

# Informe inicial

---

## Ampliació d'una llibreria en MAGMA per a codis Z2Z4 - lineals

Cristina Diéguez

06/03/2016



**Universitat Autònoma  
de Barcelona**

Es presenta la proposta detallada del treball de final de grau a realitzar, en el qual s'exposen els objectius a assolir i la metodologia que es seguirà. Al mateix temps, es detalla la planificació inicial que es pretén seguir durant tota la realització del mateix.

## Índex

1. Introducció .....	2
2. Estat de l'art .....	4
3. Objectius .....	6
4. Planificació i metodologia .....	6
4.1. Metodologia .....	6
4.2. Planificació temporal.....	8
5. Bibliografia .....	9

## 1. Introducció

Tot procés d'enviament i recepció de missatges es realitza a través d'un canal de comunicació. No obstant, el missatge rebut a la sortida del canal pot haver patit alguna alteració a causa del soroll i fa que aquest es converteixi en erroni. Aleshores, perquè el receptor pugui eliminar possibles errors es realitza el procés de codificació-descodificació.

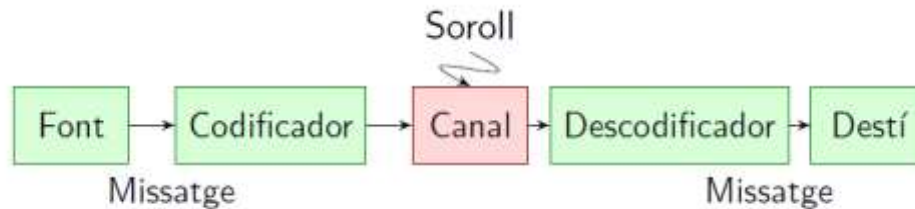


Figura 1. Procés de comunicació

En la codificació, s'assigna a cada símbol o conjunt de símbols de la font una paraula-codi afegint redundància. L'agrupació de totes aquestes paraules-codi formen un codi.

En el procés de descodificació si s'ha produït algun error en la transmissió i el descodificador detecta l'error, es realitza la correcció aprofitant la redundància afegida. Aleshores el que es vol és poder seguir enviant missatge tenint la capacitat de detecció i correcció d'errors. El segon teorema de Shannon especifica aquest fet ja que declara que és possible transmetre informació a través d'un canal amb soroll si es transmet a una taxa de transmissió de la informació<sup>1</sup> inferior a la capacitat del canal. Aquest teorema dona peu a la teoria de codificació per a la correcció d'errors que té per objectiu la construcció de codificadors i descodificadors que permetin enviar el màxim d'informació amb el mínim d'errors.

L'escenari ideal seria trobar codis on cadascun dels vectors tingués una única descodificació possible, però trobar codis amb aquestes propietats no és una tasca banal. Una manera de poder minimitzar els errors seria augmentant la redundància en el missatge per poder, després, corregir els errors. El desavantatge que presenta és que existeix una penalització de la velocitat de transmissió.

Dins de l'àmbit informàtic, el tipus de codis que presenten un gran interès per les seves característiques són els codis binaris lineals perquè experimenten les següents propietats:

1. Tancament per la suma: dues paraules codi sumades és igual a una altra paraula-codi.

<sup>1</sup> Taxa de transmissió de la informació  $R_T = \frac{\log_2 |C|}{n}$

2. Tancament per la multiplicació d'un escalar: si es multiplica una paraula-codi per un escalar, dóna una altra paraula-codi.

Gràcies a aquestes propietats, es pot trobar un conjunt de paraules-codis linealment independents que són capaces de generar totes les paraules del codi. Aleshores, s'observa que un codi pot ser representat de forma matricial per les paraules-codi linealment independents. Es crea el que s'anomena matriu generadora que facilita el treball i la representació d'aquest tipus de codis. Amb aquest fet, es redueixen els costos de memòria i còmput a l'hora de treballar amb codis.

A part d'usar codis binaris que són un tipus de representació utilitzant l'anell d'enters mòdul 2, actualment es treballa amb codis  $Z_2Z_4$ -additius, que són codis sobre l'anell  $Z_2^\alpha \times Z_4^\beta$ . Aquests codis també posseeixen el concepte de matriu generadora  $G$  que està formada per:

- $\gamma$  vectors d'ordre 2: aquells formats per coordenades a  $\{0,2\}$ .
- $\delta$  vectors d'ordre 4: aquells que contenen alguna coordenada a  $\{1,3\}$ .

