Network Coding and the Combinatorics of Codes

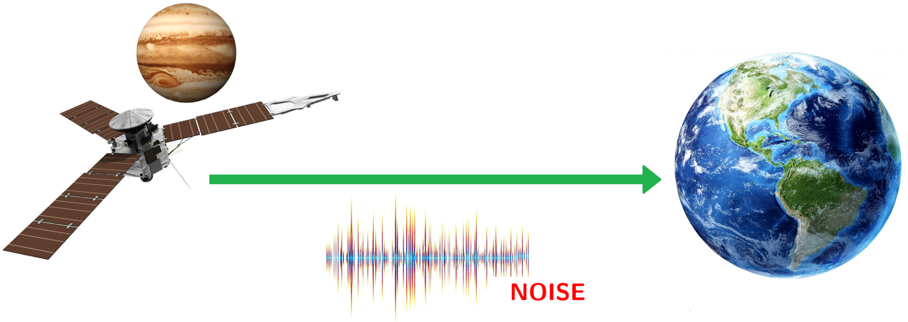
Alberto Ravagnani

UCD Algebra and Number Theory Seminar, **Oct. 2019**

What is coding theory?

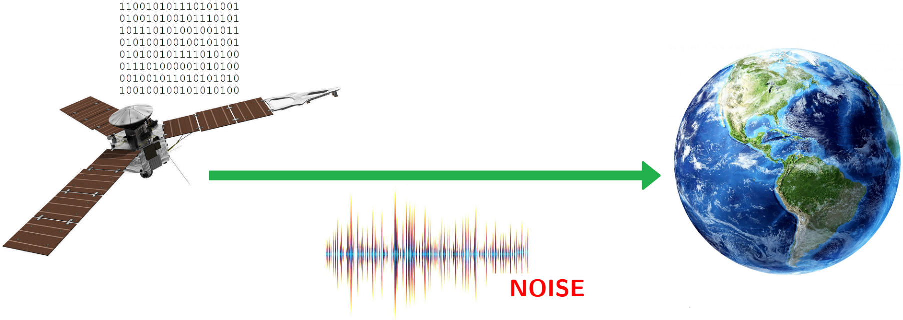
What is coding theory?

A **code** is a mathematical object that corrects the errors caused by a noise.

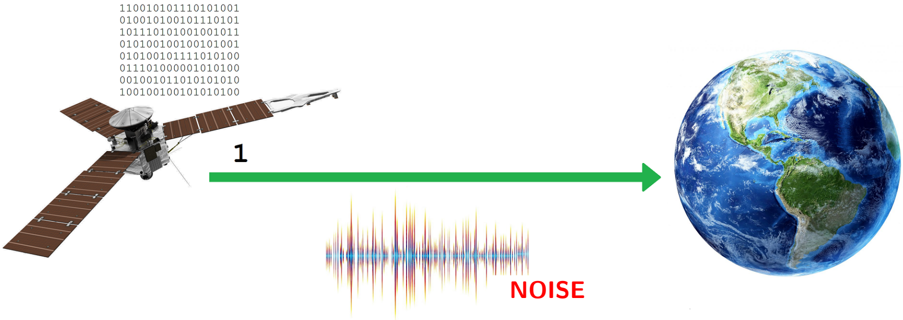


What is coding theory?

A **code** is a mathematical object that corrects the errors caused by a noise.

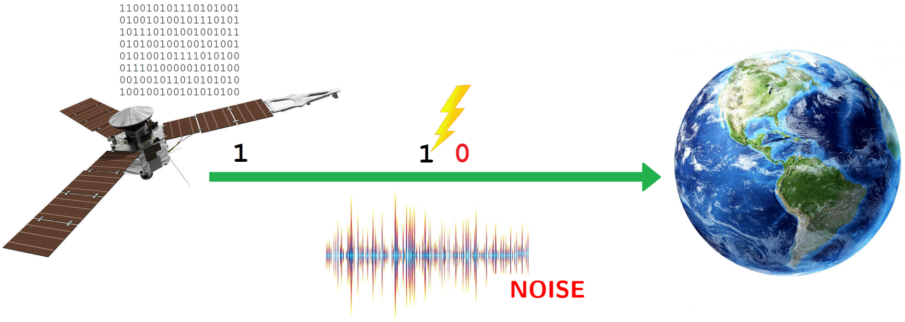


A **code** is a mathematical object that corrects the errors caused by a noise.



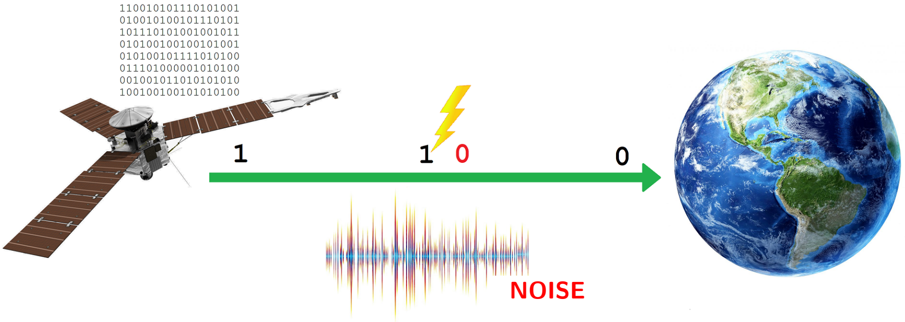
What is coding theory?

A **code** is a mathematical object that corrects the errors caused by a noise.



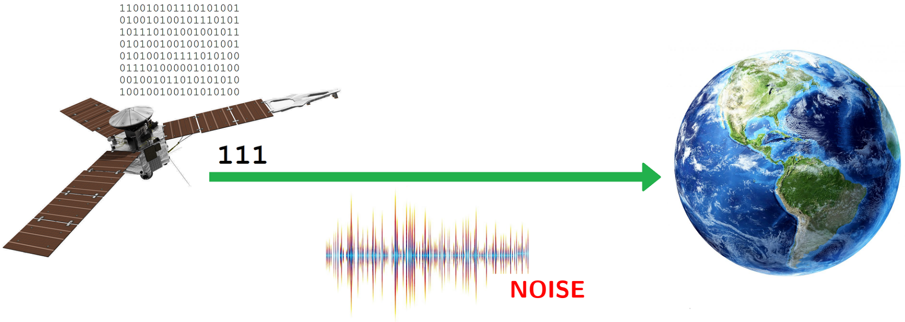
What is coding theory?

A **code** is a mathematical object that corrects the errors caused by a noise.



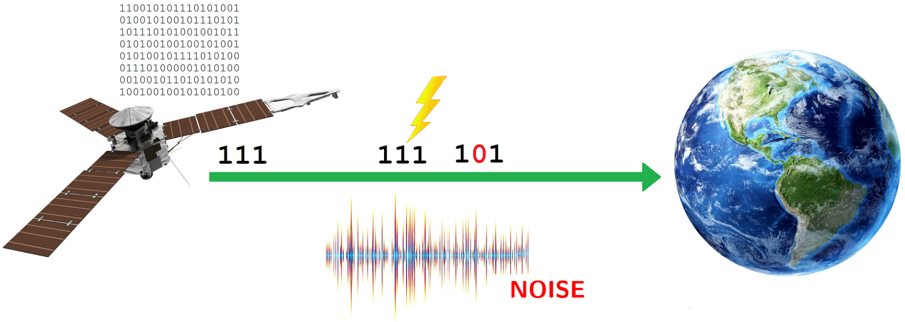
What is coding theory?

A **code** is a mathematical object that corrects the errors caused by a noise.



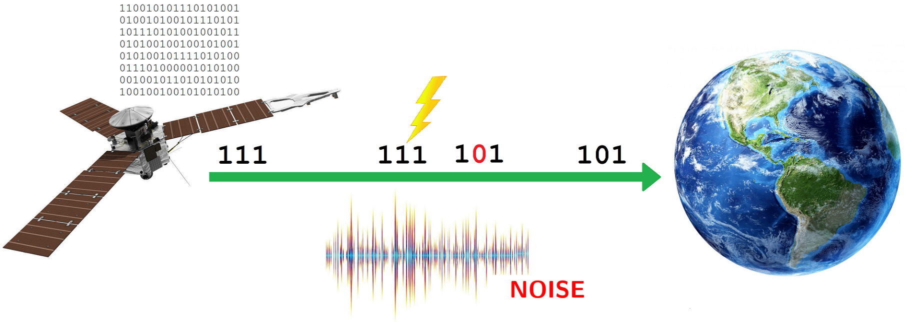
What is coding theory?

