathbf{x}}]E[e^{\mathbf{t}^{\mathrm{tr}}\mathbf{F}\mathbf{w}}] \\ &\leq e^{\mathbf{t}^{\mathrm{tr}}(\mathbf{F}\mathcal{H}-\mathbf{R})(\mathbf{F}\mathcal{H}-\mathbf{R})^{\mathrm{tr}}\mathbf{t}\cdot\sigma_s^2/2} \cdot e^{\mathbf{t}^{\mathrm{tr}}\mathbf{F}\mathbf{F}^{\mathrm{tr}}\mathbf{t}\cdot\sigma_c^2/2} \\ &= e^{\mathbf{t}^{\mathrm{tr}}[\sigma_s^2(\mathbf{F}\mathcal{H}-\mathbf{R})(\mathbf{F}\mathcal{H}-\mathbf{R})^{\mathrm{tr}}+\sigma_c^2\mathbf{F}\mathbf{F}^{\mathrm{tr}}]\mathbf{t}/2} = e^{\sigma_c^2\|\mathbf{t}\|^2/2}. \end{aligned}$$

The last step holds because the covariance matrix [5]

$$\begin{aligned} & \sigma_s^2(\mathbf{F}\mathcal{H}-\mathbf{R})(\mathbf{F}\mathcal{H}-\mathbf{R})^{\mathrm{tr}} + \sigma_c^2\mathbf{F}\mathbf{F}^{\mathrm{tr}} \\ = \quad & \sigma_s^2\rho^{-2}\mathbf{R}^{-T}\mathbf{R}^{-1} + \sigma_c^2\rho^{-2}\mathbf{R}^{-T}\mathcal{H}^{\mathrm{tr}}\mathcal{H}\mathbf{R}^{-1} \\ = \quad & \sigma_c^2\mathbf{R}^{-T}(\rho^{-1}\mathbf{I} + \mathcal{H}^{\mathrm{tr}}\mathcal{H})\mathbf{R}^{-1} = \sigma_c^2\mathbf{I}. \end{aligned}$$

For any unit vector $\mathbf{u}$, we have

$$E[e^{t\mathbf{u}^{\mathrm{tr}}\mathbf{w}'}] = E[e^{(t\mathbf{u})^{\mathrm{tr}}\mathbf{w}'}] \quad \leq \quad E[e^{\sigma_c^2\|t\mathbf{u}\|^2/2}] = e^{\sigma_c^2 t^2/2}$$

completing the proof. □

Finally, from Theorem 2, taking an universal lattice Λ from the Minkowski-Hlawka ensemble (9), the error probability vanishes as long as the VNR $\gamma_{\mathbf{R}\Lambda}(\sigma) > 2\pi e$ (as $T \to \infty$), i.e..

$$\frac{V(\mathbf{R}\Lambda)^{\frac{2}{nT}}}{\sigma_c^2} = \frac{|\mathbf{I} + \rho\mathbf{H}^{\mathrm{tr}}\mathbf{H}|^{\frac{1}{n}} V(\Lambda)^{\frac{2}{nT}}}{\sigma_c^2} > 2\pi e. \tag{21}$$

Thus, from [12, Lemma 6], any rate

$$\begin{aligned} R &= \frac{n}{2}\log(2\pi e\sigma_s^2) - \frac{1}{T}\log(V(\Lambda)) - \varepsilon \\ &= \frac{1}{2}\log\det\left(\mathbf{I} + \rho\mathbf{H}^{\mathrm{tr}}\mathbf{H}\right) - \varepsilon = C - \varepsilon, \end{aligned} \tag{22}$$

for any arbitrarily small ε is achievable. Note that the achievable rate only depends on $\mathbf{H}$ through $\det(\mathbf{I} + \rho\mathbf{H}^{\mathrm{tr}}\mathbf{H})$. Therefore, there exists a lattice Λ achieving capacity C of the compound channel.