Coneixent aquesta informació, es determina que el codi  $C$  obtingut a partir de la matriu generadora  $G$  és d'ordre  $2^\gamma \cdot 4^\delta$ . Aquesta expressió, a la vegada, proporciona el número de paraules-codi de  $C$ . A l'hora de generar-les, s'opera  $(x,y) \cdot G$  on  $x$  pertany a  $Z_2^\gamma$  i  $y$  a  $Z_4^\delta$  ja que si  $x$  correspongués a  $Z_4$  es duplicarien les mateixes paraules-codi.

Amb la finalitat de poder treballar a  $Z_2$ , existeix una funció anomenada *Gray map*  $\phi$  que proporciona les imatges de les paraules-codi en aquest conjunt; és a dir, realitza una bijecció:

$$\begin{array}{ccc} \phi: & Z_4 & \longrightarrow Z_2 \\ & 0 & \longrightarrow 00 \\ & 1 & \longrightarrow 01 \\ & 2 & \longrightarrow 11 \\ & 3 & \longrightarrow 10 \end{array}$$

Si  $v=(v_1, v_2)$  que pertany a  $Z_2^\alpha \times Z_4^\beta$ , es defineix  $\phi(v)=(v_1, \phi(v_2))$ . Aquesta bijecció només afecta, realment, a les  $\beta$  coordenades quaternàries de les paraules-codi mentre que les  $\alpha$  coordenades binàries es mantenen iguals. El codi binari  $\phi(C)$  s'anomena  $Z_2Z_4$ -lineal i, de forma general, no és lineal.

La distància de Hamming  $d_H(u, v)$  entre dues paraules-codi  $u, v$  que pertanyen a  $Z_2$  és el nombre de coordenades per les quals difereixen  $u$  i  $v$ . El pes de Hamming  $w_H(u)$  és la distància de Hamming entre la paraula-codi  $u$  i la paraula-codi formada per 0. En codis quaternaris, que són un tipus de representació utilitzant l'anell d'enters mòdul 4, s'utilitza la mètrica de Lee. El pes de Lee  $w_L(v)$  és la suma dels pesos de Lee de les seves coordenades. Els pesos d'aquestes coordenades estan definits de la següent manera:

- $w_L(0)=0$
- $w_L(1)=w_L(3)=1$
- $w_L(2)=2$

La funció Gray map  $\phi$  preserva les distàncies, transformant la distància de Lee a distància de Hamming. Gràcies a aquesta funció  $\phi(C)$  s'observa que el pes de Lee coincideix amb el pes de Hamming sobre  $Z_2$ . Un exemple d'aquest succés seria:

$$\phi(10230) = (0100111000)$$

$$w_L(10230) = w_L(1) + w_L(0) + w_L(2) + w_L(3) + w_L(0) = 4$$

$$w_H(0100111000) = 4$$

La distància mínima d'un codi  $C$  és el mínim valor de la distància de totes les paraules-codi que formen  $C$ . Així mateix, el pes mínim d'un codi  $C$  és el mínim valor dels pesos de totes les paraules-codi. En aquest projecte, es treballarà amb aquests conceptes a l'hora de plantejar i implementar funcions que ajudin en el procés de descodificació, la qual pot ser descodificació via síndrome o descodificació per mínima distància.

## 2. Estat de l'art

El grup CCSG és un equip format dins del Departament d'Enginyeria de la Informació i de les Comunicacions, d'EIC, i dedicat, parcialment, a la investigació de codis  $Z_2Z_4$ -lineals amb la finalitat dur a terme els següents objectius, entre d'altres:

- Construir i caracteritzar nous codis  $Z_2Z_4$ -lineals, i computar els seus paràmetres estructurals.

- Caracteritzar i construir famílies de codis Reed-Muller  $Z_2Z_4$ -lineals, i computar-ne el rang i la dimensió del kernel.
- Desenvolupar els algorismes necessaris per establir noves opcions en l'ocultació de dades i en l'autenticació de documents, utilitzant codis  $Z_2Z_4$ -lineals.
- Expandir el software de MAGMA implementant nous paquets amb la finalitat de treballar eficientment amb codis  $Z_2Z_4$ -lineals, codis regulars i codis no lineals.

