Space-time codes of maximal size in MIMO communication

Daniele Rotanzi

Master's project

Fall-Winter 2009-10, UBC

Under the supervision of UBC Professor **Zinovy Reichstein** and

EPFL Professor **Eva Bayer Fluckiger**.

Contents

Terminology and notations

$\mathbb{N}$	the set of non-negative integers, $\mathbb{N} = \{0, 1, 2, \ldots\}$.
$\mathbb{N}^+$	the set of positive integers, $\mathbb{N}^+ = \{1, 2, \ldots\}$.
i	an imaginary number such that $i^2 = -1$.
$\mathbb{Z}[i]$	the ring of Gaussian integers, $\mathbb{Z}[i] := \{a + ib : a, b \in \mathbb{Z}\}$.
$\Re(z), \Im(z)$	real and imaginary part of a complex number z.
$\mathbb{F}$	the field that represents either $\mathbb{R}$ or $\mathbb{C}$.
$\mathbb{R}^+$	the positive real numbers, $x \in \mathbb{R}^+$ if and only if $x > 0$.
$\mathbb{M}_{n\times m}(A)$	the ring of matrices of size $n \times m$, for $n, m \in \mathbb{N}^+$, with entries in the ring A.
$\mathbb{M}_n(A)$	the ring of $n \times n$ matrices, $\mathbb{M}_n(A) := \mathbb{M}_{n\times n}(A)$.
$\overline{z}$	the conjugate of a complex number.
X^T	the transpose of a matrix.
X^*	the conjugate transpose of a matrix. If $X = (x_{jk}) \in \mathbb{M}_n(\mathbb{C})$, then $(X^*)_{jk} = \overline{x_{kj}}$.
$\lvert\cdot\rvert$	the absolute value of a complex number. If $z \in \mathbb{C}$, where $z = x + iy$, with $x, y \in \mathbb{R}$, then $\lvert z\rvert = \sqrt{x^2 + y^2}$.
$\lVert\cdot\rVert$	the Euclidean norm of a vector of $\mathbb{F}^n$, with $n \in \mathbb{N}^+$. If $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{F}^n$, then $\lVert\mathbf{x}\rVert = \sqrt{\sum_{j=1}^n \lvert x_j\rvert^2}$. We use it also for matrices of $\mathbb{M}_n(\mathbb{F})$, viewing them as vectors in $\mathbb{F}^{n^2}$. So, if $X = (x_{jk}) \in \mathbb{M}_n(\mathbb{F})$, then $\lVert X\rVert = \sqrt{\sum_{j,k=1}^n \lvert x_{jk}\rvert^2}$.
$\#$	the cardinality of a (finite) set.
$\mathbb{B}_n(\mathbf{c}, R)$	the n-dimensional ball centered at $\mathbf{c} \in \mathbb{R}^n$ of radius R, with $0 \le R \in \mathbb{R}$. If $\mathbf{c} = (c_1, \ldots, c_n)^T$, then $\mathbb{B}_n(\mathbf{c}, R) := \{(x_1, \ldots, x_n)^T \in \mathbb{R}^n : \sum_{j=1}^n (x_j - c_j)^2 \le R^2\}$.
$\mathbb{B}_n(R)$	a more concise notation for $\mathbb{B}_n(\mathbf{0}, R)$.
$\mathcal{V}$	the volume of a set.
Γ	the Gamma function. For any $z \in \mathbb{C}$, with $\Re(z) > 0$, $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$.

Matrices will be represented with capital letters, like A, X, Y, etc.
Vectors will be represented with small letters in bold, like $\mathbf{v}, \mathbf{x}, \mathbf{y}$, etc.

Introduction

Reliability of any communication technology, such as wireless channels, can be measured by the *probability of incorrect decoding* of a sent message. The lower is this probability the more reliable is the technology. The main problem in wireless communication is that the signals are affected not only by the noise at the receiving antennas (like it is the case for communication through wires), but also by the degradation due to obstacles of the physical environment in which the signals run. This kind of degradation is commonly referred to as *fading*. In response to this problem and for other technical reasons, Multiple-Input Multiple-Output (MIMO) channels have become increasingly popular in the past decade.
A MIMO channel consists of n transmitting and m receiving antennas. At each time slot t, the transmitting antennas send one signal each and each of the receiving antennas gets a linear superposition of these transmitted signals. Assuming that the data are transmitted over a time frame of length $T = n$, the codewords, i.e., the sent messages, are represented by complex square matrices of size n. Then a code is just a (finite) subset $\mathcal{C}$ of $\mathbb{M}_n(\mathbb{C})$ and since the coding occurs in space (the n transmitting antennas) and time, we speak of *space-time codes*. As we will see in chapter 2, studies from the late nineties (see for instance [1], chapter 5 of [2] or chapters $2-3$ of [3]) showed that, under certain assumptions, the probability $\mathbb{P}(X \to Y)$ of decoding a sent message $X \in \mathcal{C}$ as a different codeword $Y \in \mathcal{C}$ depends mainly on the absolute value of the determinant of the matrix $X - Y$. More precisely, the larger is $|\det(X - Y)|$, the lower is the probability $\mathbb{P}(X \to Y)$. This result led to a criterion for the design of optimal space-time codes: as X, Y range over pairs of distinct elements of $\mathcal{C}$, the minimal value of $|\det(X-Y)|$ should be large. On the other hand there is a physical constraint on the norm of the codewords that one can send. In fact, the larger the norm of a codeword is, the more power is needed to send it and since just a finite amount of power is available, the codewords of $\mathcal{C}$ cannot exceed it.
This leads to the following natural mathematical question:

Given a subfield L of $\mathbb{C}$, for any $n \in \mathbb{N}^+$ and any $r \in \mathbb{R}^+$, find a set $\mathcal{C} \subset \mathbb{M}_n(L)$ of maximal size such that the following two conditions are satisfied:

- ***energy constraint**: $\|X\| \leq 1$, for all $X \in \mathcal{C}$,*
- ***determinant criterion**: $|\det(X - Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.*

We immediately see that there is tension between these two constraints; as we place more matrices in the unit ball, the determinant criterion will become harder to satisfy. We also note that this question has the same flavour of one of the most basic questions in classical coding theory, where the role of the energy constraint is played by the length of the code and the role of the determinant criterion is played by the minimal Hamming distance.

The rest of this thesis is structured as follows.
Chapter 1 is devoted to preliminary mathematical results, which are used in the sequel. In chapter 2 we explain the transition from the engineering problem to our mathematical question. Sections 2.1 and 2.2 are mainly based on sections 2.1 and 3.2 of [3]. In chapter 3 we review the relevant material from classical coding theory, including the sphere packing and Gilbert-Varshamov bounds. In chapters 4 and 5 we derive similar bounds on the maximal size of a space-time code, in both the real and the complex case. In chapter 6 we illustrate these bounds by building an explicit code for $n = 2$, based on a construction due to S. M. Alamouti. Our main results are as follows.

Theorem 1 (Upper bound)**.** *Let L be a subfield of $\mathbb{F}$, $n \in \mathbb{N}^+$ and $r \in \mathbb{R}^+$. Let $\mathcal{C} \subset \mathbb{M}_n(L)$ such that $\|X\| \leq 1$, for all $X \in \mathcal{C}$, and $|\det(X - Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.*

- *If $\mathbb{F} = \mathbb{R}$, then*

$$\#\mathcal{C} \leq \frac{\Gamma\left(\frac{n}{2}+1\right)^n}{\Gamma\left(\frac{n^2}{2}+1\right)} \left(\frac{2}{\sqrt[n]{r}} + \sqrt{n}\right)^{n^2}.$$

- *If $\mathbb{F} = \mathbb{C}$, then*

$$\#\mathcal{C} \leq \frac{(n!)^n}{(n^2)!} \left(\frac{2}{\sqrt[n]{r}} + \sqrt{n}\right)^{2n^2}.$$

Theorem 2 (Lower bound)**.** *Let L be a subfield of $\mathbb{F}$, $n \in \mathbb{N}^+$ and $r \in \mathbb{R}^+$ such that $\sqrt[n]{r} \in L$. There exists then a set $\mathcal{C} \subset \mathbb{M}_n(L)$ with $\|X\| \leq 1$, for all $X \in \mathcal{C}$, and $|\det(X-Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$, such that*

- *if $\mathbb{F} = \mathbb{R}$,*

$$\#\mathcal{C} \geq \pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n}{2}+1\right)^{n-1}}{\Gamma\left(\frac{n^2}{2}+1\right)} \frac{\left(\frac{1}{\sqrt[n]{r}} - n\right)^{n^2}}{\left(\frac{1}{\sqrt[n]{r}} + \sqrt{n}\right)^{n^2-n} \sum_{k=0}^{n-1} \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}+1\right)} \left(\frac{1}{\sqrt[n]{r}} + \sqrt{k}\right)^k},$$

- *if $\mathbb{F} = \mathbb{C}$,*

$$\#\mathcal{C} \geq \pi^n \frac{(n!)^{n-1}}{(n^2)!} \frac{\left(\frac{1}{\sqrt[n]{r}} - \sqrt{2}n\right)^{2n^2}}{\left(\frac{1}{\sqrt[n]{r}} + \sqrt{2n}\right)^{2n^2-2n} \sum_{k=0}^{n-1} \frac{\pi^k}{k!} \left(\frac{1}{\sqrt[n]{r}} + \sqrt{2k}\right)^{2k}}.$$

Chapter 1

Preliminary results

The purpose of this chapter is to collect mathematical results which will be repeatedly used in the sequel.

Lemma 1.1. *Let $n \in \mathbb{N}^+$, $\mathbf{c} \in \mathbb{R}^n$ and $0 \leq R \in \mathbb{R}$. Then*

$$\mathcal{V}\left(\mathbb{B}_n(\mathbf{c}, R)\right) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} R^n.$$

Proof. See section 11.33 of [6]. □

Lemma 1.2. *For any $z \in \mathbb{C}$ with $\Re(z) > 0$, we have $\Gamma(z+1) = z\Gamma(z)$. In particular $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$ and, if $n \in \mathbb{N}$, then $\Gamma(n+1) = n!$.*

Proof. See section 6.23 of [6]. □

Lemma 1.3. *Let $X \in \mathbb{M}_n(\mathbb{C})$. Then $\|X\|^2 = \mathrm{tr}(XX^*)$.*

Proof. Denote $(X)_{jk}$ by x_{jk}.
We have $\mathrm{tr}(XX^*) = \sum_{j=1}^n (XX^*)_{jj}$ and

$$(XX^*)_{jj} = \sum_{k=1}^n (X)_{jk}(X^*)_{kj} = \sum_{k=1}^n x_{jk}\overline{x_{jk}} = \sum_{k=1}^n |x_{jk}|^2.$$

Therefore $\mathrm{tr}(XX^*) = \sum_{j=1}^n \sum_{k=1}^n |x_{jk}|^2 = \|X\|^2$. □

Lemma 1.4. *Let $X = (\mathbf{x_1}|\ldots|\mathbf{x_n}) \in \mathbb{M}_n(\mathbb{C})$. If the columns of X are mutually orthogonal, then $|\det(X)| = \prod_{j=1}^n \|\mathbf{x_j}\|$.*

Proof. Let $(X)_{jk} = x_{jk}$ and denote by $\langle \cdot, \cdot \rangle$ the usual inner product of the $\mathbb{C}$-vector space $\mathbb{C}^n$. Notice that

$$(X^*X)_{jk} = \sum_{l=1}^{n} (X^*)_{jl}(X)_{lk} = \sum_{l=1}^{n} \overline{x_{lj}} x_{lk} = \langle \mathbf{x_k}, \mathbf{x_j} \rangle = \begin{cases} 0, & j \neq k \\ \|\mathbf{x_j}\|^2, & j = k. \end{cases}$$

Hence X^*X is a diagonal matrix and $\det(X^*X) = \prod_{j=1}^{n} \|\mathbf{x_j}\|^2$. On the other hand, remember that $\det(X^*) = \overline{\det(X)}$ and thus

$$\det(X^*X) = \overline{\det(X)} \det(X) = |\det(X)|^2.$$

Therefore $|\det(X)| = \prod_{j=1}^{n} \|\mathbf{x_j}\|$, as claimed. □

Lemma 1.5 (Cauchy-Schwarz inequality)**.** *Let V be a $\mathbb{C}$-vector space with inner product $\langle \cdot, \cdot \rangle$ and let $\mathbf{v}, \mathbf{w} \in V$. Then*

$$\langle \mathbf{v}, \mathbf{w} \rangle \leq \|\mathbf{v}\| \|\mathbf{w}\|.$$

Proof. For any $\alpha \in \mathbb{C}$ we have that

$$0 \leq \|\mathbf{v} - \alpha \mathbf{w}\|^2 = \langle \mathbf{v} - \alpha \mathbf{w}, \mathbf{v} - \alpha \mathbf{w} \rangle = \|\mathbf{v}\|^2 - \overline{\alpha} \langle \mathbf{v}, \mathbf{w} \rangle - \alpha \overline{\langle \mathbf{v}, \mathbf{w} \rangle} + |\alpha|^2 \|\mathbf{w}\|^2.$$

Letting $\alpha = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}$, it follows that $0 \leq \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|^2}{\|\mathbf{w}\|^2}$, yielding our result. □

Lemma 1.6 (arithmetic-geometric mean inequality)**.** *Let $x_1, x_2, \ldots, x_n$ be non-negative real numbers. We then have that*

$$\frac{x_1 + x_2 + \ldots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \cdots x_n}.$$

