Almost universal codes achieving ergodic MIMO capacity within a constant gap

Laura Luzzi, Roope Vehkalahti

Almost universal codes achieving ergodic MIMO capacity within a constant gap

Laura Luzzi and Roope Vehkalahti

***Abstract*—This work addresses the question of achieving capacity with lattice codes in multi-antenna block fading channels when the number of fading blocks tends to infinity.**
A design criterion based on the normalized minimum determinant is proposed for division algebra multi-block space-time codes over fading channels; this plays a similar role to the Hermite invariant for Gaussian channels.
Under maximum likelihood decoding, it is shown that this criterion is sufficient to guarantee transmission rates within a constant gap from capacity both for deterministic channels and ergodic fading channels. Moreover, if the number of receive antennas is greater or equal than the number of transmit antennas, the same constant gap is achieved under naive lattice decoding as well. In the case of independent identically distributed Rayleigh fading, the error probability vanishes exponentially fast.
In contrast to the standard approach in the literature which employs random lattice ensembles, the existence results in this paper are derived from number theory. First the gap to capacity is shown to depend on the discriminant of the chosen division algebra; then class field theory is applied to build families of algebras with small discriminants. The key element in the construction is the choice of a sequence of division algebras whose centers are number fields with small root discriminants.

***Index Terms*—MIMO, block fading, space-time codes, number theory, division algebras**

I. Introduction

It is well-known [3] that in ergodic multiple-input multiple-output (MIMO) fading channels with channel state information at receiver only, the maximal mutual information is achieved with Gaussian circularly symmetric random inputs. In this case the existence of capacity-achieving codes can be proven with standard random coding arguments.

It has been shown that by combining simple modulation and strong outer codes such as turbo or LDPC codes, it is possible to operate at rates close to capacity with small error probability [4, 5]. However, to the best of our knowledge, the problem of achieving ergodic capacity with *explicit codes* for all ranges of signal-to-noise ratio (SNR) is still open.

This is in strong contrast to the classical complex Gaussian single antenna channel, where the capacity is $\log(1+\mathrm{SNR})$ and it is known that several lattice code constructions achieve $\log \mathrm{SNR} - C$ rates for some constant gap C. These constructions are based on a rich theory of lattice codes developed to attack these questions. At the heart of this theory are sphere packing arguments showing that the performance of a lattice code in the classical Gaussian channel can be roughly estimated by the size of a geometrical invariant of the lattice, the *Hermite invariant*. In particular the Hermite invariant can be used to roughly measure how close to capacity a family of lattices can get. This connection has been extremely fruitful and has led to a monumental work connecting algebra, geometry and information theory [6].

In the case of fading channels the situation is quite different. While it is well-known that space-time lattice codes from division algebras [7] provide good performance over multiple antenna fading channels, and a rich algebraic theory has been developed to optimize specific code designs [8, 9], there are as yet no results connecting capacity questions and the geometry of lattices. The minimum determinant criterion [10] allows to improve the worst-case pairwise error probability in the high-SNR regime, when coding over a single fading block. Optimizing this value has been the major concern of several works in space-time coding [8, 11, 12]. However, no design criterion has been suggested to approach the MIMO capacity with explicit lattice codes.

In this paper we address this problem and show that when we are allowed to encode and decode over a growing number of fading blocks, the *normalized minimum determinant* plays a similar role to the Hermite constant in Gaussian channels. In particular it can be used to measure how close to capacity a given family of lattice codes can get.

Based on this design criterion we prove that for a MIMO channel with n transmit and n_r receive antennas, where $n_r \geq n$, there exists a family of multi-block lattice codes $L_{n,k} \subset M_{n \times nk}(\mathbb{C})$, where k goes to infinity, that achieves a constant gap to capacity both in the Gaussian MIMO case and ergodic fading case. More precisely, consider an ergodic fading MIMO channel with channel matrix $H \in M_{n_r \times n}(\mathbb{C})$, and whose capacity is $C = \mathbb{E}_H\left[\log\det(I_{n_r} + \frac{\mathrm{SNR}}{n} H^\dagger H)\right]$. Then our scheme achieves any rate

$$R < \mathbb{E}_H\left[\log\det\frac{\mathrm{SNR}}{n}H^\dagger H\right] - n\log C_L - n\log\frac{4n}{\pi e}, \quad (1)$$

where C_L is a certain geometric invariant of the family of lattices. Note that while the gap to capacity is independent of the SNR, it does depend on n and also on the channel statistics.

These rates are achieved not only with maximum likelihood (ML) decoding, but also with naive lattice decoding as

Part of this work appeared at ISIT 2015 [1, 2].

L. Luzzi is with ETIS (UMR 8051, ENSEA, Université de Cergy-Pontoise, CNRS), 95014 Cergy-Pontoise, France (e-mail: laura.luzzi@ensea.fr).

R. Vehkalahti is now with the Department of Communications and Networking, FI-02150, Aalto University, Espoo, Finland (e-mail: roope.vehkalahti@aalto.fi). While this work was in progress, he was with the Department of Mathematics and Statistics, FI-20014, University of Turku, Finland.

The research of R. Vehkalahti was funded by Academy of Finland grants #252457, #283135 and #299916 and by the Finnish Cultural Foundation.

long as $n_r \geq n$. Furthermore, the same scheme achieves positive rates of reliable communication for more general fading processes $\{H_i\}$ under the mild hypothesis that the weak law of large numbers holds for the sequence of random variables $\{\log\det H_i^\dagger H_i\}$. As far as we know, this is the first non-random coding scheme which achieves constant gap to capacity for all SNR levels in ergodic MIMO channels.

Instead of using random coding arguments we consider algebraic multi-block division algebra codes introduced in [13, 14, 15], and developed further in [16] and [17]. We use the most general form presented in [18].

We derive the existence of multi-block codes with special properties from two classical results from class field theory. First we choose the center K of the algebra from a tower of Hilbert class fields having constant root discriminant [19], and then we prove the existence of a K-central division algebra with small discriminant. Unfortunately, while the family of codes in question is well-defined and deterministic, the best known algorithms to compute Hilbert class fields of arbitrary number fields [20, 21] have very high computational complexity, and thus our construction cannot be made explicit at present.

In most works on algebraic space-time coding the code design criterion is derived from an upper bound for the pairwise error probability [10] together with the union bound [22, 23]. In our proofs we abandon this method and consider a hard sphere packing approach, classically used in lattice coding for the AWGN channel. The idea, formalized in Section III, is to exploit the special *multiplicative structure* of algebraic codes. It was observed in the context of diversity-multiplexing gain trade-off (DMT) analysis [9, 24], that for codes based on division algebras having the so-called *non-vanishing determinant* property, fading has a diminishing effect on the Euclidean distance of the received constellations only if the channel itself is bad. This property was formalized in [24], where the authors introduced *approximately universal codes* for fading channels. Our main results Theorem 4.1 and Theorem 4.7 rely on this "incompressibility" property of algebraic lattices. It follows that our codes are *almost universal* and perform within a constant gap to capacity for a wide class of channels, having only mild restrictions on fading.

While we discuss specific lattice codes from division algebras, our proofs do work for any ensemble of matrix lattices with asymptotically good normalized minimum determinant. The larger this value is, the smaller the gap to the capacity.

This work also suggests that capacity questions in fading channels are naturally linked to problems in the mathematical research area of *geometry of numbers*. Unlike the single antenna Gaussian case, many of the questions that arise have not been actively studied by the mathematical community. Hopefully, studying such questions may lead to a comprehensive geometric theory of lattices for multiple antenna fading channels.

We note that the proposed lattice code constructions are not yet practical, since they are based on number fields whose existence is proved through class field theory. Given a fixed degree, the required number fields can be found using computational algebra software, but this process is computationally taxing. Decoding of the proposed codes is also very complex and the constructions we provide still have a large gap to capacity.

On the other side, as demonstrated in Section VIII, the existence results we use are very pessimistic. For small degrees the normalized minimum determinants of the best possible lattices are considerably better than the bounds provided by our existence results.

A. Related work

While our work shows that one can achieve a constant gap to capacity in ergodic MIMO channels with a fixed family of algebraic codes, it is natural to consider the more general question of whether it is possible to achieve capacity with any lattice codes. Such a result would be a generalization of the work in [25, 26, 27] which proved the existence of random lattice code ensembles achieving rate $\log(\text{SNR})$ over the AWGN channel. By making the extra assumption that the transmitter and receiver have access to a common source of randomness in the form of a dither, the authors in [28] finally proved that the AWGN capacity is achievable with random lattice codes. An explicit multilevel construction from polar codes was recently proposed in [29].

As far as we know our work [2] was the first to give a proof that lattice codes achieve a constant gap to capacity in ergodic fading MIMO channels in the symmetric case where $n = n_r$. In the single antenna fast fading channel this problem was considered before in [30], which claims that random lattices achieve a constant gap to capacity. In [31] the authors extend their previous results and claim to give a proof that random lattices achieve capacity in single antenna ergodic fading channels. However, we believe that at least in its current form, the analysis in both works is missing some fundamental details. In particular, the gap $\Delta < 1 + \log \mathbb{E}_h \left[1/|h|^2\right]$, given in [30, Theorem 3], is infinite even when the fading process $\{h_i\}$ is i.i.d. complex Gaussian. In [31, Equation (20)] the authors state that for a given fixed fading realization, the Minkowski-Hlawka theorem implies that there exists a lattice for which the error probability is upper bounded in a certain way. However, they proceed as if there existed a single lattice that would satisfy this upper bound for any channel state. To the best of our knowledge, such a result cannot be derived from Minkowski-Hlawka.

While the main focus of our work is on the ergodic MIMO channel, in Section V-A we also consider deterministic MIMO channels. This work has at least two predecessors.

In [32, Theorem 3] the authors proved that for a given deterministic channel matrix H there exist lattice codes that achieve a constant gap to capacity. This corresponds to the deterministic model or "channel model 1" in Telatar's paper [3, page 2]. However, in their work the chosen code did depend on the channel matrix H. The authors in [33] went further and proved the existence of lattice codes that achieve a constant gap to capacity for a set of channels with the same white-input capacity. This corresponds to the "channel model 3" in [3]. Here the channel is random, but stays fixed during the time of transmission. Our results in Section V-A are somewhere in-between these two works. We have a single family of codes which achieves a constant gap to capacity for all fixed channels H, but the size of the gap does depend on H.

We point out that our results for the deterministic channel are simply side results and the main focus of this paper is on the block version of the ergodic channel model 2 in [3]. As far as we can see the approach in [32] or in [33] can not be straightforwardly extended to this model.

As far as explicit algebraic constructions are concerned, our work is indebted to several previous papers.
The idea to use division algebra codes to achieve capacity can be tracked down to the work of H.-f. Lu in [14]. While studying the diversity-multiplexing gain tradeoff (DMT) of multi-block codes he conjectured that the ensemble of multi-block division algebra codes might approach the ergodic Rayleigh fading capacity. Our work confirms that conjecture; however, we point out that it is unlikely that DMT-optimality alone is enough to approach capacity. Instead one should pick the code very carefully by maximizing the normalized minimum determinant.

The families of number fields on which our constructions are based were first brought to coding theory in [34], where the authors pointed out that the corresponding lattices have large Hermite constant. C. Xing in [35] remarked that these families of number fields provide the best known normalized product distance making them a natural candidate for achieving constant gap to capacity in fading single antenna channels.

Our geometry of numbers approach has its roots in [36], where the authors studied lattice codes in single antenna fading channels and defined the normalized product distance. They also pointed out that using this criterion reduces the lattice design to a problem in geometry of numbers.

The generalization of these ideas to the MIMO channel was developed in [8] and [37], where the code design for quasi-static MIMO channel problem was translated into lattice theoretic language and where a formal definition of normalized minimum determinant was given. However, none of these works considered the relation between geometry of numbers and capacity problems.

Recent results: Since the initial submission of this paper, there have been several advances on the topic. In a revised version of [38, Section 4.5], S. Vituri gave a proof of existence of lattice codes achieving a constant gap to capacity for ergodic SISO channels. It appears that with minor modifications this proof implies the existence of capacity-achieving lattices. In [39] the authors prove that polar lattices achieve capacity in i.i.d fading channels. This is not only an existence result, but provides an explicit low-complexity code construction as well. In [40] the authors prove the existence of lattice codes achieving capacity in the compound SISO channel, where the fading is random during the first s time units, but then gets repeated in blocks of length s. This work is most closely related to [33].

B. Organization of the paper

In Section II we introduce the multi-block channel model and recall the relevant properties of lattice codes. In Section III-A we develop a geometric design criterion for capacity approaching lattice codes for fading multiple antenna channels and define the concept of *reduced Hermite invariant* which is an analogue of the classical Hermite invariant. In Section III-B we state the existence of lattices having asymptotically good normalized minimum determinant (the proof will be given in Section VII). In Section IV we prove that the lattice codes of the previous section achieve positive rates over a very general class of channels. We then prove that they achieve a constant gap to capacity over Gaussian MIMO channels (Section V-A) and ergodic fading channels (Section V-B). In Section VI we focus on the i.i.d. Rayleigh fading channel model, and show that the error probability vanishes exponentially. In Section VII we prove the existence of asymptotically good lattices, and in Section VIII we specialize our results to the single antenna case. Finally in Section IX we explore the connection between capacity questions in fading channels and geometry of numbers. Section X discusses some perspectives and open problems.

C. Notation

Throughout the paper, capacity is measured in bits. Accordingly, we denote by $\log$ the base 2 logarithm in rate and capacity expressions; the natural logarithm will be denoted by $\ln$.

II. Multiblock lattice codes

A. Channel model

We consider a MIMO system with n transmit and n_r receive antennas, where transmission takes place over k quasi-static fading blocks of delay $T = n$. Each multi-block codeword $X \in M_{n\times nk}(\mathbb{C})$ has the form $[X_1, X_2, \ldots, X_k]$, where the submatrix $X_i \in M_n(\mathbb{C})$ is sent during the i-th block. The received signals are given by

$$Y_i = H_i X_i + W_i, \qquad i \in \{1, \ldots, k\} \tag{2}$$

where $H_i \in M_{n_r\times n}(\mathbb{C})$ and $W_i \in M_{n_r\times n}(\mathbb{C})$ are the channel and noise matrices. The coefficients of W_i are modeled as circular symmetric complex Gaussian with zero mean and unit variance per complex dimension. Perfect channel state information is available at the receiver but not at the transmitter, and decoding is performed after all k blocks have been received. We will call such a channel an *(n, n_r, k)-multi-block channel.* For the sake of simplicity, in the rest of the paper we will suppose that $n_r \geq n$ unless explicitly stated otherwise. We also assume that for all $i \geq 1$, $H_i \in M_{n_r\times n}$ is full-rank with probability 1, and that the random variable $\sum_{i=1}^{k} \frac{1}{k} \log\det(H_i^\dagger H_i)$ converges in probability to some constant when the number of blocks k tends to infinity. This channel model covers several standard MIMO channels such as the Rayleigh block fading channel and the Gaussian MIMO channel.