A **code** is a mathematical object that corrects the errors caused by a noise.



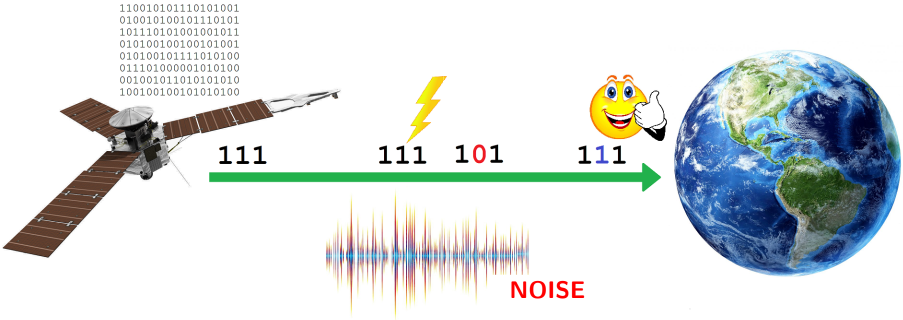
What is coding theory?

A **code** is a mathematical object that corrects the errors caused by a noise.



What is coding theory?

A **code** is a mathematical object that corrects the errors caused by a noise.



Error-correcting codes

Idea behind coding theory: add **redundancy**.

Error-correcting codes

Idea behind coding theory: add **redundancy**.

Encoder $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ injective linear map, $n \geq k$.

Error-correcting codes

Idea behind coding theory: add **redundancy**.

Encoder $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ injective linear map, $n \geq k$.

Example: binary 3-time repetition scheme

$E : \mathbb{F}_2 \to \mathbb{F}_2^3, \quad E(a) = (a, a, a)$ for all $a \in \mathbb{F}_2$.

Error-correcting codes

Idea behind coding theory: add **redundancy**.

Encoder $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ injective linear map, $n \geq k$.

Example: binary 3-time repetition scheme

$E : \mathbb{F}_2 \to \mathbb{F}_2^3, \quad E(a) = (a, a, a)$ for all $a \in \mathbb{F}_2$.

Note: the image of E is a k-dimensional subspace of $\mathbb{F}_q^n$.

Example [continued]

$E(\mathbb{F}_2) = \{(0,0,0),(1,1,1)\}$.

Error-correcting codes

Idea behind coding theory: add **redundancy**.

Encoder $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ injective linear map, $n \geq k$.

Example: binary 3-time repetition scheme

$E : \mathbb{F}_2 \to \mathbb{F}_2^3, \quad E(a) = (a, a, a)$ for all $a \in \mathbb{F}_2$.

<u>Note</u>: the image of E is a k-dimensional subspace of $\mathbb{F}_q^n$.

Example [continued]

$E(\mathbb{F}_2) = \{(0,0,0),(1,1,1)\}$.

Definition

A **code** is an $\mathbb{F}_q$-linear subspace $\mathscr{C} \leq \mathbb{F}_q^n$. Elements of $\mathscr{C}$: **codewords**.

(we often forget about E)

Error-correcting codes

In a good quality code $\mathscr{C} \leq \mathbb{F}_q^n$, vectors are "far apart"...

Definition

- The **Hamming distance** between vectors $x, y \in \mathbb{F}_q^n$ is $d_{\mathrm{H}}(x,y) = \#\{i \mid x_i \neq y_i\}$.

Error-correcting codes

In a good quality code $\mathscr{C} \leq \mathbb{F}_q^n$, vectors are "far apart"...

Definition

- The **Hamming distance** between vectors $x, y \in \mathbb{F}_q^n$ is $d_{\mathrm{H}}(x,y) = \#\{i \mid x_i \neq y_i\}$.
- The **Hamming weight** of a vector $x \in \mathbb{F}_q^n$ is $\omega_{\mathrm{H}}(x) = d_{\mathrm{H}}(x,0)$.
- The **minimum Hamming distance** of a code $\mathscr{C} \neq \{0\}$ is the integer

$$d_{\mathrm{H}}(\mathscr{C}) \;=\; \min\{d_{\mathrm{H}}(x,y) \mid x, y \in \mathscr{C}\ x \neq y\} \;=\; \min\{\omega_{\mathrm{H}}(x) \mid x \in \mathscr{C},\ x \neq 0\}.$$

Error-correcting codes

In a good quality code $\mathscr{C} \leq \mathbb{F}_q^n$, vectors are "far apart"...

Definition

- The **Hamming distance** between vectors $x, y \in \mathbb{F}_q^n$ is $d_{\mathrm{H}}(x,y) = \#\{i \mid x_i \neq y_i\}$.
- The **Hamming weight** of a vector $x \in \mathbb{F}_q^n$ is $\omega_{\mathrm{H}}(x) = d_{\mathrm{H}}(x,0)$.
- The **minimum Hamming distance** of a code $\mathscr{C} \neq \{0\}$ is the integer

$$d_{\mathrm{H}}(\mathscr{C}) \;=\; \min\{d_{\mathrm{H}}(x,y) \mid x,y \in \mathscr{C}\ x \neq y\} \;=\; \min\{\omega_{\mathrm{H}}(x) \mid x \in \mathscr{C},\, x \neq 0\}.$$

Note: a code $\mathscr{C}$ corrects up to $\lfloor (d-1)/2 \rfloor$ errors, where $d = d_{\mathrm{H}}(\mathscr{C})$.

Error-correcting codes

In a good quality code $\mathscr{C} \leq \mathbb{F}_q^n$, vectors are "far apart"...

Definition

- The **Hamming distance** between vectors $x, y \in \mathbb{F}_q^n$ is $d_{\mathrm{H}}(x,y) = \#\{i \mid x_i \neq y_i\}$.
- The **Hamming weight** of a vector $x \in \mathbb{F}_q^n$ is $\omega_{\mathrm{H}}(x) = d_{\mathrm{H}}(x,0)$.
- The **minimum Hamming distance** of a code $\mathscr{C} \neq \{0\}$ is the integer

$$d_{\mathrm{H}}(\mathscr{C}) \;=\; \min\{d_{\mathrm{H}}(x,y) \mid x,y \in \mathscr{C}\ x \neq y\} \;=\; \min\{\omega_{\mathrm{H}}(x) \mid x \in \mathscr{C},\ x \neq 0\}.$$

Note: a code $\mathscr{C}$ corrects up to $\lfloor (d-1)/2 \rfloor$ errors, where $d = d_{\mathrm{H}}(\mathscr{C})$.

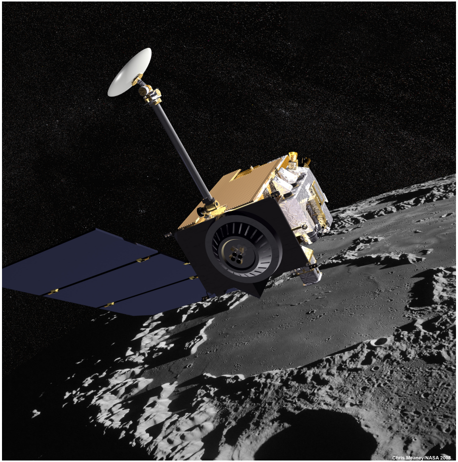
Theorem (Singleton, Komamiya)

Let $\mathscr{C} \leq \mathbb{F}_q^n$ be a non-zero code. Then $\dim(\mathscr{C}) \leq n - d_{\mathrm{H}}(\mathscr{C}) + 1$.

If $\mathscr{C}$ meets the bound with equality, then it is called an **MDS** code.