Proof. We prove this lemma by induction.
For the base case we take $n = 2$. It is then well known that

$$\frac{x_1 + x_2}{2} \geq \sqrt{x_1 x_2}, \quad \forall\, x_1, x_2 \geq 0,$$

since

$$(x_1 + x_2)^2 - 4x_1 x_2 = x_1^2 + x_2^2 - 2x_1 x_2 = (x_1 - x_2)^2 \geq 0.$$

The base case being proved, we have to demonstrate the induction step. The induction hypothesis is the following

$$\mathcal{P}(n): \quad \left(\frac{x_1 + x_2 + \ldots + x_n}{n} \right)^n \geq x_1 \cdots x_n, \quad \forall\, x_1, \ldots, x_n \geq 0.$$

We have to prove that $\mathcal{P}(n)$ implies $\mathcal{P}(n+1)$. Let then $x_1, \ldots, x_{n+1}$ be $n+1$ non-negative real numbers and let

$$\mu_{n+1} = \frac{x_1 + x_2 + \ldots + x_{n+1}}{n+1}.$$

We want to prove that $(\mu_{n+1})^{n+1} \geq x_1 \cdots x_{n+1}$.
Without loss of generality, we can assume x_n to be the largest of the x_j and x_{n+1} to be the smallest of the x_j. Hence $x_n \geq \mu_{n+1}$ and $x_{n+1} \leq \mu_{n+1}$ and thus

$$(x_n - \mu_{n+1})(\mu_{n+1} - x_{n+1}) \geq 0. \tag{1.1}$$

Consider now the non-negative real numbers $x_1, \ldots, x_{n-1}, x'_n$, where $x'_n = x_n + x_{n+1} - \mu_{n+1} \geq x_n - \mu_{n+1} \geq 0$. By the induction hypothesis

$$(\mu'_n)^n \geq x_1 \cdots x_{n-1} x'_n, \tag{1.2}$$

where $\mu'_n = \frac{x_1 + \ldots + x_{n-1} + x'_n}{n}$. Notice that

$$n\mu'_n = \underbrace{x_1 + \ldots + x_{n-1} + x_n + x_{n+1}}_{=(n+1)\mu_{n+1}} - \mu_{n+1} = n\mu_{n+1},$$

so $\mu'_n = \mu_{n+1}$. Thanks to (1.1), we also have that

$$\begin{aligned} (x_n - \mu_{n+1})(\mu_{n+1} - x_{n+1}) \geq 0 \quad &\Leftrightarrow \quad (x_n + x_{n+1})\mu_{n+1} - \mu_{n+1}^2 \geq x_n x_{n+1} \\ &\Leftrightarrow \quad x'_n \mu_{n+1} \geq x_n x_{n+1}. \end{aligned}$$

Combining all this with (1.2) we obtain:

$$(\mu_{n+1})^{n+1} = (\mu_{n+1})^n \mu_{n+1} = (\mu'_n)^n \mu_{n+1} \geq x_1 \cdots x_{n-1} x'_n \mu_{n+1} \geq x_1 \cdots x_n x_{n+1},$$

as desired. □

Theorem 1.7. *Let $X \in \mathbb{M}_n(\mathbb{C})$ and let $\mathbf{c_1}, \cdots, \mathbf{c_n}$ be the columns of X, i.e., $X = (\mathbf{c_1}| \cdots |\mathbf{c_n})$. Then $|\det(X)| \leq \prod_{k=1}^n \|\mathbf{c_k}\|$.*

Proof. First of all, if $\det(X) = 0$ the theorem is obvious.
Let then suppose that $\det(X) \neq 0$. Thus the columns of X are linearly independent. Applying the Gram-Schmidt process to $\mathbf{c_1}, \ldots, \mathbf{c_n}$ we obtain n new vectors $\mathbf{v_1}, \ldots, \mathbf{v_n}$ that are orthogonal. More precisely, for $k = 2, \ldots, n$ let $W_k = \text{span}(\mathbf{c_1}, \ldots, \mathbf{c_{k-1}})$ and let $\mathbf{w_k}$ be the orthogonal projection of $\mathbf{c_k}$ on W_k. By definition of the $\mathbf{v_1}, \ldots, \mathbf{v_n}$ we have that $\mathbf{v_1} = \mathbf{c_1}$ and, for $k = 2, \ldots, n$, $\mathbf{v_k} = \mathbf{c_k} - \mathbf{w_k}$, with $\langle \mathbf{v_k}, \mathbf{w_k} \rangle = 0$.
We then consider the matrix $Y = (\mathbf{v_1}| \cdots |\mathbf{v_n})$, whose columns are thus orthogonal.

Thanks to lemma 1.4 we have that $|\det(Y)| = \prod_{k=1}^{n} \|\mathbf{v_k}\|$.
Notice that $\det(Y) = \det(X)$, because each column of Y is the sum of the correspondent column of X with a linear combination of the previous columns of X. In fact $\mathbf{v_1} = \mathbf{c_1}$ and, for $k = 2, \ldots, n$,

$$\mathbf{v_k} = \mathbf{c_k} - \sum_{j=1}^{k-1} \frac{\langle \mathbf{c_k}, \mathbf{v_j} \rangle}{\langle \mathbf{v_j}, \mathbf{v_j} \rangle} \mathbf{v_j}.$$

Therefore the determinant remains the same.
We then have that

$$|\det(X)| = |\det(Y)| = \prod_{k=1}^{n} \|\mathbf{v_k}\|.$$

Observe also that $\|\mathbf{v_k}\| \leq \|\mathbf{c_k}\|$.
For $k = 1$ it is trivial, since we have equality. For $k = 2, \ldots, n$, $\mathbf{c_k} = \mathbf{v_k} + \mathbf{w_k}$ and therefore

$$\|\mathbf{c_k}\|^2 = \langle \mathbf{v_k} + \mathbf{w_k}, \mathbf{v_k} + \mathbf{w_k} \rangle = \langle \mathbf{v_k}, \mathbf{v_k} \rangle + \langle \mathbf{w_k}, \mathbf{w_k} \rangle = \|\mathbf{v_k}\|^2 + \|\mathbf{w_k}\|^2 \geq \|\mathbf{v_k}\|^2.$$

Hence we have that $\|\mathbf{v_k}\| \leq \|\mathbf{c_k}\|$ and so

$$|\det(X)| = \prod_{k=1}^{n} ||\mathbf{v_k}|| \leq \prod_{k=1}^{n} ||\mathbf{c_k}||,$$

as claimed. □

Chapter 2

From the engineering problem to the mathematical question

2.1 The mathematical model of a MIMO channel

A MIMO channel consists of $n \in \mathbb{N}^+$ transmitting antennas and $m \in \mathbb{N}^+$ receiving antennas. At each time slot each of the n transmitting antennas send simultaneously a signal that is modeled as a complex number. Each receiving antenna collects all the n signals sent. As we said in the introduction, fading affects these signals and at each receiving antenna there is some noise. We model this situation as follows.
At time slot $t \in \mathbb{N}^+$, let $r_{j,t}$ be the signal collected of receiving antenna j, $x_{k,t}$ be the signal sent by transmitting antenna k, $w_{j,t}$ the noise at receiving antenna j and $h^t_{j,k}$ (here t is a superscript, not a power) the fading coefficient of the path from transmitting antenna k to receiving antenna j. The two parameters $h^t_{j,k}$ and $w_{j,t}$ are modeled as complex random variables (recall that a complex random variable Z is such that $Z = X + iY$, where X and Y are two real random variables). With these notations we have that:

$$\begin{aligned} r_{j,t} &= h^t_{j,1}x_{1,t} + h^t_{j,2}x_{2,t} + \cdots + h^t_{j,n}x_{n,t} + w_{j,t} \\ &= \sum_{k=1}^{n} h^t_{j,k}x_{k,t} + w_{j,t}, \text{ for all } j = 1, \ldots, m, \end{aligned}$$

as illustrated in Figure 2.1.

Let $\mathbf{r_t} = (r_{1,t}, \ldots, r_{m,t})^T \in \mathbb{C}^m$, $\mathbf{x_t} = (x_{1,t}, \ldots, x_{n,t})^T \in \mathbb{C}^n$, $\mathbf{w_t} = (w_{1,t}, \ldots, w_{m,t})^T \in \mathbb{C}^m$ and $H_t \in \mathbb{M}_{m\times n}(\mathbb{C})$, where $H_t = \left(h^t_{j,k}\right)$. H_t is called the fading (or channel) matrix and it stores all the fading coefficients of all

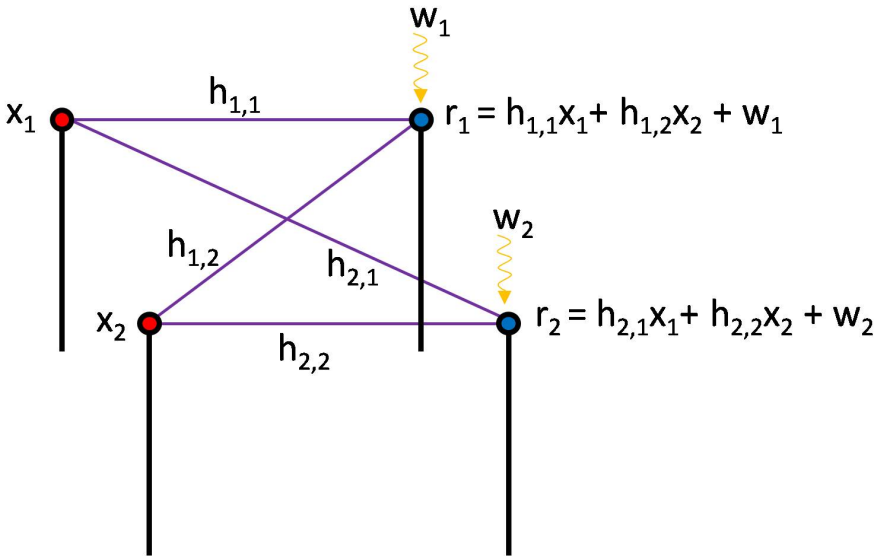

Figure 2.1: A MIMO system with $n = 2$ transmitting and $m = 2$ receiving antennas. For the sake of simplicity, we omitted the subscript/superscript t.

possible paths that a signal can follow, at time t. We can then write

$$\mathbf{r_t} = H_t\mathbf{x_t} + \mathbf{w_t}.$$

The transmission of data is usually separated into time frames of some fixed length $T \in \mathbb{N}^+$. Consequently, for a transmission frame of length T, we have

$$\mathbf{r_t} = H_t\mathbf{x_t} + \mathbf{w_t}, \quad \text{for } t = 1, \ldots, T. \tag{2.1}$$

In this model we see that the fading matrix H_t depends on time t. This is because, a priori, nothing ensures that the environment between the transmitter and the receiver remains the same. However we will assume that this change is slow. Denoting by $T' \in \mathbb{N}^+$ the period in which the fading matrix remains the same, we will then assume that $T \ll T'$. This hypothesis is called *slow fading* and in this work we will only consider this kind of situation. This means that our fading matrix can be assumed to be fixed for all $t = 1, \ldots, T$, i.e., $H_t = H$.
Let $R = (\mathbf{r_1}|\cdots|\mathbf{r_T}) \in \mathbb{M}_{m\times T}(\mathbb{C})$, $X = (\mathbf{x_1}|\cdots|\mathbf{x_T}) \in \mathbb{M}_{n\times T}(\mathbb{C})$ and $W = (\mathbf{w_1}|\cdots|\mathbf{w_T}) \in \mathbb{M}_{m\times T}(\mathbb{C})$. Equation (2.1) can now be rewritten as

$$R = HX + W. \tag{2.2}$$

With this model we can then define what a space-time code is.

Definition 2.1 (space-time code)**.** Let $n, T \in \mathbb{N}^+$, where n is the number of transmitting antennas and T the temporal length of the transmission frame. A *space-time code* is then a subset $\mathcal{C}$ of $\mathbb{M}_{n\times T}(\mathbb{C})$.
The codewords are then complex matrices of size $n \times T$.

Here we think of n (the number of transmitting antennas) as the space parameter and T as the time parameter of our space-time code.

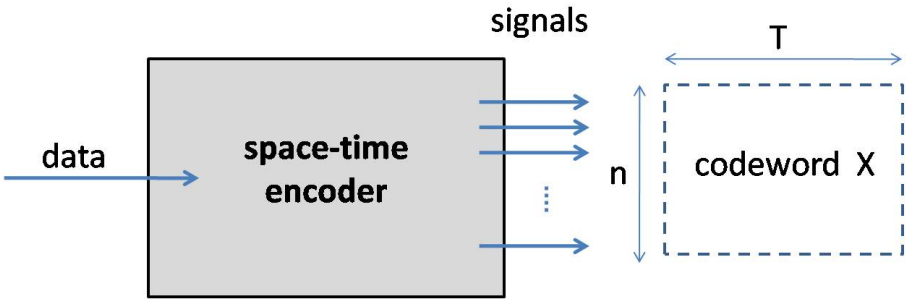

Figure 2.2: Scheme of a space-time encoder of a MIMO channel. The data, which usually are a set of bits, are the real information the transmitter wants to send.

2.2 The probability of incorrect decoding

A reliable communication is fundamental for a channel to be useful. The most natural way of measuring reliability of a communication channel is the probability of error P_e. Obviously, the smaller is P_e the more reliable is the channel. We will now compute P_e for our model (2.2) of a slow fading MIMO channel under the following assumptions:

- the noise entries $w_{j,k}$ are independent samples of a zero-mean circularly symmetric complex Gaussian random variable of variance $\sigma^2/2$ per dimension, for all $j = 1, \ldots, n$ and all $k = 1, \ldots, T$. This means that if $w_{j,k} = u_{j,k} + iv_{j,k}$, then

$$\begin{pmatrix} u_{j,k} \\ v_{j,k} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right),$$

  where $\mathcal{N}$ denotes a Gaussian (or normal) random vector.