A *multi-block code* $\mathcal{C}$ in a (n, n_r, k)-channel is a set of matrices in $M_{n\times nk}(\mathbb{C})$. In particular we will concentrate on finite codes that are drawn from lattices. Let R denote the code rate in bits per complex channel use; equivalently, $|\mathcal{C}| =$

2^{Rkn}. We assume that every matrix X in a finite code $\mathcal{C} \subset M_{n\times nk}(\mathbb{C})$ satisfies the average power constraint

$$\frac{1}{nk}\mathbb{E}[\|X\|^2] \leq P, \tag{3}$$

where $\|X\|$ is the Frobenius norm of the matrix X.

B. Lattice codes

Given a nonzero matrix $B \in M_n(\mathbb{C})$, we use the notation $\mathbb{Z}B$ for the one-dimensional $\mathbb{Z}$-module generated by B. Given two $\mathbb{Z}$-modules V and V', we denote their direct sum by $V \oplus V'$.

Definition 2.1: A *matrix lattice* $L \subseteq M_{n\times nk}(\mathbb{C})$ has the form

$$L = \mathbb{Z}B_1 \oplus \mathbb{Z}B_2 \oplus \cdots \oplus \mathbb{Z}B_m,$$

where the matrices $B_1, \ldots, B_m \in M_{n\times n}(\mathbb{C})$ are linearly independent over $\mathbb{R}$, i.e., form a lattice basis, and m is called the *rank* or the *dimension* of the lattice.

The space $M_{n\times nk}(\mathbb{C})$ is a $2n^2k$-dimensional real vector space with a real inner product

$$\langle X, Y\rangle = \Re(\mathrm{tr}(XY^\dagger)),$$

where tr is the matrix trace. This inner product also naturally defines a metric on the space $M_{n\times nk}(\mathbb{C})$ by setting $||X|| = \sqrt{\langle X, X\rangle}$.

Given an m dimensional lattice $L \subset M_{n\times nk}(\mathbb{C})$, its *Gram matrix* is defined as

$$G(L) = (\langle B_i, B_j\rangle)_{1\leq i,j\leq m},$$

where $\{B_i\}_{1\leq i\leq m}$ is a basis of L. The volume of the fundamental parallelotope of L is then defined as $\mathrm{Vol}(L) = \sqrt{|\det(G(L))|}$.

In the following we will use the notation $\mathbb{R}(L)$ for the linear space generated by the basis elements of the lattice L.

Lemma 2.2: [41] Let us suppose that L is a lattice in $M_{n\times kn}(\mathbb{C})$ and S is a Jordan measurable bounded subset of $\mathbb{R}(L)$. Then there exists $X \in M_{n\times kn}(\mathbb{C})$ such that

$$|(L + X) \cap S| \geq \frac{\mathrm{Vol}(S)}{\mathrm{Vol}(L)}.$$

Given a family of lattices $L_{n,k} \subseteq M_{n\times nk}(\mathbb{C})$, let us now show how we can design multi-block codes $\mathcal{C}$ having rate greater or equal to a prescribed constant R, and satisfying the average power constraint (3), from a scaled version $\alpha L_{n,k}$ of the lattices, where α is a suitable energy normalization constant. We denote by $B(r)$ the set of matrices in $M_{n\times nk}(\mathbb{C})$ with Frobenius norm smaller or equal to r. According to Lemma 2.2, we can choose a constant shift $X_R \in M_{n\times nk}(\mathbb{C})$ such that for $\mathcal{C} = B(\sqrt{Pkn}) \cap (X_R + \alpha L_{n,k})$ we have

$$2^{Rnk} = |\mathcal{C}| \geq \frac{\mathrm{Vol}(B(\sqrt{Pkn}))}{\mathrm{Vol}(\alpha L_{n,k})} = \frac{C_{n,k}P^{n^2k}}{\alpha^{2n^2k}\,\mathrm{Vol}(L_{n,k})},$$

where $C_{n,k} = \frac{(\pi nk)^{n^2k}}{(n^2k)!}$. We then find the following condition for the scaling constant:

$$\alpha^2 = \frac{C_{n,k}^{\frac{1}{n^2k}}P}{2^{\frac{R}{n}}\,\mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}}} \tag{4}$$

III. Design criteria for fading channels

In this section we propose a new design criterion for capacity approaching lattice codes in fading channels. We note that the design criterion derived here will finally be the familiar minimum determinant criterion. However, we hope that our alternative characterization offers more insight on the topic and can have applications in further research.

A. Reduced Hermite invariant

We recall the classical definition of the Hermite invariant, which characterizes the density of a lattice packing:

Definition 3.1: The Hermite invariant of an m-dimensional lattice $L \subset M_{n\times nk}(\mathbb{C})$ can be defined as

$$\mathrm{h}(L) = \frac{\inf\{\,||X||^2 \mid X \in L, X \neq 0\}}{\mathrm{Vol}(L)^{2/m}}.$$

On the $n \times n$ MIMO Gaussian channel such that the channel matrices $H_i = I_n\ \forall i$, the classical sphere packing approach is to choose a $2n^2k$-dimensional lattice code $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ such that $\mathrm{h}(L_{n,k})$ is as large as possible.

Let us now assume that we have a finite code $\mathcal{C}_L \subset L_{n,k} \subset M_{n\times nk}(\mathbb{C})$. We define the following notation for componentwise multiplication of multi-block matrices: given $X = [X_1, \ldots, X_k]$ and $H = [H_1, \ldots, H_k] \in M_{n\times nk}(\mathbb{C})$,

$$H * X \div [H_1X_1, \ldots, H_kX_k]. \tag{5}$$

With this notation, the channel output $Y = [Y_1, \ldots, Y_k]$ can be written as

$$Y = H * X + W, \tag{6}$$

where $X = [X_1, \ldots, X_k]$ is the transmitted multi-block codeword, $H = [H_1, \ldots, H_k]$ is the random channel realization and $W = [W_1, \ldots, W_k]$ is the multi-block noise. From the receiver's point of view, this is equivalent to an additive white Gaussian noise channel where the lattice code is

$$H * \mathcal{C}_L = \{H * X \mid X \in \mathcal{C}_L\}.$$

Even if the lattice $L_{n,k}$ (and therefore the code $\mathcal{C}_L$) has good minimum distance, there is no guarantee that the same can be said about the lattice $H * \mathcal{C}_L$. This leads us to consider matrix lattices $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ which would have good minimum distance after any (reasonable) channel. If we assume that each of the matrices H_i in equation (2) has full rank with probability 1, then the multiplication $X \mapsto H * X$ is a bijective linear mapping with probability 1. For any lattice $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ having basis $B_1, \ldots, B_{2n^2k}$ we then have that

$$\begin{aligned} H * L_{n,k} &= \{H * X \mid X \in L_{n,k}\} \\ &= \mathbb{Z}(H * B_1) \oplus \cdots \oplus \mathbb{Z}(H * B_{2n^2k}), \end{aligned}$$

is a lattice with basis $\{H * B_1, \cdots, H * B_{2n^2k}\}$, and $\mathrm{h}(H * L_{n,k})$ is well defined.

As a discrete group, $H * L_{n,k}$ has positive Hermite invariant, but even if $\mathrm{h}(L_{n,k})$ is large there is no guarantee that $\mathrm{h}(H * L_{n,k})$ is.

Given a matrix $X = [X_1, \ldots, X_k] \in M_{n\times nk}(\mathbb{C})$, we define its *product determinant*

$$\mathrm{pdet}(X) = \prod_{i=1}^{k} \det(X_i). \tag{7}$$

For convenience we first introduce a group of matrices

$$G = \{H \in M_{n\times nk}(\mathbb{C}) \mid \mathrm{pdet}(H) = 1\}. \tag{8}$$

Definition 3.2: The *reduced Hermite invariant* of an m-dimensional lattice $L \subset M_{n\times nk}(\mathbb{C})$ with respect to the group G is defined as

$$\mathrm{rh}_{\mathrm{G}}(L) = \inf_{H\in G}\{\mathrm{h}(H * L)\}.$$

For any lattice L, $\mathrm{h}(L) > 0$. The same is not true for the reduced Hermite invariant. Let us now describe the set of lattices L for which $\mathrm{rh}_{\mathrm{G}}(L) > 0$.

Definition 3.3: The *minimum determinant* of the lattice $L \subseteq M_{n\times nk}(\mathbb{C})$ is defined as

$$\det{}_{\min}(L) := \inf_{X\in L\setminus\{\mathbf{0}\}} |\mathrm{pdet}(X)|\,.$$

If $\det_{\min}(L) > 0$ we say that the lattice satisfies the *non-vanishing determinant* (NVD) property.

We can now define the *normalized minimum determinant* $\delta(L)$, which is obtained by first scaling the lattice L to have a unit size fundamental parallelotope and then taking the minimum determinant of the resulting scaled lattice. A simple computation proves the following.

Lemma 3.4: Let L be an m-dimensional matrix lattice in $M_{n\times nk}(\mathbb{C})$. We then have that

$$\delta(L) = \frac{\det_{\min}(L)}{(\mathrm{Vol}(L))^{nk/m}}. \tag{9}$$

The normalized minimum determinant provides an alternative characterization of the reduced Hermite invariant, but before that we need a well known lemma.

Lemma 3.5: Let A be an $m\times m$ complex matrix. We have the inequality

$$|\det(A)| \leq \frac{\|A\|^m}{m^{m/2}}.$$

Proof: Let λ_i, $i = 1, \ldots, m$ be the eigenvalues of $A^\dagger A$. By the arithmetic-geometric mean inequality we have

$$|\det(A)|^2 = \det(A^\dagger A) = \prod_{i=1}^{m}\lambda_i \leq \left(\frac{\sum_{i=1}^{m}\lambda_i}{m}\right)^m$$
$$= \left(\frac{\mathrm{tr}(A^\dagger A)}{m}\right)^m = \frac{\|A\|^{2m}}{m^m}. \qquad \square$$

For a matrix $X \in M_{n\times nk}(\mathbb{C})$ this immediately implies that

$$|\mathrm{pdet}(X)| \leq \frac{\|X\|^{nk}}{(nk)^{nk/2}}.$$

Proposition 3.6: If $L \subset M_{n\times nk}(\mathbb{C})$ is a $2n^2k$-dimensional lattice, then

$$nk\,(\delta(L))^{2/nk} = \mathrm{rh}_{\mathrm{G}}(L).$$

Proof: If the lattice L includes a non-zero element X such that $\mathrm{pdet}(X) = 0$, it is easy to see that $nk\,(\delta(L))^{2/nk} = \mathrm{rh}_{\mathrm{G}}(L) = 0$.

Let us now assume that $\mathrm{pdet}(X) \neq 0$, for all $X \neq 0$. If $\mathrm{pdet}(H) = 1$, Lemma 3.5 implies that

$$\|H * X\|^2 \geq nk\,|\mathrm{pdet}(H * X)|^{2/nk} = nk\,|\mathrm{pdet}(X)|^{2/nk}\,.$$

It follows that $nk\,(\delta(L))^{2/nk} \leq \mathrm{rh}_{\mathrm{G}}(L)$.

Let us now assume that we have a sequence of codewords $X^{(i)} \in L$ such that

$$\lim_{i\to\infty} nk|\mathrm{pdet}(X^{(i)})|^{2/nk} = nk\,(\delta(L))^{2/nk}\,.$$

Given $X^{(i)} = [X_1^{(i)}, \ldots, X_k^{(i)}]$, we can choose $H^{(i)} = \mathrm{pdet}(X^{(i)})^{1/nk}[(X_1^{(i)})^{-1}, \ldots, (X_k^{(i)})^{-1}]$ so that $\mathrm{pdet}(H^{(i)}) = 1$. We then have

$$||H^{(i)} * X^{(i)}||^2 = nk|\mathrm{pdet}(X^{(i)})|^{2/nk}$$

for every i and therefore

$$\lim_{i\to\infty} ||H^{(i)} * X^{(i)}||^2 = nk\,(\delta(L))^{2/nk}\,.$$

It follows that $nk\,(\delta(L))^{2/nk} = \mathrm{rh}_{\mathrm{G}}(L)$. $\square$

Remark 3.7: Our definition of the reduced Hermite invariant rh_G depends heavily on the group G. The group chosen in (8) can be seen as a block diagonal subgroup of $\mathrm{SL}_{kn}(\mathbb{C})$. We could also consider a subgroup $G_1 \subset G$ and define rh_{G_1} with respect to this group. A natural consequence of these definitions is that for two subgroups G_1, G_2 of G such that $G_1 \subseteq G_2$, we have that

$$\mathrm{rh}_{G_1}(L) \geq \mathrm{rh}_{G_2}(L).$$

Remark 3.8: While the definition of the reduced Hermite invariant is very natural, we have found very few previous works considering similar concepts. The case $n = 1$ was considered by Skriganov in [42], where the author also proved Proposition 3.6 in this special case. Our results can therefore be seen as a natural generalization of this work. For general n the authors in [43] defined the Hermite invariant for generalized ideals in division algebras in the spirit of *Arakelov theory*. Their Hermite invariant is analogous to our concept of reduced Hermite invariant.

B. Asymptotically good families of lattices

Based on the observations in the previous section, we introduce the following definitions:

Definition 3.9: A sequence of lattices $L_{n,k}$ is *asymptotically good for the AWGN channel* if $\mathrm{h}(L_{n,k}) \geq cn^2k$, for some positive fixed constant c. Similarly, a sequence of lattices is *asymptotically good for fading channels* if $\mathrm{rh}_{\mathrm{G}}(L_{n,k}) \geq cn^2k$. As seen in Proposition 3.6, this is equivalent to asking that $\delta(L_{n,k})^{2/nk} \geq cn$.

We will show in the next sections that these properties guarantee that the lattice sequences achieve constant gap to capacity over AWGN and fading channels respectively.

In order to keep the paper suitable for a larger audience we will postpone the proof of the following existence result to Section VII.

Proposition 3.10: Given n, there exists a family of $2n^2k$-dimensional lattices $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$, where k grows to infinity, and a constant $G < 92.4$ such that

$$\mathrm{Vol}(L_{n,k}) \leq 23^{\frac{kn(n-1)}{10}} \left(\frac{G}{2}\right)^{n^2k}$$

$$\det\nolimits_{\min}(L_{n,k}) = 1 \quad \text{and} \quad \delta(L_{n,k}) \geq \frac{1}{23^{\frac{k}{20}(n-1)}(G/2)^{\frac{nk}{2}}}.$$

Remark 3.11: In this section we have developed the notion of reduced Hermite invariant for the case $n_r = n$. We observe that this notion does not extend to the case $n_r < n$, because the image $H * L$ of an infinite lattice L will no longer be a lattice, and the minimum distance in $H * L$ will be zero. However, when considering finite constellations $\mathcal{C}_L$, it is still possible to find suitable lower bounds on the minimum distance of the received constellation $H*\mathcal{C}_L$, as will be shown in the following sections.