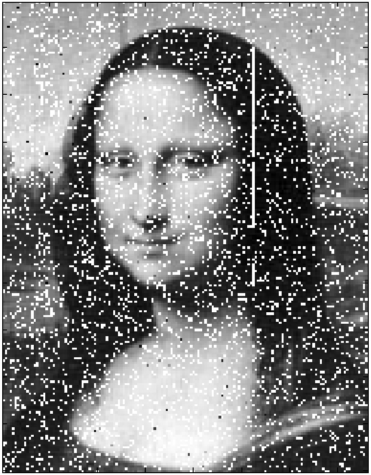
A concrete example

The LRO (Lunar Reconnaissance Orbiter) is taking pictures of the Moon...



A concrete example

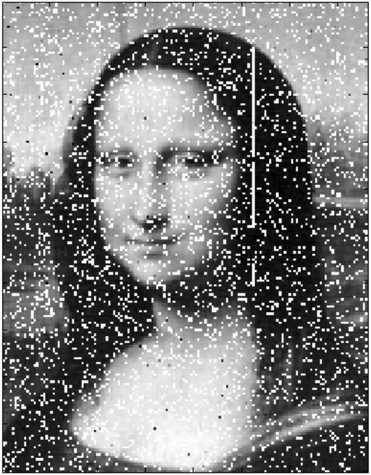
Test of quality of transmissions:



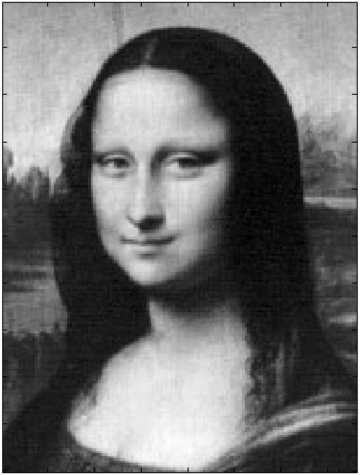
without coding

A concrete example

Test of quality of transmissions:



without coding



with coding

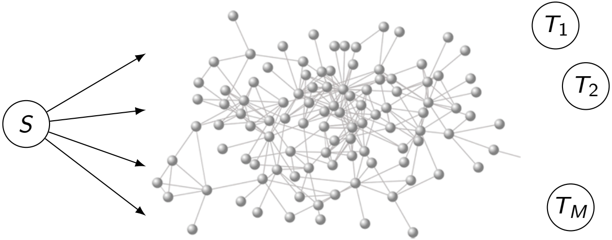
Network communication

Classical coding theory: **one** source of information, **one** terminal.

Network communication

Classical coding theory: **one** source of information, **one** terminal.

Network coding: **one/multiple** sources of information, **multiple** terminals.



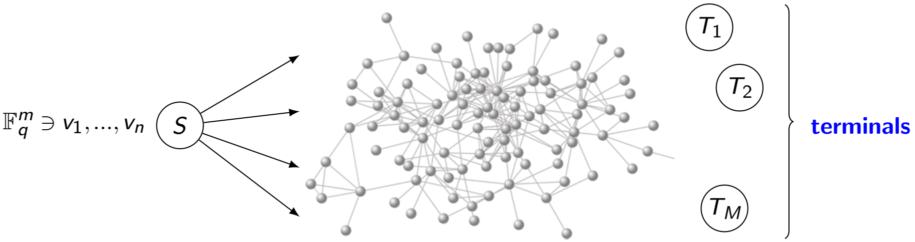
Applications: LTE (mobile phones), distributed storage, peer-to-peer, streaming,...

Network coding

Network coding: data transmission over (noisy/lossy) networks

Network coding

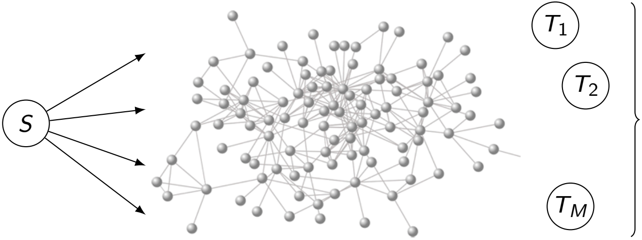
Network coding: data transmission over (noisy/lossy) networks



- One source S attempts to transmit messages $v_1, \ldots, v_n \in \mathbb{F}_q^m$.
- The terminals demand **all** the messages (multicast).

Network coding: data transmission over (noisy/lossy) networks

$\mathbb{F}_q^m \ni v_1, \dots, v_n$



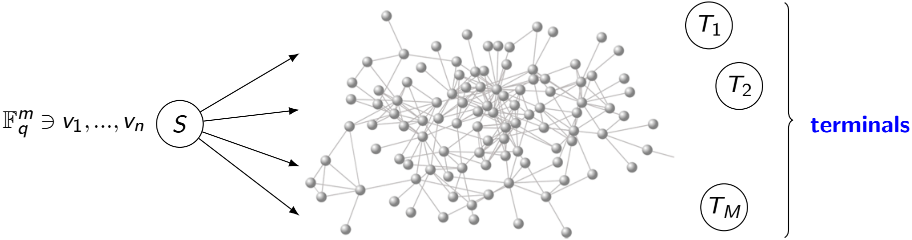
terminals

- One source S attempts to transmit messages $v_1, \dots, v_n \in \mathbb{F}_q^m$.
- The terminals demand **all** the messages (multicast).

What should the nodes do?

Network coding

Network coding: data transmission over (noisy/lossy) networks



- One source S attempts to transmit messages $v_1, \ldots, v_n \in \mathbb{F}_q^m$.
- The terminals demand **all** the messages (multicast).

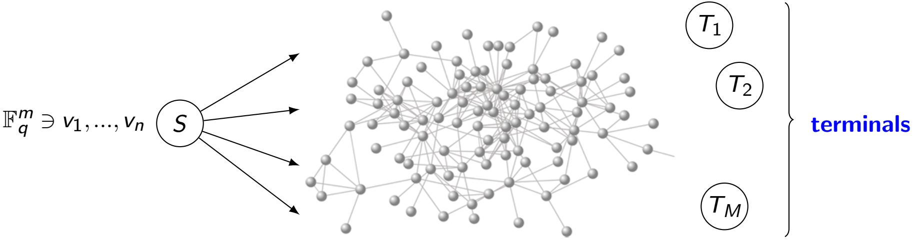
What should the nodes do?

Goal

Maximize the number of transmitted messages per channel use (**rate**).

Network coding

Network coding: data transmission over (noisy/lossy) networks



- One source S attempts to transmit messages $v_1, \dots, v_n \in \mathbb{F}_q^m$.
- The terminals demand **all** the messages (multicast).

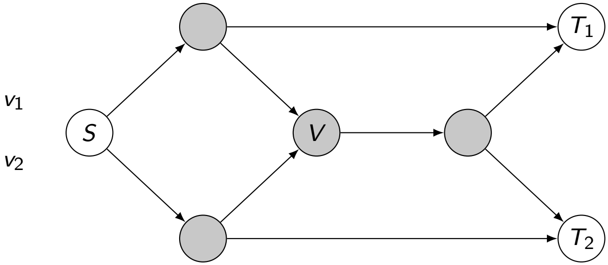
What should the nodes do?

Goal

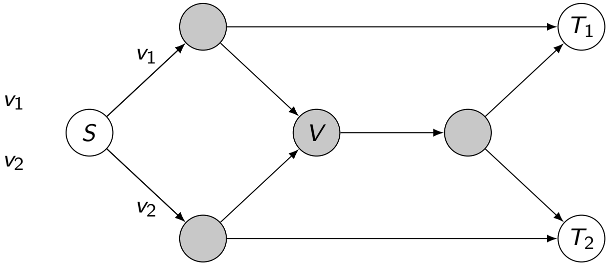
Maximize the number of transmitted messages per channel use (**rate**).

IDEA (Ahlswede-Cai-Li-Yeung 2000): allow the nodes to recombine packets.