- the fading coefficients $h_{j,k}$ are independent samples of a zero-mean circularly symmetric complex Gaussian random variable of variance 0.5 per dimension, for all $j = 1, \ldots, m$ and all $k = 1, \ldots, n$. As before, if $h_{j,k} = f_{j,k} + ig_{j,k}$, this means that

$$\begin{pmatrix} f_{j,k} \\ g_{j,k} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

- Maximum-likelihood decoding is used. Under our assumptions it means the following. Suppose $X \in \mathcal{C}$ is the sent codeword and thus $R_X = HX + W$ is the received matrix. The decoder decides in favour of the codeword $Y \in \mathcal{C}$ such that $\|R_X - HY\|$ is minimal.

Remark 2.2. Notice that if U, V are two independent random variables such that $U, V \sim \mathcal{N}(0, \tau^2)$ then $\sqrt{U^2 + V^2}$ has Rayleigh distribution. Since here the fading coefficients $h_{j,k}$ are such that $|h_{j,k}|$ has Rayleigh distribution, the channel is called Rayleigh fading channel.

Remark 2.3. It is important to remember that a good code needs to have as many codewords as possible, so that to have a wider range of possible signals to transmit.

Let us now compute the probability of incorrect decoding P_e. First of all we have to define it properly. Denote with $\mathbb{P}$ the measure of probability of an event. For a given space-time code $\mathcal{C}$, we choose to define the probability of error as follows:

$$P_e := \sum_{X \in \mathcal{C}} \mathbb{P}(X \text{ is decoded incorrectly}, X \text{ is sent}).$$

Using conditional probability we obtain

$$P_e = \sum_{X \in \mathcal{C}} \mathbb{P}(X \text{ is decoded incorrectly} | X \text{ is sent}) \mathbb{P}(X \text{ is sent}).$$

We assume that the codewords are all equally likely to be sent. Hence

$$P_e = \frac{1}{\#\mathcal{C}} \sum_{X \in \mathcal{C}} \mathbb{P}(X \text{ is decoded incorrectly} | X \text{ is sent}).$$

For $X, Y \in \mathcal{C}$, denote by $X \to Y$ the event that Y is decoded when X is sent. For all $X \in \mathcal{C}$ we then have that

$$\mathbb{P}(X \text{ is decoded incorrectly} | X \text{ is sent}) = \sum_{\substack{Y \in \mathcal{C} \\ Y \neq X}} \mathbb{P}(X \to Y | X \text{ is sent}).$$

The probability $\mathbb{P}(X \to Y | X \text{ is sent})$ is commonly referred to as the *pairwise error probability* (PEP). Putting everything together we obtain

$$P_e = \frac{1}{\#\mathcal{C}} \sum_{X \in \mathcal{C}} \sum_{\substack{Y \in \mathcal{C} \\ Y \neq X}} \mathbb{P}(X \to Y | X \text{ is sent}).$$

Therefore P_e depends only on the PEP and thus minimizing it means minimize P_e. Jafarkhani, in section 3.2 of [3], gives an upper bound of this PEP. He does it in the following way. He first conditions on H, i.e., he finds an upper bound for $\mathbb{P}(X \to Y | X \text{ is sent}, H = h)$. Then he uses the fact that

$$\begin{aligned} \mathbb{P}(X \to Y | X \text{ is sent}) &= E_H \left[\mathbb{P}(X \to Y | X \text{ is sent}, H) \right] \\ &= \int_{\text{domain of } H} \mathbb{P}(X \to Y | X \text{ is sent}, H = h) f_H(h) dh, \end{aligned}$$

where $f_H(\cdot)$ is the probability density function of H.
He then shows that

$$\mathbb{P}(X \to Y | X \text{ is sent}) \leq \frac{1}{\prod_{k=1}^{n} \left(1 + \frac{\lambda_k}{4\sigma^2}\right)^m}, \tag{2.3}$$

where $\lambda_1, \ldots, \lambda_n$ are the non-negative real eigenvalues of the matrix $A(X, Y) = (X - Y)(X - Y)^*$.
Without loss of generality we can assume that $\lambda_1, \ldots, \lambda_r$ are the non-zero eigenvalues of $A(X, Y)$, for some $r \in \{0, 1, \ldots, n\}$. Hence (2.3) can be written as

$$\mathbb{P}(X \to Y | X \text{ is sent}) \leq \frac{1}{\prod_{k=1}^{r} \left(1 + \frac{\lambda_k}{4\sigma^2}\right)^m}.$$

Since $(1 + a)^{-1} \leq a^{-1}$, for any $a \in \mathbb{R}^+$, we can write

$$\mathbb{P}(X \to Y | X \text{ is sent}) \leq \prod_{k=1}^{r} \left(\frac{\lambda_k}{4\sigma^2} \right)^{-m}.$$

We want to choose our code $\mathcal{C}$ so that the quantity on the right hand side of this inequality is small. This happens for example if the determinant of $A(X, Y)$ is non-zero and large, for any pair of codewords $X, Y \in \mathcal{C}$, with $X \neq Y$.

2.3 The mathematical question

From now on we suppose that $T = n$, so that the codewords are square matrices of size n. In this case

$$\det(A(X-Y)) = \det(X-Y)\overline{\det(X-Y)} = |\det(X-Y)|^2.$$

Hence, maximizing $\det(A(X,Y))$ is equivalent to maximize $|\det(X-Y)|$. We can therefore derive one criterion for the design of an optimal space-time code $\mathcal{C}$ in terms of reliability:

- **determinant criterion:** For a given $r \in \mathbb{R}^+$, $|\det(X-Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.

At first glance it could seem easy to find optimal space-time codes following this criterion. In fact, letting I_n be the identity matrix of $\mathbb{M}_n(\mathbb{C})$, we could consider the scalar matrices, i.e., for a given $r \in \mathbb{R}^+$, we could denote $X_1 = \sqrt[n]{r}I_n, X_2 = 2\sqrt[n]{r}I_n, X_3 = 3\sqrt[n]{r}I_n, \ldots$ and consider $\mathcal{C} = \{X_1, X_2, X_3, \ldots\}$. Then for any $X_j, X_k \in \mathcal{C}$, with $j \neq k$, $j, k \geq 1$,

$$|\det(X_j - X_k)| = \left|\det\left((j-k)\sqrt[n]{r}I_n\right)\right| = \underbrace{|j-k|^n}_{\geq 1} r \geq r.$$

So we could build a space-time code of infinite size such that P_e would be as small as desired.
This would be too easy. Indeed there is an important condition to take into account in this problem. The power needed to send a signal $z \in \mathbb{C}$ is directly proportional to its square norm $|z|^2$ and the available power at the transmitter is limited. Therefore every codeword X of a space-time code $\mathcal{C}$ has to be such that its square Euclidean norm $\|X\|^2$ is bounded by some energy constraint $0 \leq e^2 \in \mathbb{R}$. Hence, returning to the example of the scalar matrices, for a given r the code would be of fixed size. More precisely its size could not exceed $\frac{2e}{\sqrt[n]{r}\sqrt{n}}$, since $\|X_j\| = j\sqrt[n]{r}\sqrt{n}$, where $j \geq 1$.
Therefore we have to consider also this other condition when looking for an optimal space-time code $\mathcal{C}$:

- **energy constraint:** For some $e \in \mathbb{R}^+$, $\|X\| \leq e$, for all $X \in \mathcal{C}$.

We have thus arrived at the following mathematical question:

Given a subfield L of $\mathbb{C}$, for any $n \in \mathbb{N}^+$ and any $r \in \mathbb{R}^+$, find a set $\mathcal{C} \subset \mathbb{M}_n(L)$ of maximal size such that the following two conditions are satisfied:

- ***energy constraint**: $\|X\| \leq 1$, for all $X \in \mathcal{C}$,*
- ***determinant criterion**: $|\det(X - Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.*

Remark 2.4. Note that we have normalized the energy constraint to 1. In fact, of the two parameters e and r, where $\|X\| \leq e$ and $|\det(X - Y)| \geq r$, only one of them is necessary. Indeed we can rescale the matrices by a factor of $1/e$ in order to obtain $\|X\| \leq 1$ and $|\det(X - Y)| \geq \frac{r}{e^n} = r'$ (provided that $e \in L$). We could also normalize r to 1 and leave e as a parameter.

Chapter 3

Preliminaries from classical coding theory

In this chapter we give a brief summary of the definitions and results from classical coding theory which will motivate the arguments in the subsequent chapters. For a more detailed treatment of classical coding theory, we refer the reader to [4] or [5].

Definition 3.1 (code, codeword)**.** Let $\mathbb{F}_q$ be a finite field of cardinality $q = p^m$, for some prime p and some $m \in \mathbb{N}^+$. A *code* of length $n \in \mathbb{N}^+$ is a subset $\mathcal{C}$ of $(\mathbb{F}_q)^n$.
A *codeword* is an element of $\mathcal{C}$. So each codeword has the same length.

Definition 3.2 (Hamming distance)**.** Let $\mathcal{C} \subset (\mathbb{F}_q)^n$ be a code and let $\mathbf{x}, \mathbf{y}$ be two codewords of $\mathcal{C}$. The *Hamming distance* of $\mathbf{x}$ and $\mathbf{y}$, denoted by $d_H(\mathbf{x}, \mathbf{y})$, is the number of entries in which $\mathbf{x}$ and $\mathbf{y}$ differ. In other words

$$d_H(\mathbf{x}, \mathbf{y}) := \#\{k : x_k \neq y_k, k = 1, \ldots, n\}.$$

Observe that d_H is a metric on $(\mathbb{F}_q)^n$, as stated in the following lemma.

Lemma 3.3. *Let* $\mathbf{x}, \mathbf{y}$ *and* $\mathbf{z} \in (\mathbb{F}_q)^n$*. They satisfy*

1. $d_H(\mathbf{x}, \mathbf{y}) \geq 0$*, with equality if and only if* $\mathbf{x} = \mathbf{y}$*;*
2. $d_H(\mathbf{x}, \mathbf{y}) = d_H(\mathbf{y}, \mathbf{x})$*;*
3. $d_H(\mathbf{x}, \mathbf{z}) \leq d_H(\mathbf{x}, \mathbf{y}) + d_H(\mathbf{y}, \mathbf{z})$.

Proof. The first two properties are trivial.
For the third one, let $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ be elements of $(\mathbb{F}_q)^n$ and define

$$D_H(\mathbf{a}, \mathbf{b}) := \{k : a_k \neq b_k, k = 1, \ldots, n\}.$$

Note that $d_H(\mathbf{a}, \mathbf{b}) = \#D_H(\mathbf{a}, \mathbf{b})$.
We claim that $D_H(\mathbf{x}, \mathbf{z}) \subset D_H(\mathbf{x}, \mathbf{y}) \cup D_H(\mathbf{y}, \mathbf{z})$.
In order to prove this, let $k \in D_H(\mathbf{x}, \mathbf{z})$. This means that $x_k \neq z_k$. If $y_k \neq x_k$, then $k \in D_H(\mathbf{x}, \mathbf{y})$. Otherwise $y_k = x_k$, which means that $y_k \neq z_k$ and thus $k \in D_H(\mathbf{y}, \mathbf{z})$. In any case $k \in D_H(\mathbf{x}, \mathbf{y}) \cup D_H(\mathbf{y}, \mathbf{z})$, hence proving our claim.
Now, since for any finite set A and B we have that $\#A \cup B \leq \#A + \#B$, we obtain

$$\begin{aligned} d_H(\mathbf{x}, \mathbf{z}) = \#D_H(\mathbf{x}, \mathbf{z}) \leq \#D_H(\mathbf{x}, \mathbf{y}) \cup D_H(\mathbf{y}, \mathbf{z}) &\leq \#D_H(\mathbf{x}, \mathbf{y}) + \#D_H(\mathbf{y}, \mathbf{z}) \\ &= d_H(\mathbf{x}, \mathbf{y}) + d_H(\mathbf{y}, \mathbf{z}), \end{aligned}$$

as requested. □

Definition 3.4 (minimum Hamming distance)**.** Let $\mathcal{C} \subset (\mathbb{F}_q)^n$ be a code. The *minimum Hamming distance* of $\mathcal{C}$ is

$$d_H(\mathcal{C}) := \min\{d_H(\mathbf{x}, \mathbf{y}) : \mathbf{x} \neq \mathbf{y},\ \mathbf{x}, \mathbf{y} \in \mathcal{C}\}.$$

This parameter is very important, especially in the design of *error correcting codes*. In fact, suppose we have a code $\mathcal{C} \subset (\mathbb{F}_q)^n$, where d is its minimum Hamming distance. Then we can correct up to $\left\lfloor \frac{d-1}{2} \right\rfloor$ errors, as stated in following proposition.