IV. Achievable rates for general channels

Suppose that we have an infinite family of lattices $L_k \in \mathbb{C}^k$ with Hermite invariants satisfying $\frac{\mathrm{h}(L_k)}{k} \geq c$, for some positive constant c. Then a classical result in information theory states that with this family of lattices, all rates satisfying

$$R < \log P - \log\left(\frac{4}{\pi e}\right) + \log c,$$

are achievable in the additive complex Gaussian channel [6, Chapter 3]. This means that we can attach a single number $\mathrm{h}(L_k)$ to each lattice $L_k \in \mathbb{C}^k$, which roughly describes its performance and in particular estimates how close to the capacity a family of lattices can get. The following theorem can be seen as an analogue of this result for fading channels.

Theorem 4.1: Suppose that $n_r \geq n$, and let $\{H_i\}_{i\in\mathbb{Z}}$ be a fading process such that $H_i \in M_{n_r\times n}$ is full-rank with probability 1, and such that the weak law of large numbers holds for the random variables $\{\log\det(H_i^\dagger H_i)\}$, i.e. $\exists \mu > 0$ such that $\forall \epsilon > 0$,

$$\lim_{k\to\infty} \mathbb{P}\left\{\left|\frac{1}{k}\sum_{i=1}^{k} \log\det(H_i^\dagger H_i) - \mu\right| > \epsilon\right\} = 0. \tag{10}$$

Let $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional multi-block lattice codes such that

$$\det\nolimits_{\min}(L_{n,k}) = 1, \quad \text{and} \quad \mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L \tag{11}$$

for some constant $C_L > 0$. Then, any rate

$$R < \mu + n\left(\log P - \log\frac{4n^2}{\pi e} - \log C_L\right) \tag{12}$$

is achievable using the codes $L_{n,k}$ both with ML decoding and naive lattice decoding.

For all the fading processes that satisfy equation (10) with the same μ we achieve the rate (12) with the *same* code, hence our codes achieve in this scenario universally the same rate. However, as we will see later (Remarks 5.2 and 5.8), the gap to the capacity of the channel might depend on the fading process.

Remark 4.2: We note that existence of a family of lattices with

$$C_L \leq 23^{\frac{(n-1)}{10n}}\left(\frac{G}{2}\right),$$

was given in Proposition 3.10.

Remark 4.3: This theorem is stated by giving two conditions (11) for the lattices $L_{n,k}$. However, according to Lemma 3.4 we could have captured both of these conditions by an equivalent assumption $\delta(L_{n,k})^{2/nk} \geq \frac{1}{C_L}$. Proposition 3.6 then transforms this condition to

$$\mathrm{rh}_{\mathrm{G}}(L_{n,k}) = nk\left(\delta(L_{n,k})\right)^{2/nk} \geq \frac{nk}{C_L}.$$

As only k is growing, we can further write that $\mathrm{rh}_{\mathrm{G}}(L_{n,k}) \geq n^2kC'_L$, where $C'_L = n/C_L$. We can therefore see that conditions (11) assure that the family of lattices $L_{n,k}$ is asymptotically good in the sense of Section III-B.

The achievable rate R in Theorem 4.1 can then be seen as a complete analogue to the classical sphere packing result in AWGN channels.

Remark 4.4: The condition (10) holds in particular for ergodic stationary fading channels and for constant MIMO channels. These special cases will be analyzed further in Section V, where we will show that the codes in Theorem 4.1 achieve a constant gap to channel capacity.

To prove Theorem 4.1, we need the following Lemma:

Lemma 4.5: Consider the finite code $\mathcal{C} = B(\sqrt{Pkn}) \cap (X_R + \alpha L_{n,k})$ defined in Section II-B. Suppose that the receiver performs maximum likelihood decoding or "naive" lattice decoding (closest point search in the infinite lattice). Then, under the hypotheses of Theorem 4.1, $\forall\epsilon > 0$ the error probability is bounded by

$$P_e \leq 2e^{-\frac{kn^2\epsilon^2}{8}} + \mathbb{P}\left\{\frac{\alpha^2}{4n}\prod_{i=1}^{k}\det(H_i^\dagger H_i)^{\frac{1}{nk}} < 1+\epsilon\right\} \tag{13}$$

Proof: We distinguish two cases: the symmetric case where $n_r = n$, and the asymmetric case with $n_r > n$.

a) Case $n_r = n$: Suppose that $\bar{X} \in \mathcal{C}$ is the transmitted multi-block codeword and that $Y = H*\bar{X}+W$ is the received multi-block signal, where we use the notation (5). Here $W = [W_1, W_2, \ldots, W_k]$ denotes the multi-block noise. The outputs of the maximum likelihood (ML) decoder and naive lattice decoder (NLD) are given respectively by

$$\hat{X}_{\mathrm{ML}} = \underset{X\in\mathcal{C}}{\mathrm{argmin}}\, \|Y - H*X\|,$$

$$\hat{X}_{\mathrm{NLD}} = \underset{X\in\alpha L_{n,k}}{\mathrm{argmin}}\, \|Y - H*X\|.$$

Let d_H denote the minimum Euclidean distance in the received lattice:

$$d_H^2 = \min_{\substack{X,\bar{X}\in L_{n,k}\\ X\neq\bar{X}}} \left\|H*(X-\bar{X})\right\|^2$$

$$= \min_{\substack{X,\bar{X}\in L_{n,k}\\ X\neq\bar{X}}} \sum_{i=1}^{k}\left\|H_i(X_i-\bar{X}_i)\right\|^2.$$

We note that if $\|W\| < d_H/2$, then both decoders will output the correct codeword $\bar{X}$. In fact, under this assumption, $\forall X \in \alpha L_{n,k}$ such that $X \neq \bar{X}$ we have

$$\begin{aligned} \|W\| &= \left\|Y - H * \bar{X}\right\| < \frac{1}{2}\left\|H * (\bar{X} - X)\right\| \\ &\leq \frac{1}{2}\left\|Y - H * \bar{X}\right\| + \frac{1}{2}\left\|Y - H * X\right\| \end{aligned}$$

and so $\left\|Y - H * \bar{X}\right\| < \|Y - H * X\|$. Thus for both decoders the error probability is bounded by

$$P_e \leq \mathbb{P}\left\{\|W\|^2 \geq \left(\frac{d_H}{2}\right)^2\right\}.$$

By the law of total probability, $\forall \epsilon > 0$ we have

$$P_e \leq \mathbb{P}\left\{\frac{\|W\|^2}{kn^2} \geq 1 + \epsilon\right\} + \mathbb{P}\left\{\frac{d_H^2}{4kn^2} < 1 + \epsilon\right\}. \tag{14}$$

Note that $2\|W\|^2 \sim \chi^2(2kn^2)$, and the tail of the chi-square distribution is bounded as follows for $\epsilon \in (0,1)$ [44]:

$$\mathbb{P}\left\{\frac{\|W\|^2}{kn^2} \geq 1 + \epsilon\right\} \leq 2e^{-\frac{kn^2\epsilon^2}{8}}. \tag{15}$$

Thus, the first term in equation (14) vanishes exponentially fast as $k \to \infty$.

In order to provide an upper bound for the second term, we consider a lower bound for the minimum distance in the received lattice. We have

$$\begin{aligned} d_H^2 &\geq \alpha^2 nk \min_{X \in L_{n,k} \setminus \{0\}} \prod_{i=1}^{k} |\det(H_i X_i)|^{\frac{2}{nk}} \\ &\geq \alpha^2 nk \prod_{i=1}^{k} |\det(H_i)|^{\frac{2}{nk}}, \end{aligned}$$

where the first bound comes from Lemma 3.5 and the second from the hypothesis that $\det_{\min}(L_{n,k}) = 1$. Therefore, the second term in (14) is upper bounded by

$$\mathbb{P}\left\{\frac{\alpha^2}{4n} \prod_{i=1}^{k} |\det(H_i)|^{\frac{2}{nk}} < 1 + \epsilon\right\}. \tag{16}$$

b) Case $n_r > n$: In this case, the lattice $H * L_{n,k}$ is $2n^2k$-dimensional but is contained in a $2n_r nk$-dimensional space. For all $i = 1, \ldots, k$, consider the QR decomposition

$$H_i = Q_i R_i, \quad Q_i \in M_{n_r \times n_r}(\mathbb{C}), \quad R_i \in M_{n_r \times n}(\mathbb{C}),$$

where Q_i is unitary and R_i is upper triangular. We have $Q_i = [Q_i' \; Q_i'']$, where $Q_i' \in M_{n_r \times n}(\mathbb{C})$ is such that $(Q_i')^\dagger Q_i' = I_n$, and $R_i = \begin{bmatrix} R_i' \\ 0 \end{bmatrix}$, with $R_i' \in M_n(\mathbb{C})$ upper triangular. Note that the "thin" QR decomposition $H_i = Q_i' R_i'$ also holds.

Multiplying the channel equation (2) by $Q_i^\dagger$, we obtain the equivalent system

$$\tilde{Y}_i = Q_i^\dagger Y_i = R_i X_i + Q_i^\dagger W_i$$

for all $i = 1, \ldots, k$. Note that

$$\tilde{Y}_i = \begin{bmatrix} Y_i' \\ Y_i'' \end{bmatrix} = \begin{bmatrix} R_i' X_i + (Q_i')^\dagger W_i \\ (Q_i'')^\dagger W_i \end{bmatrix}.$$

Thus, the second component contains only noise and no information. The output of the naive lattice decoder can be written as

$$\begin{aligned} \hat{X}_{\text{NLD}} &= \operatorname*{argmin}_{X' \in \alpha L_{n,k}} \sum_{i=1}^{k} \|Y_i - H_i X_i'\|^2 \\ &= \operatorname*{argmin}_{X' \in \alpha L_{n,k}} \sum_{i=1}^{k} \left\|\tilde{Y}_i - R_i X_i'\right\|^2 \\ &= \operatorname*{argmin}_{X' \in \alpha L_{n,k}} \sum_{i=1}^{k} \left(\|Y_i' - R_i' X_i'\|^2 + \left\|(Q_i'')^\dagger W_i\right\|^2\right) \\ &= \operatorname*{argmin}_{X' \in \alpha L_{n,k}} \sum_{i=1}^{k} \|Y_i' - R_i' X_i'\|^2, \end{aligned}$$

since the second component does not depend on the lattice point X'. Thus, the naive lattice decoder for the original system declares an error if and only if the naive lattice decoder for the (n, n, k) multi-block system with components

$$Y_i' = R_i' X_i + (Q_i')^\dagger W_i = R_i' X_i + W_i'$$

does. (Note that $W_i' = (Q_i')^\dagger W_i$ is an $n \times n$ matrix with i.i.d. Gaussian entries of variance 1 per complex dimension.) A similar reasoning holds for the ML decoder.

Let $d_{R'}$ be the minimum distance in the $2n^2k$-dimensional lattice generated by $R' = [R_1', \ldots, R_k']$:

$$d_{R'} = \min_{\substack{X, \bar{X} \in \alpha L_{n,k} \\ X \neq \bar{X}}} \sum_{i=1}^{k} \left\|R_i'(X_i - \bar{X}_i)\right\|^2.$$

Observe that $\forall i = 1, \ldots, k$,

$$\left\|H_i(X_i - \bar{X}_i)\right\|^2 = \left\|Q_i' R_i'(X_i - \bar{X}_i)\right\|^2 = \left\|R_i'(X_i - \bar{X}_i)\right\|^2$$

Thus, $d_H = d_R'$. Moreover, $\det(H_i^\dagger H_i) = \det((R_i')^\dagger R_i') = |\det(R_i')|^2$. Similarly to the symmetric case, the error probability of the naive lattice decoder and of the ML decoder can be bounded by

$$P_e \leq \mathbb{P}\left\{\|W'\|^2 \geq \left(\frac{d_{R'}}{2}\right)^2\right\},$$

where $W' = [W_1', \ldots, W_k']$. We can write

$$d_{R'}^2 \geq \alpha^2 nk \prod_{i=1}^{k} |\det(R_i')|^{\frac{2}{nk}} = \alpha^2 nk \prod_{i=1}^{k} \det(H_i^\dagger H_i)^{\frac{1}{nk}}.$$

The proof then follows exactly the same steps as in the symmetric case. □

Proof of Theorem 4.1: The second term in (13) can be rewritten as

$$\mathbb{P}\left\{\frac{1}{k} \sum_{i=1}^{k} \frac{1}{n} \log\det(H_i^\dagger H_i) < \log\left(\frac{4n(1+\epsilon)}{\alpha^2}\right)\right\},$$

and will vanish as long as

$$\log\left(\frac{4n(1+\epsilon)}{\alpha^2}\right) < \frac{\mu}{n}.$$

Recalling that the normalization constant α^2 in equation (4) satisfies

$$\alpha^2 \geq \frac{C_{n,k}^{1/n^2k} P}{2^{R/n} C_L}$$

under the hypothesis that $\mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L$, a sufficient condition to have vanishing error probability is

$$\frac{R}{n} < \log P + \frac{\mu}{n} - \log(4n(1+\epsilon)) - \log C_L + \frac{1}{n^2k} \log C_{n,k}$$

From Stirling's approximation, for large k we have

$$(C_{n,k})^{\frac{1}{n^2k}} \approx \pi e / (n(2\pi n^2 k)^{\frac{1}{2n^2k}}). \qquad (17)$$

Since $\frac{1}{2nk} \log 2\pi n^2 k \to 0$ when $k \to \infty$, any rate

$$R < \mu + n(\log P - \log(4n(1+\epsilon)) - \log C_L + \log \pi e - \log n)$$

is achievable. This holds $\forall \epsilon > 0$, and concludes the proof. □

Remark 4.6: The two-sided convergence in probability in equation (10) is actually not required in the proof of Theorem 4.1. The theorem still holds provided that $\forall \epsilon > 0$,

$$\lim_{k \to \infty} \mathbb{P} \left\{ \mu - \frac{1}{k} \sum_{i=1}^{k} \log \det(H_i^\dagger H_i) > \epsilon \right\} = 0. \qquad (18)$$

Moreover, if we have exponentially fast convergence in (18), then the error probability P_e also vanishes exponentially fast when $k \to \infty$.

As a final remark, we note that for the ML decoder we can prove an analogue of Theorem 4.1 also in the case $n_r < n$, although the bound on achievable rates is more involved:

Theorem 4.7: Suppose that $n_r < n$, and let $\{H_i\}_{i \in \mathbb{Z}}$ be a fading process such that $H_i \in M_{n_r \times n}$ is full-rank with probability 1. Suppose that the weak law of large numbers holds for the random variables $\log \det(H_i H_i^\dagger)$, i.e. $\exists \mu > 0$ such that $\forall \epsilon > 0$,

$$\lim_{k \to \infty} \mathbb{P} \left\{ \left| \frac{1}{k} \sum_{i=1}^{k} \log \det(H_i H_i^\dagger) - \mu \right| > \epsilon \right\} = 0. \qquad (19)$$

Let $L_{n,k} \subset M_{n \times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional multi-block lattice codes satisfying (11). Then, any rate

$$R < \mu + n_r(\log P - 1) + (n - n_r)\log(n - n_r) - n \log \frac{2n^2 C_L}{\pi e}$$

is achievable using the codes $L_{n,k}$ with ML decoding.