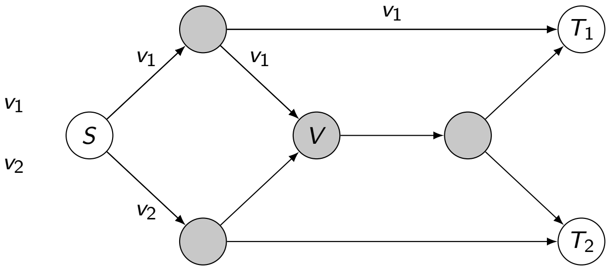
The "Butterfly" network



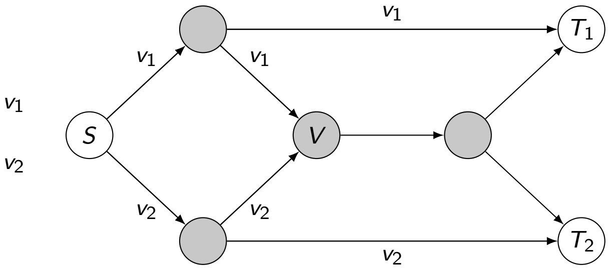
The “Butterfly” network



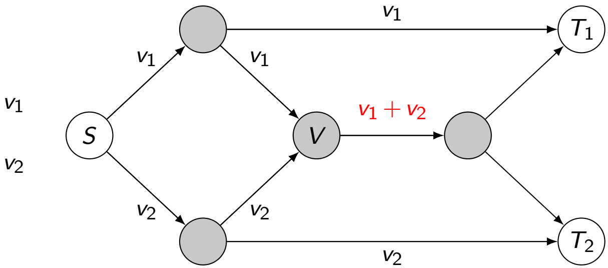
The “Butterfly” network



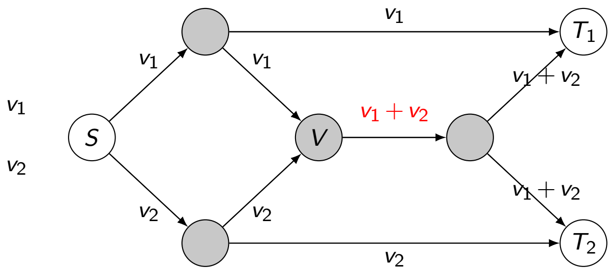
The "Butterfly" network



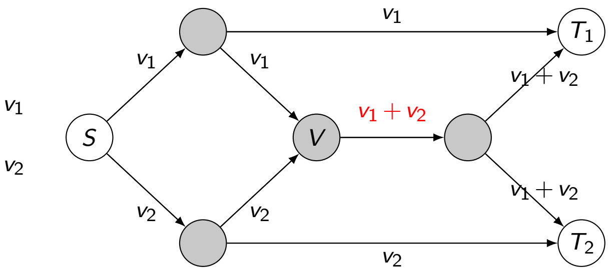
The “Butterfly” network



The "Butterfly" network



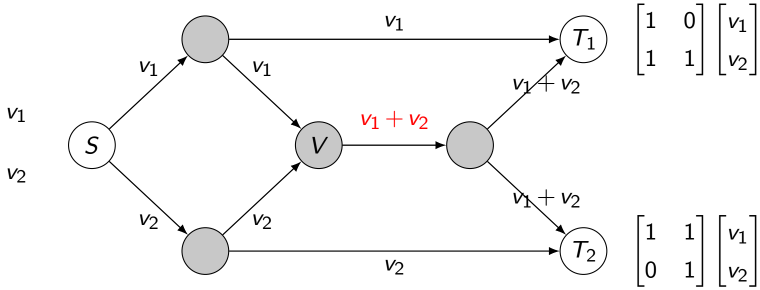
The "Butterfly" network



$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

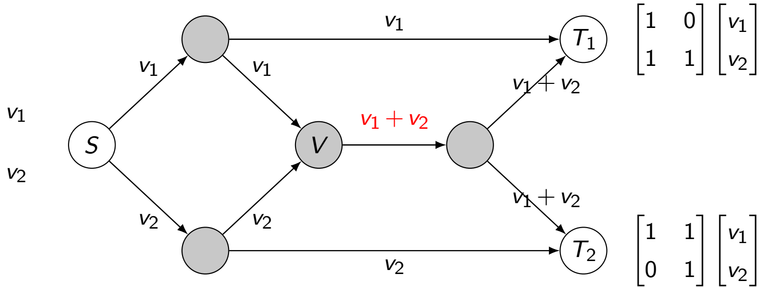
$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

The "Butterfly" network



Note: This strategy is better than routing.

The “Butterfly” network



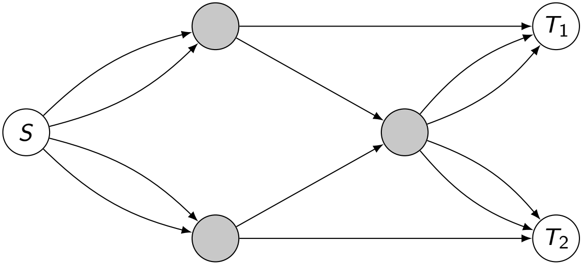
Note: This strategy is better than routing.

Theorem (Li-Yeung-Cai 2002, Koetter-Médard 2003)

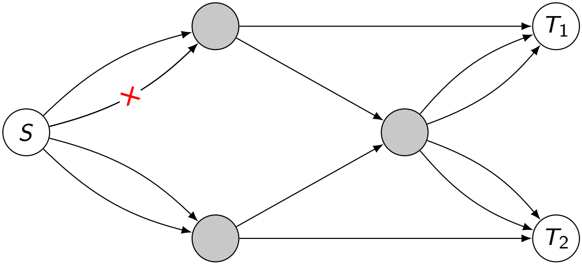
This strategy (**linear** network coding) applies to general networks and is capacity achieving (w.r. to certain models), provided that $q \gg 0$.

Also, efficient algorithms to design the network operations are known.

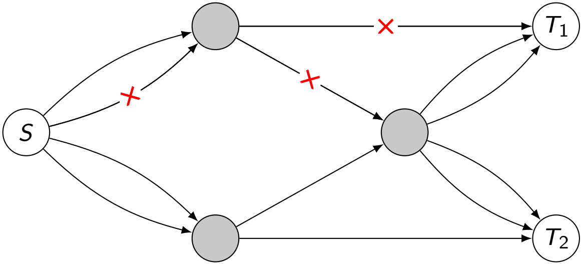
Error correction in networks



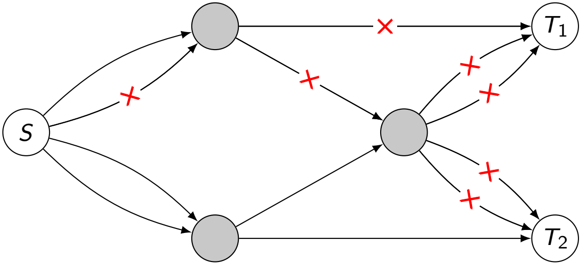
Error correction in networks



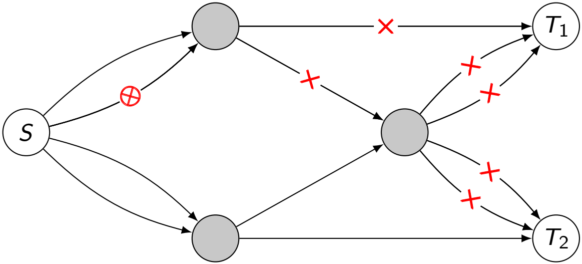
Error correction in networks



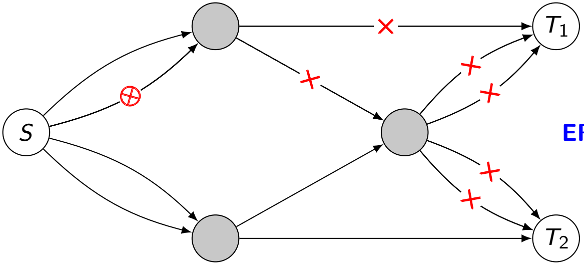
Error correction in networks



Error correction in networks

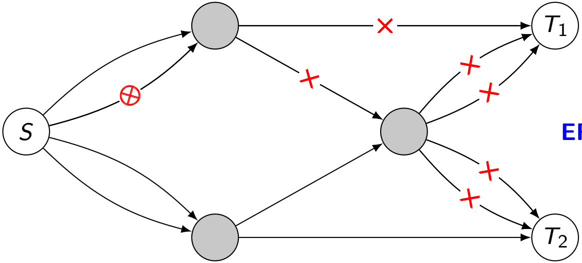


Error correction in networks



ERROR AMPLIFICATION

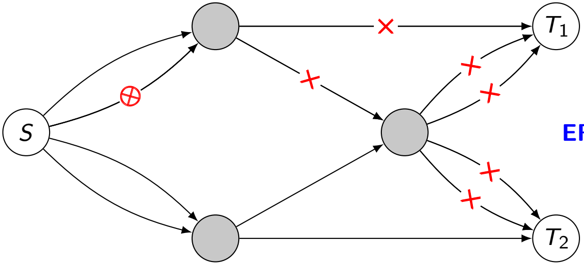
Error correction in networks



ERROR AMPLIFICATION

Natural solution: design the node operations carefully (decoding at intermediate nodes).