Proposition 3.5. *Let $\mathcal{C} \subset (\mathbb{F}_q)^n$ be a code of minimum Hamming distance $\geq d$. If a codeword is sent with at most $\left\lfloor \frac{d-1}{2} \right\rfloor$ errors, these errors can be corrected.*

Proof. Let $t = \left\lfloor \frac{d-1}{2} \right\rfloor$ and let $\mathbf{x} \in \mathcal{C}$ be a sent codeword. By hypothesis the received word $\mathbf{y} \in (\mathbb{F}_q)^n$ contains at most t errors. Then $d_H(\mathbf{x}, \mathbf{y}) \leq t$.
We now claim that for any other codeword $\mathbf{z} \in \mathcal{C}$, with $\mathbf{z} \neq \mathbf{x}$, we have that $d_H(\mathbf{z}, \mathbf{y}) > t$. In fact, if this was not the case, i.e., if there were a codeword $\mathbf{z} \in \mathcal{C}$, $\mathbf{z} \neq \mathbf{x}$, such that $d_H(\mathbf{z}, \mathbf{y}) \leq t$, we would obtain that

$$d_H(\mathbf{x}, \mathbf{z}) \leq d_H(\mathbf{x}, \mathbf{y}) + d_H(\mathbf{y}, \mathbf{z}) \leq 2t = 2\left\lfloor \frac{d-1}{2} \right\rfloor \leq d - 1.$$

But this is a contradiction with the fact that $\mathcal{C}$ has minimum distance $\geq d$. Thus our claim is proved. It then follows that $\mathbf{x}$ is the only codeword at a distance at most t from $\mathbf{y}$, therefore $\mathbf{y}$ can be decoded correctly, hence correcting all the errors. □

The theory of error correcting codes is a much studied topic since the second half of the XX century. Thanks to proposition 3.5, the basic goal is

to find codes that have as many codewords as possible with a fixed desired minimum distance. Obviously $d_H((\mathbb{F}_q)^n) = 1$ and it is clear that if we keep adding elements of $(\mathbb{F}_q)^n$ to $\mathcal{C}$, its minimum distance will tend to decrease. Hence the following mathematical question is natural in this context:

> *Given a prime power q, for any $n \in \mathbb{N}^+$ and $d \geq 0$, find a set $\mathcal{C} \subset (\mathbb{F}_q)^n$ of maximal size such that $\mathcal{C}$ has minimum Hamming distance $\geq d$.*

An answer is not known in general but various upper and lower bounds have been proved on the size of $\mathcal{C}$. The simplest of these are the *sphere packing bound* and the *Gilbert-Varshamov bound*, respectively, which we recall here.

Before stating these bounds, we first define two useful sets. For $\mathbf{x} \in (\mathbb{F}_q)^n$ and $\rho \in \{0, 1, \ldots, n\}$, define

$$S(\mathbf{x}, \rho) = \{\mathbf{z} \in (\mathbb{F}_q)^n : d_H(\mathbf{z}, \mathbf{x}) = \rho\}$$

and

$$B(\mathbf{x}, \rho) = \{\mathbf{z} \in (\mathbb{F}_q)^n : d_H(\mathbf{z}, \mathbf{x}) \leq \rho\} = \bigcup_{k=0}^{\rho} S(\mathbf{x}, k).$$

The set $B(\mathbf{x}, \rho)$ may be viewed as a ball in $(\mathbb{F}_q)^n$ centered at $\mathbf{x}$ and of radius ρ, relatively to the metric d_H.

Lemma 3.6. *For $\mathbf{x} \in (\mathbb{F}_q)^n$ and $\rho \in \{0, 1, \ldots, n\}$,*

$$\#B(\mathbf{x}, \rho) = \sum_{k=0}^{\rho} \binom{n}{k} (q-1)^k.$$

Proof. Clearly $\#B(\mathbf{x}, \rho) = \#\bigcup_{k=0}^{\rho} S(\mathbf{x}, k) = \sum_{k=0}^{\rho} \#S(\mathbf{x}, k)$, because the $S(\mathbf{x}, k)$ are disjoint. Since $\#S(\mathbf{x}, k)$ is the number of elements of $(\mathbb{F}_q)^n$ which differ from $\mathbf{x}$ in exactly k places,

$$\#S(\mathbf{x}, k) = \binom{n}{k} (q-1)^k$$

and the lemma follows. □

Theorem 3.7 (sphere-packing bound). *Let $\mathcal{C} \subset (\mathbb{F}_q)^n$ be a code with minimum Hamming distance $\geq d$. Then*

$$\#\mathcal{C} \leq \frac{q^n}{\sum_{k=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{k} (q-1)^k}.$$

Proof. Let $t = \lfloor \frac{d-1}{2} \rfloor$.
Since the Hamming distance is a metric, the balls $B(\mathbf{x}, t)$ are pairwise disjoint, as $\mathbf{x}$ ranges over $\mathcal{C}$. Thanks to lemma 3.6, we then obtain

$$q^n = \#(\mathbb{F}_q)^n \geq \# \bigcup_{\mathbf{x} \in \mathcal{C}} B(\mathbf{x}, t) = \sum_{\mathbf{x} \in \mathcal{C}} \#B(\mathbf{x}, t) = \#\mathcal{C} \cdot \sum_{k=0}^{t} \binom{n}{k} (q-1)^k,$$

as desired. □

Theorem 3.8 (Gilbert-Varshamov bound). *Let $n, d \in \mathbb{N}^+$ and q be a prime power. There exists then a code $\mathcal{C} \subset (\mathbb{F}_q)^n$ with minimum Hamming distance $\geq d$, such that*

$$\#\mathcal{C} \geq \frac{q^n}{\sum_{k=0}^{d-1} \binom{n}{k} (q-1)^k}.$$

Proof. Start with $\mathcal{C} = \{\mathbf{x_1}\}$, where $\mathbf{x_1}$ is any element of $(\mathbb{F}_q)^n$. Then apply the following procedure:

1. If there is an element $\mathbf{z} \in (\mathbb{F}_q)^n$ such that $\mathbf{z} \notin \bigcup_{\mathbf{x} \in \mathcal{C}} B(\mathbf{x}, d-1)$, then add $\mathbf{z}$ to $\mathcal{C}$, otherwise stop.

2. Repeat step 1 until you have to stop.

Since the total number of elements of $(\mathbb{F}_q)^n$ is finite, this process will stop in a finite number of steps and the resulting code $\mathcal{C}$ will have minimum distance at least d by construction.
Moreover, at that point $\mathcal{C}$ will be such that

$$\bigcup_{\mathbf{x} \in \mathcal{C}} B(\mathbf{x}, d-1) = (\mathbb{F}_q)^n,$$

because otherwise we could have continued with the process. Hence, again using lemma 3.6, we obtain

$$q^n = \#(\mathbb{F}_q)^n = \# \bigcup_{\mathbf{x} \in \mathcal{C}} B(\mathbf{x}, d-1) \leq \sum_{\mathbf{x} \in \mathcal{C}} \#B(\mathbf{x}, d-1) = \#\mathcal{C} \cdot \sum_{k=0}^{d-1} \binom{n}{k} (q-1)^k,$$

as desired. □

Chapter 4

An upper bound

Recall that, given a subfield L of $\mathbb{C}$, for $n \in \mathbb{N}^+$ and $r \in \mathbb{R}^+$, we are looking for a set $\mathcal{C} \subset \mathbb{M}_n(L)$ of maximal size such that the following two conditions are satisfied:

- *energy constraint:* $\|X\| \leq 1$, for all $X \in \mathcal{C}$,
- *determinant criterion:* $|\det(X - Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.

In this chapter we will prove an upper bound on the size of $\mathcal{C}$ by mimicking the proof of the sphere packing bound in classical coding theory (see theorem 3.7).

Since $\mathbb{M}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$, any matrix $X \in \mathbb{M}_n(\mathbb{R})$ can be seen as an element of the $\mathbb{R}$ vector space $\mathbb{R}^{n^2}$ and vice versa. Similarly, since $\mathbb{M}_n(\mathbb{C}) \cong \mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$, any matrix $X \in \mathbb{M}_n(\mathbb{C})$ can be seen as an element of the $\mathbb{R}$ vector space $\mathbb{R}^{2n^2}$.
Recall that $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$.
For any matrix $X = (\mathbf{x_1}|\dots|\mathbf{x_n}) \in \mathbb{M}_n(\mathbb{F})$ and any $0 \leq \rho \in \mathbb{R}$ we define the following set:

$$B_n(X, \rho) := \{A = (\mathbf{a_1}|\cdots|\mathbf{a_n}) \in \mathbb{M}_n(\mathbb{F}) : \|\mathbf{a_k} - \mathbf{x_k}\| \leq \rho,\ k = 1, \dots, n\}.$$

If $\mathbb{F} = \mathbb{R}$, $B_n(X, \rho)$ is the direct product of n copies of n-dimensional balls of radius ρ. If $\mathbb{F} = \mathbb{C}$, it is the direct product of n copies of $2n$-dimensional balls of radius ρ.
Let S be a subset of $\mathbb{R}^n$, with $\mathrm{int}(S)$ we mean the interior of S. Clearly

$$\mathrm{int}(B_n(X, \rho)) = \{A = (\mathbf{a_1}|\cdots|\mathbf{a_n}) \in \mathbb{M}_n(\mathbb{F}) : \|\mathbf{a_k} - \mathbf{x_k}\| < \rho,\ k = 1, \dots, n\}.$$

Proposition 4.1. *Let X and Y be two distinct matrices of $\mathbb{M}_n(\mathbb{F})$ such that $|\det(X - Y)| \geq r$, for some $r \in \mathbb{R}^+$. Then*

$$\mathrm{int}(B_n(X, \sqrt[n]{r}/2)) \cap \mathrm{int}(B_n(Y, \sqrt[n]{r}/2)) = \emptyset.$$

Proof. Let $X = (\mathbf{x_1}|\dots|\mathbf{x_n})$ and $Y = (\mathbf{y_1}|\dots|\mathbf{y_n})$. By contradiction, suppose that the intersection of $\mathrm{int}(B_n(X, \sqrt[n]{r}/2))$ and $\mathrm{int}(B_n(Y, \sqrt[n]{r}/2))$ in nonempty, i.e., there exists a matrix $A = (\mathbf{a_1}|\cdots|\mathbf{a_n}) \in \mathbb{M}_n(\mathbb{F})$ that belongs to both of them. For all $k = 1, \dots, n$, we then have that

$$\|\mathbf{x_k} - \mathbf{y_k}\| = \|\mathbf{x_k} - \mathbf{a_k} + \mathbf{a_k} - \mathbf{y_k}\| \leq \|\mathbf{x_k} - \mathbf{a_k}\| + \|\mathbf{a_k} - \mathbf{y_k}\| < \frac{\sqrt[n]{r}}{2} + \frac{\sqrt[n]{r}}{2} = \sqrt[n]{r}.$$

Applying theorem 1.7 to the matrix $X - Y$ we obtain that

$$r \leq |\det(X - Y)| \leq \prod_{k=1}^{n} \|\mathbf{x_k} - \mathbf{y_k}\| < \prod_{k=1}^{n} \sqrt[n]{r} = r,$$

which is a contradiction, thus proving our proposition. □

Figure 4.1 illustrates the next lemma.

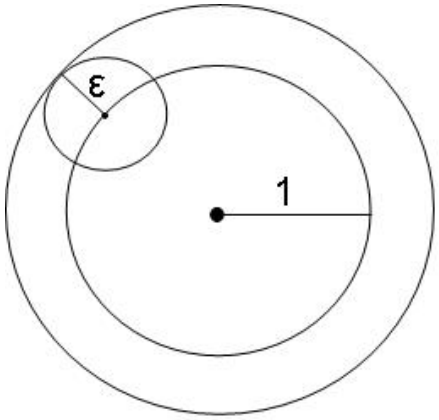

Figure 4.1: The disk of radius 1 represent the energy constraint: every element of $\mathcal{C}$ must lie inside it. The disk of radius $\varepsilon = \sqrt{n}\sqrt[n]{r}/2$ is meant to represent the set $B_n(X, \sqrt[n]{r}/2)$, for some X on the boundary of the disk or radius 1. We see that these sets, for $X \in \mathcal{C}$, have to lie in a ball of radius $1 + \varepsilon$ centered at the origin.

Lemma 4.2. *Let $X \in \mathbb{M}_n(\mathbb{F})$ such that $\|X\| \leq 1$.*

- *If $\mathbb{F} = \mathbb{R}$, then $B_n(X, \sqrt[n]{r}/2) \subseteq \mathbb{B}_{n^2}\left(1 + \frac{\sqrt{n}\sqrt[n]{r}}{2}\right)$.*
- *If $\mathbb{F} = \mathbb{C}$, then $B_n(X, \sqrt[n]{r}/2) \subseteq \mathbb{B}_{2n^2}\left(1 + \frac{\sqrt{n}\sqrt[n]{r}}{2}\right)$.*

Proof. Let $X = (\mathbf{x_1}|\dots|\mathbf{x_n})$ and let $A \in B_n(X, \sqrt[n]{r}/2)$, with $A = (\mathbf{a_1}|\dots|\mathbf{a_n})$. We then have that $\|\mathbf{a_k} - \mathbf{x_k}\| \leq \sqrt[n]{r}/2$, for all $k = 1, \dots, n$. Therefore, we have that

$$\|A - X\|^2 = \sum_{k=1}^{n} \|\mathbf{a_k} - \mathbf{x_k}\|^2 \leq n \left(\frac{\sqrt[n]{r}}{2}\right)^2.$$

Thus

$$\|A\| \leq \|A - X\| + \underbrace{\|X\|}_{\leq 1} \leq \frac{\sqrt{n}\sqrt[n]{r}}{2} + 1.$$

Hence, if $\mathbb{F} = \mathbb{R}$, $A \in \mathbb{B}_{n^2}\left(1 + \frac{\sqrt{n}\sqrt[n]{r}}{2}\right)$ and if $\mathbb{F} = \mathbb{C}$, $A \in \mathbb{B}_{2n^2}\left(1 + \frac{\sqrt{n}\sqrt[n]{r}}{2}\right)$. □

We can now give an upper bound for the size of our set $\mathcal{C}$.