Un dels principals articles escrits sobre codis  $Z_2Z_4$ -lineals és [1] en què s'estudien els codis  $Z_2Z_4$ -additius a partir dels quals s'extreu la corresponent imatge binària i s'obtenen els codis  $Z_2Z_4$ -lineals. D'aquests nous codis, es detallen les seves propietats i les formes estàndard de les matrius generadores i de control. En aquest article se'n fa la primera definició. També esmentar l'article [2] ja que s'hi demostren les propietats necessàries pel càlcul de rangs i kernels dels codis  $Z_2Z_4$ -lineals.

Degut que en aquest projecte serà necessari treballar amb les distàncies de Hamming dels codis ja mencionats, una de les referències a destacar és [6] perquè defineix i implementa mètodes algorítmics per la computació de la distància mínima de Hamming per codis  $Z_2Z_4$ -lineals.

El llenguatge MAGMA porta incorporat llibreries sobre  $Z_2$  i sobre  $Z_4$  cosa que facilita el treball sobre aquests anells. Gràcies als membres d'aquest grup d'investigació CCSG, des de fa uns anys, es duu a terme el desenvolupament d'una nova llibreria en MAGMA per permetre treballar amb codis  $Z_2Z_4$ -lineals, fent ús d'algunes llibreries ja existents en aquest llenguatge. Es pretén que aquesta estigui totalment integrada amb la llibreria per a codis lineals.

Com s'ha esmentat anteriorment, s'usarà el llenguatge MAGMA amb la finalitat de realitzar la part pràctica del projecte. Gràcies a l'estructura  $Z_2Z_4$ -lineals es pot treballar amb aquest llenguatge de manera lineal a  $Z_2^\alpha \times Z_4^\beta$  i també es poden relacionar aquests codis amb codis sobre  $Z_2$  i sobre  $Z_4$ .

Les referències existents sobre l'ús d'aquest llenguatge en l'entorn desitjat són les proporcionades per MAGMA, [9] i [10]. Addicionalment, també es consultaran les guies que facilita el CCSG per MAGMA [8] i per entendre els paquets de codis  $Z_2Z_4$ -lineals, [3], i codis  $Z_2Z_4$ -additius, [4].

El fet d'haver de dissenyar i implementar unes funcions que s'integraran dins de MAGMA implica que s'han de seguir els estàndards marcats per tal mantenir el format. Com a conseqüència, es consultarà la guia [7] ja que proporciona aquests estàndards.

### 3. Objectius

Els objectius que s'assoliran durant el transcurs d'aquest treball seran els esmentats a continuació:

- Realització d'un estudi dels codis  $Z_2Z_4$ -lineals tant a nivell conceptual com de les propietats que els defineixen.
- Aprenentatge d'un nou llenguatge i entorn de programació anomenat MAGMA.
- Disseny i desenvolupament de tests de prova i tests d'integració de les funcions realitzades prèviament.
- Creació de noves funcions de codis  $Z_2Z_4$ -lineals per realitzar els càlcul del pes mínim de Lee, la distància mínima de Lee, el nombre mínim de paraules-codi i la distribució de pesos .
- Ampliació de la llibreria de MAGMA existent sobre teoria de codis  $Z_2Z_4$ -lineals amb les noves funcions desenvolupades per tal de facilitar la descodificació via síndrome.

### 4. Planificació i metodologia

#### 4.1. Metodologia

Aquest projecte consta d'un component teòric rellevant a partir del qual es poden desenvolupar les funcions necessàries per aquesta ampliació de la llibreria de MAGMA. L'expansió de coneixement esmentada implica un fil de continuïtat constant al llarg de tot el projecte ja que és necessari aquest suport a l'hora de tractar els codis  $Z_2Z_4$ -lineals. Com a conseqüència, aquesta obtenció de coneixement provindrà de les publicacions al respecte, per exemple [1,2,5,6], i de l'ajuda i assessorament de la tutora del projecte Cristina Fernández. També es seguiran els formats establerts per codis lineals binaris i quaternaris referenciats en els articles consultats.