Error correction in networks



ERROR AMPLIFICATION

Natural solution: design the node operations carefully (decoding at intermediate nodes).

Other solution: use rank-metric codes.

Rank-metric codes

Definition

A **rank-metric code** is an $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n\times m}$. If $\mathscr{C} \neq \{0\}$, then its **minimum rank distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C}, X \neq 0\}.$$

Rank-metric codes

Definition

A **rank-metric code** is an $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n\times m}$. If $\mathscr{C} \neq \{0\}$, then its **minimum rank distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C},\, X \neq 0\}.$$

In standard scenarios, communication schemes based on rank-metric codes are:

(1) capacity-achieving (for $q \gg 0$)

(2) compatible with linear network coding

Rank-metric codes

Definition

A **rank-metric code** is an $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$. If $\mathscr{C} \neq \{0\}$, then its **minimum rank distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C},\, X \neq 0\}.$$

In standard scenarios, communication schemes based on rank-metric codes are:

(1) capacity-achieving (for $q \gg 0$)

(2) compatible with linear network coding

Remark 1: for some scenarios, there is no communication scheme based on rk-metric codes with both (1) and (2). E.g., geographically restricted errors, erasures, ...

Kschischang, R., *Adversarial Network Coding*, IEEE Trans. Inf. Th. 2018.

Rank-metric codes

Definition

A **rank-metric code** is an $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$. If $\mathscr{C} \neq \{0\}$, then its **minimum rank distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C},\, X \neq 0\}.$$

In standard scenarios, communication schemes based on rank-metric codes are:

(1) capacity-achieving (for $q \gg 0$)

(2) compatible with linear network coding

Remark 1: for some scenarios, there is no communication scheme based on rk-metric codes with both (1) and (2). E.g., geographically restricted errors, erasures, ...

Kschischang, R., *Adversarial Network Coding*, IEEE Trans. Inf. Th. 2018.

Remark 2: wireless networks are a very different story

Gorla, R., *An Algebraic Framework for End-to-End PLNC*, IEEE Trans. Inf. Th. 2018.

Rank-metric codes

Definition

A **rank-metric code** is an $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$. If $\mathscr{C} \neq \{0\}$, then its **minimum rank distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C},\, X \neq 0\}.$$

- Introduced and studied by Delsarte ('78) for combinatorial interest
- Re-discovered by Gabidulin ('85), Roth ('91), and Cooperstein ('98)
- Re-discovered by Silva-Kschischang-Koetter ('08) for network error amplification

Rank-metric codes

Definition

A **rank-metric code** is an $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n\times m}$. If $\mathscr{C} \neq \{0\}$, then its **minimum rank distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C},\, X \neq 0\}.$$

- Introduced and studied by Delsarte ('78) for combinatorial interest
- Re-discovered by Gabidulin ('85), Roth ('91), and Cooperstein ('98)
- Re-discovered by Silva-Kschischang-Koetter ('08) for network error amplification

NOTATION: $2 \leq n \leq m$ integers.

Rank-metric codes

Definition

A **rank-metric code** is an $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n\times m}$. If $\mathscr{C} \neq \{0\}$, then its **minimum rank distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C},\, X \neq 0\}.$$

- Introduced and studied by Delsarte ('78) for combinatorial interest
- Re-discovered by Gabidulin ('85), Roth ('91), and Cooperstein ('98)
- Re-discovered by Silva-Kschischang-Koetter ('08) for network error amplification

NOTATION: $2 \leq n \leq m$ integers.

There is a rank-analogue of the Singleton bound:

Theorem (Delsarte)

Let $\mathscr{C} \leq \mathbb{F}_q^{n\times m}$ be a non-zero rank-metric code. We have

$$\dim(\mathscr{C}) \leq m(n - d_{\mathrm{rk}}(\mathscr{C}) + 1).$$

Rank-metric codes

Definition

A **rank-metric code** is an $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$. If $\mathscr{C} \neq \{0\}$, then its **minimum rank distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C},\, X \neq 0\}.$$

- Introduced and studied by Delsarte ('78) for combinatorial interest
- Re-discovered by Gabidulin ('85), Roth ('91), and Cooperstein ('98)
- Re-discovered by Silva-Kschischang-Koetter ('08) for network error amplification

NOTATION: $2 \leq n \leq m$ integers.

There is a rank-analogue of the Singleton bound:

Theorem (Delsarte)

Let $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$ be a non-zero rank-metric code. We have

$$\dim(\mathscr{C}) \leq m(n - d_{\mathrm{rk}}(\mathscr{C}) + 1).$$

A code $\mathscr{C}$ is **MRD** if it meets the bound with equality $\quad (\Longrightarrow \dim(\mathscr{C}) \equiv 0 \mod m)$.

Classes of codes

Hamming space

- $\mathbb{F}_q^n, \quad d_{\mathrm{H}}(x,y) = |\{i \mid x_i \neq y_i\}|$
- Code: $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^n$
- Bound: $\dim(\mathscr{C}) \leq n - d_{\mathrm{H}}(\mathscr{C}) + 1$
- Codes meeting the bound: **MDS** codes

Matrix rank-metric space

- $\mathbb{F}_q^{n\times m}$ with $n \leq m, \quad d_{\mathrm{rk}}(X,Y) = \mathrm{rk}(X - Y)$
- Code: $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n\times m}$
- Bound: $\dim(\mathscr{C}) \leq m(n - d_{\mathrm{rk}}(\mathscr{C}) + 1)$
- Codes meeting the bound: **MRD** codes

Classes of codes

Hamming space

- $\mathbb{F}_q^n, \quad d_{\mathrm{H}}(x,y) = |\{i \mid x_i \neq y_i\}|$
- Code: $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^n$
- Bound: $\dim(\mathscr{C}) \leq n - d_{\mathrm{H}}(\mathscr{C}) + 1$
- Codes meeting the bound: **MDS** codes

Matrix rank-metric space

- $\mathbb{F}_q^{n\times m}$ with $n \leq m, \quad d_{\mathrm{rk}}(X,Y) = \mathrm{rk}(X-Y)$
- Code: $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n\times m}$
- Bound: $\dim(\mathscr{C}) \leq m(n - d_{\mathrm{rk}}(\mathscr{C}) + 1)$
- Codes meeting the bound: **MRD** codes

Vector rank-metric space

- $\mathbb{F}_{q^m}^n$ with $m \geq n, \quad d_{\mathrm{rk}}(x,y) = \dim_{\mathbb{F}_q} \mathrm{span}\{x_1 - y_1, \dots, x_n - y_n\}$
- Code: $\mathbb{F}_{q^m}$-subspace $\mathscr{C} \leq \mathbb{F}_{q^m}^n$
- Bound: $\dim_{\mathbb{F}_{q^m}}(\mathscr{C}) \leq n - d_{\mathrm{rk}}(\mathscr{C}) + 1$
- Codes meeting the bound: (vector) **MRD** codes

Density of MDS codes

A randomly chosen k-dimensional code is MDS with high probability, if $q \gg 0$.

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n}$$

Density of MDS codes

A randomly chosen k-dimensional code is MDS with high probability, if $q \gg 0$.

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n}$$

Density of MDS codes

A randomly chosen k-dimensional code is MDS with high probability, if $q \gg 0$.

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1$$

Density of MDS codes

A randomly chosen k-dimensional code is MDS with high probability, if $q \gg 0$.

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1$$

We say that MDS codes are **dense** within the set of k-dimensional codes in $\mathbb{F}_q^n$.