Theorem 4.3. *Let L be a subfield of $\mathbb{F}$, $n \in \mathbb{N}^+$ and $r \in \mathbb{R}^+$. Let $\mathcal{C} \subset \mathbb{M}_n(L)$ such that $\|X\| \leq 1$, for all $X \in \mathcal{C}$, and $|\det(X - Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.*

- *If $\mathbb{F} = \mathbb{R}$, then*

$$\#\mathcal{C} \leq \frac{\Gamma\left(\frac{n}{2} + 1\right)^n}{\Gamma\left(\frac{n^2}{2} + 1\right)} \left(\frac{2}{\sqrt[n]{r}} + \sqrt{n}\right)^{n^2}.$$

- *If $\mathbb{F} = \mathbb{C}$, then*

$$\#\mathcal{C} \leq \frac{(n!)^n}{(n^2)!} \left(\frac{2}{\sqrt[n]{r}} + \sqrt{n}\right)^{2n^2}.$$

Proof. Proposition 4.1 tells us that

$$\mathcal{V}\left(\bigcup_{X \in \mathcal{C}} \operatorname{int}(B_n(X, \sqrt[n]{r}/2))\right) = \sum_{X \in \mathcal{C}} \mathcal{V}\left(\operatorname{int}(B_n(X, \sqrt[n]{r}/2))\right),$$

since this union is disjoint. Clearly $\mathcal{V}(\operatorname{int}(B_n(X, \sqrt[n]{r}/2))) = \mathcal{V}(B_n(X, \sqrt[n]{r}/2))$ and

$$\bigcup_{X \in \mathcal{C}} \operatorname{int}(B_n(X, \sqrt[n]{r}/2)) \subseteq \bigcup_{X \in \mathcal{C}} B_n(X, \sqrt[n]{r}/2).$$

Therefore

$$\sum_{X \in \mathcal{C}} \mathcal{V}\left(B_n(X, \sqrt[n]{r}/2)\right) \leq \mathcal{V}\left(\bigcup_{X \in \mathcal{C}} B_n(X, \sqrt[n]{r}/2)\right).$$

Applying lemma 4.2 to all the elements of $\mathcal{C}$ we have that

$$\mathcal{V}\left(\bigcup_{X\in\mathcal{C}} B_n\left(X, \frac{\sqrt[n]{r}}{2}\right)\right) \leq \begin{cases} \mathcal{V}\left(\mathbb{B}_{n^2}\left(1+\frac{\sqrt{n}\sqrt[n]{r}}{2}\right)\right) = \frac{\pi^{\frac{n^2}{2}}}{\Gamma\left(\frac{n^2}{2}+1\right)}\left(1+\frac{\sqrt{n}\sqrt[n]{r}}{2}\right)^{n^2}, & \mathbb{F}=\mathbb{R} \\ \mathcal{V}\left(\mathbb{B}_{2n^2}\left(1+\frac{\sqrt{n}\sqrt[n]{r}}{2}\right)\right) = \frac{\pi^{n^2}}{(n^2)!}\left(1+\frac{\sqrt{n}\sqrt[n]{r}}{2}\right)^{2n^2}, & \mathbb{F}=\mathbb{C}. \end{cases}$$

For any $X \in \mathbb{M}_n(\mathbb{F})$, notice that the volume of $B_n(X, \sqrt[n]{r}/2)$ does not depend on X, since it is the direct product of n balls of radius $\sqrt[n]{r}/2$. These balls are n-dimensional if $\mathbb{F}=\mathbb{R}$ and $2n$-dimensional if $\mathbb{F}=\mathbb{C}$. Therefore

$$\sum_{X\in\mathcal{C}} \mathcal{V}\left(B_n(X, \sqrt[n]{r}/2)\right) = \#\mathcal{C}\cdot\mathcal{V}\left(B_n(X, \sqrt[n]{r}/2)\right)$$

and

$$\mathcal{V}\left(B_n(X, \sqrt[n]{r}/2)\right) = \begin{cases} \mathcal{V}\left(\mathbb{B}_n\left(\frac{\sqrt[n]{r}}{2}\right)\right)^n = \frac{\pi^{\frac{n^2}{2}}}{\Gamma\left(\frac{n}{2}+1\right)^n}\left(\frac{\sqrt[n]{r}}{2}\right)^{n^2}, & \text{if } \mathbb{F}=\mathbb{R} \\ \mathcal{V}\left(\mathbb{B}_{2n}\left(\frac{\sqrt[n]{r}}{2}\right)\right)^n = \frac{\pi^{n^2}}{(n!)^n}\left(\frac{\sqrt[n]{r}}{2}\right)^{2n^2}, & \text{if } \mathbb{F}=\mathbb{C}. \end{cases}$$

We then obtain

$$\#\mathcal{C} \leq \begin{cases} \mathcal{V}\left(\mathbb{B}_{n^2}\left(1+\frac{\sqrt{n}\sqrt[n]{r}}{2}\right)\right)/\mathcal{V}\left(\mathbb{B}_n\left(\frac{\sqrt[n]{r}}{2}\right)\right)^n, & \text{if } \mathbb{F}=\mathbb{R} \\ \mathcal{V}\left(\mathbb{B}_{2n^2}\left(1+\frac{\sqrt{n}\sqrt[n]{r}}{2}\right)\right)/\mathcal{V}\left(\mathbb{B}_{2n}\left(\frac{\sqrt[n]{r}}{2}\right)\right)^n, & \text{if } \mathbb{F}=\mathbb{C}, \end{cases}$$

which yields the desired results. □

We give here the upper bound for the special case $n=2$.

Example 4.4. Let L be a subfield of $\mathbb{F}$ and $r \in \mathbb{R}^+$. Let $\mathcal{C} \subset \mathbb{M}_n(L)$ such that $\|X\| \leq 1$, for all $X \in \mathcal{C}$, and $|\det(X-Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.

- If $\mathbb{F}=\mathbb{R}$, then
$$\#\mathcal{C} \leq 2\left(\sqrt{\frac{2}{r}}+1\right)^4.$$
- If $\mathbb{F}=\mathbb{C}$, then
$$\#\mathcal{C} \leq \frac{8}{3}\left(\sqrt{\frac{2}{r}}+1\right)^8.$$

Chapter 5

A lower bound

Recall that, given a subfield L of $\mathbb{C}$, $n \in \mathbb{N}^+$ and $r \in \mathbb{R}^+$, we are looking for a set $\mathcal{C} \subset \mathbb{M}_n(L)$ of maximal size such that the two following conditions are satisfied:

- *energy constraint:* $\|X\| \leq 1$, for all $X \in \mathcal{C}$,
- *determinant criterion:* $|\det(X-Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.

After rescaling the codewords of $\mathcal{C}$ by a factor of $1/\sqrt[n]{r}$, i.e., considering the set $\mathcal{C}' = \mathcal{C}/\sqrt[n]{r} \subset \mathbb{M}_n(L)$, the two conditions, for all distinct $X', Y' \in \mathcal{C}'$, are

$$\|X'\| = \frac{1}{\sqrt[n]{r}}\|X\| \leq \frac{1}{\sqrt[n]{r}} \quad \text{and} \quad |\det(X'-Y')| = \frac{1}{r}|\det(X-Y)| \geq 1.$$

In order to do this the only condition is that $\sqrt[n]{r}$ belongs to L, otherwise $\mathcal{C}'$ would not be a subset of $\mathbb{M}_n(L)$. Letting $\mathbb{W}$ represent either $\mathbb{Z}$ or $\mathbb{Z}[i]$, we also want the entries of the codewords $X \in \mathcal{C}'$ to belong to $\mathbb{W}$.

Letting $R = \frac{1}{\sqrt[n]{r}}$, we then concentrate on the equivalent problem of finding a set $\mathcal{C} \subset \mathbb{M}_n(\mathbb{W})$ of maximal size, such that, for some fixed $R \in \mathbb{R}^+$, the following two conditions hold:

- *energy constraint:* $\|X\| \leq R$, for all $X \in \mathcal{C}$ and
- *determinant criterion:* $|\det(X-Y)| \geq 1$, for all distinct $X, Y \in \mathcal{C}$.

The interesting point in this formulation is that if two matrices $X, Y \in \mathbb{M}_n(\mathbb{W})$ do not satisfy the determinant criterion, then $\det(X-Y) = 0$, because $|\det(X-Y)| \in \mathbb{N}$.

The purpose of this chapter is to prove a lower bound on the size of a space-time code by mimicking the proof of the Gilbert-Varshamov bound in classical coding theory (see theorem 3.8).
The following notations will be useful.

- Let $m \in \mathbb{N}^+$, $0 \leq R \in \mathbb{R}$ and $d \in \mathbb{R}^+$. The lattice $(d\mathbb{Z})^m$ cuts the m-dimensional euclidean space $\mathbb{R}^m$ into hypercubes of side d. We define $\gamma_m(d, R)$ to be the number of these hypercubes that have a non-empty intersection with $\mathbb{B}_m(R)$. If $d = 1$, we denote it simply by $\gamma_m(R)$.

- For $m \in \mathbb{N}^+$ and $0 \leq R \in \mathbb{R}$, define

$$\beta_m(R) := \#\mathbb{Z}^m \cap \mathbb{B}_m(R),$$

  the number of integer points inside the m-dimensional ball of radius R centered at the origin.

- For $n \in \mathbb{N}^+$ and $0 \leq R \in \mathbb{R}$, define

$$B_n(R) := \{X \in \mathbb{M}_n(\mathbb{W}) : \|X\| \leq R\}$$

  and $b_n(R) := \#B_n(R)$.
  Note that

$$b_n(R) = \begin{cases} \beta_{n^2}(R), & \text{if } \mathbb{W} = \mathbb{Z} \\ \beta_{2n^2}(R), & \text{if } \mathbb{W} = \mathbb{Z}[i]. \end{cases}$$

- For $n \in \mathbb{N}^+$, $0 \leq R \in \mathbb{R}$ and $X \in B_n(R)$, define

$$F_n(X, R) := \{A \in B_n(R) : \det(A - X) = 0\}.$$

Lemma 5.1. *Let $m \in \mathbb{N}^+$, $0 \leq R \in \mathbb{R}$ and $d \in \mathbb{R}^+$. Then*

$$\mathcal{V}\left(\mathbb{B}_m\left(\frac{R}{d}\right)\right) \leq \gamma_m(d, R) \leq \mathcal{V}\left(\mathbb{B}_m\left(\frac{R}{d} + \sqrt{m}\right)\right).$$

Proof. Note that the volume of these $\gamma_m(d, R)$ hypercubes is $\gamma_m(d, R) \cdot d^m$. Since we are counting any hypercube that has a non-empty intersection with $\mathbb{B}_m(R)$, it is clear that the union of all these hypercubes contains this ball. In particular we then have that

$$\mathcal{V}\left(\mathbb{B}_m\left(R\right)\right) \leq \gamma_m(d, R) \cdot d^m \quad \Leftrightarrow \quad \gamma_m(d, R) \geq \mathcal{V}\left(\mathbb{B}_m\left(\frac{R}{d}\right)\right).$$

On the other hand, inside a hypercube of side d, two points can be at a distance which is at most $d\sqrt{m}$. Therefore if we expand our ball $\mathbb{B}_m(R)$ to have radius $R + d\sqrt{m}$, we will be sure that this new sphere will contain all the $\gamma_m(d, R)$ hypercubes. Hence

$$\gamma_m(d, R) \cdot d^m \leq \mathcal{V}\left(\mathbb{B}_m\left(R + d\sqrt{m}\right)\right) \Leftrightarrow \gamma_m(d, R) \leq \mathcal{V}\left(\mathbb{B}_m\left(\frac{R}{d} + \sqrt{m}\right)\right),$$

as claimed. □

Lemma 5.2. *Let $m \in \mathbb{N}^+$ and $0 \leq R \in \mathbb{R}$. Then*

$$\gamma_m(R - \sqrt{m}) \leq \beta_m(R) \leq \gamma_m(R).$$

Proof. We consider the lattice $\mathbb{Z}^m$ that divides $\mathbb{R}^m$ into hypercubes of side 1. Each of these hypercubes has 2^m vertices, which are integer points. To the hypercube $[a_1, a_1 + 1] \times \ldots \times [a_m, a_m + 1]$ we associate the point $(a_1, \ldots, a_m)$, for all $(a_1, \ldots, a_m) \in \mathbb{Z}^m$, as shown in figure 5.1.

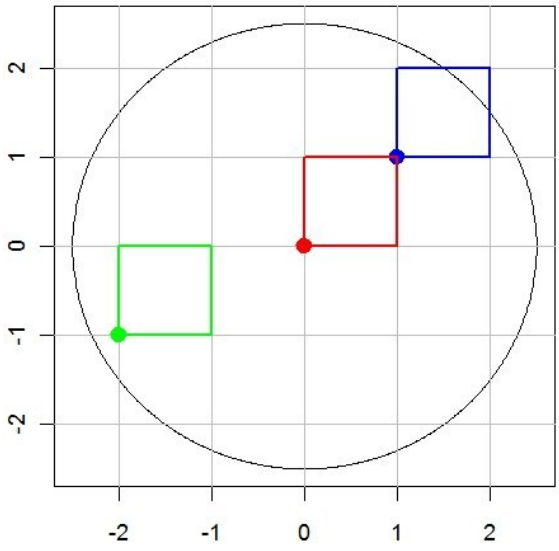

Figure 5.1: Here $m = 2$. To each square we associate its lower left vertex. The black circle is the boundary of $\mathbb{B}_2(R)$ for $R = 2.5$.