We remark that the result does *not* extend to the naive lattice decoder. The proof of Theorem 4.7 can be found in Appendix A.

V. Achieving constant gap to capacity for Gaussian MIMO channels and ergodic channels

A. Gaussian MIMO channel

We now consider a deterministic model, where $H_i = H$ is constant. When H is known both at the transmitter and receiver, the channel capacity is given by [3]

$$C(P) = \max_{Q_{\mathbf{x}} \geq 0, \mathrm{tr}(Q_{\mathbf{x}}) \leq P} \log \det(I_{n_r} + H Q_{\mathbf{x}} H^\dagger), \qquad (20)$$

where $Q_{\mathbf{x}}$ is the covariance matrix of the input $\mathbf{x}$ for a single channel use. However, if the channel is known at the receiver but not at the transmitter, the transmitter cannot use optimal power allocation and waterfilling, and can only achieve the *white-input capacity* corresponding to uniform power allocation $Q_{\mathbf{x}} = \frac{P}{n} I_n$:

$$C_{\mathrm{WI}} = \log \det \left(I_{n_r} + \frac{P}{n} H H^\dagger \right) = \log \det \left(I_n + \frac{P}{n} H^\dagger H \right).$$

This is for example the case for an open-loop broadcast channel where the transmitter cannot perform rate adaptation for all the users.

Clearly, Theorems 4.1 and 4.7 apply to the deterministic channel scenario since the law of large numbers holds. Moreover, the convergence of the error probability to zero will be exponential, since the second term in equation (13) is actually zero. The following corollary then shows that a constant gap to white-input capacity is achievable:

Corollary 5.1: Consider a deterministic channel such that $H_i = H$ for all $i \geq 1$, and let $L_{n,k} \subset M_{n \times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional multi-block lattice codes such that $\det_{\min}(L_{n,k}) = 1$ and $\mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L$. Then, this coding scheme can achieve any rate

$$R < \log \det \frac{P}{n} H^\dagger H - n \log C_L - n \log \frac{4n}{\pi e}$$

if $n_r \geq n$, and any rate

$$R < \log \det \frac{P}{n} H H^\dagger - 2n_r - (n - n_r) \log \frac{n}{n - n_r} - n \log \frac{n C_L}{\pi e}$$

if $n_r < n$.

Remark 5.2: In the case $n_r \geq n$, let λ_i, $i = 1, \ldots, n$ be the singular values of H. Then the channel capacity can be written as $C(P) = \sum_{i=1}^{n} \log\left(1 + \frac{P}{n}\lambda_i\right)$. The previous corollary shows that the achievable rate is of the form

$$R(P) = \max \left(0, \log \det \frac{P}{n} H^\dagger H - c \right)$$

for some constant $c > 0$. Let $P_{\min}$ be the smallest value of P such that $R(P) > 0$ if $P > P_{\min}$. Then, for $P \leq P_{\min}$ we have that $C(P) - R(P) = C(P) \leq C(P_{\min})$, while for $P > P_{\min}$, $C(P) - R(P) = \sum_{i=1}^{n} \log\left(1 + \frac{n}{P\lambda_i}\right) + c$ which is a strictly decreasing function of P and tends to c when $P \to \infty$. Thus, for all $P > 0$ we have that $C(P) - R(P) \leq C(P_{\min})$. This shows that the gap is bounded by a constant for all SNR, however the value of $P_{\min}$ and therefore the constant depends on the channel H. As a consequence, the supremum of the gap over all deterministic channels H is not bounded. This is an artifact of the spherical shaping technique which incurs the loss of "+1" in the capacity formula.

A similar argument holds for $n_r < n$.

Example 5.3 (AWGN channel): It follows from Corollary 5.1 that any rate

$$R < \log P - \log \frac{2G}{\pi e} \qquad (21)$$

is achievable with the proposed scheme on the single antenna AWGN channel. In this case, we have $P_{\min} = \frac{2G}{\pi e} = 21.63$ (or equivalently, 13.35 dB). The maximum gap to capacity is $C(P_{\min}) = \log(1 + P_{\min}) \approx 4.50$ bits. The achievable rate as a function of SNR (in dB) is plotted in Figure 1.

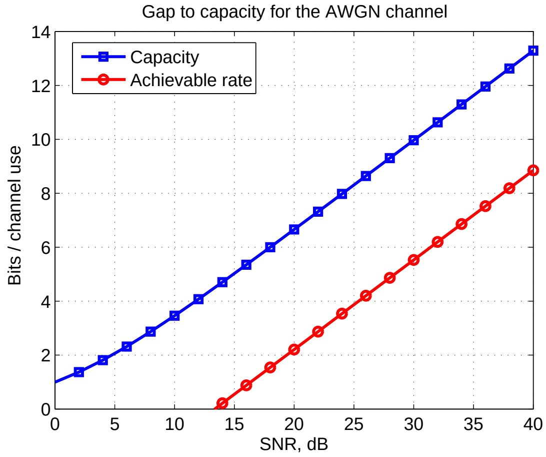

Fig. 1. Achievable rate on a single-antenna AWGN channel.

Similarly, one can compute the penalty in dB which is incurred when using the proposed scheme compared to a capacity-achieving scheme. To obtain a rate $C(P)$ with our scheme, we need power P_{eq} such that $\log(1+P) = \log P_{\mathrm{eq}} - \log \frac{2G}{\pi e}$, or equivalently $P_{\mathrm{eq}} = (1+P)\frac{2G}{\pi e}$. The penalty $P_{\mathrm{eq}}(\mathrm{dB}) - P(\mathrm{dB}) = 10\log_{10}\left(\frac{2G}{\pi e}\left(1+\frac{1}{P}\right)\right)$ is a strictly decreasing function of P, and is equal to $16.36\,\mathrm{dB}$ for $P(\mathrm{dB}) = 0\,\mathrm{dB}$. When $P \to \infty$, the penalty tends to $10\log_{10}(\frac{2G}{\pi e}) = 13.35\,\mathrm{dB}$.

B. Stationary ergodic channels

We now specialize the results of Section IV to the case where the fading process $\{H_i\}$ is *ergodic* and *stationary*. For the sake of completeness, we review the relevant definitions here.

Let $\mathcal{I}$ be the set $\mathbb{Z}$ or $\mathbb{N}$, and consider a random process $X^{\mathcal{I}} = \{X_i\}_{i\in\mathcal{I}}$ on a probability space $(\Omega, \mathcal{B}, \mathbb{P})$ where each random variable X_i takes values in a separable Banach space $\mathcal{X}$. The sequence space $(\mathcal{X}^{\mathcal{I}}, \mathcal{B}(\mathcal{X}^{\mathcal{I}}))$ with the Borel sigma-algebra inherits a probability measure m_X from the underlying probability space, defined by

$$m_X(A) = \mathbb{P}\left\{\omega :\ X^{\mathcal{I}}(\omega) \in A\right\} \quad \forall A \in \mathcal{B}(\mathcal{X}^{\mathcal{I}}). \tag{22}$$

Definition 5.4: The process $\{X_i\}$ is called *stationary* if $\forall t, k \in \mathbb{N}$, $\forall i_1, i_2, \ldots, i_k \in \mathcal{I}$, the joint distribution of $(X_{i_1}, X_{i_2}, \ldots, X_{i_k})$ is the same as that of $(X_{i_1+t}, X_{i_2+t}, \ldots, X_{i_k+t})$.

In this case it is well-known [45, p. 494] that the measure m_X is invariant with respect to the shift map $T : \mathcal{X}^{\mathcal{I}} \to \mathcal{X}^{\mathcal{I}}$ such that $T(\{x_i\}) = \{x_{i+1}\}$.

Definition 5.5: The process $\{X_i\}$ is called *ergodic* if $\forall A \in \mathcal{B}(\mathcal{X}^{\mathcal{I}})$ such that $T^{-1}(A) = A$, we have that $m_X(A)$ is equal to 0 or 1.

We now go back to the channel model (2). For the sake of simplicity, we suppose that $n_r \geq n$. If the fading process $\{H_i\}$ is stationary and ergodic, it is not hard to see that the random process $\{X_i\} = \left\{\log\det(H_i^\dagger H_i)\right\}$ taking values in $\mathcal{X} = \mathbb{R}$ is also stationary and ergodic, and the shift $T : \mathbb{R}^{\mathcal{I}} \to \mathbb{R}^{\mathcal{I}}$ preserves the measure m_X defined in (22).

For an ergodic process such that the shift T is measure-preserving, Birkhoff's theorem [46] guarantees that for any $f \in L^1(\mathcal{X}^{\mathcal{I}}, \mathcal{B}(\mathcal{X}^{\mathcal{I}}), m_X)$, the sample means with respect to f converge almost everywhere: for almost all $\{x_i\} \in \mathcal{X}^{\mathcal{I}}$,

$$\lim_{k\to\infty} \frac{1}{k} \sum_{n=1}^{k} f(T^n(\{x_i\})) = \int_{\mathcal{X}^{\mathcal{I}}} f dm_X. \tag{23}$$

In particular, the projection $\Pi : \mathbb{R}^{\mathcal{I}} \to \mathbb{R}$ on the first coordinate is L^1 according to the image measure m_X if and only if $\mathbb{E}\left[\left|\log\det H^\dagger H\right|\right] < \infty$. Under this hypothesis, Birkhoff's theorem implies the law of large numbers:

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} X_i = \int_{\mathbb{R}^{\mathcal{I}}} \Pi(\{x_i\}) dm_X(\{x_i\}) = \int_{\Omega} \Pi \circ X^{\mathcal{I}} d\mathbb{P}$$

$$= \int_{\Omega} X_1 d\mathbb{P} = \mathbb{E}[X] \quad \text{a.e.} \tag{24}$$

In other words,

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} \log\det(H_i^\dagger H_i) = \mathbb{E}_H\left[\log\det(H^\dagger H)\right] \tag{25}$$

almost everywhere.

In the ergodic stationary case, it is well-known [3, 47] that the ergodic capacity of the channel is well-defined and does not depend on the channel correlation with respect to time, but only on its first order statistics. Given a power constraint P in equation (3), the ergodic capacity (per channel use) is equal to

$$C(P) = \max_{Q_{\mathbf{x}} \geq 0, \operatorname{tr}(Q_{\mathbf{x}}) \leq P} \mathbb{E}_H\left[\log\det(I_{n_r} + H Q_{\mathbf{x}} H^\dagger)\right],$$

where H is a random matrix with the same first-order distribution of the process $\{H_i\}$, which is independent of time by stationarity, and $Q_{\mathbf{x}}$ is the covariance matrix of the input $\mathbf{x}$ for one channel use[1].

If we suppose that the channel is *isotropically invariant*, i.e. the distribution of H is invariant under right multiplication by unitary matrices, then under the assumption of no CSI at the transmitter, the optimal input covariance matrix is $Q_{\mathbf{x}} = \frac{P}{n} I_n$ [3] and we have

$$C(P) = \mathbb{E}_H\left[\log\det\left(I_{n_r} + \frac{P}{n} H H^\dagger\right)\right].$$

Since $\det(I_{n_r} + \frac{P}{n} H H^\dagger) = \det(I_n + \frac{P}{n} H^\dagger H)$, we can also write

$$C(P) = \mathbb{E}_H\left[\log\det\left(I_n + \frac{P}{n} H^\dagger H\right)\right].$$

The following Corollary to Theorem 4.1 shows that in this case, the proposed multi-block codes can achieve a constant gap to ergodic capacity.

Corollary 5.6: Suppose that $n_r \geq n$ and that the fading process $\{H_i\}$ is ergodic, stationary and isotropically invariant. Moreover, suppose that $\mathbb{E}\left[\left|\log\det H^\dagger H\right|\right] < \infty$. Let $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional

[1] We note that the capacity (per channel use) of the block fading MIMO channel of finite block length T with perfect channel state information at the receiver is independent of T [48, eq. (9)]. So the previous result still holds in the multi-block case.

multi-block lattice codes such that $\det_{\min}(L_{n,k}) = 1$ and $\mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L$. Then, any rate

$$R < \mathbb{E}_H \left[\log \det \frac{P}{n} H^\dagger H \right] - n \log C_L - n \log \frac{4n}{\pi e}$$

is achievable using the codes $L_{n,k}$ both with ML decoding and naive lattice decoding.

Proof: From equation (25), we have that the hypotheses of Theorem 4.1 are satisfied (actually, only the weak law of large numbers was required). Consequently, any rate

$$\begin{aligned} R &< n \log P + \mathbb{E}_H \left[\log \det H^\dagger H \right] - n \log C_L - n \log \frac{4n^2}{\pi e} \\ &= \log \left(\frac{P}{n} \right)^n + \mathbb{E}_H \left[\log \det H^\dagger H \right] - n \log C_L - n \log \frac{4n}{\pi e} \\ &= \mathbb{E}_H \left[\log \det \frac{P}{n} H^\dagger H \right] - n \log C_L - n \log \frac{4n}{\pi e} \end{aligned}$$

is achievable. □

A similar corollary to Theorem 4.7 holds in the case $n_r < n$:

Corollary 5.7: Suppose that $n_r < n$ and that the fading process $\{H_i\}$ is ergodic, stationary and isotropically invariant. Moreover, suppose that $\mathbb{E}\left[\left| \log \det HH^\dagger \right| \right] < \infty$. Let $L_{n,k} \subset M_{n \times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional multi-block lattice codes such that $\det_{\min}(L_{n,k}) = 1$ and $\mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L$. Then, any rate R lower than

$$\mathbb{E}_H \left[\log \det \frac{P}{n} HH^\dagger \right] - 2n_r - (n - n_r) \log \frac{n}{n - n_r} - n \log \frac{nC_L}{\pi e}$$

is achievable using the codes $L_{n,k}$ with ML decoding.

Remark 5.8: Using the same argument as in Remark 5.2, we can show that the achievable rate is within a constant gap from capacity, although this constant will depend on the channel statistics.

Under the hypothesis $\mathbb{E}_H \left[\left| \log \det H^\dagger H \right| \right]$, we have $\left| \mathbb{E}_H \left[\log \det H^\dagger H \right] \right| < \infty$. The achievable rate is of the form $R(P) = \max\left(0, n \log \frac{P}{n} + \mathbb{E}_H[\log \det H^\dagger H] - c\right)$ for some constant $c > 0$. Let $P_{\min}$ be the smallest value of P such that $R(P) > 0$ if $P > P_{\min}$. For $P \leq P_{\min}$ we have that $C(P) - R(P) = C(P) \leq C(P_{\min})$.