We study "density questions" in coding theory in:

Byrne, R., *Partition-Balanced Families of Codes and Asymptotoc Enumeration in Coding Theory*, J. Combinatorial Theory A, to appear.

The notion of density

Definition

Let $S \subseteq \mathbb{N}$ be an infinite set. Let $(\mathscr{F}_s \mid s \in S)$ be a sequence of finite non-empty sets indexed by S, and let $(\mathscr{F}'_s \mid s \in S)$ be a sequence of sets with $\mathscr{F}'_s \subseteq \mathscr{F}_s$ for all $s \in S$.

The **density function** $S \to \mathbb{Q}$ of $\mathscr{F}'_s$ in $\mathscr{F}_s$ is $\quad s \mapsto |\mathscr{F}'_s|/|\mathscr{F}_s|$.

If

$$\lim_{s \to +\infty} |\mathscr{F}'_s|/|\mathscr{F}_s| = \delta,$$

then $\mathscr{F}'_s$ has **density** δ in $\mathscr{F}_s$.

- $\delta = 1$: $\mathscr{F}'_s$ is **dense** in $\mathscr{F}_s$
- $\delta = 0$: $\mathscr{F}'_s$ is **sparse** in $\mathscr{F}_s$

The notion of density

Definition

Let $S \subseteq \mathbb{N}$ be an infinite set. Let $(\mathscr{F}_s \mid s \in S)$ be a sequence of finite non-empty sets indexed by S, and let $(\mathscr{F}'_s \mid s \in S)$ be a sequence of sets with $\mathscr{F}'_s \subseteq \mathscr{F}_s$ for all $s \in S$.

The **density function** $S \to \mathbb{Q}$ of $\mathscr{F}'_s$ in $\mathscr{F}_s$ is $\quad s \mapsto |\mathscr{F}'_s|/|\mathscr{F}_s|$.

If
$$\lim_{s \to +\infty} |\mathscr{F}'_s|/|\mathscr{F}_s| = \delta,$$
then $\mathscr{F}'_s$ has **density** δ in $\mathscr{F}_s$.

- $\delta = 1$: $\mathscr{F}'_s$ is **dense** in $\mathscr{F}_s$
- $\delta = 0$: $\mathscr{F}'_s$ is **sparse** in $\mathscr{F}_s$

Example

$S = \mathbb{N}_{\geq 1} \qquad \mathscr{F}_s = \{n \in \mathbb{N} \mid 1 \leq n \leq s\} \qquad \mathscr{F}'_s = \{p \in \mathbb{N} \mid p \leq s,\ p \text{ prime}\}$.

Then: $\qquad |\mathscr{F}'_s|/|\mathscr{F}_s| \to 0, \qquad |\mathscr{F}'_s|/|\mathscr{F}_s| \sim 1/\log(s)$

(Hadamard, de la Vallée-Poussin, 1896)

Density of MDS codes

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1.$$

Density of MDS codes

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1.$$

Sketch of proof

- The k-dimensional MDS codes in $\mathbb{F}_q^n$ are in bijection with the non-zeros of a polynomial $p \in \mathbb{F}_q[z_1, \dots, z_N]$, where $N = k(n-k)$
- $\deg(p) \leq k\binom{n}{k}$
- Using the **Schwartz-Zippel Lemma**, one has

Density of MDS codes

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1.$$

Sketch of proof

- The k-dimensional MDS codes in $\mathbb{F}_q^n$ are in bijection with the non-zeros of a polynomial $p \in \mathbb{F}_q[z_1, \dots, z_N]$, where $N = k(n-k)$
- $\deg(p) \leq k\binom{n}{k}$
- Using the **Schwartz-Zippel Lemma**, one has

$$\frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} \geq \frac{q^{k(n-k)}\left(1 - \frac{k}{q}\binom{n}{k}\right)}{\begin{bmatrix} n \\ k \end{bmatrix}_q}$$

Density of MDS codes

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1.$$

Sketch of proof

- The k-dimensional MDS codes in $\mathbb{F}_q^n$ are in bijection with the non-zeros of a polynomial $p \in \mathbb{F}_q[z_1, \ldots, z_N]$, where $N = k(n-k)$
- $\deg(p) \leq k\binom{n}{k}$
- Using the **Schwartz-Zippel Lemma**, one has

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} \geq \lim_{q \to +\infty} \frac{q^{k(n-k)}\left(1 - \frac{k}{q}\binom{n}{k}\right)}{\begin{bmatrix} n \\ k \end{bmatrix}_q}$$

Density of MDS codes

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1.$$

Sketch of proof

- The k-dimensional MDS codes in $\mathbb{F}_q^n$ are in bijection with the non-zeros of a polynomial $p \in \mathbb{F}_q[z_1, \dots, z_N]$, where $N = k(n-k)$
- $\deg(p) \leq k\binom{n}{k}$
- Using the **Schwartz-Zippel Lemma**, one has

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} \geq \lim_{q \to +\infty} \frac{q^{k(n-k)}\left(1 - \frac{k}{q}\binom{n}{k}\right)}{\begin{bmatrix} n \\ k \end{bmatrix}_q} = 1$$

Density problems in coding theory

We study density problems in general:

- Ambient space: Hamming space, matrix rk-metric space, vector rk-metric space
- Various properties related to: minimum distance, covering radius, maximality

Density problems in coding theory

We study density problems in general:

- Ambient space: Hamming space, matrix rk-metric space, vector rk-metric space
- Various properties related to: minimum distance, covering radius, maximality

classical arguments (based on Schwartz-Zippel Lemma) often fail.

Density problems in coding theory

We study density problems in general:

- Ambient space: Hamming space, matrix rk-metric space, vector rk-metric space
- Various properties related to: minimum distance, covering radius, maximality

⚠ classical arguments (based on Schwartz-Zippel Lemma) often fail.

Idea

Look at **families** of codes that exhibit regularity properties with respect to partitions of the ambient space $X \in \{\mathbb{F}_q^n,\ \mathbb{F}_q^{n\times m},\ \mathbb{F}_{q^m}^n\}$.

Definition

Let $\mathscr{P} = \{P_1, P_2, \dots, P_\ell\}$ be a partition of X.

A family $\mathscr{F}$ of codes in X is $\mathscr{P}$**-balanced** if for all $x \in X$ the number

$$|\{\mathscr{C} \in \mathscr{F} \mid x \in \mathscr{C}\}|$$

only depends on the class of x with respect to the partition $\mathscr{P}$.

Density problems in coding theory

We study density problems in general:

- Ambient space: Hamming space, matrix rk-metric space, vector rk-metric space
- Various properties related to: minimum distance, covering radius, maximality

classical arguments (based on Schwartz-Zippel Lemma) often fail.

Idea

Look at **families** of codes that exhibit regularity properties with respect to partitions of the ambient space $X \in \{\mathbb{F}_q^n,\ \mathbb{F}_q^{n\times m},\ \mathbb{F}_{q^m}^n\}$.

Definition

Let $\mathscr{P} = \{P_1, P_2, \dots, P_\ell\}$ be a partition of X.

A family $\mathscr{F}$ of codes in X is **$\mathscr{P}$-balanced** if for all $x \in X$ the number

$$|\{\mathscr{C} \in \mathscr{F} \mid x \in \mathscr{C}\}|$$

only depends on the class of x with respect to the partition $\mathscr{P}$.

We use $\mathscr{P}$-balanced families to estimate the number of codes with a certain property.