If we now consider those hypercubes that have non-empty intersection with $\mathbb{B}_m(R)$, we see that their number will exceed the number of integer points inside $\mathbb{B}_m(R)$. In fact, using the one-to-one correspondence we just defined between integer points and hypercubes, we have that the hypercube associated to some integer point inside $\mathbb{B}_m(R)$ will have a non-empty intersection with the ball (the point itself for instance). Therefore we have that the set of hypercubes associated to the integer points inside $\mathbb{B}_m(R)$ is included in the the set of hypercubes that have a non-empty intersection with $\mathbb{B}_m(R)$. Hence $\beta_m(R) \leq \gamma_m(R)$.
On the other hand, consider the set of hypercubes that have a non-empty intersection with the m-dimensional ball $\mathbb{B}_m(R - \sqrt{m})$. These hypercubes are completely included in the m-dimensional ball $\mathbb{B}_m(R)$, since the greatest

distance between two points inside one of these hypercubes is $\sqrt{m}$. Therefore also their left-most vertex is included in $\mathbb{B}_m(R)$ and thus, using again the one-to-one relation between integer points and hypercubes, $\gamma_m(R - \sqrt{m}) \leq \beta_m(R)$. □

Corollary 5.3. *Let $m \in \mathbb{N}^+$ and $0 \leq R \in \mathbb{R}$. Then*

$$\mathcal{V}\left(\mathbb{B}_m\left(R - \sqrt{m}\right)\right) \leq \beta_m(R) \leq \mathcal{V}\left(\mathbb{B}_m\left(R + \sqrt{m}\right)\right).$$

Proof. Follows combining lemma 5.1 and lemma 5.2. □

Proposition 5.4. *Let $m \in \mathbb{N}^+$, $0 \leq R \in \mathbb{R}$ and $k \in \mathbb{N}$ such that $k < m$. Let $\mathbf{v_1}, \ldots, \mathbf{v_k}, \mathbf{p} \in \mathbb{Z}^m$ and define the set*

$$H_{k,m} := \left\{ \mathbf{a} \in \mathbb{Z}^m : \mathbf{a} = \mathbf{p} + \sum_{t=1}^{k} \lambda_t \mathbf{v_t}, \text{ where } \lambda_t \in \mathbb{R}, \forall\, t = 1, \ldots, k \right\}.$$

We then have that $\# H_{k,m} \cap \mathbb{B}_m(R) \leq \beta_k(R)$.

Figure 5.2 illustrates the proposition.

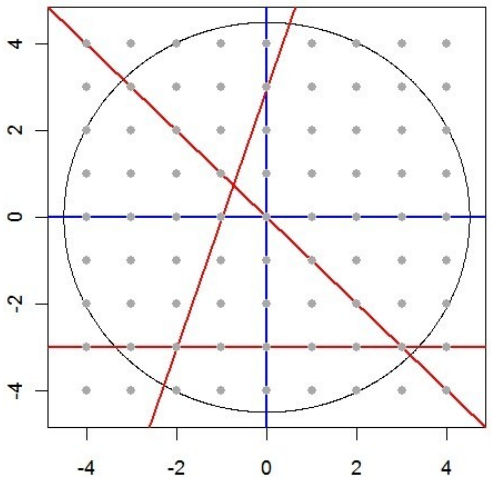

Figure 5.2: Here $m = 2$ and $R = 4.5$. We see that no matter how a line is chosen (the red lines), those that will contain the highest number of integer points inside $\mathbb{B}_m(R)$ are the two lines (in blue) that pass through the origin with direction $\mathbf{e_1} = (1, 0)$ and $\mathbf{e_2} = (0, 1)$. As one can see, the integer points that lie on the x or y axis inside $\mathbb{B}_2(R)$ are the integer points inside $\mathbb{B}_1(R)$.

Proof. We have k vectors $\mathbf{v_1}, \ldots, \mathbf{v_k} \in \mathbb{Z}^m$, where $\mathbf{v_t} = (v_{1t}, \ldots, v_{mt})^T$, for all $t = 1, \ldots, k$. For some $d \in \{1, \ldots, k\}$, the matrix $V = (\mathbf{v_1}| \ldots |\mathbf{v_k})$ will have rank d. This means that among $\mathbf{v_1}, \ldots, \mathbf{v_k}$, the maximum number of linearly independent vectors is d. So $H_{k,m}$ is a d-dimensional lattice in $\mathbb{R}^m$. Without loss of generality, let us assume that these d linearly independent vectors are $\mathbf{v_1}, \ldots, \mathbf{v_d}$ and so

$$H_{k,m} = \left\{ \mathbf{a} \in \mathbb{Z}^m : \mathbf{a} = \mathbf{p} + \sum_{t=1}^{d} \lambda_t \mathbf{v_t}, \text{ where } \lambda_t \in \mathbb{R}, \forall\, t = 1, \ldots, d \right\}.$$

Consequently, the $m \times d$ matrix $A = (\mathbf{v_1}| \ldots |\mathbf{v_d})$ has rank d. Since $d < m$, because by hypothesis $k < m$, this means that we can find a non-zero minor of size $d \times d$, i.e., we can find d rows among the m of A such that the corresponding square matrix has non-zero determinant. Let us denote this rows by $1 \leq j_1 < \ldots < j_d \leq n$. Hence the matrix

$$B = \begin{pmatrix} v_{j_1,1} & \ldots & v_{j_1,d} \\ \vdots & \ddots & \vdots \\ v_{j_d,1} & \ldots & v_{j_d,d} \end{pmatrix}$$

is such that $\det(B) \neq 0$.
Let $H^0_{j_1,\ldots,j_d}$ be the lattice generated by the basis vectors $\mathbf{e_{j_1}}, \ldots, \mathbf{e_{j_d}}$ and passing from the origin:

$$H^0_{j_1,\ldots,j_d} := \left\{ \mathbf{a} \in \mathbb{Z}^m : \mathbf{a} = \sum_{t=1}^{d} \lambda_t \mathbf{e_{j_t}},\ \lambda_t \in \mathbb{R},\ t = 1, \ldots, d \right\}.$$

Note that $H^0_{j_1,\ldots,j_d} \cong \mathbb{Z}^d$.
Now consider the orthogonal projection of $H_{k,m}$ on $H^0_{j_1,\ldots,j_d}$:

$$\begin{aligned} \pi_{j_1,\ldots,j_d} : \quad H_{k,m} \subset \mathbb{Z}^m \quad &\to H^0_{j_1,\ldots,j_d} \subset \mathbb{Z}^m \\ (x_1, \ldots, x_m) \quad &\mapsto (\tilde{x}_1, \ldots, \tilde{x}_m), \text{ where } \tilde{x}_j = \begin{cases} x_j, & j \in \{j_1, \ldots, j_d\} \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

We claim that this map is one-to-one.
In order to prove this, let $\mathbf{x} = (x_1, \ldots, x_m)^T, \mathbf{y} = (y_1, \ldots, y_m)^T \in H_{k,m}$ such that $\pi_{j_1,\ldots,j_d}(\mathbf{x}) = \pi_{j_1,\ldots,j_d}(\mathbf{y})$. This means that $x_j = y_j \Leftrightarrow x_j - y_j = 0$, for all $j \in \{j_1, \ldots, j_d\}$. On the other hand there are $\lambda_1, \ldots, \lambda_d$ and $\mu_1, \ldots, \mu_d \in \mathbb{R}$ such that $\mathbf{x} = \mathbf{p} + \sum_{t=1}^{d} \lambda_t \mathbf{v_t}$ and $\mathbf{y} = \mathbf{p} + \sum_{t=1}^{d} \mu_t \mathbf{v_t}$. The fact that $x_j - y_j = 0$, for all $j \in \{j_1, \ldots, j_d\}$, means that

$$\sum_{t=1}^{d} (\lambda_t - \mu_t) v_{jt} = 0, \qquad \text{for all } j \in \{j_1, \ldots, j_d\}. \tag{5.1}$$

Letting $\boldsymbol{\lambda} = (\lambda_1 - \mu_1, \ldots, \lambda_d - \mu_d)^T, \mathbf{0_d} = (0, \ldots, 0)^T \in \mathbb{R}^d$, (5.1) can be rewritten as

$$B\boldsymbol{\lambda} = \mathbf{0_d}.$$

Since $\det(B) \neq 0$ we can invert B and thus this linear system has exactly one solution, i.e., $\boldsymbol{\lambda} = \mathbf{0_d}$, which means that $\lambda_t = \mu_t$, for all $t = 1, \ldots, d$, implying that $\mathbf{x} = \mathbf{y}$, as desired.
Now let $\mathbf{a} \in H_{k,m} \cap \mathbb{B}_m(R)$, then clearly $\pi_{j_1,\ldots,j_d}(\mathbf{a}) \in H^0_{j_1,\ldots,j_d} \cap \mathbb{B}_m(R)$, since $\|\mathbf{a}\| \geq \|\pi_{j_1,\ldots,j_d}(\mathbf{a})\|$. Therefore, since the projection is one-to-one, we must have that

$$\#\, H_{k,m} \cap \mathbb{B}_m(R) \leq \#\, H^0_{j_1,\ldots,j_d} \cap \mathbb{B}_m(R) = \#\mathbb{Z}^d \cap \mathbb{B}_m(R) = \beta_d(R).$$

Since obviously $\beta_s(R) \leq \beta_{s'}(R)$, for all positive integers $s \leq s'$, we have that $\beta_d(R) \leq \beta_k(R)$, thus proving our proposition. □

Theorem 5.5. *Let $n \in \mathbb{N}^+$, $R \in \mathbb{R}^+$ and $\mathbb{W}$ be either $\mathbb{Z}$ or $\mathbb{Z}[i]$. There exists then a set $\mathcal{C} \subset \mathbb{M}_n(\mathbb{W})$ with $\|X\| \leq R$, for all $X \in \mathcal{C}$, and $|\det(X - Y)| \geq 1$, for all distinct $X, Y \in \mathcal{C}$, such that*

- *if $\mathbb{W} = \mathbb{Z}$,*

$$\#\mathcal{C} \geq \pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n}{2}+1\right)^{n-1}}{\Gamma\left(\frac{n^2}{2}+1\right)} \frac{(R-n)^{n^2}}{(R+\sqrt{n})^{n^2-n} \sum_{k=0}^{n-1} \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}+1\right)} \left(R+\sqrt{k}\right)^k},$$

- *if $\mathbb{W} = \mathbb{Z}[i]$,*

$$\#\mathcal{C} \geq \pi^n \frac{(n!)^{n-1}}{(n^2)!} \frac{(R-\sqrt{2}n)^{2n^2}}{\left(R+\sqrt{2n}\right)^{2n^2-2n} \sum_{k=0}^{n-1} \frac{\pi^k}{k!} \left(R+\sqrt{2k}\right)^{2k}}.$$

Proof. Start with $\mathcal{C} = \{X_1\}$, for any $X_1 \in B_n(R)$, and apply the following procedure:

1. If there exists $A \in B_n(R)$ such that $A \notin \bigcup_{X \in \mathcal{C}} F_n(X, R)$, add A to $\mathcal{C}$, otherwise stop.

2. Repeat step 1 until you have to stop.

This process will stop in a finite number of step, since $B_n(R)$ contains only a finite number of elements. By construction, the resulting set $\mathcal{C}$ is such that

$\|X\| \leq R$, for all $X \in \mathcal{C}$, and $|\det(X-Y)| \geq 1$, for all distinct $X, Y \in \mathcal{C}$. Moreover, at this point $\mathcal{C}$ is such that

$$B_n(R) = \bigcup_{X \in \mathcal{C}} F_n(X, R),$$

because otherwise we could have continued with the process.
We then obtain that

$$b_n(R) = \#B_n(R) = \# \bigcup_{X \in \mathcal{C}} F_n(X, R) \leq \sum_{X \in \mathcal{C}} \#F_n(X, R).$$

Define

$$B'_n(R) := \{X = (\mathbf{x_1}| \dots |\mathbf{x_n}) \in \mathbb{M}_n(\mathbb{W}) : \|\mathbf{x_j}\| \leq R, \ \forall j = 1, \dots, n\}.$$

Note that $B_n(R) \subset B'_n(R)$. For $X \in B'_n(R)$, define also

$$F'_n(X, R) := \{A \in B'_n(R) : \det(A - X) = 0\}.$$

Since $B_n(R) \subset B'_n(R)$, clearly $F_n(X, R) \subset F'_n(X, R)$ and so $\#F_n(X, R) \leq \#F'_n(X, R)$, for all $X \in B'_n(R)$.
Let $A = (\mathbf{a_1}| \dots |\mathbf{a_n}), X = (\mathbf{x_1}| \dots |\mathbf{x_n}) \in \mathbb{M}_n(\mathbb{W})$ and note that

$$\begin{aligned} \det(A - X) = 0 \quad \Leftrightarrow \quad & \mathbf{a_1} - \mathbf{x_1} = \mathbf{0} \text{ or } \exists\, k \in \{2, \dots, n\} \text{ such that} \\ & \mathbf{a_k} - \mathbf{x_k} \text{ is a linear combination } (LC) \text{ of } \mathbf{a_1} - \\ & \mathbf{x_1}, \dots, \mathbf{a_{k-1}} - \mathbf{x_{k-1}}. \end{aligned}$$