Let λ_i, $i = 1, \ldots, n$ be the (random) singular values of H. For $P > P_{\min}$, $C(P) - R(P) = \sum_{i=1}^n \mathbb{E}_H \left[\log \left(1 + \frac{n}{P\lambda_i} \right) \right] + c$ which is a strictly decreasing function of P and tends to c when $P \to \infty$. This shows that the gap is uniformly bounded by a constant which depends on the channel statistics.

VI. Achievable rates and error probability bounds for i.i.d. Rayleigh fading channels

We now suppose that the entries of H_i are i.i.d. circular symmetric complex Gaussian with zero mean and unit variance per complex dimension, and that the fading blocks H_i are independent. In this case, the achievable rate can be computed explicitly, and we can prove that the error probability vanishes exponentially fast.

Let $\psi(x) = \frac{d}{dx} \ln \Gamma(x)$ denote the Digamma function. Then we have the following:

Proposition 6.1: Let $L_{n,k} \subset M_{n \times nk}(\mathbb{C})$ be a family of $2n^2k$-dimensional multi-block lattice codes such that $\det_{\min}(L_{n,k}) = 1$ and $\mathrm{Vol}(L_{n,k})^{\frac{1}{n^2k}} \leq C_L$. Then, over the (n, n_r, k) multi-block channel, these codes achieve any rate

$$R < \mathbb{E}_H \left[\log \det \frac{P}{n} H^\dagger H \right] - n \log C_L - n \log \frac{4n}{\pi e},$$

where

$$\mathbb{E}_H \left[\log \det \frac{P}{n} H^\dagger H \right] = n \log \frac{P}{n} e^{\frac{1}{n} \sum\limits_{i=n_r-n+1}^{n_r} \psi(i)}. \tag{26}$$

Moreover, the error probability vanishes exponentially fast.

Proof of Proposition 6.1: The first statement follows from Corollary 5.6. The next step is to prove equation (26). It is well-known [49, 50] that if H is an $n_r \times n$ matrix with i.i.d. complex Gaussian entries having variance per real dimension $1/2$, the random variable $\det(H^\dagger H)$, corresponding to the determinant of the Wishart matrix $H^\dagger H$, is distributed as the product

$$V_{n,n_r} = Z_{n_r-n+1} Z_{n_r-n+2} \cdots Z_{n_r}$$

of n independent variables, such that $\forall j = n_r - n + 1, \ldots, n_r$, $2Z_j$ is a chi square random variable with $2j$ degrees of freedom. The density of Z_j is $p_{Z_j}(x) = \frac{x^{j-1}e^{-x}}{\Gamma(j)}$. We have

$$\mathbb{E}[\ln Z_j] = \frac{1}{\Gamma(j)} \int_0^\infty x^{j-1} e^{-x} \ln x \, dx = \psi(j),$$

$$M_{n,n_r} = \mathbb{E}[\ln V_{n,n_r}] = \sum_{j=n_r-n+1}^{n_r} \psi(j) = \mathbb{E}_H \left[\ln \det H^\dagger H \right].$$

Then if we consider the base 2 logarithm, we find

$$\begin{aligned} \mathbb{E}_H \left[\log \det H^\dagger H \right] &= \mathbb{E}[\log V_{n,n_r}] = \frac{M_{n,n_r}}{\ln 2} \\ &= \frac{\sum_{j=n_r-n+1}^{n_r} \psi(j)}{\ln 2} = n \log e^{\frac{1}{n} \sum_{j=n_r-n+1}^{n_r} \psi(j)} \end{aligned}$$

which concludes the proof of equation (26).

In order to show that the error probability converges exponentially fast, by Remark 4.6 it is enough to show that we have exponential convergence in equation (18).

Consider a sequence of i.i.d. random variables $\ln V_{n,n_r}^{(i)}$, $i = 1, \ldots, k$, with the same distribution as $\ln V_{n,n_r}$. Using the Chernoff bound [51], given $\delta > 0$, $\forall v > 0$ we have

$$\begin{aligned} &\mathbb{P} \left\{ \frac{M_{n,n_r}}{\ln 2} - \frac{1}{k} \sum_{i=1}^k \log \det H_i^\dagger H_i \geq \frac{\delta}{\ln 2} \right\} \\ &= \mathbb{P} \left\{ M_{n,n_r} - \frac{1}{k} \sum_{i=1}^k \ln \det H_i^\dagger H_i \geq \delta \right\} \\ &= \mathbb{P} \left\{ M_{n,n_r} - \frac{1}{k} \sum_{i=1}^k \ln V_{n,n_r}^{(i)} \geq \delta \right\} \\ &\leq e^{kv(M_{n,n_r} - \delta)} \left(\mathbb{E}[e^{-v \ln V_{n,n_r}}] \right)^k \end{aligned} \tag{27}$$

The tightest bound in (27) is obtained for v_δ such that

$$\mathbb{E}[-\ln V_{n,n_r} e^{-v_\delta \ln V_{n,n_r}}] = (\delta - M_{n,n_r}) \mathbb{E}[e^{-v_\delta \ln V_{n,n_r}}].$$

Observe that

$$\mathbb{E}[Z_j^{-v}] = \frac{1}{\Gamma(j)} \int_0^\infty x^{j-1-v} e^{-x} dx = \frac{\Gamma(j-v)}{\Gamma(j)}, \tag{28}$$

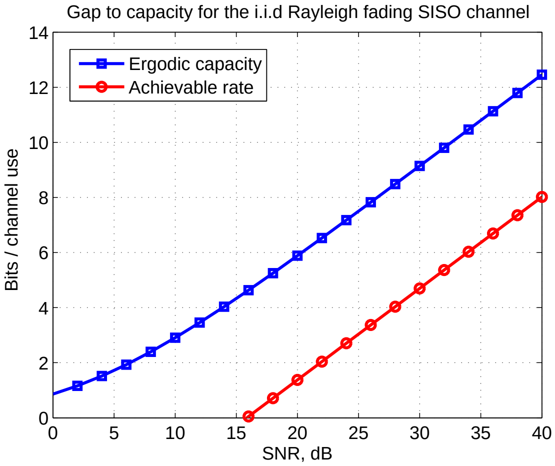

Fig. 2. Achievable rate and channel capacity for the single antenna i.i.d. Rayleigh fading channel.

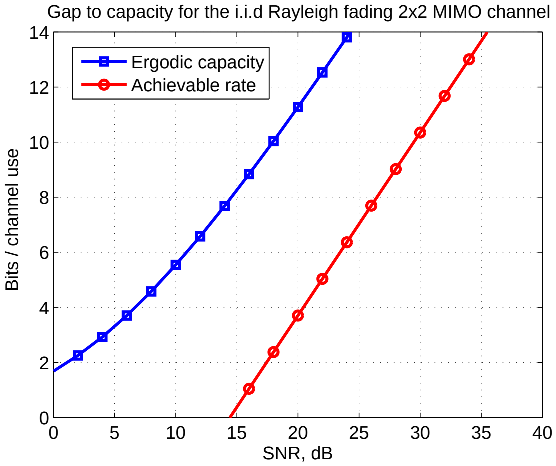

Fig. 3. Achievable rate and channel capacity for the 2×2 MIMO i.i.d. Rayleigh fading channel.

$$\mathbb{E}[Z_j^{-v} \ln Z_j] = \frac{1}{\Gamma(j)} \int_0^\infty x^{j-1-v} e^{-x} \ln x \, dx$$
$$= \frac{\Gamma(j-v)}{\Gamma(j)} \psi(j-v). \tag{29}$$

Thus we find

$$\mathbb{E}\left[e^{-v \ln V_{n,n_r}}\right] = \mathbb{E}\left[V_{n,n_r}^{-v}\right] = \prod_{j=n_r-n+1}^{n_r} \mathbb{E}\left[Z_j^{-v}\right]$$
$$= \prod_{j=n_r-n+1}^{n_r} \frac{\Gamma(j-v)}{\Gamma(j)},$$
$$\mathbb{E}\left[-\ln V_{n,n_r} e^{-v \ln V_{n,n_r}}\right] = \mathbb{E}\left[-V_{n,n_r}^{-v} \ln V_{n,n_r}\right]$$
$$= \sum_{j=n_r-n+1}^{n_r} \mathbb{E}\left[-\ln Z_j \prod_{l=n_r-n+1}^{n_r} Z_l^{-v}\right]$$
$$= \sum_{j=n_r-n+1}^{n_r} \left(\prod_{l \neq j} \mathbb{E}[Z_l^{-v}]\right) \mathbb{E}[-Z_j^{-v} \ln Z_j]$$
$$= -\sum_{j=n_r-n+1}^{n_r} \left(\prod_{l \neq j} \frac{\Gamma(l-v)}{\Gamma(l)}\right) \frac{\Gamma(j-v)}{\Gamma(j)} \psi(j-v)$$
$$= -\prod_{l=n_r-n+1}^{n_r} \frac{\Gamma(l-v)}{\Gamma(l)} \sum_{j=n_r-n+1}^{n_r} \psi(j-v).$$

Consequently, the tightest bound in (27) is achieved when

$$\delta = \sum_{l=n_r-n+1}^{n_r} (\psi(l) - \psi(l-v_\delta)). \tag{30}$$

Note that as $\delta \to 0$, $v_\delta \to 0$. The right-hand side in equation (27) for $v = v_\delta$ can be rewritten as

$$e^{k v_\delta (-\delta + \sum_{j=n_r-n+1}^{n_r} \psi(j))} \left(\prod_{l=n_r-n+1}^{n_r} \frac{\Gamma(l-v_\delta)}{\Gamma(l)}\right)^k$$
$$= e^{k(-v_\delta \delta + \sum_{j=n_r-n+1}^{n_r} (v_\delta \psi(j) + \ln \Gamma(j-v_\delta) - \ln \Gamma(j)))}$$
$$= e^{k \sum_{j=n_r-n+1}^{n_r} (v_\delta \psi(j-v_\delta) - \ln \Gamma(j) + \ln \Gamma(j-v_\delta))}$$

using (30).
Recall that $\Gamma(x)$ is monotone decreasing for $0 < x < a_0 = 1.461632\ldots$ and monotone increasing for $x > a_0$. Using the mean value theorem for the function $\ln \Gamma(x)$ in the interval $[i-v_\delta, i]$ we get that for $i = 1$, $v_\delta \psi(1-v_\delta) + \ln \Gamma(1-v_\delta) \leq 0$, and for $i \geq 2$, $v_\delta \psi(i-v_\delta) \leq \ln \Gamma(i) - \ln \Gamma(i-v_\delta)$. Thus, the exponent is negative both for $n = n_r$ and for $n > n_r$. We can conclude that

$$\mathbb{P}\left\{ \frac{M_{n,n_r}}{\ln 2} - \frac{1}{k} \sum_{i=1}^{k} \log \det H_i^\dagger H_i \geq \frac{\delta}{\ln 2} \right\} \leq e^{-k K_{n,n_r,\delta}}$$

for some positive constant $K_{n,n_r,\delta}$. Since the bound holds $\forall \delta > 0$, using Remark 4.6 with $\mu = \frac{M_{n,n_r}}{\ln 2}$, we find that the error probability tends to 0 exponentially fast for any rate

$$R < n \left(\log \frac{P}{n} e^{\frac{1}{n} \sum_{i=n_r-n+1}^{n_r} \psi(i)} - \log C_L - \log \frac{4n}{\pi e} \right). \quad \square$$

Corollary 6.2: Over the (n, n, k) multi-block channel, reliable communication is guaranteed when $k \to \infty$ for rates

$$R < n \left(\log \frac{P}{n} e^{\frac{1}{n} \sum_{i=1}^{n} \psi(i)} - \log \frac{2n}{\pi e} - \log 23^{\frac{1}{10}\left(1-\frac{1}{n}\right)} G \right)$$

using the multi-block code construction in Proposition 3.10.

Example 6.3: The achievable rates for the single antenna and 2×2 MIMO i.i.d. Rayleigh fading channel are plotted in Figures 2 and 3 respectively. The maximum gap to capacity is approximately 4.5 bits for the SISO case and 8 bits for the 2×2 MIMO case.

VII. Existence of asymptotically good lattices

All of our capacity results depend on the existence of lattices with asymptotically good normalized minimum determinants, which was claimed in Section III-B. In this section we will prove this result.
We will first recall the construction of single-block space-time codes from cyclic division algebras (see for example [23]). Due to space constraints, we refer the reader to [52] for algebraic definitions.

Definition 7.1: Let K be an algebraic number field of degree m and assume that E/K is a cyclic Galois extension of degree n with Galois group $Gal(E/K) = \langle \sigma \rangle$. We can define an associative K-algebra

$$\mathcal{A} = (E/K, \sigma, \gamma) = E \oplus uE \oplus u^2 E \oplus \cdots \oplus u^{n-1} E,$$

where $u \in \mathcal{A}$ is an auxiliary generating element subject to the relations $xu = u\sigma(x)$ for all $x \in E$ and $u^n = \gamma \in K^*$. We call the resulting algebra a *cyclic algebra*.
Here K is the center of the algebra $\mathcal{A}$.

Definition 7.2: We call $\sqrt{[\mathcal{A} : K]}$ the *degree* of the algebra $\mathcal{A}$. It is easily verified that the degree of $\mathcal{A}$ is equal to n.

We consider $\mathcal{A}$ as a right vector space over E and note that every element $a = x_0 + ux_1 + \cdots + u^{n-1}x_{n-1} \in \mathcal{A}$ has the following representation as a matrix:

$$\phi(a) = \begin{pmatrix} x_0 & \gamma\sigma(x_{n-1}) & \gamma\sigma^2(x_{n-2}) & \cdots & \gamma\sigma^{n-1}(x_1) \\ x_1 & \sigma(x_0) & \gamma\sigma^2(x_{n-1}) & & \gamma\sigma^{n-1}(x_2) \\ x_2 & \sigma(x_1) & \sigma^2(x_0) & & \gamma\sigma^{n-1}(x_3) \\ \vdots & & & & \vdots \\ x_{n-1} & \sigma(x_{n-2}) & \sigma^2(x_{n-3}) & \cdots & \sigma^{n-1}(x_0) \end{pmatrix}$$

The mapping ϕ is called the *left regular representation* of $\mathcal{A}$ and allows us to embed any cyclic algebra into $M_n(\mathbb{C})$. Under such an embedding $\phi(\mathcal{A})$ forms an mn^2-dimensional $\mathbb{Q}$-vector space.

We are particularly interested in algebras $\mathcal{A}$ for which $\phi(a)$ is invertible for all non-zero $a \in \mathcal{A}$.

Definition 7.3: A cyclic K-algebra $\mathcal{D}$ is a *division algebra* if every non-zero element of $\mathcal{D}$ is invertible.