MRD vector rk-metric codes

Using the Schwartz-Zippel lemma:

Theorem (Neri-Trautmann-Randrianarisoa-Rosenthal, 2017)

For vector-rank-metric codes ($\mathbb{F}_{q^m}$-linear)

$$\frac{\#\text{ of } k\text{-dim MRD codes in } \mathbb{F}_{q^m}^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_{q^m}^n} \geq q^{mk(n-k)} \begin{bmatrix} n \\ k \end{bmatrix}_{q^m}^{-1} \left(1 - \sum_{r=0}^{k} \begin{bmatrix} k \\ k-r \end{bmatrix}_q \begin{bmatrix} n-k \\ r \end{bmatrix}_q q^{r^2} q^{-m}\right)$$

MRD vector rk-metric codes

Using the Schwartz-Zippel lemma:

Theorem (Neri-Trautmann-Randrianarisoa-Rosenthal, 2017)

For vector-rank-metric codes ($\mathbb{F}_{q^m}$-linear)

$$\frac{\#\text{ of }k\text{-dim MRD codes in }\mathbb{F}_{q^m}^n}{\#\text{ of }k\text{-dim codes in }\mathbb{F}_{q^m}^n} \geq q^{mk(n-k)} \begin{bmatrix} n \\ k \end{bmatrix}_{q^m}^{-1} \left(1 - \sum_{r=0}^{k} \begin{bmatrix} k \\ k-r \end{bmatrix}_q \begin{bmatrix} n-k \\ r \end{bmatrix}_q q^{r^2} q^{-m}\right)$$

We can improve this bound as follows:

Theorem (Byrne-R.)

For vector-rank-metric codes ($\mathbb{F}_{q^m}$-linear)

$$\frac{\#\text{ of }k\text{-dim MRD codes in }\mathbb{F}_{q^m}^n}{\#\text{ of }k\text{-dim codes in }\mathbb{F}_{q^m}^n} \geq 1 - \frac{q^{mk}-1}{(q^m-1)(q^{mn}-1)} \left(-1 + \sum_{i=0}^{d-1} \begin{bmatrix} n \\ i \end{bmatrix}_q \prod_{j=0}^{i-1} (q^m - q^j)\right)$$

MRD matrix rk-metric codes

MRD codes: rank-analogue of MDS codes. So one might expect them to be dense...

MRD matrix rk-metric codes

MRD codes: rank-analogue of MDS codes. So one might expect them to be dense...

However, MRD matrix codes are <u>not</u> dense

MRD matrix rk-metric codes

MRD codes: rank-analogue of MDS codes. So one might expect them to be dense...

However, MRD matrix codes are <u>not</u> dense

Theorem (Byrne-R.)

Let $m \geq n \geq 2$ and let $1 \leq k \leq mn-1$ be integers.

- If m does not divide k, then there is no k-dimensional MRD code $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$.
- If m divides k, then

$$\frac{\#\text{ of } k\text{-dim non-MRD codes in } \mathbb{F}_q^{n\times m}}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^{n\times m}} \geq$$

$$q \begin{bmatrix} mn \\ k \end{bmatrix}^{-1} \left(\sum_{h=1}^{m(n-k)} \begin{bmatrix} t \\ h \end{bmatrix} \sum_{s=h}^{m(n-k)} \begin{bmatrix} m(n-k)-h \\ s-h \end{bmatrix} \begin{bmatrix} mn-s \\ mn-k \end{bmatrix} (-1)^{s-h} q^{\binom{s-h}{2}} \right) \cdot$$

$$\cdot \left(1 - \frac{\left(q^k-1\right)\left(q^{mn-k}-1\right)}{2\left(q^{mn}-q^{mn-k}\right)} \right).$$

The RHS goes to $1/2$ as $q \to +\infty$ and to $1/2(q/(q-1)-(q-1)^2)$ as $m \to +\infty$.

Non-density of MRD matrix codes

Corollary (Byrne-R.)

Let $m \geq n \geq 2$ and let $1 \leq k \leq mn-1$ be integers.

- If m does not divide k, then there is no k-dimensional MRD code $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$.
- If m divides k, then

$$\liminf_{q \to +\infty} \frac{\# \text{ of } k\text{-dim non-MRD codes in } \mathbb{F}_q^{n \times m}}{\# \text{ of } k\text{-dim codes in } \mathbb{F}_q^{n \times m}} \geq 1/2.$$

$$\liminf_{m \to +\infty} \frac{\# \text{ of } k\text{-dim non-MRD codes in } \mathbb{F}_q^{n \times m}}{\# \text{ of } k\text{-dim codes in } \mathbb{F}_q^{n \times m}} \geq \frac{1}{2}\left(\frac{q}{q-1} - (q-1)^{-2}\right) \geq 1/2.$$

Matrix MRD codes are <u>not</u> dense

Non-density of MRD matrix codes

Corollary (Byrne-R.)

Let $m \geq n \geq 2$ and let $1 \leq k \leq mn-1$ be integers.

- If m does not divide k, then there is no k-dimensional MRD code $\mathscr{C} \leq \mathbb{F}_q^{n\times m}$.
- If m divides k, then

$$\liminf_{q\to+\infty} \frac{\#\text{ of } k\text{-dim non-MRD codes in } \mathbb{F}_q^{n\times m}}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^{n\times m}} \;\geq\; 1/2.$$

$$\liminf_{m\to+\infty} \frac{\#\text{ of } k\text{-dim non-MRD codes in } \mathbb{F}_q^{n\times m}}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^{n\times m}} \;\geq\; \frac{1}{2}\left(\frac{q}{q-1}-(q-1)^{-2}\right) \;\geq\; 1/2.$$

Matrix MRD codes are not dense

Non-density was also shown by Antrobus/Gluesing-Luerssen with different methods.

Other results

We study:

- Density of codes that are **optimal** (MDS, MRD, MRD)
- Density of codes of bounded **minimum distance**
- Density of codes that meet the *redundancy bound* for their **covering radius**
- Density of matrix codes that meet the *initial set bound* for their covering radius
- Density of optimal codes within **maximal** codes (with respect to inclusion)
- ...

R., *Whitney numbers of combinatorial geometries and higher-weight Dowling lattices*, arXiv 1909.10249.

Codes with the Hamming metric and geometric lattices

R., *Whitney numbers of combinatorial geometries and higher-weight Dowling lattices*, arXiv 1909.10249.

Example of question

How many codes $\mathscr{C} \leq \mathbb{F}_q^n$ are there of dimension k and $d_{\mathrm{H}}(\mathscr{C}) > d$?

R., *Whitney numbers of combinatorial geometries and higher-weight Dowling lattices*, arXiv 1909.10249.

Example of question

How many codes $\mathscr{C} \leq \mathbb{F}_q^n$ are there of dimension k and $d_{\mathrm{H}}(\mathscr{C}) > d$?

Why?

- valid math question
- applications to density problems
- randomized constructions of codes

Codes with the Hamming metric and geometric lattices

R., *Whitney numbers of combinatorial geometries and higher-weight Dowling lattices*, arXiv 1909.10249.

Example of question

How many codes $\mathscr{C} \leq \mathbb{F}_q^n$ are there of dimension k and $d_{\mathrm{H}}(\mathscr{C}) > d$?

Why?

- valid math question
- applications to density problems
- randomized constructions of codes

Theorem (Dowling 1971, Zaslavsky 1987)

Counting codes $\longleftarrow$ computing the ch. polynomials of certain geometric lattices.

Codes with the Hamming metric and geometric lattices

R., *Whitney numbers of combinatorial geometries and higher-weight Dowling lattices*, arXiv 1909.10249.

Example of question

How many codes $\mathscr{C} \leq \mathbb{F}_q^n$ are there of dimension k and $d_{\mathrm{H}}(\mathscr{C}) > d$?

Why?

- valid math question
- applications to density problems
- randomized constructions of codes

Theorem (Dowling 1971, Zaslavsky 1987)

Counting codes $\longleftarrow$ computing the ch. polynomials of certain geometric lattices.

In particular, of **higher-weight Dowling lattices** (abbreviated **HWDLs**).

HWDLs

$\mathscr{H}(q,n,d)$ is a sublattice of the lattice of subspaces of $\mathbb{F}_q^n$.

Definition

$\mathscr{H}(q,n,d)$ consists of those subspaces of $\mathbb{F}_q^n$ that have a basis made of vectors of Hamming weight $\leq d$, ordered by inclusion.

HWDLs

$\mathscr{H}(q,n,d)$ is a sublattice of the lattice of subspaces of $\mathbb{F}_q^n$.

Definition

$\mathscr{H}(q,n,d)$ consists of those subspaces of $\mathbb{F}_q^n$ that have a basis made of vectors of Hamming weight $\leq d$, ordered by inclusion.

- Introduced by Dowling in 1971
- Studied by Dowling, Zaslavsky, Bonin, Kung, Brini, Games, ...
- To date, still very little is known about HWDLs
- Closely related to Segre's conjecture, open since the 50s.