El component pràctic adoptarà la metodologia de treball de MAGMA sobre codis lineals en els diferents anells aprofitant les llibreries que aquest ja presenta. La seva codificació seguirà l'estil estrictament marcat pels estàndards facilitant, així, la seva integració a la llibreria. Les funcions previstes a desenvolupar durant tota el procés són les següents:

- ZZZ4MinimumLeeWeight
- ZZZ4MinimumLeeDistance
- ZZZ4MinimumWord
- ZZZ4MinimumWords
- WeightDistribution

El procés de desenvolupament del projecte seguirà una pràctica de programació d'enginyeria de software anomenada *Test-Driven Development (TDD)*. Aquesta és base en la idea d'elaborar primer les proves, després escriure el codi font que serà avaluat amb els tests fets i, per últim, fer la reestructuració o refacció del codi escrit. Al realitzar la metodologia en aquest ordre descrit en la figura 1, obliga a fer un exercici previ d'anàlisi dels requisits i dels diversos escenaris possibles. També implica una reducció del codi al mínim imprescindible per resoldre la prova que hi té associada minimitzant errors. Degut a l'important paper que juguen els test, és imprescindible la realització d'un conjunt de proves unitàries que cobreixin tots els casos a avaluar. Al mateix temps, l'etapa de reestructuració també és rellevant ja que facilita la comprensió del codi i millora l'estructura interna, sense canviar el comportament extern. Amb aquesta pràctica s'aconsegueixen codis més robusts, més segurs, més llegibles, amb més facilitat per mantenir-los al llarg del temps i una major rapidesa en el desenvolupament.

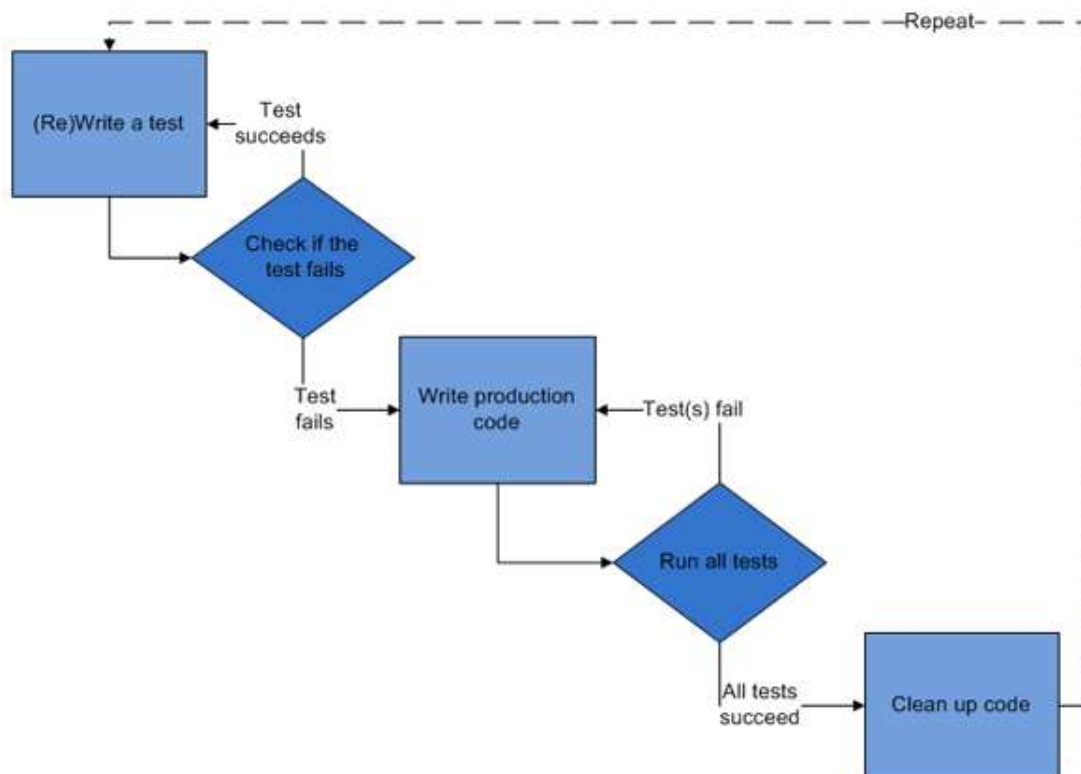


Figura 2: Cicle de vida de *Test-Driven Development*.




