If we assume that $\mathcal{D}$ is a division algebra, then ϕ is an injective mapping to $M_n(\mathbb{C})$ and every non-zero element in $\phi(\mathcal{D})$ is invertible. However, $\phi(\mathcal{D})$ is not a lattice. Therefore we will instead consider a suitable subset of $\mathcal{D}$.

Definition 7.4: A *$\mathbb{Z}$-order* Λ in $\mathcal{D}$ is a subring of $\mathcal{D}$ having the same identity element as $\mathcal{D}$, and such that Λ is a finitely generated module over $\mathbb{Z}$ which generates $\mathcal{D}$ as a linear space over $\mathbb{Q}$.

With the previous definition, the set $\phi(\Lambda)$ is a matrix lattice that can be used for coding over a single space-time block.

A generalization of the embedding ϕ to the multi-block case was proposed in [13, 14] for division algebras whose center K contains an imaginary quadratic field. In this paper we consider a more general multi-block construction developed in [18], which applies to any totally complex center K.
We say that a degree $2k$ number field K is totally complex if for every $\mathbb{Q}$-embedding $\beta_i : K \hookrightarrow \mathbb{C}$ the image $\beta_i(K)$ includes complex elements. The field K has $2k$ distinct $\mathbb{Q}$-embeddings $\beta_i : K \hookrightarrow \mathbb{C}$. As we assumed that K is totally complex, each of these embeddings is part of a complex conjugate pair. We will denote by $\overline{\beta_i}$ the embedding given by $x \mapsto \overline{\beta_i(x)}$.

For each β_i we can find an embedding $\alpha_i : E \hookrightarrow \mathbb{C}$ such that $\alpha_i|_K = \beta_i$. This choice can be made in such a way that $\overline{\alpha_i}|_K = \overline{\beta_i}$. We will suppose that the embeddings $\{\alpha_1, \ldots, \alpha_{2k}\}$ have been ordered in such a way that $\alpha_i = \overline{\alpha_{i+k}}$, for $1 \leq i \leq k$. Let a be an element of $\mathcal{D}$ and $A = \phi(a)$. Consider the mapping $\varphi : \mathcal{D} \mapsto M_{n \times nk}(\mathbb{C})$ given by

$$a \mapsto (\alpha_1(A), \ldots, \alpha_k(A)), \qquad (31)$$

where each α_i is extended to an embedding $\alpha_i : M_n(E) \hookrightarrow M_n(\mathbb{C})$.
The following result was proven in [18, Proposition 5]:

Proposition 7.5: Let Λ be a $\mathbb{Z}$-order in $\mathcal{D}$ and φ the previously defined embedding. Then $\varphi(\Lambda)$ is a $2kn^2$-dimensional lattice in $M_{n \times nk}(\mathbb{C})$ which satisfies

$$\det_{\min}(\varphi(\Lambda)) = 1, \;\; \mathrm{Vol}(\varphi(\Lambda)) = 2^{-kn^2}\sqrt{|d(\Lambda/\mathbb{Z})|}$$

and

$$\delta(\varphi(\Lambda)) = \left(\frac{2^{2kn^2}}{|d(\Lambda/\mathbb{Z})|}\right)^{1/4n}.$$

Here $d(\Lambda/\mathbb{Z})$ is the $\mathbb{Z}$-discriminant of the order Λ. It is a non-zero integer we can associate to any $\mathbb{Z}$-order of $\mathcal{D}$. We refer the reader to [52] for the relevant definitions.

We can now see that in order to maximize the minimum determinant of a multi-block code, we have to minimize the $\mathbb{Z}$-discriminant of the corresponding $\mathbb{Z}$-order Λ.

The first step to attack this question is to assume that Λ has some extra structure. Let $\mathcal{O}_K$ be the ring of algebraic integers of K. If we assume that Λ is also an $\mathcal{O}_K$ module, then the $\mathcal{O}_K$-discriminant of Λ is well-defined [52], and will be denoted by $d(\Lambda/\mathcal{O}_K)$. The following formula holds:

$$d(\Lambda/\mathbb{Z}) = N_{K/\mathbb{Q}}(d(\Lambda/\mathcal{O}_K))(d_K)^{n^2}, \qquad (32)$$

where $N_{K/\mathbb{Q}}$ is the algebraic norm in K and d_K is the discriminant of the field K.

In the case of fixed center K, [12] addressed the problem of finding the division algebras with the smallest $\mathcal{O}_K$-discriminant, yielding the densest MIMO lattices. The main construction is based on the following result (Theorem 6.14 in [12]):

Theorem 7.6: Let K be a number field of degree $2k$ and P_1 and P_2 be two prime ideals of K. Then there exists a degree n division algebra $\mathcal{D}$ having an $\mathcal{O}_K$-order Λ with discriminant

$$d(\Lambda/\mathbb{Z}) = (N_{K/\mathbb{Q}}(P_1)N_{K/\mathbb{Q}}(P_2))^{n(n-1)}(d_K)^{n^2}. \qquad (33)$$

Theorem 7.6 suggests that in order to build families of (n, n, k) multi-block codes with the largest normalized minimum determinant, we should proceed in four steps:

a) choose a sequence of center fields K of degree $2k$ such that their discriminants d_K grow as slowly as possible;
b) given the center K, choose an algebra $\mathcal{D}$ satisfying (33), where P_1 and P_2 are the prime ideals in K with the smallest norms[2];
c) find an order Λ of $\mathcal{D}$ which satisfies (33);
d) produce an explicit representation of $\mathcal{D}$ as a matrix lattice.

We now discuss the choice of a suitable sequence of center fields. The following theorem by Martinet [19] proves the existence of infinite sequences of totally complex number fields K with small discriminants d_K. As we will see, choosing such a field as the center of the algebra $\mathcal{D}$ is a key element to obtain a good normalized minimum determinant.

[2] However, we note [18] that *a priori* there may be a trade-off between these two choices, so that minimizing the two terms in (32) separately may be suboptimal.

Theorem 7.7 (Martinet): There exists an infinite tower of totally complex number fields $\{K_k\}$ of degree $2k$, where $2k = 5 \cdot 2^{2+t}$, such that

$$|d_{K_k}|^{\frac{1}{2k}} = G, \tag{34}$$

for $G \approx 92.368$.

The following Lemma shows that the number fields in the Martinet family have suitable primes of small norm yielding a good bound in Theorem 7.6.

Lemma 7.8: Every number field K_k in the Martinet family has ideals P_1 and P_2 such that

$$N_{K/\mathbb{Q}}(P_1) \leq 23^{k/10} \text{ and } N_{K/\mathbb{Q}}(P_2) \leq 23^{k/10}.$$

Proof: Every field K_k has a subfield $F = \mathbb{Q}(\cos(2\pi/11), \sqrt{2}, \sqrt{-23})$, where $[F : \mathbb{Q}] = 20$ (see for example [53, p. 395]). The field F has prime ideals B_1 and B_2 such that $N_{F/\mathbb{Q}}(B_i) = 23$. Let us now suppose that P_1 and P_2 are such prime ideals of K_k that $P_i \cap \mathcal{O}_F = B_i$. Transitivity of the norm then gives us that

$$N_{K_k/\mathbb{Q}}(P_i) = N_{F/\mathbb{Q}}(N_{K_k/F}(P_i)) \leq N_{F/\mathbb{Q}}(B_i)^{2k/20}. \quad \square$$

Armed with this observation, we can finally prove Proposition 3.10.

Proof of Proposition 3.10: Suppose that we have a degree $2k$ field extension K in the Martinet family of totally complex fields such that (34) holds. We know that this field K has some primes P_1 and P_2 such that $N_{K/\mathbb{Q}}(P_1) \leq 23^{k/10}$ and $N_{K/\mathbb{Q}}(P_2) \leq 23^{k/10}$. Then, there exists a central division algebra $\mathcal{D}$ of degree n over K, and a maximal order Λ of $\mathcal{D}$, such that

$$\begin{aligned} d(\Lambda/\mathbb{Z}) &= (N_{k/\mathbb{Q}}(P_1)N_{K/\mathbb{Q}}(P_2))^{n(n-1)}(d_K)^{n^2} \\ &\leq (23^{k/5})^{(n(n-1))}(G^{2k})^{n^2}. \end{aligned} \quad \square$$

Remark 7.9: The number field towers in Theorem 7.7 are not the best known possible. It was shown in [54] that one can construct a family of totally complex fields such that $G < 82.2$, but this choice would add some notational complications.

Remark 7.10: The existence of Martinet's family of number fields is based on the work of Golod and Shafarevich [55], where the authors prove that there exist fields K having infinite class field towers

$$K = K^{(0)} \subset K^{(1)} \subset \cdots \subset K^{(i)} \subset K^{(i+1)} \subset \cdots$$

Such towers are constructed recursively from a base field $K = K^{(0)}$, by considering its Hilbert class field $K^{(1)}$ and then repeating this process, in such a way that $K^{(i+1)}$ is always the Hilbert class field of $K^{(i)}$. Due to the properties of Hilbert class fields, the fields in such towers always have constant root discriminants.

Remark 7.11: Let us now detail the explicit steps needed to realize the proposed algebraic constructions as matrix lattices. As mentioned in the introduction, there exist algorithms to find the number fields K from the Martinet family, although the complexity of this task may turn out to be prohibitive. The second step is then to build an algebra $\mathcal{D}$ with the properties described in equation (33). We will not elaborate on this topic, but one can follow similar steps as in [12, Section VI]. The next step is to find a suitable order from the division algebra $\mathcal{D}$. Here one can use the algorithms given in [56]. The explicit presentations needed to turn these algebraic structures into lattices are obtained from cyclic representations as in equation (31).

Since the first step required to construct the center is generally too taxing, in practice for fixed (small) k one should not choose a field from the Martinet family, but instead choose a degree $2k$ totally complex field with the smallest known discriminant and then build the division algebra on top of this field. Although Lemma 7.8 will not necessarily hold, a trivial observation is that every number field of degree $2k$ has prime ideals P_1 and P_2 such that $N_{K/\mathbb{Q}}(P_1) \leq 2^{2k}$ and $N_{K/\mathbb{Q}}(P_2) \leq 3^{2k}$.

VIII. Corollaries for the single antenna fading channel

The single antenna fast fading channel is one of the special cases of the general channel model (2). It is particularly illuminating as the connection to the classical AWGN lattice coding is most striking. In this case the abstract matrix lattices of Section VII correspond to number field codes that have been studied for twenty years [22]. Due to the familiarity and simplicity of this model we can most easily compare our work to previous research on the topic.

In the single antenna case the channel model (2) gets simplified to

$$y_i = h_i \cdot x_i + w_i, \tag{35}$$

where x_i are the transmitted symbols, and $\forall i = 1, \ldots, k$, w_i are i.i.d. complex Gaussian random variables with variance 1 per complex dimension and $\{h_i\}$ is some complex fading process such that $\sum_{i=1}^{k} \frac{1}{k} \log |h_i|^2$ converges in probability to some constant when the number of blocks k tends to infinity.

This scenario has received considerable interest in the case of an i.i.d. complex Gaussian fading process $\{h_i\}$, and several works have focused on the design of lattice codes for this model [36, 57]. The analysis of the union bound for the pairwise error probability for a lattice code $L \subset \mathbb{C}^k$ leads to a design criterion based on the maximization of the *normalized product distance*

$$\mathrm{Nd}_{p,\min}(L) = \inf_{\mathbf{x} \in L \setminus \{0\}} \frac{\prod_{i=1}^{k} |x_i|}{\mathrm{Vol}(L)^{\frac{1}{2}}}.$$

Note that the normalized product distance is a special case (for $n = 1$) of the normalized minimum determinant in (9). Most of the works in the literature have focused on the optimization of the product distance for lattice signal constellations with a fixed number of blocks k; few authors [35, 58] have also studied the upper and lower bounds for $\mathrm{Nd}_{p,\min}$ over all lattices when k grows to infinity.

However, there has been no general consensus on whether significant gain could be achieved from coding over an extensive number of fading realizations. For example the authors in [59] state that: "increasing the diversity does not necessarily increase to the same extent the performance: in fact, the minimum product distance decreases and the product kissing number increases. Simulations show that most of the gain is obtained for diversity orders up to 16”. In fact, the analysis of the distribution of pairwise errors in the union bound as in [60] shows that the *product kissing number* [57], or number of worst case occurrences, will grow fast and *a priori* might eat away the product distance gain. However, this issue seems to be due to the suboptimality of the union bound rather than to the codes themselves.

In fact, let us consider an infinite family of $2k$-dimensional lattices $L_k \subset \mathbb{C}^k$ with normalized product distance satisfying $(\mathrm{Nd}_{\mathrm{p,min}}(L_k))^{2/k} \geq c$, for some positive constant c.

According to Theorem 4.1 and Remark 4.3 we then have the following.

Corollary 8.1: Any rate R

$$R < \mathbb{E}_h \left[\log P \left|h\right|^2\right] - \log \frac{4}{\pi e} + \log c,$$

is achievable with the family L_k of lattices over the fading channel (35).

This result proves that indeed we gain by coding over an increasing number of blocks, assuming that we have a family of lattices L_k with the described product distances. According to Proposition 3.6, the condition $(\mathrm{Nd}_{\mathrm{p,min}}(L_k))^{2/k} \geq c$ implies that $\mathrm{rh}_{\mathrm{G}}(L_k) \geq kc$. It reveals that families of lattice codes with large product distance do not only have large Hermite invariants, but also that their reduced Hermite invariants are large as well. Thus, the product distance is not only relevant in capacity considerations or in the high SNR scenario, but also plays a role when coding over a finite number of fading realizations for low SNR.

A. Approaching capacity with number field codes

Using the normalized product distance as a code design criterion led to lattice constructions based on number fields in [61, 36, 57, 22]. However, none of these works considered capacity questions.

Let us now show how the construction in Proposition 7.5, when specialized to the single antenna case, is just the standard method used to build lattice codes from number fields [57] and how this method can be used to approach capacity in fast fading channels.

Let $K/\mathbb{Q}$ be a totally complex extension of degree $2k$ and $\{\sigma_1, \ldots, \sigma_k\}$ be a set of $\mathbb{Q}$-embeddings, such that we have chosen one from each complex conjugate pair. Then we can define a *relative canonical embedding* of K into $\mathbb{C}^n$ by

$$\varphi(x) = (\sigma_1(x), \ldots, \sigma_k(x)).$$

The ring of algebraic integers $\mathcal{O}_K$ has a $\mathbb{Z}$-basis $W = \{w_1, \ldots, w_{2k}\}$ and $\varphi(W)$ is a $\mathbb{Z}$-basis for the full lattice $\varphi(\mathcal{O}_K)$ in $\mathbb{C}^k$.

Proposition 7.5 now simplifies to the following.