Remark 1. *Note that we do not need a condition on the SNR as in [1] anymore, thanks to sub-Gaussianity. The only condition we need is that D_{Λ,σ_s} behaves like a continuous Gaussian distribution, namely, $\epsilon_\Lambda(\sigma_s)$ is negligible. When this is the case, the signal power $P \approx \sigma^2$ and $\rho \approx \mathrm{SNR}$.*

VI. Acknowledgments

The work of Antonio Campello was funded by FAPESP under grant 2014/20602-8. This work was supported in part by FP7 project PHYLAWS (EU FP7-ICT 317562).

References

[1] Cong Ling and J.-C. Belfiore. Achieving AWGN Channel Capacity With Lattice Gaussian Coding. *IEEE Transactions on Information Theory*, 60(10):5918–5929, Oct 2014.

[2] Frédérique Oggier and Emanuele Viterbo. Algebraic Number Theory and Code Design for Rayleigh Fading Channels. *Commun. Inf. Theory*, 1(3):333–416, December 2004.

[3] O. Ordentlich and U. Erez. Precoded Integer-Forcing Universally Achieves the MIMO Capacity to Within a Constant Gap. *IEEE Transactions on Information Theory*, 61(1):323–340, Jan 2015.

[4] Laura Luzzi and Roope Vehkalahti. Almost universal codes achieving ergodic MIMO capacity within a constant gap. *CoRR*, abs/1507.07395, 2015.

[5] H. El Gamal, G. Caire, and M.O. Damen. Lattice Coding and Decoding Achieve the Optimal Diversity-Multiplexing tradeoff of MIMO channels. *IEEE Transactions on Information Theory*, 50(6):968–985, June 2004.

[6] Y. Yona and M. Feder. Fundamental Limits of Infinite Constellations in MIMO Fading Channels. *Information Theory, IEEE Transactions on*, 60(2):1039–1060, Feb 2014.

[7] M. Punekar, J.J. Boutros, and E. Biglieri. A Poltyrev outage limit for lattices. In *IEEE International Symposium on Information Theory (ISIT),*, pages 456–460, June 2015.

[8] Shlomi Vituri. Dispersion Analysis of Infinite Constellations in Ergodic Fading Channels. *CoRR*, abs/1309.4638, 2013.

[9] W. L. Root and P. P. Varaiya. Capacity of Classes of Gaussian Channels. *SIAM Journal on Applied Mathematics*, 16(6):1350–1393, 1968.

[10] Jun Shi and R.D. Wesel. A study on universal codes with finite block lengths. *IEEE Transactions on Information Theory*, 53(9):3066–3074, Sept 2007.

[11] E. Viterbo G. Rekaya, J-C. Belfiore. A very efficient lattice reduction tool on fast fading channels. In *Proceedings of the Internation Symposium on Information Theory and its Applications (ISITA), Parma, Italy*, 2004.

[12] Cong Ling, L. Luzzi, J.-C. Belfiore, and D. Stehle. Semantically Secure Lattice Codes for the Gaussian Wiretap Channel. *IEEE Transactions on Information Theory*, 60(10):6399–6416, Oct 2014.

[13] H.-A. Loeliger. Averaging bounds for lattices and linear codes. *IEEE Transactions on Information Theory*, 43(6):1767–1773, Nov 1997.

[14] W. Kositwattanarerk, Soon Sheng Ong, and F. Oggier. Construction A of Lattices Over Number Fields and Block Fading (Wiretap) Coding. *IEEE Transactions on Information Theory*, 61(5):2273–2282, May 2015.

[15] Yu-Chih Huang, Krishna R. Narayanan, and Ping-Chung Wang. Adaptive compute-and-forward with lattice codes over algebraic integers. *CoRR*, abs/1501.07740, 2015.

[16] J. Neukirch. *Algebraic Number Theory*, volume 322 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg, 1999.

[17] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(1):625–635, 1993.

[18] R. Vershynin. *Introduction to the non-asymptotic analysis of random matrices*. Chapter 5, "Compressed Sensing, Theory and Applications". Edited by Y. Eldar and G. Kutyniok, Cambridge University Press, 2012.