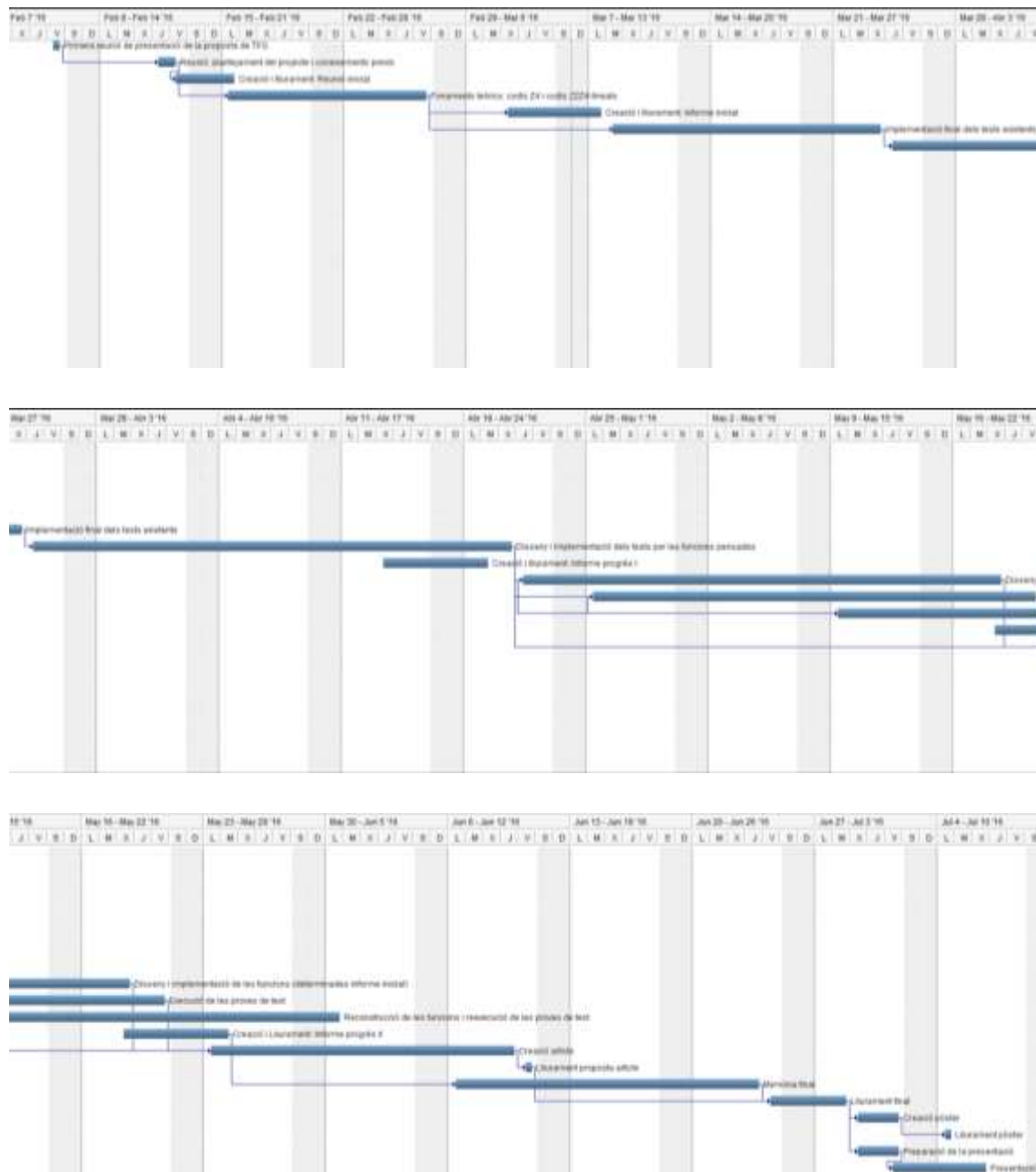
En la fase d'elaboració de proves, es durà a terme la descripció i implementació de tests de caixa negra programats on s'especificaran quins són els resultats correctes desitjats per poder realitzar la comparació amb els obtinguts. D'aquesta forma es podran detectar i observar errors en les funcions programades, si en presenten, amb l'objectiu de modificar-les i reavaluar-les.

## 4.2. Planificació temporal

A la planificació que es mostra a continuació s'ha quantificat la durabilitat de tot el projecte, dividint aquest en tasques principals que s'han de realitzar en els períodes marcats per la bona execució del treball.

Aquesta planificació temporal s'ha realitzat mitjançant un diagrama de Gantt on s'han especificat els recursos del projecte així com la interdependència de tasques. Juntament amb aquest informe, s'afegeix el diagrama amb el seu format per una millor apreciació de la planificació.