Corollary 8.2: Let φ be the previously defined embedding and K a degree $2k$ totally complex number field. Then $\varphi(\mathcal{O}_K)$ is a $2k$-dimensional lattice in $\mathbb{C}^k$ which satisfies

$$\det\nolimits_{\min}\left(\varphi(\mathcal{O}_K)\right) = 1,\ \mathrm{Vol}(\varphi(\mathcal{O}_K)) = 2^{-k}\sqrt{|d_K|}$$

and

$$\delta(\varphi(\mathcal{O}_K)) = \left(\frac{2^{2k}}{|d_K|}\right)^{1/4}.$$

Using Martinet’s family of fields K_k from Theorem 7.7 and setting $L_{1,k} = \varphi(\mathcal{O}_{K_k})$ we have

$$\mathrm{Vol}(L_{1,k}) \leq \left(\frac{G}{2}\right)^k \text{ and } \det\nolimits_{\min}(L_{1,k}) = 1,$$

where $G \approx 92.368$. Specializing to the case where the fading process is i.i.d complex Gaussian we have that any rate

$$R < \log(Pe^{-\gamma}) - \log\left(\frac{2G}{\pi e}\right), \tag{36}$$

where $e^{-\gamma} = \mathbb{E}_h[\log |h|^2]$, is achievable.

B. Known bounds on discriminants and Hermite invariants

Equation (36) reveals that the codes based on the Martinet family have a rather large gap to capacity. However, the right-hand side of (36) is just a lower bound on the maximum achievable rate with lattice codes, and might be improved with a better error probability estimate and/or a better choice of the lattice sequence.

The Odlyzko bound [62] states that when $k \to \infty$ we have that $|d_K|^{1/2k} \geq 22.3$. However it is not known whether it is possible to reach this lower bound. For small values of k, there exist number fields having considerably smaller root discriminants. Table I [62] lists the best known root discriminants for totally complex number fields of degree $2k$. The first four values are known to be optimal.

TABLE I
BEST KNOWN ROOT DISCRIMINANTS FOR TOTALLY COMPLEX NUMBER FIELDS K OF SMALL DEGREE $2k$.

k	$\lvert d_K\rvert^{1/2k}$
1	1.732..
2	3.289..
3	4.622..
4	5.787..
5	6.793..

As seen in Corollary 8.1, we are only interested in the normalized product distance of the lattices under consideration. For example, instead of considering the image of the ring of integers $\mathcal{O}_K$ under the embedding φ, one can use an ideal of this ring of integers [63, 58, 1] or more generally any lattices with good normalized product distance.

The Minkowski-Hlawka theorem provides a non-constructive proof of the existence of $2k$-dimensional lattices $L_k \subset \mathbb{C}^k$ having Hermite invariants $\mathrm{h}(L_k) \sim \frac{k}{\pi e}$ [6]. It is an open question whether it is possible to obtain also $\mathrm{rh}_{\mathrm{G}}(L_k) \sim \frac{k}{\pi e}$ or equivalently $(\mathrm{Nd}_{\mathrm{p,min}}(L_k)) \sim \left(\frac{1}{\pi e}\right)^{k/2}$.

IX. Geometry of numbers for fading channels

In the previous sections we have shown that the normalized minimum determinant provides a design criterion to build capacity-approaching lattice codes for block fading multiple antenna channels. Let us now see how this approach fits into a more general context and can be regarded as a natural generalization of the classical theory of lattices for Gaussian channels. Finally we show how the code design problems, both in Gaussian and fading channels, can be seen as instances of the same problem in the mathematical theory of *geometry of numbers* [41] .

We denote with $\mathcal{L}_{(n,k)}$ the set of all $2n^2k$-dimensional lattices in the space $M_{n\times nk}(\mathbb{C})$ having volume one.

In the single antenna case, we can measure how a lattice $L \in \mathcal{L}_{(1,k)}$ roughly performs over the AWGN channel by analyzing how the function

$$f_1(x_1,\dots,x_k) = |x_1|^2 + |x_2|^2 + \cdots + |x_k|^2 \qquad (37)$$

behaves on the lattice. We have shown in this paper that the corresponding function for the fast fading channel is

$$f_2(x_1,x_2,\dots,x_k) = k|x_1x_2\cdots x_k|^{2/k}. \qquad (38)$$

Similarly, the function $f_3 : \mathcal{L}_{(n,k)} \to \mathbb{R}$ defined by

$$f_3(X_1,X_2,\dots,X_k) = nk\prod_{i=1}^{k} |\det(X_i)|^{2/nk},$$

can be used to analyze how a multi-block lattice $L \in \mathcal{L}_{(n,k)}$ performs over the block fading MIMO channel.

We immediately note that all these functions share common characteristics.

Definition 9.1: A continuous function F: $M_{n\times kn}(\mathbb{C}) \to \mathbb{R}$ is called a homogeneous form of degree $\sigma > 0$ if

$$|F(\alpha X)| = |\alpha|^\sigma |F(X)| \quad \forall \alpha \in \mathbb{R}, \forall X \in M_{n\times kn}(\mathbb{C}).$$

We can now see that all the functions f_i are homogeneous forms of degree 2. With this observation we can place our study into a more general context in geometry of numbers.

Definition 9.2: Let us consider the body $S(F) = \{X \mid X \in M_{n\times kn}(\mathbb{C}), |F(X)| \leq 1\}$, and a $2kn^2$ dimensional lattice L with a fundamental parallelotope of volume one. We define the *homogeneous minima* $\lambda(F,L)$ of F with respect to the lattice L by

$$\lambda(F,L) = (\inf\{\lambda \mid \lambda > 0, \dim(\mathbb{R}(\lambda S(F)\cap L)) \geq 1\})^\sigma,$$

where $\mathbb{R}(\lambda S(F)\cap L)$ is the $\mathbb{R}$-linear space generated by the elements in $\lambda S(F)\cap L$.

Using this notation we can now see that the Hermite invariant and reduced Hermite invariants are homogeneous minima $\lambda(f_1,L) = \mathrm{h}(L)$, $\lambda(f_2,L) = \mathrm{rh}_G(L)$ and $\lambda(f_3,L) = \mathrm{rh}_G(L)$. Given $i \in \{1,2,3\}$, suppose that $\{L_{n,k}^{(i)}\}$ is a sequence of lattices with the property that $\lambda(f_i, L_{n,k}^{(i)}) \geq c_i$. In this paper we proved that any rate R_i such that

$$\begin{aligned} R_1 &< \log_2(P) - \log_2\left(\frac{4}{\pi e}\right) + \log_2 c_1, \\ R_2 &< \mathbb{E}_h\left[\log_2 P\,|h|^2\right] - \log_2\frac{4}{\pi e} + \log_2 c_2, \\ R_3 &< \mathbb{E}_H\left[\log_2\det\frac{P}{n}H^\dagger H\right] - n\log_2\frac{4n}{\pi e} + n\log_2 c_3, \end{aligned}$$

is achievable for $\{L_{n,k}^{(i)}\}$ over the corresponding channel. Using this notation, characterizations of achievable rates using lattice codes have now been transformed into purely geometrical questions about the existence of lattices with certain properties.

A natural question is how close to capacity we can get with these methods by taking the best possible lattice sequences in terms of their homogeneous minimum. This leads us to the concept of *absolute homogeneous minimum*

$$\lambda(F) = \sup_{L\in\mathcal{L}_{(n,k)}} \lambda(F,L).$$

This remark suggests that there is a very general connection between information theory and geometry of numbers for different channel models. It seems that given a fading channel model, there exists a form whose absolute homogeneous minimum provides a lower bound for the achievable rate using lattice codes.

Here $\lambda(f_1)$ is the Hermite constant. The value of the Hermite constant $H(k)$, for different values of k, has been studied in mathematics for hundreds of years and there exists an extensive literature on the topic. In particular good upper and lower bounds are available and it has been proven that we can get quite close to Gaussian channel capacity with this approach [6, Chapter 3].

In the case of $\lambda(f_2)$, the problem has been considered in the context of algebraic number fields and some upper bounds have been provided. As far as we know the best lower bounds come from the existence results provided by number field constructions [35] and [1].

The properties of $\lambda(f_3)$ seem to be be far less researched in the literature. Simple upper bounds can be derived from bounds for Hermite constants as pointed out in [37], and lower bounds are obtained from division algebra constructions as described in this paper, but the mathematical literature doesn't seem to offer any ready-made results for this problem.

Remark 9.3: The definitions for the geometry of numbers given in this section were stated for lattices in the space $M_{n\times nk}(\mathbb{C})$, while usually the definitions are given in the space $\mathbb{R}^m$. This is however, just to keep our notation simple. The space $M_{n\times nk}(\mathbb{C})$ can be identified with the space $\mathbb{R}^{2n^2k}$ and we could have given the definitions also in the traditional form using this identification.

X. Discussion and questions for further research

In this work we proved the existence of lattice codes achieving constant gap to capacity in ergodic fading channels. Unlike the case of existence results based on random coding, our finite codes are always built from the same family of lattices, irrespective of the SNR and even of the fading statistics. Hence, using the minimum determinant as a design principle leads to extremely robust codes. In particular division algebra and number field codes have this robustness property.

However, our codes still have a considerable gap to capacity and further research is needed. Let us now point out a few directions this research can take next.

In the case of single user channels the clearest goal is to improve our methods and close the gap to capacity. We note that this gap depends on several factors. First of all, the normalized minimum determinant affects the value of the gap. Second, our bound for the error probability is based on sphere packing and might be suboptimal.

Thus, the possible improvements to our construction are two-fold. In the first place, one could try to find families of lattices $L_{n,k} \subset M_{n\times nk}(\mathbb{C})$ with larger normalized minimum determinant, for instance by replacing the centers in our constructions with families of number fields having smaller discriminants. One can also consider more general examples of lattices than those arising from orders in division algebras: for example ideals of orders, or in the case of number field codes, ideals of the ring of algebraic integers. In the second place, in this paper we have not considered the issue of shaping. Improving the shaping properties of our lattices might lead to a better error probability bound.

Another approach is to relax our minimum determinant code design criterion. Our codes are extremely robust and quite universal in the sense that they respond very well to any non pathological fading realization. This universality is of course a strength, but it could also lead to a situation where the codes are rather good for every channel, but not optimal for any. If we fix a channel model, it may be possible to weaken the design principle. This might allow us to consider larger ensembles of lattices and possibly to close the gap to capacity in this fixed channel model.

In this paper we have considered block fading MIMO channels, but we hope that the methods developed here can be applied also in a more general setting. Let us now sketch an outline for possible generalizations.

The reduced Hermite invariant is a natural analogue of the classical Hermite invariant for fading channels. This concept can likely be generalized to other fading channel models, such as for example intersymbol interference channels. Given a fading channel we can ask what would be the group (or set) G that would represent the action of the channel, and define the corresponding reduced Hermite invariant h_G. The next question is then to find lattices that would maximize this value. In the case of the block fading channel, the problem was made more accessible by Proposition 3.6, where we proved that h_G can be seen as the minimum of a certain homogeneous form. This line of thought suggests a general approach to turn the chase for capacity into a problem in geometry of numbers for different channel models. It also raises several questions. For example we can ask which are the channel models where this approach can be applied and for which groups G the reduced Hermite invariant corresponds to some homogeneous form.

Finally, the lattice codes proposed in this paper can have applications to other problems in information theory, such as coding for multiple access fading channels and for information theoretic security[3].

Acknowledgements

The authors would like to thank Roberto Pinciroli for helpful discussions about ergodic dynamical systems, and the three anonymous reviewers for their detailed comments and suggestions which helped to improve the paper.

Appendix

A. Proof of Theorem 4.7

With a similar approach as in the proof of Theorem 4.1, we consider the following upper bound:

$$\begin{aligned} P_e &\le \mathbb{P}\left\{ \|W\|^2 \ge \left(\frac{d_H}{2}\right)^2 \right\} \\ &\le \mathbb{P}\left\{ \frac{\|W\|^2}{knn_r} \ge 1+\epsilon \right\} + \mathbb{P}\left\{ \frac{d_H^2}{4knn_r} < 1+\epsilon \right\}, \end{aligned} \tag{39}$$

where

$$d_H^2 = \min_{\substack{X,\bar{X}\in\mathcal{C} \\ X\neq\bar{X}}} \sum_{i=1}^{k} \left\| H_i(X_i - \bar{X}_i) \right\|^2$$

is the minimum distance in the *finite* received constellation. The first term in equation (39) tends to zero exponentially fast when $k \to \infty$ since $2\|W\|^2 \sim \chi^2(2knn_r)$. We now focus on the second term in equation (39), and begin by finding a lower bound on d_H.
For all $i \in \{1,\dots,k\}$, let $\lambda_{i,j}$, $j = 1,\dots,n$ be the singular values of $H_i^\dagger H_i$ with

$$0 = \lambda_{i,1} = \cdots = \lambda_{i,n-n_r} < \lambda_{i,n-n_r+1} \le \cdots \le \lambda_{i,n},$$

and $l_{i,j}$ the singular values of $(X_i - \bar{X}_i)(X_i - \bar{X}_i)^\dagger$ with

$$l_{i,1} \ge l_{i,2} \ge \cdots \ge l_{i,n}.$$

Using the mismatched eigenvalue bound [65, 9], we have

$$\left\| H_i(X_i - \bar{X}_i) \right\|^2 \ge \sum_{j=1}^{n} \lambda_{i,j} l_{i,j} = \sum_{j=n-n_r+1}^{n} \lambda_{i,j} l_{i,j}.$$

Consequently, we find that

$$\begin{aligned} d_H^2 &\ge \alpha^2 \min_{\substack{X,\bar{X}\in\mathcal{C} \\ X\neq\bar{X}}} \sum_{i=1}^{k} \sum_{j=n-n_r+1}^{n} \lambda_{i,j} l_{i,j} \\ &\ge \alpha^2 n_r k \prod_{i=1}^{k} \prod_{j=n-n_r+1}^{n} \left(\lambda_{i,j} l_{i,j}\right)^{\frac{1}{n_r k}}. \end{aligned} \tag{40}$$

Using the NVD property of the code, we get

$$\prod_{i=1}^{k}\prod_{j=1}^{n} l_{i,j} = \prod_{i=1}^{k} \left|\det(X_i - \bar{X}_i)\right|^2 \ge 1$$

[3] An application to the wiretap channel is considered in our recent work [64].