HWDLs

$\mathscr{H}(q,n,d)$ is a sublattice of the lattice of subspaces of $\mathbb{F}_q^n$.

Definition

$\mathscr{H}(q,n,d)$ consists of those subspaces of $\mathbb{F}_q^n$ that have a basis made of vectors of Hamming weight $\leq d$, ordered by inclusion.

- Introduced by Dowling in 1971
- Studied by Dowling, Zaslavsky, Bonin, Kung, Brini, Games, ...
- To date, still very little is known about HWDLs
- Closely related to Segre's conjecture, open since the 50s.

Examples

- For $d=1$, $\mathscr{H}(q,n,d)$ is the Boolean algebra of subsets of $\{1,\dots,n\}$
- For $d=2$, $\mathscr{H}(q,n,d)$ is isomorphic to the q-analogue of the partition lattice (Dowling '73)

Studying these lattices is an old open problem in combinatorics.

HWDLs

Theorem (R., 2019)

The following are *equivalent*:

- (partial) knowledge of the number of codes with $d_{\mathrm{H}}(\mathscr{C}) > d$
- (partial) knowledge of the Whitney numbers of HWDL's

HWDLs

Theorem (R., 2019)

The following are *equivalent*:

- (partial) knowledge of the number of codes with $d_{\mathrm{H}}(\mathscr{C}) > d$
- (partial) knowledge of the Whitney numbers of HWDL's

More precisely, let $\alpha_k(q,n,d) = \#\{\mathscr{C} \leq \mathbb{F}_q^n \mid \dim(C) = k,\ d_{\mathrm{H}}(\mathscr{C}) > d\}$. Then

$$\alpha_k(q,n,d) = \sum_{i=0}^{k} w_i(q,n,d) \begin{bmatrix} n-i \\ k-i \end{bmatrix}_q \quad \text{for } 0 \leq k \leq n$$

$$w_i(q,n,d) = \sum_{k=0}^{i} \alpha_k(q,n,d) \begin{bmatrix} n-k \\ i-k \end{bmatrix}_q (-1)^{i-k} q^{\binom{i-k}{2}} \quad \text{for } 0 \leq i \leq n$$

Recall: the i-th Whitney number of $\mathscr{L} = \mathscr{H}(q,n,d)$ is

$$w_i(q,n,d) = \sum_{\mathrm{rk}(x)=i} \mu_{\mathscr{L}}(0,x)$$

HWDLs

Theorem (R., 2019)

For all $n \geq 9$ we have

$$
\begin{aligned}
-w_3(2,n,3) \;=\; & \sum_{1\leq \ell_1<\ell_2<\ell_3\leq n-2} \left(\prod_{j=1}^{3} \binom{n-\ell_j-9+3j)}{2} \right) + 8\binom{n}{3} \sum_{s=3}^{8} \binom{n-3}{n-s} (-1)^{s-3} \\
& + 106\binom{n}{4} \sum_{s=4}^{8} \binom{n-4}{n-s} (-1)^{s-4} + 820\binom{n}{5} \sum_{s=5}^{8} \binom{n-5}{n-s} (-1)^{s-5} \\
& + 4565\binom{n}{6} \sum_{s=6}^{8} \binom{n-6}{n-s} (-1)^{s-6} \\
& + 19810\binom{n}{8} \sum_{s=7}^{8} \binom{n-7}{n-s} (-1)^{s-7} + 70728\binom{n}{8}.
\end{aligned}
$$

HWDLs

Theorem (R., 2019)

Let $n \geq 6$. We have

$$\begin{aligned} w_2(q,n,3) &= \frac{1}{72}q^4n^6 - \frac{1}{12}q^4n^5 + \frac{1}{18}q^4n^4 + \frac{1}{2}q^4n^3 - \frac{77}{72}q^4n^2 + \frac{7}{12}q^4n - \frac{1}{18}q^3n^6 \\ &+ \frac{5}{12}q^3n^5 - \frac{49}{72}q^3n^4 - \frac{7}{6}q^3n^3 + \frac{269}{72}q^3n^2 - \frac{9}{4}q^3n + \frac{1}{12}q^2n^6 - \frac{3}{4}q^2n^5 \\ &+ 2q^2n^4 - \frac{7}{12}q^2n^3 - \frac{43}{12}q^2n^2 + \frac{17}{6}q^2n - \frac{1}{18}qn^6 + \frac{7}{12}qn^5 - \frac{157}{72}qn^4 \\ &+ \frac{19}{6}qn^3 - \frac{55}{72}qn^2 - \frac{3}{4}qn + \frac{1}{72}n^6 - \frac{1}{6}n^5 + \frac{29}{36}n^4 - \frac{23}{12}n^3 + \frac{157}{72}n^2 - \frac{11}{12}n. \end{aligned}$$

HWDLs

Theorem (R., 2019)

For all integers $n \geq d \geq 2$ and any prime power q,

$$w_2(q,n,d) = (q^{n-1}-1)\sum_{j=1}^{d}\binom{n}{j}(q-1)^{j-2} - \sum_{1\leq \ell_1<\ell_2\leq n}\Bigg[q^{n-\ell_1-1}\left(\sum_{j=0}^{d-1}\binom{n-\ell_2}{j}(q-1)^j\right)$$

$$+\sum_{j=d}^{n-\ell_2}\sum_{h=0}^{d-1}\binom{n-\ell_2}{j}\binom{n-\ell_1-1}{h}(q-1)^{j+h}$$

$$+\sum_{s=d}^{n-\ell_2}\sum_{t=0}^{d-2}\binom{n-\ell_2}{s}\binom{n-\ell_1-1-s}{t}(q-1)^{s+t}\sum_{v=d-t}^{s}\gamma_q(s,s-d+t+2,v)\Bigg],$$

where the $\gamma_a(b,c,v)$'s are the *agreement numbers*.

Theorem (R., 2019)

For all integers $n \geq d \geq 2$ and any prime power q,

$$w_2(q,n,d) = (q^{n-1}-1)\sum_{j=1}^{d}\binom{n}{j}(q-1)^{j-2} - \sum_{1\leq \ell_1<\ell_2\leq n}\Bigg[q^{n-\ell_1-1}\left(\sum_{j=0}^{d-1}\binom{n-\ell_2}{j}(q-1)^j\right)$$
$$+\sum_{j=d}^{n-\ell_2}\sum_{h=0}^{d-1}\binom{n-\ell_2}{j}\binom{n-\ell_1-1}{h}(q-1)^{j+h}$$
$$+\sum_{s=d}^{n-\ell_2}\sum_{t=0}^{d-2}\binom{n-\ell_2}{s}\binom{n-\ell_1-1-s}{t}(q-1)^{s+t}\sum_{v=d-t}^{s}\gamma_q(s,s-d+t+2,v)\Bigg],$$

where the $\gamma_a(b,c,v)$'s are the *agreement numbers*.

$\gamma_a(b,c,v)$ is a polynomial in a (for any b, c and v) whose coefficients are expressions involving the Bernoulli numbers:

$$\frac{x}{e^x-1} = \sum_{n=0}^{+\infty} B_n \frac{x^n}{n!}.$$

$\rightarrow$ **polynomiality** in q of $w_2(q,n,d)$ (R. 2019)

Back to density results

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1.$$

Back to density results

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1.$$

Theorem (R., 2019)

Let $n \geq k \geq 1$ be integers. We have

$$\frac{\#\text{ of } k\text{-dim non-MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} \sim \binom{n}{k} q^{-1} \qquad \text{as } q \to +\infty.$$

Back to density results

Theorem (Folklore)

Let $n \geq k \geq 1$ be integers. We have

$$\lim_{q \to +\infty} \frac{\#\text{ of } k\text{-dim MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} = 1.$$

Theorem (R., 2019)

Let $n \geq k \geq 1$ be integers. We have

$$\frac{\#\text{ of } k\text{-dim non-MDS codes in } \mathbb{F}_q^n}{\#\text{ of } k\text{-dim codes in } \mathbb{F}_q^n} \sim \binom{n}{k} q^{-1} \qquad \text{as } q \to +\infty.$$

Thank you very much!