		Nombre	Duración	Inicio	Fin	Predecesoras	Recursos
1		Primera reunió de presentació de la proposta de TFG	1d	05/02/2016	05/02/2016		Cristina Diéguez
2		Reunió: plantejament del projecte i coneixements previs	1d	11/02/2016	12/02/2016	1	Cristina Diéguez
3		Creació i lliurament: Reunió inicial	2d?	12/02/2016	15/02/2016	2	Cristina Diéguez
4		Fonaments teòrics: codis Z4 i codis Z2Z4-lineals	10d	15/02/2016	26/02/2016	2	Cristina Diéguez
5		Creació i lliurament: Informe inicial	4d?	02/03/2016	07/03/2016	4	Cristina Diéguez
6		Implementació final dels tests existents	12d?	08/03/2016	23/03/2016	4	Cristina Diéguez
7		Disseny i implementació dels tests per les funcions pensades	20d?	24/03/2016	20/04/2016	6	Cristina Diéguez
8		Creació i lliurament: Informe progrés I	4d?	13/04/2016	19/04/2016		Cristina Diéguez
9		Disseny i implementació de les funcions (determinades informe)	20d?	21/04/2016	18/05/2016	7	Cristina Diéguez
10		Execució de les proves de test	20d?	25/04/2016	20/05/2016	7	Cristina Diéguez
11		Reconstrucció de les funcions i reexecució de les proves de test	16d?	09/05/2016	30/05/2016	9II, 10II	Cristina Diéguez
12		Creació i lliurament: Informe progrés II	4d?	18/05/2016	24/05/2016		Cristina Diéguez
13		Creació article	14d?	23/05/2016	09/06/2016	7, 9, 10	Cristina Diéguez
14		Lliurament proposta article	1d?	10/06/2016	10/06/2016	13	Cristina Diéguez
15		Memòria final	14d?	06/06/2016	23/06/2016	12	
16		Lliurament final	3d?	24/06/2016	28/06/2016	14, 15	Cristina Diéguez
17		Creació pòster	3d?	29/06/2016	01/07/2016	16	Cristina Diéguez
18		Lliurament pòster	1d?	04/07/2016	04/07/2016	17	Cristina Diéguez
19		Preparació de la presentació	3d?	29/06/2016	01/07/2016	16	Cristina Diéguez
20		Presentació	4d?	01/07/2016	06/07/2016	19	Cristina Diéguez



## 5. Bibliografía

- [1] J. Borges, C. Fernández-Córdoba, J. Pujol, J. Rifà and M. Villanueva, "Z<sub>2</sub>Z<sub>4</sub>-linear codes: generator matrices and duality," *Designs, Codes and Cryptography*, vol. 54, pp. 167-179, 2010.
- [2] C. Fernández-Córdoba, J. Pujol, and M. Villanueva, "Z<sub>2</sub>Z<sub>4</sub>-linear codes: rank and kernel," *Designs, Codes and Cryptography* (July 2010) vol. 56. no. 1, pp. 43-59, ISSN: 0925-1022. DOI: 10.1109/TIT.2011.2119465.

- [3] J. Borges, C. Fernández, J. Pujol, J. Rifà and M. Villanueva, " $\mathbb{Z}_2\mathbb{Z}_4$ -linear codes. A Magma package," Universitat Autònoma de Barcelona, 2007.
- [4] J. Borges, C. Fernández, B. Gastón, J. Pujol, J. Rifà and M. Villanueva, " $\mathbb{Z}_2\mathbb{Z}_4$ -Additive Codes. A MAGMA Package", Univeristat Autònoma de Barcelona, 2009.
- [5] B. Gastón, "Codis  $\mathbb{Z}_2\mathbb{Z}_4$ -Additius en MAGMA", Projecte de final de carrera, Universitat Autònoma de Barcelona, 2008.
- [6] M. Pujol, "Computing the Minimum Hamming Distance for  $\mathbb{Z}_2\mathbb{Z}_4$ -linear codes", Projecte de final de màster, Universitat Autònoma de Barcelona, 2012.
- [7] CCSG Development Group, "CCSG Style Guide", Univeristat Autònoma de Barcelona, 2011.
- [8] CCSG Development Group, "Breu Introducció al MAGMA", Univeristat Autònoma de Barcelona, 2004.
- [9] "Handbook MAGMA", <https://magma.maths.usyd.edu.au/magma/handbook/>
- [10] Computational Algebra Group, "Overview og MAGMA V2.19 Features", University of Sydney, 2016, <https://magma.maths.usyd.edu.au/magma/overview/pdf/overv219.pdf>.