Therefore

$$\begin{aligned} & F'_n(X, R) = \{A = (\mathbf{a_1}| \dots |\mathbf{a_n}) \in B'_n(R) : \mathbf{a_1} = \mathbf{x_1}\} \cup \\ & \bigcup_{k=2}^{n} \{A = (\mathbf{a_1}| \dots |\mathbf{a_n}) \in B'_n(R) : \mathbf{a_k} - \mathbf{x_k} = LC(\mathbf{a_1} - \mathbf{x_1}, \dots, \mathbf{a_{k-1}} - \mathbf{x_{k-1}})\}. \end{aligned}$$

We now have to count how many matrices there are in these sets. For this purpose, we have to separate the two cases $\mathbb{W} = \mathbb{Z}$ and $\mathbb{W} = \mathbb{Z}[i]$.
Let us begin with $\mathbb{W} = \mathbb{Z}$.
We have that

$$\#\{A = (\mathbf{a_1}| \dots |\mathbf{a_n}) \in B'_n(R) : \mathbf{a_1} = \mathbf{x_1}\} = 1 \cdot \beta_n(R)^{n-1}$$

since for $\mathbf{a_1}$ we have only one possible choice and for the other $n-1$ columns we have $\beta_n(R)$ possible choices each, since $\mathbf{a_k} \in \mathbb{Z}^n \cap \mathbb{B}_n(R)$, for all $k =$

$1, \dots, n$.
For $k \in \{2, \dots, n\}$

$$\#\{A = (\mathbf{a_1}|\dots|\mathbf{a_n}) \in B'_n(R) : \mathbf{a_k} - \mathbf{x_k} = LC(\mathbf{a_1} - \mathbf{x_1}, \dots, \mathbf{a_{k-1}} - \mathbf{x_{k-1}})\}$$
$$\leq \beta_k(R)\beta_n(R)^{n-1}.$$

In fact, for column $\mathbf{a_k}$ we have at most $\beta_k(R)$ possibilities, thanks to proposition 5.4, where here $\mathbf{p} = \mathbf{x_k}$ and $\mathbf{v_j} = \mathbf{a_j} - \mathbf{x_j}$, for $j = 1, \dots, k-1$. For the others columns of A we have no restriction, so we have $\beta_n(R)$ for each of them, as before.
Therefore

$$\#F'_n(X, R) \leq \beta_n(R)^{n-1} \left(1 + \sum_{k=1}^{n-1} \beta_k(R)\right).$$

Exactly in the same way, for $\mathbb{W} = \mathbb{Z}[i]$ we have that

$$\#F'_n(X, R) \leq \beta_{2n}(R)^{n-1} \left(1 + \sum_{k=1}^{n-1} \beta_{2k}(R)\right),$$

since $\mathbf{a_k} \in \mathbb{Z}^{2n} \cap \mathbb{B}_{2n}(R)$, for all $k = 1, \dots, n$, because $\mathbf{a_k} \in \mathbb{Z}[i]^n \cong \mathbb{Z}^{2n}$ and $\|\mathbf{a_k}\| \leq R$.
Resuming, we have that

$$b_n(R) \leq \sum_{X \in \mathcal{C}} \#F'_n(X, R) \leq \begin{cases} \#\mathcal{C} \cdot \beta_n(R)^{n-1} \left(1 + \sum_{k=1}^{n-1} \beta_k(R)\right), & \mathbb{W} = \mathbb{Z} \\ \#\mathcal{C} \cdot \beta_{2n}(R)^{n-1} \left(1 + \sum_{k=1}^{n-1} \beta_{2k}(R)\right), & \mathbb{W} = \mathbb{Z}[i]. \end{cases}$$

Remembering that $b_n(R) = \beta_{n^2}(R)$ if $\mathbb{W} = \mathbb{Z}$ and $b_n(R) = \beta_{2n^2}(R)$ if $\mathbb{W} = \mathbb{Z}[i]$, we then have that

$$\#\mathcal{C} \geq \begin{cases} \frac{\beta_{n^2}(R)}{\beta_n(R)^{n-1}\left(1+\sum_{k=1}^{n-1} \beta_k(R)\right)}, & \mathbb{W} = \mathbb{Z} \\ \frac{\beta_{2n^2}(R)}{\beta_{2n}(R)^{n-1}\left(1+\sum_{k=1}^{n-1} \beta_{2k}(R)\right)}, & \mathbb{W} = \mathbb{Z}[i]. \end{cases}$$

Thanks to corollary 5.3 these fractions can be bounded from below using only volumes of balls. In fact

$$\begin{cases} \beta_{n^2}(R) \geq \mathcal{V}\left(\mathbb{B}_{n^2}(R-n)\right) = \frac{\pi^{\frac{n^2}{2}}}{\Gamma\left(\frac{n^2}{2}+1\right)}(R-n)^{n^2} \\ \beta_{2n^2}(R) \geq \mathcal{V}\left(\mathbb{B}_{2n^2}\left(R-\sqrt{2}n\right)\right) = \frac{\pi^{n^2}}{(n^2)!}\left(R-\sqrt{2}n\right)^{2n^2} \end{cases}$$

and, for all $k = 1, \dots, n$,

$$\begin{cases} \beta_k(R) \leq \mathcal{V}\left(\mathbb{B}_k\left(R+\sqrt{k}\right)\right) = \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}+1\right)}\left(R+\sqrt{k}\right)^k \\ \beta_{2k}(R) \leq \mathcal{V}\left(\mathbb{B}_{2k}\left(R+\sqrt{2k}\right)\right) = \frac{\pi^k}{k!}\left(R+\sqrt{2k}\right)^{2k}. \end{cases}$$

Putting everything together we complete the proof. □

Theorem 5.6. *Let L be a subfield of $\mathbb{F}$, $n \in \mathbb{N}^+$ and $r \in \mathbb{R}^+$ such that $\sqrt[n]{r} \in L$. There exists then a set $\mathcal{C} \subset \mathbb{M}_n(L)$ with $\|X\| \leq 1$, for all $X \in \mathcal{C}$, and $|\det(X-Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$, such that*

- *if $\mathbb{F} = \mathbb{R}$,*

$$\#\mathcal{C} \geq \pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n}{2}+1\right)^{n-1}}{\Gamma\left(\frac{n^2}{2}+1\right)} \frac{\left(\frac{1}{\sqrt[n]{r}} - n\right)^{n^2}}{\left(\frac{1}{\sqrt[n]{r}} + \sqrt{n}\right)^{n^2-n} \sum_{k=0}^{n-1} \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}+1\right)} \left(\frac{1}{\sqrt[n]{r}} + \sqrt{k}\right)^k},$$

- *if $\mathbb{F} = \mathbb{C}$,*

$$\#\mathcal{C} \geq \pi^n \frac{(n!)^{n-1}}{(n^2)!} \frac{\left(\frac{1}{\sqrt[n]{r}} - \sqrt{2}n\right)^{2n^2}}{\left(\frac{1}{\sqrt[n]{r}} + \sqrt{2n}\right)^{2n^2-2n} \sum_{k=0}^{n-1} \frac{\pi^k}{k!} \left(\frac{1}{\sqrt[n]{r}} + \sqrt{2k}\right)^{2k}}.$$

Proof. Again $\mathbb{W}$ will represent either $\mathbb{Z}$ or $\mathbb{Z}[i]$, depending if $\mathbb{F}$ represents $\mathbb{R}$ or $\mathbb{C}$ respectively.
Letting $R = \frac{1}{\sqrt[n]{r}}$, thanks to theorem 5.5 we know that there exists a set $\mathcal{C}' \subset \mathbb{M}_n(\mathbb{W})$ with $\|X\| \leq 1/\sqrt[n]{r}$, for all $X \in \mathcal{C}'$, and $|\det(X-Y)| \geq 1$, for all distinct $X, Y \in \mathcal{C}$, such that

- if $\mathbb{W} = \mathbb{Z}$ (i.e., if $\mathbb{F} = \mathbb{R}$),

$$\#\mathcal{C}' \geq \pi^{\frac{n}{2}} \frac{\Gamma\left(\frac{n}{2}+1\right)^{n-1}}{\Gamma\left(\frac{n^2}{2}+1\right)} \frac{\left(\frac{1}{\sqrt[n]{r}} - n\right)^{n^2}}{\left(\frac{1}{\sqrt[n]{r}} + \sqrt{n}\right)^{n^2-n} \sum_{k=0}^{n-1} \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}+1\right)} \left(\frac{1}{\sqrt[n]{r}} + \sqrt{k}\right)^k},$$

- if $\mathbb{W} = \mathbb{Z}[i]$ (i.e., if $\mathbb{F} = \mathbb{C}$),

$$\#\mathcal{C}' \geq \pi^n \frac{(n!)^{n-1}}{(n^2)!} \frac{\left(\frac{1}{\sqrt[n]{r}} - \sqrt{2}n\right)^{2n^2}}{\left(\frac{1}{\sqrt[n]{r}} + \sqrt{2n}\right)^{2n^2-2n} \sum_{k=0}^{n-1} \frac{\pi^k}{k!} \left(\frac{1}{\sqrt[n]{r}} + \sqrt{2k}\right)^{2k}}.$$

Letting $\mathcal{C} := \sqrt[n]{r}\mathcal{C}'$, we have that $\mathcal{C} \subset \mathbb{M}_n(L)$, because $\sqrt[n]{r} \in L$ and $\mathbb{W} \subset L$. Moreover, for all $X \in \mathcal{C}$

$$\|X\| = \|\sqrt[n]{r}X'\| = \sqrt[n]{r}\|X'\| \leq \sqrt[n]{r}\frac{1}{\sqrt[n]{r}} = 1,$$

and, for all distinct $X, Y \in \mathcal{C}$,

$$|\det(X-Y)| = |\det(\sqrt[n]{r}(X'-Y'))| = r|\det(X'-Y')| \geq r \cdot 1 = r.$$

Since $\#\mathcal{C} = \#\mathcal{C}'$, the theorem is proved. □

We give here the lower bound for the special case $n = 2$.

Example 5.7. Let L be a subfield of $\mathbb{F}$ and $r \in \mathbb{R}^+$ such that $\sqrt{r} \in L$. There exists then a set $\mathcal{C} \subset \mathbb{M}_2(L)$ with $\|X\| \leq 1$, for all $X \in \mathcal{C}$, and $|\det(X - Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$, such that

- if $\mathbb{F} = \mathbb{R}$,
$$\#\mathcal{C} \geq \frac{\pi}{2} \frac{\left(\frac{1}{\sqrt{r}} - 2\right)^4}{\left(\frac{1}{\sqrt{r}} + \sqrt{2}\right)^2 \left(\frac{2}{\sqrt{r}} + 3\right)},$$
- if $\mathbb{F} = \mathbb{C}$,
$$\#\mathcal{C} \geq \frac{\pi^2}{12} \frac{\left(\frac{1}{\sqrt{r}} - 2\sqrt{2}\right)^8}{\left(\frac{1}{\sqrt{r}} + 2\right)^4 \left(\pi \left(\frac{1}{\sqrt{r}} + \sqrt{2}\right)^2 + 1\right)}.$$

Chapter 6

An example: the Alamouti code

In 1998 the electrical engineer Siavash M. Alamouti proposed in his paper [7] the following space-time code for two transmitting antennas:

$$\mathcal{C}_A = \left\{ \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix} : (z,w) \in A \right\} \subset \mathbb{M}_2(\mathbb{C}),$$

where A is any subset of $\mathbb{C}^2$. This kind of code has become very popular because of its simplicity and of its good performance.
Letting $X, X' \in \mathcal{C}_A$ we have that

$$\begin{aligned} \det(X-X') = \begin{vmatrix} z-z' & -(\overline{w}-\overline{w'}) \\ w-w' & \overline{z}-\overline{z'} \end{vmatrix} &= (z-z')(\overline{z-z'}) + (w-w')(\overline{w-w'}) \\ &= |z-z'|^2 + |w-w'|^2 \geq 0 \end{aligned}$$

and it is zero if and only if $z = z'$ and $w = w'$, i.e., when $X = X'$. Moreover

$$\|X\|^2 = 2\left(|z|^2 + |w|^2\right).$$

The non-trivial question that follows is how to choose A in order to have a code of maximal size that satisfies the energy constraint and the determinant criterion. Here we want to give an example of an Alamouti code in order to test the bounds we proved in chapters 4 and 5.
We then fix $n = 2$ and $L = \mathbb{C}$. Let

$$C_2\left(\sqrt{2}/4\right) := \left\{ z \in \mathbb{C} : |\Re(z)|, |\Im(z)| \leq \sqrt{2}/4 \right\}$$

and define the Alamouti code

$$\mathcal{C}_A = \left\{ \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix} : z, w \in \sqrt{r}\mathbb{Z}[i] \cap C_2\left(\sqrt{2}/4\right) \right\},$$

where we set $A = \left(\sqrt{r}\mathbb{Z}[i] \cap C_2\left(\sqrt{2}/4\right)\right)^2$.
Let $X \in \mathcal{C}_A$. We have that

$$\begin{aligned} \|X\|^2 = 2\left(|z|^2 + |w|^2\right) &= 2\left(\Re(z)^2 + \Im(z)^2 + \Re(w)^2 + \Im(w)^2\right) && (6.1) \\ &\leq 2\left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) = 1, && (6.2) \end{aligned}$$

and so $\|X\| \leq 1$, for all $X \in \mathcal{C}_A$.
Let also $X, X' \in \mathcal{C}_A$, with $X \neq X'$. The fact that X and X' are distinct means that either $z \neq z'$ or $w \neq w'$, which implies that either $|z - z'| \geq \sqrt{r}$ or $|w - w'| \geq \sqrt{r}$. Hence

$$|\det(X - X')| = |z - z'|^2 + |w - w'|^2 \geq r,$$

for all distinct $X, X' \in \mathcal{C}_A$.
Therefore $\mathcal{C}_A$ is a code that satisfies the energy constraint and the determinant criterion. Its cardinality is easily computed, since both z and w lie in the same square grid in the complex plane:

$$\#\mathcal{C}_A = \left(\left(\left\lfloor \frac{2 \cdot \frac{\sqrt{2}}{4}}{\sqrt{r}} \right\rfloor + 1\right)^2\right)^2 = \left(\left\lfloor \frac{1}{\sqrt{2r}} \right\rfloor + 1\right)^4.$$

Let us now see how this fits with the upper and lower bound we proved.