Therefore, we have the lower bound

$$\begin{aligned}\prod_{i=1}^{k}\prod_{j=n-n_r+1}^{n} l_{i,j} &\geq \left(\prod_{i=1}^{k}\prod_{j=1}^{n-n_r} l_{i,j}\right)^{-1}\\ &\geq \left(\frac{1}{(n-n_r)k}\sum_{i=1}^{k}\sum_{j=1}^{n-n_r} l_{i,j}\right)^{-(n-n_r)k}\\ &\geq \left(\frac{2Pn}{\alpha^2(n-n_r)}\right)^{-k(n-n_r)},\end{aligned}$$

where we have used the arithmetic-geometric mean inequality and the power constraint

$$\alpha^2\left\|X-\bar{X}\right\|^2 = \alpha^2\sum_{j=1}^{n} l_{i,j} \leq 2Pkn.$$

Replacing the previous expression in (40), we obtain

$$\begin{aligned}d_H^2 &\geq \frac{\alpha^2 n_r k\prod_{i=1}^{k}\prod_{j=n-n_r+1}^{n}\lambda_{i,j}^{\frac{1}{n_r k}}}{\left(\frac{2Pn}{\alpha^2(n-n_r)}\right)^{\frac{n-n_r}{n_r}}}\\ &= (\alpha^2)^{\frac{n}{n_r}}\left(\frac{n-n_r}{2Pn}\right)^{\frac{n-n_r}{n_r}} n_r k\prod_{i=1}^{k}\det(H_iH_i^\dagger)^{\frac{1}{n_r k}}.\end{aligned}$$

The second term in (39) can thus be upper bounded by

$$\begin{aligned}&\mathbb{P}\left\{\prod_{i=1}^{k}\det(H_iH_i^\dagger)^{\frac{1}{n_r k}} < 4(1+\epsilon)\left(\frac{n}{\alpha^2}\right)^{\frac{n}{n_r}}\left(\frac{2P}{n-n_r}\right)^{\frac{n-n_r}{n_r}}\right\}\\ &= \mathbb{P}\left\{\frac{1}{k}\sum_{i=1}^{k}\log\det H_iH_i^\dagger < \log\frac{(4(1+\epsilon))^{n_r}n^n(2P)^{n-n_r}}{\alpha^{2n}(n-n_r)^{n-n_r}}\right\}\end{aligned}$$

By hypothesis the weak law of large numbers (19) holds, i.e. $\frac{1}{k}\sum_{i=1}^{k}\log\det H_iH_i^\dagger \to \mu$ as $k \to \infty$. Thus, the error probability will vanish provided that for sufficiently large k,

$$\log\frac{(4(1+\epsilon))^{n_r}n^n(2P)^{n-n_r}}{\alpha^{2n}(n-n_r)^{n-n_r}} < \mu$$

Recalling that $\alpha^2 \geq \frac{C_{n,k}^{\frac{1}{n^2 k}}P}{2^{\frac{R}{n}}C_L}$, the condition can be rewritten as

$$\begin{aligned}R < \mu + n_r\log P - n_r\log 2(1+\epsilon) - n\log 2nC_L + \frac{\log C_{n,k}}{nk}\\ + (n-n_r)\log(n-n_r).\end{aligned}$$

Using Stirling's approximation (17), for large k we have

$$\frac{\log C_{n,k}}{nk} \approx n\log\pi e - n\log n - \frac{1}{2nk}\log 2\pi n^2 k.$$

Asymptotically, we find that any rate

$$R < \mu + n_r\log\frac{P}{2(1+\epsilon)} - n\log\frac{2n^2C_L}{\pi e} + (n-n_r)\log(n-n_r)$$

is achievable. Since this is true for all $\epsilon > 0$, this concludes the proof. □

References

[1] R. Vehkalahti and L. Luzzi, "Number field lattices achieve Gaussian and Rayleigh channel capacity within a constant gap", in *IEEE Int. Symp. Inform. Theory* (ISIT), pp. 436–440, Hong Kong, China, June 2015

[2] L. Luzzi and R. Vehkalahti, "Division algebra codes achieve MIMO block fading channel capacity within a constant gap", *IEEE Int. Symp. Inf. Theory* (ISIT), pp. 446–450 Hong Kong, China, June 2015.

[3] E. Telatar, "Capacity of multi-antenna Gaussian channels", *Europ. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

[4] B.M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel", *IEEE Trans. commun.* Vol.51, Issue 3, pp. 389–399, March 2003.

[5] A. Sanderovich, M. Peleg, and S. Shamai, "LDPC coded MIMO multiple access with iterative joint decoding", *IEEE Trans. Inf. Theory*, vol. 51, n. 4, pp. 1437–1450, April 2005.

[6] J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, Springer-Verlag, New York, 1988.

[7] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high-rate space-time block codes from division algebras", *IEEE Trans. Inf. Theory* , vol. 49, no. 10, pp. 2596–2616, Oct. 2003.

[8] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect Space-Time Block Codes", *IEEE Trans. Inf. Theory*, vol. 52 n.9, Sept. 2006.

[9] P. Elia, K. R. Kumar, P. V. Kumar, H.-F. Lu, and S. A. Pawar, "Explicit Space-Time Codes Achieving the Diversity-Multiplexing Gain Tradeoff", *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3869–3884, September 2006.

[10] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communications: Performance Criterion and Code Construction", *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, March 1998.

[11] G. Wang and X.-G. Xia, "On Optimal Multi-Layer Cyclotomic Space-Time Code Designs", *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 1102–1135, March 2005.

[12] R. Vehkalahti, C. Hollanti, J. Lahtonen, and K. Ranto, "On the densest MIMO lattices from cyclic division algebras," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3751–3780, Aug. 2009.

[13] S. Yang and J.-C. Belfiore, "Optimal space-time codes for the MIMO amplify-and-forward cooperative channel", *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 647–663, Feb. 2007.

[14] H.-F. Lu, "Constructions of multi-block space-time coding schemes that achieve the diversity-multiplexing tradeoff", *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3790–3796, Aug. 2008.

[15] P. Elia and P. Vijay Kumar, "Approximately-universal space-time codes for the parallel, multi-block and cooperative dynamic-decode-and-forward channels", available at http://arxiv.org/abs/0706.3502.

[16] C. Hollanti and H.-f. Lu, "Construction methods for asymmetric and multi-block space-time codes," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1086 –1103, Mar. 2009.

[17] R. Vehkalahti, C. Hollanti, and F. Oggier, "Fast-decodable asymmetric space-time codes from division algebras", *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2362–2384, Apr. 2012.

[18] B. Linowitz, M. Satriano and R. Vehkalahti, "A non-commutative analogue of the Odlyzko bounds and bounds on performance for space-time lattice codes", *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1971–1984, Apr. 2015.

[19] J. Martinet, "Tours de corps de classes et estimations de discriminants", *Inventiones Mathematicae* n. 44, 1978, pp. 65–73

[20] H. Cohen, F. Diaz y Diaz and M. Olivier, "Computing ray class groups, conductors and discriminants", *Mathematics of Computation*, vol. 67 n. 222, pp. 773-795, 1998

[21] C. Fieker, "Computing class fields via the Artin map", *Mathematics of Computation*, vol. 70 n. 235, pp. 1293–1303, 2001

[22] F. Oggier and E. Viterbo, "Algebraic number theory and code design for Rayleigh fading channels", *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 3, pp. 333–415, Dec. 2004.

[23] F. E. Oggier, J.-C. Belfiore, and E. Viterbo, "Cyclic division algebras: A tool for space-time coding", *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 1, pp. 1–95, 2007.

[24] S. Tavildar and P. Viswanath, "Approximately Universal Codes Over Slow-Fading Channels", *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233–3258, July, 2006.

[25] R. de Buda, "Some optimal codes have structure", *IEEE J. Select. Areas Commun.*, vol. 7, pp. 893–899, Aug. 1989.

[26] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel", *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 273–278, Jan. 1998.

[27] H. A. Loeliger, "Averaging bounds for lattices and linear codes", *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1767–1773, Nov. 1997.

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

[29] Y. Yan, C. Ling, and X. Wu, "Polar lattices: Where Arikan meets Forney", *IEEE Int. Symp. Inf. Theory* (ISIT), pp. 1292–1296, Istanbul, Turkey, Jul. 2013

[30] A. Hindy and A. Nosratinia, "Approaching the ergodic capacity with lattice coding," *IEEE Global Communications Conference (GLOBECOM)*, pp. 1492–1496, Austin, USA, Dec. 2014.

[31] A. Hindy and A. Nosratinia, "Achieving the ergodic capacity with lattice codes", *IEEE Int. Symp. Inform. Theory* (ISIT), Hong Kong, China, June 2015.

[32] H. El Gamal, G. Caire, M. Damen, "Lattice Coding and Decoding Achieve the Optimal Diversity-Multiplexing Tradeoff of MIMO Channels", *IEEE Trans. Inform. Theory*, vol. 50, no.6 pp. 968–985, Jun. 2004.

[33] O. Ordentlich and U. Erez, "Precoded Integer-Forcing Universally Achieves the MIMO Capacity to Within a Constant Gap", *IEEE Trans. Inf. Theory*, vol. 61, no.1 pp. 323–340, Jan. 2015.

[34] S.N. Litsyn and M.A. Tsfasman, "Constructive high-dimensional sphere packings", *Duke Math. J.* 54 (1987), no. 1, pp. 147–161.

[35] C. Xing, "Diagonal Lattice Space-Time Codes From Number Fields and Asymptotic Bounds", *IEEE Trans. Inf. Theory*, vol.53, no. 11, pp. 3921–3926, Nov. 2007.

[36] X. Giraud and J.-C. Belfiore, "Constellations matched to the Rayleigh fading channel", *IEEE Trans. Inf. Theory*, vol. 42, no.1 , pp. 106–115, Jan. 1996.

[37] J. Lahtonen and R. Vehkalahti, "Dense MIMO matrix lattices - a meeting point for class field theory and invariant theory", *Proc. Applied Algebra, Algebraic Algorithms, and Error Correcting Codes* (AAECC-17), Bangalore, India, 2007.

[38] S. Vituri, "Dispersion Analysis of Infinite Constellations in Ergodic Fading Channels", Master thesis, Tel Aviv University, 2013, available at http://arxiv.org/abs/1309.4638

[39] L. Liu and C. Ling "Polar Codes and Polar Lattices for Independent Fading Channels", *IEEE Int. Symp. Inf. Theory* (ISIT), Barcelona, Spain, June 2016.

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

[41] P. M. Gruber and C. G. Lekkerkerker, *Geometry of Numbers*, Elsevier, Amsterdam, The Netherlands, 1987.

[42] M. M. Skriganov, "Constructions of uniform distributions in terms of geometry of numbers", *Algebra i Analiz*, Volume 6, Issue 3, pp. 200–230, 1994.

[43] E. Bayer-Fluckiger, J.-P. Cerri and J. Chaubert, "Euclidean minima and central division algebras", *International Journal of Number Theory*, Vol. 5, No. 7, pp. 1155–1168, 2009.

[44] B. Laurent and P. Massart, "Adaptive estimation of a quadratic functional by model selection", *Annals of Statistics*, vol. 28, pp. 1302–1338, 2000

[45] P. Billingsley, *Probability and measure*, 3rd edition, Wiley, 1995

[46] M. Pollicott and M. Yuri, *Dynamical Systems and Ergodic Theory*, Cambridge University Press, 1998.

[47] G. Caire, P. Elia, and K.R. Kumar, "Space-time coding: An overview", *J. Commun. Softw. Syst.*, vol. 2, no. 3, pp. 212–227, Sept. 2006.

[48] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading", *IEEE Trans. Inf. Theory*, vol. 45, no.1, Jan. 1999.

[49] N. R. Goodman, *The distribution of the determinant of a complex Wishart distributed matrix*, Ann. Math. Statist., 34, pp. 178–180.

[50] A. Edelman, "Eigenvalues and condition numbers of random matrices", Ph.D. Thesis, MIT 1989

[51] J. Proakis, *Digital communications*, 4th edition, McGraw-Hill 2001

[52] I. Reiner, *Maximal Orders*, Academic Press, New York 1975.

[53] M. A. Tsfasman and S. G. Vlăduţ, "Infinite Global Fields and the Generalized Brauer-Siegel Theorem", *Moscow Mathematical Journal*, vol. 2, no. 2, pp. 329–402, April-June, 2002.

[54] F. Hajir and C. Maire, "Asymptotically good towers of global fields", *Proc. European Congress of Mathematics*, pp. 207–218, Birkhäuser Basel, 2001.

[55] E. S. Golod and I. R. Shafarevich, "On the class field tower", *Izv. Akad. Nauk* SSSSR 28: 261-272, 1964 (in Russian)

[56] G. Ivanyos and L. Rónyai, "Finding maximal orders in semisimple algebras over Q", *Computational Complexity* n. 3, pp. 245–261, 1993.

[57] J. Boutros, E. Viterbo, C. Rastello and J.-C. Belfiore, "Good lattice constellations for both Rayleigh fading and Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, March 1996.

[58] E. Bayer-Fluckiger, F. Oggier, and E. Viterbo, "Algebraic Lattice Constellations: Bounds on Performance", *IEEE Trans. Inf. Theory*, vol. 52, n. 1, pp. 319–327, Jan. 2006.

[59] E. Biglieri, J. Proakis and S. Shamai (Shitz), "Fading channels: information theoretic and communications aspects", *IEEE Trans. Inf. Theory*, Vol. 44, No. 6, pp. 2619–2692, Oct. 1998.

[60] R. Vehkalahti, H.-f. Lu and L. Luzzi, "Inverse determinant sums and connections between fading channel information theory and algebra", *IEEE Trans. Inf. Theory*, vol 59, no. 9, pp. 6060–6082, Sep. 2013.

[61] K. Boullé and J. C. Belfiore, "Modulation schemes designed for the Rayleigh channel", in *Proc. CISS '92*, pp. 288–293, Princeton, NJ, Mar. 1992.

[62] A. M. Odlyzko, "Bounds for discriminants and related estimates for class numbers, regulators and zeros of zeta functions: a survey of recent results", *Sém. Théor. Nombres Bordeaux* , vol. 2, no. 1, pp. 119–141, 1990.

[63] F. Oggier, "Algebraic methods for channel coding", PhD thesis, EPFL, Lausanne, 2005.

[64] L. Luzzi, C. Ling and R. Vehkalahti, "Almost universal codes for fading wiretap channels", *IEEE Int. Symp. Inf. Theory* (ISIT), Barcelona, Spain, June 2016.

[65] C. Köse and R. D. Wesel, "Universal space-time trellis codes", *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2717-–2727, Oct. 2003.

Laura Luzzi received the degree (Laurea) in Mathematics from the University of Pisa, Italy, in 2003 and the Ph.D. degree in Mathematics for Technology and Industrial Applications from Scuola Normale Superiore, Pisa, Italy, in 2007. From 2007 to 2012 she held postdoctoral positions in Télécom-ParisTech and Supélec, France, and a Marie Curie IEF Fellowship at Imperial College London, United Kingdom. She is currently an Assistant Professor at ENSEA de Cergy, Cergy-Pontoise, France, and a researcher at ETIS (ENSEA -

Université de Cergy-Pontoise- CNRS).
Her research interests include algebraic space-time coding and decoding for wireless communications and physical layer security.

Roope Vehkalahti received the M.Sc. and Ph.D. degrees from the University of Turku, Finland, in 2003 and 2008, respectively, both in pure mathematics.

He was with the Department of Mathematics, University of Turku, Finland 2003-2016. In 2011-2012 he was visiting Swiss Federal Institute of Technology, Lausanne (EPFL). He is currently with the Department of Communications and Networking, Aalto University, Espoo, Finland.

His research interest include applications of algebra and number theory to information theory.