Example 4.4 says that, for $n = 2$ and $L = \mathbb{C}$, for any code $\mathcal{C} \subset \mathbb{M}_2(\mathbb{C})$ that satisfies the energy constraint and the determinant criterion,

$$\#\mathcal{C} \leq \frac{8}{3}\left(\sqrt{\frac{2}{r}} + 1\right)^8.$$

Indeed we have that

$$\#\mathcal{C}_A = \left(\left\lfloor \frac{1}{\sqrt{2r}} \right\rfloor + 1\right)^4 \leq \left(\frac{1}{\sqrt{2r}} + 1\right)^4 < \frac{8}{3}\left(\sqrt{\frac{2}{r}} + 1\right)^8.$$

On the other hand, in example 5.7, for $n = 2$ and $L = \mathbb{C}$, we found that there exists a code $\mathcal{C} \subset \mathbb{M}_n(\mathbb{C})$ that satisfies the energy constraint and the determinant criterion such that

$$\#\mathcal{C} \geq \frac{\pi^2}{12} \frac{\left(\frac{1}{\sqrt{r}} - 2\sqrt{2}\right)^8}{\left(\frac{1}{\sqrt{r}} + 2\right)^4 \left(\pi\left(\frac{1}{\sqrt{r}} + \sqrt{2}\right)^2 + 1\right)}.$$

Again, we have

$$\begin{aligned}
\frac{\pi^2}{12}\frac{\left(\frac{1}{\sqrt{r}}-2\sqrt{2}\right)^8}{\left(\frac{1}{\sqrt{r}}+2\right)^4\left(\pi\left(\frac{1}{\sqrt{r}}+\sqrt{2}\right)^2+1\right)} &< \frac{\pi^2}{12}\frac{\left(\frac{1}{\sqrt{r}}\right)^8}{\left(\frac{1}{\sqrt{r}}\right)^4\left(\pi\left(\frac{1}{\sqrt{r}}\right)^2\right)} = \frac{\pi}{12}\left(\frac{1}{\sqrt{r}}\right)^2 \\
&< \frac{1}{2}\left(\frac{1}{\sqrt{r}}\right)^2 = \left(\frac{1}{\sqrt{2r}}\right)^2 \\
&\leq \left(\left\lfloor\frac{1}{\sqrt{2r}}\right\rfloor+1\right)^2 \\
&\leq \left(\left\lfloor\frac{1}{\sqrt{2r}}\right\rfloor+1\right)^4 = \#\mathcal{C}_A.
\end{aligned}$$

So the size of our code is better than the lower bound.

The Alamouti code is closely related to the theory of quaternion algebras. For further applications of the theory of central simple algebras to the design of MIMO codes, see, e.g., [8], [9], [10].

Conclusion

While the use of multiple receiving antennas in wireless communication is a much more old topic, the theoretical engineering framework on MIMO channels rose only in the late nineties (see e.g., [1], [12], [13], [14], [15]). Very quickly this area of research turned out to be very attractive also for mathematics, because of the challenges that MIMO communication was bringing. Hence, since the beginning of the new century, there have been several mathematical studies on this topic (see e.g., [8], [9], [10], [16], [17]).
A novel feature of our approach is the precise mathematical question which is deduced from the engineering problem in chapter 2: *Find the largest size of a space-time code over a subfield L of the complex numbers satisfying the energy constraint and the determinant criterion.* This question is analogous to the main problem of classical error-correcting coding theory (see chapter 3); here L replaces the choice of a finite field (the alphabet) in classical coding theory, the energy constraint is analogous to choosing the length of a classical error-correcting code and the determinant criterion is analogous to a lower bound on its minimal (Hamming) distance. We used this analogy to derive upper and lower bounds on the maximal size of space-time codes in chapter 4 and 5 respectively.

Our proofs are entirely geometric; for this reason our bounds are primarily of interest in the cases where L is the field of real or complex numbers. It would be then interesting to try to sharpen them for other fields L. If we had more time, it would also have been fascinating to deepen in the subject of division algebras (see e.g., [8], [9], [10], [11]), which give the tools of constructing interesting space-time codes not only for $n = 2$ - as we did very briefly in chapter 6 -, but also for higher dimensions.
Another interesting line of research is to try to sharpen our lower bound by using the theory of random matrices. In the proof of theorem 5.5 the key point is the estimate of the number of matrices A that belong to the ball of radius R centered at the origin with integer entries and satisfying the inequality $\det(A - X) = 0$, where X is a fixed $n \times n$ matrix. Allowing A to

have real or complex entries, this is analogous to computing the volume of the set such that $\|A\| \leq R$ and satisfying $|\det(A - X)| < 1$, for some fixed $X \in \mathbb{M}_n(\mathbb{C})$. Conjecturally, for fixed R and n, the largest volume occurs if X is the zero matrix. In this case there are chances that one may be able to use the theory of random matrices and their eigenvalue distribution to give a stronger upper bound on this volume, which would improve the lower bound on the maximal size of a code. Though, this idea needs to be inspected much more in order to tell if it can lead to some concrete result.

Acknowledgments

I owe the spirit and the main ideas of this work to Professor Zinovy Reichstein, whom I want to thank also for his kindness and availability to help me whenever I needed it. I would also like to thank Professors Brian Marcus and Lutz Lampe who helped me with the engineering problem. Finally, I am very grateful to Professor Eva Bayer Fluckiger, who agreed to supervise my Master's project and gave me the opportunity to live this great experience at the University of British Columbia in Vancouver.

Appendix

In this section we just show that if we want to test the upper and the lower bound that we proved in chapter 4 and 5, only a certain interval of r is relevant.

Lemma 6.1. *Let $n \in \mathbb{N}^+$, $0 \leq R \in \mathbb{R}$ and $X, Y \in \mathbb{M}_n(\mathbb{C})$ two matrices such that $\|X\|, \|Y\| \leq R$. We then have that*

$$|\det(X-Y)| \leq \left(\frac{2R}{\sqrt{n}}\right)^n.$$

Proof. Let $\mathbf{x_1}, \ldots, \mathbf{x_n}$ and $\mathbf{y_1}, \ldots, \mathbf{y_n}$ be the columns of X and Y respectively, i.e., $X = (\mathbf{x_1}|\ldots|\mathbf{x_n})$ and $Y = (\mathbf{y_1}|\ldots|\mathbf{y_n})$.
Applying theorem 1.7 and lemma 1.6 we have that

$$|\det(X-Y)| \leq \prod_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\| = \left[\left(\prod_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\|\right)^{1/n}\right]^n \leq \left(\frac{\sum_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\|}{n}\right)^n.$$

Let $\mathbf{v} = (\|\mathbf{x_1}-\mathbf{y_1}\|, \ldots, \|\mathbf{x_n}-\mathbf{y_n}\|)^T$, $\mathbf{1_n} = (1, \ldots, 1)^T \in \mathbb{R}^n$ and denote by $\langle \cdot, \cdot \rangle$ the usual inner product of the $\mathbb{R}$-vector space $\mathbb{R}^n$. Using Cauchy-Schwarz inequality, we know that $\langle \mathbf{v}, \mathbf{1_n} \rangle \leq \|\mathbf{v}\|\|\mathbf{1_n}\|$, i.e.,

$$\sum_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\| \leq \sqrt{\sum_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\|^2}\sqrt{n}.$$

We also have that

$$\sum_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\|^2 = \|X-Y\|^2 \leq (\|X\|+\|Y\|)^2 \leq (R+R)^2 = (2R)^2.$$

So $\sqrt{\sum_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\|^2} \leq 2R$.
Therefore

$$|\det(X-Y)| \leq \left(\frac{\sum_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\|}{n}\right)^n \leq \left(\frac{\sqrt{\sum_{k=1}^{n} \|\mathbf{x_k}-\mathbf{y_k}\|^2}\sqrt{n}}{n}\right)^n \leq \left(\frac{2R}{\sqrt{n}}\right)^n,$$

as claimed. □

Remark 6.2. The bound that we obtained in lemma 6.1 is tight for any $n \in \mathbb{N}^+$ and $R \geq 0$. Indeed, choose n mutually orthogonal column vectors $\mathbf{a_1}, \ldots, \mathbf{a_n} \in \mathbb{C}^n$ of length $R/\sqrt{n}$ and set $A = (\mathbf{a_1}|\cdots|\mathbf{a_n})$ and $B = -A$. Clearly $\|A\| = \|B\| \leq R$ and then, thanks to lemma 1.4,

$$|\det(A-B)| = |\det(2A)| = 2^n|\det(A)| = 2^n \prod_{j=1}^{n} \|\mathbf{a_j}\| = \left(\frac{2R}{\sqrt{n}}\right)^n.$$

Corollary 6.3. *Let L be a subfield of $\mathbb{C}$, $n \in \mathbb{N}^+$ and $r \in \mathbb{R}^+$. Let $\mathcal{C} \subset \mathbb{M}_n(L)$ be a set such that $\|X\| \leq 1$, for all $X \in \mathcal{C}$, and $|\det(X-Y)| \geq r$, for all distinct $X, Y \in \mathcal{C}$.*
If $\mathcal{C}$ has more than one element, then $r \leq \left(\frac{2}{\sqrt{n}}\right)^n$ and in particular

1. *$r \leq 2$ and*
2. *if $n \geq 4$, $r \leq 1$.*

Proof. By hypothesis, we know that $\mathcal{C}$ contains at least two distinct elements X and Y. Thanks to lemma 6.1 for $R = 1$, we have that $r \leq |\det(X-Y)| \leq (2/\sqrt{n})^n$ and hence $r \leq (2/\sqrt{n})^n$. Since $(2/\sqrt{n})^n \leq 2$, for all $n \in \mathbb{N}^+$, $r \leq 2$, and if $n \geq 4$, then $(2/\sqrt{n})^n \leq 1$ and thus $r \leq 1$. □

Bibliography

[1] Vahid Tarokh, Nambi Seshadri, A.R. Calderbank. *Space-time codes for high data rate wireless communication: performance criterion and code construction*, IEEE Transaction on Information Theory, vol. 44, pp. 744-765, March 1998.

[2] Claude Oesteges, Bruno Clerckx. *MIMO wireless communications: From real-world propagation to space-time design*. Elsevier, Great Britain, 2007.

[3] Hamid Jafarkhani. *Space-time coding: theory and practice*. Cambridge University Press, New York, 2005.

[4] Raymond Hill. *A first course in coding theory*. The Clarendon Press, Oxford University Press, New York, 1986.

[5] F. J. MacWilliams, N. J. Sloane. *The theory of error-correcting codes*, V. I-II. North-Holland Publishing Co., Amsterdam - New York - Oxford, 1977.

[6] Tom M. Apostol. *Calculus, Volume II: Multi-variable calculus and linear algebra, with applications to differential equations and probability*. Second edition Blaisdell Publishing Co., Ginn and Co., Waltham, Mass.-Toronto, Ont.-London, 1969.

[7] Siavash M. Alamouti. *A simple transmit diversity technique for wireless communications*. IEEE Journal on Select Areas in Communications, vol. 16, pp. 1451-1458, October 1998.

[8] Grégory Berhuy, Frédérique Oggier. *Introduction to central simple algebras and their applications to wireless communication*.

[9] B. A. Sethuraman, B. S. Rajan, V. Shashidhar. *Full-diversity, high-rate space-time block codes from division algebras*. IEEE Transactions on Information Theory, vol. 49, October 2003.

[10] B. A. Sethuraman, Petros Elia, P. Vijaykumar. *Perfect space-time codes for any number of antennas.* IEEE Transactions on Information Theory, vol. 53, pp. 3853-3868, November 2007.

[11] B.A. Sethuraman. *Division algebras and wireless communication.* arXiv:0906.0997v2 [math.RA], June 2009.

[12] N. Seshadri, V. Tarokh, A. R. Calderbank. *Space-time codes for wireless communication: code construction.* IEEE 47th Vehicular Technology Conference (VTC '97), pp. 637-641, Phoenix, Ariz, USA, May 1997.

[13] A. F. Naguib, V. Tarokh, N. Seshadri, A. R. Calderbank. *A space-time coding modem for high-data-rate wireless communications.* IEEE J. Select. Areas Commun., vol. 16, no. 8, pp. 1459-1478, 1998.

[14] G.J. Foschini, M.Gans. *On the limits of wireless communication in a fading environment when using multiple antennas.* Wireless Personal Communication, March 1998.

[15] J.-C. Guey, M. P. Fitz, M. R. Bell, W.-Y. Kuo. *Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels.* Proc. IEEE VTC '96, pp. 136-140, 1996.

[16] S. Galliou, J.-C. Belfiore. *A new family of full rate diverse space-time code based on Galois theory.* Proc. IEEE Int. Symp. Information Theory, Lausanne, Switzerland, Jun.-Jul. 2002, p. 419.

[17] M. O. Damen, A. Tewfik, J.-C. Belfiore. *A construction of a space-time code based on number theory.* IEEE Trans. Inf. Theory, vol. 48, no. 3, pp. 753-761, March 2002.